

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

Кваліфікаційна наукова
праця на правах рукопису

Ізонін Іван Вікторович

УДК 004.85:004.32.26:004.032.2](043)

ДИСЕРТАЦІЯ

**АНСАМБЛІ МОДЕЛЕЙ НЕІТЕРАТИВНОГО МАШИННОГО
НАВЧАННЯ ДЛЯ АНАЛІЗУ БІОМЕДИЧНИХ ДАНИХ МАЛИХ
ОБСЯГІВ**

05.13.23 – Системи та засоби штучного інтелекту

Подається на здобуття наукового ступеня доктора технічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело



Іван ІЗОНІН

Науковий консультант:
професор, д.т.н.

Ткаченко Роман Олексійович

Львів – 2025

АНОТАЦІЯ

Ізонін І.В. Ансамблі моделей неітеративного машинного навчання для аналізу біомедичних даних малих обсягів. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.23 – Системи та засоби штучного інтелекту. – Національний університет “Львівська політехніка” Міністерства освіти і науки України, Львів, 2025.

У задачах аналізу структурованих біомедичних даних на вибірках малих обсягів використання окремих моделей машинного навчання, зокрема, традиційних штучних нейронних мереж обмежується через високу ймовірність перенавчання, недостатню репрезентативність вибірок для коректного налаштування параметрів моделі та низьку здатність до узагальнення. Особливо це стосується нелінійних моделей, які мають велику кількість параметрів та потребують значного обсягу інформації для навчання. Використання ансамблів моделей, зокрема, ансамблів ШНМ, може частково вирішити проблему перенавчання завдяки зменшенню дисперсії оцінок і покращенню властивостей узагальнення моделі. Однак синтез ансамблів гетерогенних ШНМ, який вимагає наявності базових моделей з хоча б прийнятною для практики точністю, є дуже складним, а іноді - практично недосяжним, зокрема, у випадку аналізу екстремально коротких вибірок даних. Окрім цього, ітеративні методи навчання існуючих ШНМ є ресурсомісткими, повільними та схильними до залежності від початкової ініціалізації параметрів. Це знижує відтворюваність отриманих рішень і ускладнює їх використання в умовах обмежених обчислювальних можливостей. Сукупно ці фактори обумовлюють необхідність побудови нових, точних, стійких до перенавчання і обчислювально

ефективних ансамблів моделей неітеративного машинного навчання для практичного аналізу структурованих біомедичних вибірок даних малих обсягів.

У дисертаційній роботі на основі виконаних теоретичних та експериментальних досліджень вирішено важливу теоретичну проблему розвитку теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів. При цьому отримано такі основні результати:

уперше розроблено:

- методологію розв'язання задач регресії й бінарної класифікації на основі декартового квадрата множини вхідних компонент векторів, яка містить базовий, адитивний, каскадний методи, а також методи формування комбінованої поверхні відгуку та формування аугментованої вибірки даних необхідного обсягу, що забезпечує підвищення як точності роботи нелінійних моделей машинного навчання, так і їхні властивості узагальнення в умовах навчання на екстремально-коротких вибірках біомедичних даних;
- метод синтезу ансамблю лінійних нейроподібних структур, який завдяки агрегуванню кусково-лінійно апроксимованих поверхонь відгуку кожного його елемента забезпечує підвищення точності моделей регресії та класифікації;
- метод синтезу ансамблів гомогенних нелінійних штучних нейронних мереж, який шляхом агрегування апроксимованих поверхонь відгуку, сформованих на незалежних вибірках із використанням попарних симетричних випадкових зміщень, підвищує точність моделей регресії та класифікації;
- метод часткової компенсації систематичної похибки нейронної мережі узагальненої регресії, який за рахунок використання ансамблю з двох таких нейронних мереж забезпечує підвищення точності розв'язання задач регресії у випадку аналізу біомедичних даних малих обсягів;

- метод нелінійного розширення простору незалежних ознак біомедичної вибірки даних, який внаслідок доповнення їх вихідними сигналами шару підсумовування ймовірнісної нейронної мережі забезпечує підвищення точності класифікації даних на лінійних неітеративних нейроподібних структурах;

- прямий та зворотний методи компенсації похибки моделі регресії, у яких за рахунок перетворень на основі раціональних дробів досягнуто підвищення точності та робастності апроксимації в задачах аналізу біомедичних даних малих обсягів з можливістю додаткового ітеративного уточнення результатів;

удосконалено:

- метод глобально-локальної апроксимації біомедичних даних малих обсягів, який на відміну від наявного використовує незалежне виділення глобальної та локальної компонент сигналу з використанням ансамблю з двох нейронних мереж узагальненої регресії та надає змогу додатково ітеративно уточнити шукане значення;

- метод формування комбінованої поверхні відгуку з застосуванням кластерного аналізу, який на відміну від наявного доповнює простір ознак значенням залежного атрибута центру кластера, що забезпечує підвищення точності моделей регресії та класифікації на вибірках даних малих обсягів без втрати узагальнення;

набули подальшого розвитку:

- теорія синтезу нейроподібних ансамблевих структур із неітеративним машинним навчанням, яка ґрунтується на розроблених методах нелінійного збільшення вимірності простору ознак, аугментації даних, декореляції вибірок та компенсації похибок для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів;

- метод кусково-лінійної апроксимації, який за рахунок ізоморфного подання кожного значення ознаки у вигляді її локальних компонент збільшує

вимірність простору вхідних даних та забезпечує локальну корекцію параметрів багатовимірної поверхні відгуку лінійними неітеративними нейроподібними структурами.

Практична цінність роботи полягає у тому, що розроблені та вдосконалені методи і програмні засоби є ефективними інструментами для інтелектуального аналізу біомедичних даних обмеженого обсягу. Ці методи відзначаються високими показниками точності, робастності до шумів та аномалій, а також обчислювальною ефективністю. Використання розроблених та удосконалених методів забезпечує:

- суттєвий приріст точності лінійних моделей машинного навчання, зберігаючи їх здатність до узагальнення за рахунок незначного збільшення вимірності простору незалежних ознак кожного вектора даних;
- підвищення точності нелінійних моделей машинного навчання до 59% та їх стійкості до спотворень у даних шляхом використання розробленої методології розв'язання задач регресії та бінарної класифікації на основі декартового квадрата множини вхідних компонент векторів;
- підвищення точності на 5–6% в умовах аналізу даних малих обсягів на основі розроблених ансамблів лінійних і нелінійних штучних нейронних мереж з застосуванням різних способів декореляції вхідних даних та агрегування результатів;
- зменшення середньоквадратичної похибки на 4 – 8% залежно від способу агрегування результатів роботи методу часткового усунення систематичної похибки нейронної мережі узагальненої регресії;
- зменшення середньоквадратичної похибки результатів розв'язання задачі регресії на 13% з використанням удосконаленого методу глобально локальної апроксимації та на 9% та 11% відповідно з використанням прямого і зворотного методів компенсації похибок на основі раціональних дробів.

Розроблені програмні засоби, які захищені авторським правом на комп'ютерні програми, забезпечили можливість експериментального оцінювання

ефективності усіх розроблених методів як на модельних, так і на реальних даних біомедичних досліджень. Отримані результати успішно апробовані при розв’язанні низки прикладних задач у різних напрямках біомедичних досліджень, Зокрема, в біоматеріалознавстві, ревматології, травматології, стоматології та трансплантології. Їх впровадження здійснювалося у співпраці з медичними закладами, науково-освітніми установами, а також в межах виконання наукового-дослідних робіт в Україні та міжнародних грантових проєктів за кордоном, що засвідчено відповідними актами впровадження та використання. Зокрема, результати проведених досліджень впроваджено: у Військово-медичному клінічному центрі Західного регіону в період з січня 2024 по січень 2025 року, що підтверджено відповідним Актом від 22.04.2025 року; у Комунальному некомерційному підприємстві Харківської обласної ради “Обласний медичний клінічний центр урології і нефрології ім. В. І. Шаповала” у 2022–2024 роках, що підтверджено відповідним Актом впровадження; на кафедрі стоматології дитячого віку та імплантології Харківського національного медичного університету у 2022–2024 роках, що підтверджено відповідним Актом про впровадження від 04.03.2024 року; у Комунальному некомерційному підприємстві Харківської обласної ради “Обласна клінічна лікарня” у 2022–2024 роках, що підтверджено Актом від 02.12.2024 року; в межах виконання міжнародного наукового гранту EURIZON “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” (№ 871072, 2024–2025), що підтверджено Актом використання наукових результатів №17244/3 від 17.03.2025 року, виданим Українським науково-технологічним центром; в межах виконання НДР в Україні, які фінансувалися Міністерством освіти і науки України і Національним Фондом Досліджень України, що підтверджено відповідними Актами.

Окрім цього, завдяки універсальності закладених підходів, розроблені у дисертації методи можуть бути успішно адаптовані та ефективно використані для аналізу даних у суміжних прикладних галузях, де також спостерігається малий

обсяг вибірки даних для ефективної реалізації методів навчання існуючих нелінійних штучних нейронних мереж.

Ключові слова: біомедичні дані малих обсягів, штучний інтелект, ансамблеве навчання, штучна нейронна мережа, неітеративне навчання, аугментація даних, нелінійне розширення простору вхідних даних, каскадні методи, компенсація похибок моделі регресії, глобально-локальна апроксимація, раціональні дроби.

ABSTRACT

Izonin I.V. Ensembles of non-iterative machine learning models for small biomedical data analysis. – On the rights of the manuscript.

The thesis (dissertation) for the degree of Doctor of Technical Sciences in specialty 05.13.23 – Systems and tools of Artificial Intelligence – Lviv Polytechnic National University of the Ministry of Education and Science of Ukraine, Lviv, 2025.

In tasks involving the analysis of structured biomedical data on small-sample datasets, the use of individual machine learning models, particularly traditional Artificial Neural Networks (ANNs), is constrained by a high risk of overfitting, insufficient sample representativeness for proper parameter tuning, and poor generalization capability. This is especially true for nonlinear models, which contain many parameters and require substantial training data. Employing ensemble of model, notably ANN ensembles, can partially mitigate overfitting by reducing estimate variance and enhancing generalization. However, synthesizing ensembles of heterogeneous ANNs, which demands base models with at least practically acceptable accuracy, is highly complex and often practically unattainable, especially when analyzing extremely-small samples. Moreover, existing ANN training methods are iterative, resource-intensive, slow, and sensitive to initial parameter initialization, which undermines result reproducibility and complicates their use under limited computational resources. Taken together, these factors underscore the need to develop new, accurate, overfitting-resistant, and computationally efficient ensembles of non-iterative machine learning models for the analysis of structured biomedical data on small samples.

In this dissertation, based on the theoretical and experimental studies conducted, the significant theoretical problem of advancing the theory of synthesizing neural-like ensemble structures with non-iterative machine learning algorithms to improve the accuracy of regression and classification tasks on small-sample biomedical data has been solved. The following main results and conclusions have been obtained:

For the first time:

- a methodology has been developed for solving regression and binary classification tasks based on the Cartesian square of the set of input vector components, which includes basic, additive, and cascade methods, as well as methods for forming a combined response surface and for generating an augmented data sample of the required size, and which ensures improvements both in the accuracy of nonlinear machine learning models and in their generalization properties when trained on extremely small biomedical data samples;
- a method for synthesizing an ensemble of linear neural-like structures has been developed that, by aggregating piecewise-linear approximations of each member's response surface, provides increased accuracy for regression and classification models;
- a method for synthesizing ensembles of homogeneous nonlinear artificial neural networks has been developed that, by aggregating approximated response surfaces formed on independent samples using pairwise symmetric random shifts, yields improved accuracy for regression and classification models;
- a method for partial compensation of the systematic error of a General Regression Neural Network has been developed which, through the use of an ensemble of two such networks, enhances regression accuracy when analyzing small-sample biomedical data;
- a method for nonlinear expansion of the feature space of a biomedical data sample has been developed which, by augmenting features with the output signals of the summation layer of a Probabilistic Neural Network, improves classification accuracy on linear non-iterative neural-like structures.
- direct and inverse methods for compensating regression model error have been developed, in which transformations based on rational fractions achieve increased accuracy and robustness of approximation in small-sample biomedical data analysis, with the option for additional iterative refinement of results;

Improved:

- the globally-local approximation method for small-sample biomedical data, which, unlike existing approaches, uses independent extraction of global and local signal components via an ensemble of two generalized regression neural networks, and allows for additional iterative refinement of the estimated value;
- the method for forming a combined response surface using cluster analysis, which, unlike existing methods, augments the feature space with the target value of the cluster centroid, thereby increasing accuracy of regression and classification models on small-sample datasets without loss of generalization;

Further developed:

- the theory of synthesizing neural-like ensemble structures with non-iterative machine learning algorithm, based on the developed methods of nonlinear feature-space expansion, data augmentation, sample decorrelation, and error compensation, for improving accuracy in regression and classification tasks on small-sample biomedical data;
- the piecewise-linear approximation method, which, through an isomorphic representation of each feature value as its local components, expands the input data space and provides local parameter correction of a multidimensional response surface via linear non-iterative neural-like structures.

Practical relevance. The completed doctoral research has substantial scientific value and is aimed at developing the foundations of the theory for synthesizing neural-like ensemble structures with non-iterative machine learning algorithm, focused on the effective analysis of structured biomedical data of small volume. The developed models and methods lay the groundwork for solving current tasks in intelligent data analysis under conditions of information scarcity, high dimensionality, instability, and heterogeneity of data samples, characteristic of modern biomedical engineering, bioinformatics, personalized medicine, and related fields. The research results have been incorporated into separate sections of the lecture courses “Fundamentals of Artificial Intelligence” for

Bachelor's degree students and "Modeling and Data Analysis in Information Systems" for Master's degree students in Computer Science (specialty 122). The implementation is confirmed by the corresponding act of deployment from the Department of Artificial Intelligence Systems at Lviv Polytechnic National University.

The practical value of the work lies in the fact that the developed and refined methods and software tools are effective instruments for intelligent analysis of limited-volume biomedical data. These methods are characterized by high accuracy, robustness to noise and anomalies, and computational efficiency. Use of the developed and refined methods ensures:

- a significant increase in the accuracy of linear machine learning models while preserving their generalization capability through only a slight increase in the dimensionality of each data vector's feature space;
- improvement of nonlinear machine learning model accuracy by up to 59% and their resistance to data distortions through the methodology developed for solving regression and binary classification problems based on the Cartesian square of the set of input vector components;
- an increase in accuracy by 5 – 6% under small-sample conditions using the developed ensembles of linear and nonlinear artificial neural networks with various data-decorrelation and output-aggregation techniques;
- a reduction in mean squared error by 4 – 8%, depending on the aggregation method, using the partial systematic-error compensation method for General Regression Neural Networks;
- a reduction in mean squared error of regression results by 13% using the improved globally-local approximation method, and by 9% and 11%, respectively, using the direct and inverse error-compensation methods based on rational fractions.

The developed software tools, protected by copyright for computer programs, have enabled experimental evaluation of all developed methods on both synthetic and real

biomedical research data. The results have been successfully validated in solving a range of applied tasks in various biomedical research areas, including biomaterials science, rheumatology, traumatology, dentistry, and transplantology. Their implementation was carried out in collaboration with medical institutions and research organizations, as well as within the framework of research projects in Ukraine and international grant projects abroad, as evidenced by the corresponding acts of deployment and use. In particular, the research results have been implemented: at the Western Region Military Medical Clinical Center from January 2024 to January 2025, confirmed by the Act dated 22.04.2025; at the V. I. Shapoval Regional Medical Clinical Center of Urology and Nephrology of the Kharkiv Regional Council in 2022–2024, confirmed by the corresponding Act of Deployment; at the Department of Pediatric Dentistry and Implantology at Kharkiv National Medical University in 2022–2024, confirmed by the Act of Deployment dated 04.03.2024; at the Regional Clinical Hospital of the Kharkiv Regional Council in 2022–2024, confirmed by the Act dated 02.12.2024; under the international scientific grant EURIZON “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” (No. 871072, 2024–2025), confirmed by the Act of Use of Scientific Results No. 17244/3 dated 17.03.2025, issued by the Ukrainian Science and Technology Center; within the framework of research and development in Ukraine funded by the Ministry of Education and Science of Ukraine as well as the National Research Foundation of Ukraine, confirmed by the respective Acts.

In addition, due to the universality of the underlying approaches, the methods developed in this dissertation can be successfully adapted and effectively employed for data analysis in related applied fields where sample-size deficiency also hinders effective application of existing nonlinear ANNs.

Keywords: small-scale biomedical data, artificial intelligence, ensemble learning, artificial neural network, non-iterative training, data augmentation, nonlinear input space expansion, cascade methods, regression error compensation, global-local approximation.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Статті у наукових виданнях, включених до Переліку наукових фахових видань України

1. Izonin I. V. An unsupervised-supervised ensemble technology with non-iterative training algorithm for small biomedical data analysis // *Computer systems and information technologies*. 2023. No. 4. P. 67–74. DOI: 10.31891/csit-2023-4-9. (фахове видання України, категорія Б).
2. Tkachenko R. O., Izonin I. V., Danylyk V. M., Mykhalevych V. Yu. Stacking of the SGTM neural-like structure with РБФ layer based on generation of a random curtain of its hyperparameters for prediction tasks // *Ukrainian Journal of Information Technology*. 2021. Vol. 3, No. 1. P. 49–55. DOI: 10.23939/ujit2021.03.049. (фахове видання України, категорія Б).
3. Izonin I. V. An ensemble method for the regression model parameter adjustments: direct approach // *Mathematical and computer modelling. Series: Technical sciences*. 2023. Vol. 24. P. 35–44. DOI: 10.32626/2308-5916.2023-24.35-44. (фахове видання України, категорія Б).
4. Вітинський П. В., Ткаченко Р. О., Ізонін І. В., Кустра Н. О. Ансамблі нейроподібних структур МПГП з РБФ розширенням входів для задач регресії та класифікації // *Вісник Хмельницького національного університету. Серія: Технічні науки*. 2019. № 4(275). С. 72–79. (фахове видання України, категорія Б).
5. Izonin I. V., Tkachenko R. O., Semchyshyn O. L. An ensemble method for the analysis of small biomedical data based on a neural network without training // *Elektronnoe modelirovanie*. 2023. Vol. 45, No. 6. P. 65–76. DOI: 10.15407/emodel.45.06.065. (фахове видання України, категорія Б).
6. Вітинський П. В., Ткаченко Р. О., Ізонін І. В. Ансамбль мереж GRNN для розв’язання задач регресії з підвищеною точністю // *Науковий вісник НЛТУ*

України. 2019. Т. 29, № 8. Стаття № 8. DOI: 10.36930/40290822.
(фахове видання України, категорія Б).

Статті у закордонних виданнях, проіндексованих у базі даних Scopus

7. Tkachenko R., Izonin I., Vitynskyi P., Lotoshynska N., Pavlyuk O. Development of the non-iterative supervised learning predictor based on the Ito decomposition and SGTm neural-like structure for managing medical insurance costs // *Data*. 2018. Vol. 3, No. 4. P. 46. DOI: 10.3390/data3040046.
(Індексується у Scopus).

8. Izonin I., Tkachenko R., Gregus M., Duriagina Z., Shakhovska N. PNN-SVM approach of Ti-based powder's properties evaluation for biomedical implants production // *Computers, Materials & Continua*. 2022. Vol. 71, No. 3. P. 5933–5947. DOI: 10.32604/cmc.2022.022582.
(Індексується у Scopus, кuartиль Q2).

9. Izonin I., Tkachenko R., Shakhovska N., Ilchyshyn B., Singh K. K. A two-step data normalization approach for improving classification accuracy in the medical diagnosis domain // *Mathematics*. 2022. Vol. 10, No. 11. P. 1942. DOI: 10.3390/math10111942.
(Індексується у Scopus, кuartиль Q2).

10. Izonin I., Tkachenko R., Duriagina Z., Shakhovska N., Kovtun V., Lotoshynska N. Smart web service of Ti-based alloy's quality evaluation for medical implants manufacturing // *Applied Sciences*. 2022. Vol. 12, No. 10. P. 5238. DOI: 10.3390/app12105238.
(Індексується у Scopus, кuartиль Q2).

11. Izonin I., Tkachenko R., Dronyuk I., Tkachenko P., Gregus M., Rashkevych M. Predictive modeling based on small data in clinical medicine: PBF-based additive input-doubling method // *Mathematical Biosciences and Engineering*. 2021. Vol. 18, No.

3. P. 2599–2613. DOI: 10.3934/mbe.2021132.
(Індексується у Scopus, кuartиль Q2).
12. Izonin I., Tkachenko R., Shakhovska N., Lotoshynska N. The additive input-doubling method based on the SVR with nonlinear kernels: small data approach // *Symmetry*. 2021. Vol. 13, No. 4. P. 612. DOI: 10.3390/sym13040612.
(Індексується у Scopus, кuartиль Q2).
13. Izonin I., Tkachenko R., Berezsky O., Krak I., Kováč M., Fedorchuk M. Improvement of the ANN-based prediction technology for extremely small biomedical data analysis // *Technologies*. 2024. Vol. 12, No. 7. P. 112. DOI: 10.3390/technologies12070112.
(Індексується у Scopus, кuartиль Q1).
14. Izonin I., Tkachenko R., Yendyk P., Pliss I., Bodyanskiy Y., Gregus M. Enhanced input-doubling method leveraging response surface linearization to improve classification accuracy in small medical data processing // *Computation*. 2024. Vol. 12, No. 10. P. 203. DOI: 10.3390/computation12100203.
(Індексується у Scopus, кuartиль Q2).
15. Izonin I., Tkachenko R., Hovdysh N., Berezsky O., Yemets K., Tsmots I. Cascade-based input-doubling classifier for predicting survival in allogeneic bone marrow transplants: small data case // *Computation*. 2025. Vol. 13, No. 4. P. 80. DOI: 10.3390/computation13040080.
(Індексується у Scopus, кuartиль Q2).
16. Izonin I., Tkachenko R., Bliakhar R., Kovac M., Bodyanskiy Y., Chala O. An improved ANN-based global-local approximation for small medical data analysis // *EAI Endorsed Transactions on Pervasive Health and Technology*. 2023. Vol. 9. DOI: 10.4108/eetpht.9.3320 .
(Індексується у Scopus, кuartиль Q2).
17. Izonin I., Tkachenko R., Yemets K., Havryliuk M. An interpretable ensemble structure with a non-iterative training algorithm to improve the predictive accuracy of

healthcare data analysis // *Scientific Reports*. 2024. Vol. 14, No. 1. P. 12947. DOI: 10.1038/s41598-024-61776-y.

(Індексується у Scopus, кuartиль Q1).

18. Izonin I., Muzyka R., Tkachenko R., Dronyuk I., Yemets K., Mitoulis S.-A. A method for reducing training time of ML-based cascade scheme for large-volume data analysis // *Sensors*. 2024. Vol. 24, No. 15. P. 4762. DOI: 10.3390/s24154762.

(Індексується у Scopus, кuartиль Q1).

19. Izonin I., Tkachenko R., Krak I., Berezsky O., Shevchuk I., Shandilya S. K. A cascade ensemble-learning model for the deployment at the edge: case on missing IoT data recovery in environmental monitoring systems // *Frontiers in Environmental Science*. 2023. Vol. 11. P. 1295526. DOI: 10.3389/fenvs.2023.1295526.

(Індексується у Scopus, кuartиль Q2).

20. Izonin I., Tkachenko R., Gurbych O., Kovac M., Rutkowski L., Holoven R. A non-linear SVR-based cascade model for improving prediction accuracy of biomedical data analysis // *Mathematical Biosciences and Engineering*. 2023. Vol. 20, No. 7. P. 13398–13414. DOI: 10.3934/mbe.2023597.

(Індексується у Scopus, кuartиль Q2).

21. Izonin I., Tkachenko R., Verhun V., Zub K. An approach towards missing data management using improved GRNN-SGTM ensemble method // *Engineering Science and Technology, an International Journal*. 2021. Vol. 24, No. 3. P. 749–759. DOI: 10.1016/j.jestch.2020.10.005.

(Індексується у Scopus, кuartиль Q1).

22. Tkachenko R., Izonin I., Kryvinska N., Dronyuk I., Zub K. An approach towards increasing prediction accuracy for the recovery of missing IoT data based on the GRNN-SGTM ensemble // *Sensors*. 2020. Vol. 20, No. 9. Art. No. 9. DOI: 10.3390/s20092625.

(Індексується у Scopus, кuartиль Q2).

Розділи у колективних монографіях

23. Izonin I., Tkachenko R. Universal intraensemble method using nonlinear AI techniques for regression modeling of small medical data sets // *Cognitive and Soft Computing Techniques for the Analysis of Healthcare Data*. (Collective monograph). Amsterdam: Elsevier, 2022. P. 123–150. DOI: 10.1016/B978-0-323-85751-2.00002-5. (Індексується у Scopus).

24. Ізонін І. В., Ткаченко Р. О., Зуб Х. В. Застосування бустингу в ансамблях нейронних мереж узагальненої регресії для підвищення точності розв’язання задач апроксимації // *На шляху до індустрії 4.0: інформаційні технології, моделювання, штучний інтелект, автоматизація*. (Колективна монографія). Одеса: Астропринт, 2021. С. 294–409.

Публікації, які засвідчують апробацію матеріалів дисертації

25. Ізонін І. В., Ільчишин Б., Гаврилюк М., Бойчук О. Ансамбль однорідних методів машинного навчання з усередненням результату // *Інформаційні технології і автоматизація – 2023: матеріали XVI Міжнародної науково-практичної конференції*. Одеса, жовтень 2023. С. 339–340.

26. Ізонін І. В., Бейда Я. Р., Федорчук М. А., Машталір Д. Б., Підкостельний Р. Р. Інтелектуальний аналіз коротких вибірок даних: методи аугментації даних // *Економічний і соціальний розвиток України в XXI столітті: національна візія та виклики глобалізації: збірник тез доповідей XX Міжнародної науково-практичної конференції молодих вчених*. Тернопіль, травень 2023. С. 705–707.

27. Ізонін І. В., Бляхар Р. С., Ткаченко Р. О. Нейромережевий метод послідовної глобально локальної апроксимації коротких наборів медичних даних // *Нейромережні технології та їх застосування: збірник наукових праць XXII Міжнародної наукової конференції НМТЗ-2023*. Краматорськ–Вінниця–Тернопіль, 2023. С. 87–89.

28. Izonin I., Tkachenko R., Greguš M. I-PNN: An Improved Probabilistic Neural Network for Binary Classification of Imbalanced Medical Data // In: Strauss C., Cuzzocrea A., Kotsis G., Tjoa A.M., Khalil I. (eds) Database and Expert Systems Applications. DEXA 2022. Lecture Notes in Computer Science, vol 13427. Cham: Springer, 2022. P. 147–157. DOI: 10.1007/978-3-031-12426-6_12. (*Індексується у Scopus; квартиль Q3*).

29. Izonin I., Tkachenko R., Ryvak L., Zub K., Rashkevych M., Pavliuk O. Addressing Medical Diagnostics Issues: Essential Aspects of the PNN-based Approach // Proceedings of the 3rd International Conference on Informatics & Data-Driven Medicine (IDDM 2020). CEUR Workshop Proceedings, vol. 2753. Växjö, Sweden, 19–21 November 2020. P. 209–218. (*Індексується у Scopus*).

30. Izonin I., Tkachenko R., Sydor M., Chretien S., Pliss I., Vovk T. Modification of Combined Unsupervised-Supervised Cascade Scheme for Small Biomedical Data Classification // Proceedings of the 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT). Lviv, Ukraine: IEEE, September 2023. P. 107–112. DOI: 10.1109/ELIT61488.2023.10310830. (*Індексується у Scopus*).

31. Izonin I., Tkachenko R., Shakhovska N., Ilchyshyn B., Gregus M., Strauss C. Towards Data Normalization Task for the Efficient Mining of Medical Data // Proceedings of the 12th International Conference on Advanced Computer Information Technologies (ACIT). Ruzomberok, Slovakia: IEEE, 2022. P. 480–484. DOI: 10.1109/ACIT54803.2022.9913112. (*Індексується у Scopus*).

32. Izonin I., Tepla T., Danylyuk D., Tkachenko R., Duriagina Z., Lemishka I. Towards an Intelligent Decision Making of Ti-based Powders Selection for Medical Manufacturing // Proceedings of the 2020 International Conference on Decision Aid

Sciences and Application (DASA). Sakheer, Bahrain: IEEE, November 2020. P. 1–5. DOI: 10.1109/DASA51403.2020.9316866.

(Індексується у Scopus).

33. Izonin I., Tkachenko R., Kryvinska N., Tkachenko P., Greguš M. Multiple Linear Regression Based on Coefficients Identification Using Non-iterative SGTM Neural-like Structure // In: Rojas I., Joya G., Catala A. (eds) Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science, vol 11506. Cham: Springer, 2019. P. 467–479. DOI: 10.1007/978-3-030-20521-8_39. *(Індексується у Scopus, квартиль Q2).*

34. Izonin I., Kryvinska N., Tkachenko R., Zub K., Vitynskyi P. An Extended-Input GRNN and its Application // Procedia Computer Science, vol. 160. Proceedings of the 10th EUSPN and 9th ICTH Conferences. Coimbra, Portugal: Elsevier, November 2019. P. 578–583. DOI: 10.1016/j.procs.2019.11.044.

(Індексується у Scopus).

35. Tkachenko R., Izonin I. Model and Principles for the Implementation of Neural-Like Structures Based on Geometric Data Transformations // In: Hu Z., Petoukhov S., Dychka I., He M. (eds) Advances in Computer Science for Engineering and Education. ICCSEE 2019. Advances in Intelligent Systems and Computing, vol 754. Cham: Springer, 2019. P. 578–587. DOI: 10.1007/978-3-319-91008-6_58.

(Індексується у Scopus, квартиль Q3).

36. Izonin I., Tkachenko R., Fedushko S., Koziy D., Zub K., Vovk O. ПФФ-Based Input Doubling Method for Small Medical Data Processing // In: Hu Z., Zhang Q., Petoukhov S., He M. (eds) Advances in Artificial Systems for Logistics Engineering. Lecture Notes on Data Engineering and Communications Technologies, vol 82. Cham: Springer, 2021. P. 23–31. DOI: 10.1007/978-3-030-80475-6_3.

(Індексується у Scopus).

37. Izonin I., Tkachenko R., Gregus M., Zub K., Lotoshynska N. Input Doubling Method based on SVR with ПФФ kernel in Clinical Practice: Focus on Small Data //

Procedia Computer Science, vol. 184. Warsaw, Poland: Elsevier, March 2021. P. 606–613. DOI: 10.1016/j.procs.2021.03.075.

(Індексується у Scopus).

38. Izonin I., Tkachenko R., Gregus M., Zub K., Tkachenko P. A GRNN-based Approach towards Prediction from Small Datasets in Medical Application // Procedia Computer Science, vol. 184. Warsaw, Poland: Elsevier, March 2021. P. 242–249. DOI: 10.1016/j.procs.2021.03.033.

(Індексується у Scopus).

39. Izonin I., Tkachenko R., Pidkostelnyi R., Pavliuk O., Khavalko V., Batyuk A. Experimental Evaluation of the Effectiveness of ANN-based Numerical Data Augmentation Methods for Diagnostics Tasks // Proceedings of the 4th International Conference on Informatics & Data-Driven Medicine (IDDM 2021). CEUR Workshop Proceedings, vol. 3038. Valencia, Spain, November 2021. P. 223–232.

(Індексується у Scopus).

40. Izonin I., Tkachenko R., Horbal N., Greguš M., Verhun V., Tolstyak Y. An Approach Toward Numerical Data Augmentation and Regression Modeling Using Polynomial-Kernel-Based SVR // In: Saraswat M., Roy S., Chowdhury C., Gandomi A.H. (eds) Proceedings of International Conference on Data Science and Applications. Lecture Notes in Networks and Systems, vol 288. Singapore: Springer, 2022. P. 771–781. DOI: 10.1007/978-981-16-5120-5_58.

(Індексується у Scopus, кuartиль Q4).

41. Izonin I., Tkachenko R., Yemets K., Gregus M., Tomashy Y., Pliss I. An Approach Towards Reducing Training Time of the Input Doubling Method via Clustering for Middle-Sized Data Analysis // Procedia Computer Science, vol. 241. Huntington, WV, USA: Elsevier, 2024. P. 32–39. DOI: 10.1016/j.procs.2024.08.007.

(Індексується у Scopus).

42. Izonin I., Tkachenko R., Trush B., Nahirnyi S. Homogeneous ensemble models based on ANNs with noniterative training: classification and regression tasks //

Book of Abstracts of Workshops RECI 2022 and ACeSYRI 2022. Delft, Netherlands, 15 November 2022. P. 27. URL: <https://ki.fri.uniza.sk/RECI2022/Abstracts.pdf>.

43. Izonin I., Tkachenko R., Vitynskyi P., Zub K., Tkachenko P., Dronyuk I. Stacking-based GRNN-SGTM Ensemble Model for Prediction Tasks // Proceedings of the 2020 International Conference on Decision Aid Sciences and Application (DASA). Sakheer, Bahrain: IEEE, November 2020. P. 326–330. DOI: 10.1109/DASA51403.2020.9317124.

(Індексується у Scopus).

44. Tkachenko R., Izonin I., Greguš M., Tkachenko P., Dronyuk I. Committee of the SGTM Neural-Like Structures with Extended Inputs for Predictive Analytics in Insurance // In: Younas M., Awan I., Benbernou S. (eds) Big Data Innovations and Applications. Communications in Computer and Information Science, vol 1054. Cham: Springer, 2019. P. 121–132. DOI: 10.1007/978-3-030-27355-2_9.

(Індексується у Scopus, квартиль Q3).

45. Tkachenko R., Tkachenko P., Izonin I., Vitynskyi P., Kryvinska N., Tsymbal Y. Committee of the Combined ПФ-SGTM Neural-Like Structures for Prediction Tasks // In: Awan I., Younas M., Ünal P., Aleksy M. (eds) Mobile Web and Intelligent Information Systems. Lecture Notes in Computer Science, vol 11673. Cham: Springer, 2019. P. 267–277. DOI: 10.1007/978-3-030-27192-3_21.

(Індексується у Scopus, квартиль Q2).

46. Izonin I., Tkachenko R., Kryvinska N., Gregus M., Tkachenko P., Vitynskyi P. Committee of SGTM Neural-Like Structures with ПФ kernel for Insurance Cost Prediction Task // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON). Lviv, Ukraine: IEEE, July 2019. P. 1037–1040. DOI: 10.1109/UKRCON.2019.8879905.

(Індексується у Scopus).

47. Tkachenko R., Izonin I., Kryvinska N., Chopyak V., Lotoshynska N., Danylyuk D. Piecewise-linear Approach for Medical Insurance Costs Prediction using

SGTM Neural-Like Structure // Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018). CEUR Workshop Proceedings, vol. 2255. Lviv, Ukraine, November 2018. P. 170–179.

(Індексується у Scopus).

48. Izonin I., Tkachenko R., Shcherbii O., Berezhsky O., Krak I., Oliinyk M. An Approximation Cascade Scheme via Rational Fractions for Biomedical Data Analysis // Proceedings of the 2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT). Lviv, Ukraine: IEEE, October 2023. P. 1–4. DOI: 10.1109/CSIT61576.2023.10324122.

(Індексується у Scopus).

49. Izonin I., Tkachenko R., Shcherbii O., Muzyka R., Faramarzi A., Mitoulis S.A. Enhanced Cascade Schemes for Advancing Machine Learning-Based Prediction of Heating and Cooling Loads in Residential Buildings // Proceedings of the 2024 Mediterranean Smart Cities Conference (MSCC). Martil–Tetuan, Morocco: IEEE, May 2024. P. 1–5. DOI: 10.1109/MSCC62288.2024.10697024.

(Індексується у Scopus).

50. Izonin I., Tkachenko R. An approach towards the response surface linearization via ANN-based cascade scheme for regression modeling in Healthcare // Procedia Computer Science, vol. 198. Leuven, Belgium: Elsevier, 2022. P. 724–729. DOI: 10.1016/j.procs.2021.12.313.

(Індексується у Scopus).

Публікації, які додатково відображають наукові результати дисертації

51. Ізонін І. В., Ткаченко Р. О., Томаши Є. О., Гентош Л. І. Бібліотека методів інтелектуального аналізу малих даних на основі декартового добутку впорядкованих множин входів: комп'ютерна програма. Свідectvo про реєстрацію авторського права № 127079. Зареєстровано 4 червня 2024 р. – Заявник: Ізонін І. В. – Номер заявки: c202403578; дата подання: 23 квітня 2024 р.

52. Вітинський П. Б., Ткаченко Р. О., Ізонін І. В. Комбінована нейроподібна система для задач аналітики в умовах коротких вибірок даних: комп'ютерна програма. Свідоцтво про реєстрацію авторського права № 93094. Зареєстровано 16 жовтня 2019 р. – Заявник: Вітинський П. Б., Ткаченко Р. О., Ізонін І. В. – Номер заявки: 94679; дата подання: 25 вересня 2019 р.

53. Ізонін І. В., Нагірний С. Б., Ткаченко Р. О. Гомогенні стекінгові ансамблі штучних нейронних мереж без навчання (Multi-GRNN): комп'ютерна програма. Свідоцтво про реєстрацію авторського права № 114033. Зареєстровано 9 серпня 2022 р. – Заявник: Ізонін І. В. – Номер заявки: с202203246; дата подання: 25 липня 2022 р.

ЗМІСТ

<u>АНОТАЦІЯ</u>	<u>2</u>
<u>ABSTRACT</u>	<u>8</u>
<u>СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА</u>	<u>13</u>
<u>ЗМІСТ</u>	<u>24</u>
<u>ПЕРЕЛІК ПРИЙНЯТИХ СКОРОЧЕНЬ</u>	<u>29</u>
<u>ВСТУП.....</u>	<u>30</u>
<u>РОЗДІЛ 1</u>	<u>47</u>
<u>АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ІНТЕЛЕКТУАЛЬНОГО ОПРАЦЮВАННЯ</u> <u>БІОМЕДИЧНИХ ДАНИХ МАЛИХ ОБСЯГІВ.....</u>	<u>47</u>
1.1 Поняття біомедичних даних та їх особливості	48
1.2 Визначення малих даних, їх різновиди.....	51
1.3 Проблема інтелектуального аналізу біомедичних даних	56
МАЛИХ ОБСЯГІВ.....	56
1.4 Аналіз моделей та методів для розв’язання задач регресії/класифікації ..	60
1.5 Особливості синтезу ансамблів ШНМ для розв’язання задач регресії/класифікації.....	76
1.6 Формулювання мети і постановка задач дослідження.....	83
1.7. Висновки до розділу	84
<u>РОЗДІЛ 2</u>	<u>88</u>
<u>МЕТОДИ НЕЛІНІЙНИХ ПЕРЕТВОРЕНЬ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ</u> <u>ДЛЯ ФОРМУВАННЯ ПРОСТОРІВ ОЗНАК ВИЩОЇ ВИМІРНОСТІ</u>	<u>88</u>
2.1 Розширення простору вхідних компонент векторів даних для нейроподібних структур МПГП	88
2.2 Доповнення входів нейроподібних структур МПГП нелінійними членами поліному Колмогорова-Габора.....	90
2.3 Метод розширення входів нейроподібних засобів на основі вихідних сигналів шару сумування ймовірнісної нейронної мережі.....	97

2.4	МЕТОД ФОРМУВАННЯ КОМБІНОВАНОЇ ПОВЕРХНІ ВІДГУКУ НА ОСНОВІ ВИДІЛЕНИХ КЛАСТЕРІВ	107
2.5	МЕТОД КУСКОВО-ЛІНІЙНОЇ АПРОКСИМАЦІЇ БАГАТОВИМІРНИХ ПОВЕРХОНЬ ВІДГУКУ	117
2.6.	Висновки до розділу	121
<u>РОЗДІЛ 3</u>		<u>124</u>
<u>МЕТОДОЛОГІЯ АУГМЕНТАЦІЇ ДАНИХ НА ОСНОВІ ДЕКАРТОВОГО КВАДРАТА МНОЖИНИ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ</u>		<u>124</u>
3.1	АУГМЕНТАЦІЯ ДАНИХ НА ОСНОВІ ДЕКАРТОВОГО КВАДРАТА МНОЖИНИ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ	125
3.2	МЕТОД РОЗВ'ЯЗАННЯ ЗАДАЧ РЕГРЕСІЇ ТА КЛАСИФІКАЦІЇ НА ОСНОВІ ДЕКАРТОВОГО КВАДРАТА МНОЖИНИ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ.....	129
3.3	АДИТИВНИЙ МЕТОД РОЗВ'ЯЗАННЯ ЗАДАЧІ РЕГРЕСІЇ З ЗАСТОСУВАННЯМ ДЕКАРТОВОГО КВАДРАТА МНОЖИНИ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ.....	133
3.4	МЕТОД РОЗВ'ЯЗАННЯ ЗАДАЧ РЕГРЕСІЇ ТА КЛАСИФІКАЦІЇ З ВИКОРИСТАННЯМ ПРИНЦИПІВ КАСКАДУВАННЯ.....	137
3.5	МЕТОД РОЗВ'ЯЗАННЯ ЗАДАЧ РЕГРЕСІЇ ТА КЛАСИФІКАЦІЇ НА ОСНОВІ ФОРМУВАННЯ КОМБІНОВАНОЇ ПОВЕРХНІ ВІДГУКУ	143
3.6	МЕТОД ФОРМУВАННЯ АУГМЕНТОВАНОЇ ВИБІРКИ ДАНИХ НЕОБХІДНОГО ОБСЯГУ ЗА РАХУНОК КЛАСТЕРИЗАЦІЇ.....	149
3.7	Висновки до розділу	154
<u>РОЗДІЛ 4</u>		<u>157</u>
<u>МЕТОДИ СИНТЕЗУ АНСАМБЛІВ ШНМ НА ОСНОВІ КОМБІНАЦІЙ АПРОКСИМОВАНИХ ПОВЕРХОНЬ ВІДГУКУ</u>		<u>157</u>
4.1	КОМБІНУВАННЯ АПРОКСИМОВАНИХ ПОВЕРХОНЬ ВІДГУКУ ДЛЯ ПІДВИЩЕННЯ ТОЧНОСТІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ МАЛИХ ОБСЯГІВ	158
4.2	АНСАМБЛЬ ЛІНІЙНИХ ОДНОРІДНИХ ШНМ З ВИКОРИСТАННЯМ МЕТОДУ КУСКОВО-ЛІНІЙНОЇ АПРОКСИМАЦІЇ БАГАТОВИМІРНОЇ ПОВЕРХНІ ВІДГУКУ	165
4.3	АНСАМБЛЬ ОДНОРІДНИХ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ НА ОСНОВІ МЕТОДУ ВИПАДКОВИХ ЗМІЩЕНЬ.....	173

4.4 АНСАМБЛЬ НЕІТЕРАТИВНИХ НЕЙРОПОДІБНИХ СТРУКТУР З РБФ-РОЗШИРЕННЯМ ВХОДІВ НА ОСНОВІ ВИПАДКОВИХ СИМЕТРИЧНИХ ЗМІЩЕНЬ	180
4.5 Ієрархічні методи розв’язання задач регресії і класифікації з додатковою кластеризацією даних	190
4.6 Висновки до розділу	194
<u>РОЗДІЛ 5</u>	<u>197</u>
<u>МЕТОДИ ЧАСТКОВОЇ КОМПЕНСАЦІЇ ПОХИБОК ШНМ НА ОСНОВІ</u>	
<u>КАСКАДНИХ АНСАМБЛІВ В ЗАДАЧАХ РЕГРЕСІЇ.....</u>	<u>197</u>
5.1 Компенсація похибок штучної нейронної мережі для задач регресії	198
5.2 Метод часткової компенсації похибки роботи нейронної мережі узагальненої регресії	201
5.3 Удосконалений метод розв’язання задач регресії/класифікації на основі виділення глобальної та локальної компонент поверхні відгуку	208
5.4 Методи компенсації похибки моделі регресії з використанням раціональних дробів	213
5.4.1. Прямий метод компенсації похибки моделі регресії.....	215
5.4.2. Зворотний метод компенсації похибки моделі регресії.....	220
5.5 Висновки до розділу	226
<u>РОЗДІЛ 6</u>	<u>228</u>
<u>ВЕРИФІКАЦІЯ РОЗРОБЛЕНИХ АНСАМБЛЕВИХ МЕТОДІВ НА МОДЕЛЬНИХ</u>	
<u>ДАНИХ.....</u>	<u>228</u>
6.1 Експериментальне дослідження методів нелінійного перетворення вхідних компонент векторів для формування простору ознак вищої вимірності.....	229
6.1.1 Оцінювання результатів методу нелінійного доповнення входів нейроподібних структур МПГП членами полінома Колмогорова-Габора	229
6.1.2 Оцінювання результатів методу класифікації даних на основі формування комбінованої поверхні відгуку в області кардіології	234
6.2 Експериментальне дослідження методів розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів	239

6.2.1 Оцінювання результатів базового методу розв'язання задачі регресії на основі декартового квадрата множини входних компонент векторів	239
6.2.2 Оцінювання результатів адитивного методу розв'язання задачі регресії в галузі ревматології	242
6.2.3 Оцінювання результатів методу класифікації даних з використанням принципів каскадування в області трансплантології	248
6.2.4 Оцінювання результатів методу формування аугментованої вибірки даних обмеженого обсягу шляхом кластеризації	255
6.3 Експериментальне дослідження методів часткової компенсації похибок роботи моделей машинного навчання на основі каскадних ансамблів в задачі регресії	259
6.3.1 Оцінювання результатів методу часткової компенсації похибки роботи нейронної мережі узагальненої регресії	259
6.3.2 Оцінювання результатів методу глобально локальної апроксимації даних малих обсягів в галузі біомедичної інформатики	262
6.3.3 Оцінювання результатів методів компенсації похибок моделі регресії з використанням раціональних дробів	265
6.4 Висновки до розділу	267
<u>РОЗДІЛ 7</u>	<u>272</u>
<u>ТЕСТУВАННЯ І РЕЗУЛЬТАТИ ВПРОВАДЖЕННЯ РОЗРОБЛЕНИХ МЕТОДІВ НА ДАНИХ БІОМЕДИЧНИХ ДОСЛІДЖЕНЬ.....</u>	<u>272</u>
7.1 Апробація методу розширення входів лінійної нейроподібної структури з використанням вихідних сигналів ймовірнісної нейронної мережі в галузі біоматеріалознавства	273
7.2 Апробація методу кусково-лінійної апроксимації багатовимірних поверхонь відгуку в галузі ревматології	279
7.3 Апробація методу розв'язання задачі регресії на основі декартового квадрата множини входних компонент векторів з використанням сформованої комбінованої поверхні відгуку в галузі медичної стоматології.....	284

7.4 АПРОБАЦІЯ МЕТОДУ КЛАСИФІКАЦІЇ ДАНИХ НА ОСНОВІ АНСАМБЛЮ ЙМОВІРНІСНИХ НЕЙРОННИХ МЕРЕЖ З ВИКОРИСТАННЯМ СИМЕТРИЧНИХ ВИПАДКОВИХ ЗМІЩЕНЬ У ГАЛУЗІ ТРАВМАТОЛОГІЇ.....	291
7.5 АПРОБАЦІЯ МЕТОДУ КЛАСИФІКАЦІЇ ДАНИХ НА ОСНОВІ АНСАМБЛЮ НЕІТЕРАТИВНИХ НЕЙРОПОДІБНИХ СТРУКТУР З РБФ-РОЗШИРЕННЯМ ВХОДІВ З ВИКОРИСТАННЯМ СИМЕТРИЧНИХ ВИПАДКОВИХ ЗМІЩЕНЬ У ГАЛУЗІ ТРАНСПЛАНТОЛОГІЇ.....	297
7.6 ЕТИЧНІ АСПЕКТИ ЗАСТОСУВАННЯ РОЗРОБЛЕНИХ ТА УДОСКОНАЛЕНИХ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ БІОМЕДИЧНИХ ДАНИХ МАЛИХ ОБСЯГІВ	302
7.7 Висновки до розділу	303
<u>ВИСНОВКИ.....</u>	<u>307</u>
<u>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</u>	<u>312</u>
<u>ДОДАТКИ.....</u>	<u>352</u>
ДОДАТОК А	352
МЕТРИКИ ОЦІНЮВАННЯ ТОЧНОСТІ РОЗВ’ЯЗАННЯ ЗАДАЧ РЕГРЕСІЇ ТА КЛАСИФІКАЦІЇ МЕТОДАМИ МАШИННОГО НАВЧАННЯ	352
ДОДАТОК Б.....	354
ОПТИМАЛЬНІ ПАРАМЕТРИ РЕАЛІЗАЦІЇ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ РОЗВ’ЯЗАННЯ ЗАДАЧ РЕГРЕСІЇ ТА КЛАСИФІКАЦІЇ.....	354
ДОДАТОК В.....	357
КРИТЕРІЇ ЗУПИНКИ ІТЕРАЦІЙНИХ ПРОЦЕДУР МЕТОДІВ КОМПЕНСАЦІЇ ПОХИБКИ МОДЕЛІ РЕГРЕСІЇ.....	357
ДОДАТОК Г.....	359
АКТИ ВПРОВАДЖЕННЯ ТА ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ РОБОТИ.....	359
ДОДАТОК Д.....	369
СВІДОЦТВА ПРО РЕЄСТРАЦІЮ АВТОРСЬКОГО ПРАВА НА КОМП’ЮТЕРНІ ПРОГРАМИ	369

ПЕРЕЛІК ПРИЙНЯТИХ СКОРОЧЕНЬ

ШНМ – Штучна нейронна мережа

МН – Моделі машинного навчання

ЛНПС – Лінійна нейроподібна структура

МППП – Модель послідовних геометричних перетворень

РБФ – Радіально-базисна функція

НМУР – Нейронна мережа узагальненої регресії

ЙНМ – Ймовірнісна нейронна мережа

SMOTE – Synthetic Minority Oversampling Technique

ГЗМ – Генеративно-змагальна мережа

ВА – Варіаційний автоенкодер

k-means – Метод k-середніх

БШП – Багатошаровий перцептрон

РБФ-мережа – Нейронна мережа на радіально-базисних функціях

ВСТУП

Актуальність теми. Актуальність теми дослідження зумовлена зростаючою потребою в ефективних та стійких методах аналізу біомедичних даних, які в реальних умовах часто подано у вигляді вибірок даних малого обсягу. Останні характеризуються високою вимірністю, гетерогенністю, наявністю шумів, пропусків і аномалій. Класичні статистичні методи та більшість сучасних моделей машинного навчання, включаючи штучні нейронні мережі, не забезпечують достатнього рівня властивості узагальнення на нових даних в таких умовах функціонування. Це значно обмежує їх використання у практиці біомедичних досліджень та клінічної діагностики. Особливої важливості ця проблема набуває в задачах побудови моделей регресії чи класифікації на основі біомедичних вибірок даних малих обсягів, що характерно для рідкісних захворювань, експериментальних методів лікування або персоналізованих протоколів терапії.

Проблема аналізу малих біомедичних вибірок особливо актуальна у зв'язку з високою вартістю, часозатратністю, етичними обмеженнями, рідкісністю досліджуваних станів або складністю експериментального відтворення клінічних даних. Водночас, існуючі методи нелінійного розширення ознак, аугментації даних, синтезу ансамблів штучних нейронних мереж та компенсації похибок останніх не завжди забезпечують достатню точність, стійкість до варіативності даних, а інколи – і можливість їх практичного використання в умовах аналізу біомедичних даних малого обсягу. Це зумовлює необхідність удосконалення існуючих та розроблення нових методів аналізу таких даних. Значної уваги потребують неітеративні моделі, які дозволяють зменшити обчислювальну складність без втрати точності та забезпечити прозорість і повторюваність результатів. У цьому напрямку особливу цінність мають методи, які забезпечать можливість розв'язання задач регресії та класифікації в умовах обмеженого обсягу даних для реалізації методів навчання, зокрема, штучних нейронних мереж.

Проблемами інтелектуального аналізу даних, зокрема, з використанням нейромережових засобів активно займалися такі українські вчені: С. Г. Антощук, С. А. Бабічев, О. В. Бармак, О. М. Березький, О. В. Бісікало, Є. В. Бодянський, О. А. Винокурова, Т. О. Говорущенко, В. А. Гуляєв, Ю. П. Зайченко, О. Г. Івахненко, Р. М. Камінський, В. В. Ковтун, Ю. В. Крак, В. І. Литвиненко, Н. І. Мельникова, А. О. Олійник, І. Г. Перова, С. Г. Стіренко, С. О. Субботін, Р. О. Ткаченко, І. Г. Цмоць, А. Ю. Шафроненко, Н. Б. Шаховська, В. С. Яковина, а також ряд закордонних вчених: D. Estrin, E. V. Nekler, J. Jiang, R. Kitchin, P. Kokol, P. Klasnja, C. Li, T. Shaikhina, P. Xu. Зокрема, О. Г. Івахненко здійснив фундаментальні дослідження у галузі синтезу нейронних мереж і став автором методу групового урахування аргументів, який поклав початок автоматичному моделюванню складних систем; методами та засобами опрацювання діагностичної інформації в реальному часі займався В. А. Гуляєв; неітеративними методами синтезу нейромережових моделей – Р. О. Ткаченко, С. О. Субботін; проблемами апаратної реалізації таких нейронних мереж – І. Г. Цмоць; теоретичним обґрунтуванням необхідності розвитку парадигми малих даних – Е. В. Nekler та Р. Klasnja; проблемами розпізнавання образів та кластеризації на даних малих обсягів – Є. В. Бодянський, Р. М. Камінський, О. А. Винокурова, І. Г. Перова, А. Ю. Шафроненко; алгебраїчним підходом до аналізу даних малих обсягів – Ю. В. Крак; задачами оцінювання параметрів стохастичних моделей на основі вибірок даних малих обсягів – О. В. Бісікало, В. В. Ковтун; проблемами аналізу малих даних засобами глибокого навчання та пояснювального штучного інтелекту – О. В. Бармак, Ю. В. Крак; проблемами аналізу високовимірних даних у галузі біоінформатики – В. І. Литвиненко, С. А. Бабічев; проблемами персоналізованої медицини – Ю. П. Зайченко, Н. І. Мельникова; проблемами аугментації малих обсягів даних та їх класифікації – О. М. Березький, С. Г. Стіренко; проблемами ансамблевого навчання – Н. Б. Шаховська, С. Г. Стіренко, С. Г. Антощук, В. С. Яковина.

У задачах аналізу структурованих біомедичних даних на вибірках малих обсягів використання окремих моделей машинного навчання, зокрема, класичних штучних нейронних мереж обмежене через високу ймовірність перенавчання, недостатню репрезентативність вибірок для коректного налаштування параметрів моделі та низьку здатність до узагальнення. Особливо це стосується нелінійних моделей, які мають велику кількість параметрів та потребують значного обсягу інформації для ефективної реалізації процедур навчання. Використання ансамблів моделей, зокрема, ансамблів ШНМ, може частково вирішити проблему перенавчання завдяки зменшенню дисперсії оцінок і покращенню властивостей узагальнення моделі. Однак синтез ансамблів гетерогенних ШНМ, який вимагає наявності різних базових моделей з хоча б прийнятною точністю, є дуже складним, а іноді – практично недосяжним, зокрема, у випадку аналізу екстремально коротких вибірок даних. Окрім цього, ітеративні методи навчання існуючих ШНМ є ресурсомісткими, повільними та залежними від початкової рандомізації, що знижує відтворюваність отриманих рішень і ускладнює їх використання в умовах обмежених обчислювальних можливостей. Сукупно ці фактори обумовлюють необхідність удосконалення існуючих та розроблення нових, точних, стійких до перенавчання і обчислювально ефективних ансамблів моделей неітеративного машинного навчання для аналізу структурованих біомедичних вибірок даних малих обсягів.

Аналіз існуючих рішень продемонстрував наявність низки протиріч, що виникають при побудові моделей машинного навчання для аналізу біомедичних даних малих обсягів. Ці протиріччя обумовлюють складність забезпечення високої точності, стійкості та властивості узагальнення моделей в умовах реальних застосувань:

- у сучасних задачах аналізу біомедичних даних часто спостерігається критичний дефіцит даних для реалізації методів навчання, тоді як побудова точних та надійних моделей машинного навчання потребує великого обсягу даних для

забезпечення статистичного узагальнення результатів. Це створює протиріччя між обмеженістю біомедичних вибірок і вимогами до достовірності моделей;

- біомедичні дані характеризуються високою вимірністю, гетерогенністю, а також наявністю шумів, аномалій і пропусків, що істотно ускладнює застосування класичних методів машинного навчання, в тому числі штучних нейронних мереж. Це породжує протиріччя між складністю реальних біомедичних даних і недостатньою адаптивністю існуючих моделей до таких умов;

- гетерогенні ансамблі штучних нейронних мереж демонструють значний потенціал у підвищенні точності, стійкості та узагальнювальної здатності моделей, однак ефективний підбір різномірних базових моделей у випадку аналізу екстремально коротких навчальних вибірок даних вкрай обмежений або навіть неможливий. Це створює протиріччя між перевагами гетерогенних ансамблів і практичними складнощами їх синтезу в умовах гострого дефіциту даних;

- ітеративні методи машинного навчання часто демонструють високу точність, але є обчислювально затратні, а також не забезпечують повторюваності результатів навчання, що обмежує їх застосування в реальних біомедичних системах з обмеженими ресурсами. Це створює протиріччя між необхідністю оперативної та стійкої побудови моделей і складністю застосування ітеративних методів у практичних умовах.

Взаємопов'язаний характер цих протиріч породжує актуальну теоретичну проблему розвитку основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язування задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

Зв'язок роботи з науковими програмами, планами та темами. Дисертаційна робота виконана відповідно до пріоритетних напрямів розвитку Національного університету "Львівська політехніка" у галузі систем штучного інтелекту, а також дослідницької тематики Інституту комп'ютерних наук та інформаційних технологій, зокрема, кафедри систем штучного інтелекту в рамках

напряму “Методи та засоби штучного інтелекту для підтримки прийняття рішень”. Основні теоретичні та прикладні результати дисертації отримано в процесі участі (як керівник чи виконавець) у виконанні ряду науково-дослідних робіт як національного, так і міжнародного рівня, а саме: наукового гранту EURIZON: “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” (№ 871072, 01.04.2024 – 31.03.2025, фінансування: Європейський Союз, програма Horizon Europe); прикладного дослідження “Методи та засоби дослідження маркерів старіння та їх впливу на постковідні ефекти для подовження працездатного періоду” (№ 2021.01/0103, 2022 – 2024, фінансування: Національний фонд досліджень України); науково-технічної (експериментальної) розробки “Методи та засоби інформаційної безпеки та гігієни на основі інтерпретованого штучного інтелекту” (№ 0121U107809, 2022 – 2024, фінансування: Міністерство освіти і науки України); науково-технічної (експериментальної) розробки “Технології та системи оброблення і зберігання персоналізованих військових медичних даних” (№ 0121U107809, 2022 – 2023, фінансування: Міністерство освіти і науки України); науково-технічної (експериментальної) розробки “Розроблення інформаційної технології оцінювання та прогнозування надійності програмного забезпечення методами машинного навчання” (№ 0121U109527, 2021 – 2022, фінансування: Міністерство освіти і науки України); наукового гранту “Integrated care for next generation” (iCare4NextG) (C2020/1-8, 04.2021 – 03.2024, фінансування: Європейський Союз, програма EUREKA); прикладного дослідження “Система підтримки прийняття рішень моделювання поширення вірусних інфекцій” (№ 211/01.220, 2020 – 2021, фінансування: Національний фонд досліджень України); прикладного дослідження “Національний реєстр пацієнтів на первинні імунodefіцити” (№ 0121U107803, 2021 – 2023, фінансування: Міністерство охорони здоров’я України); прикладного дослідження “Інформаційна технологія опрацювання персоналізованої медичної інформації” (№ 0119U002257, 2019 – 2020, фінансування: Міністерство освіти і науки України).

Мета і завдання дослідження. Метою роботи є розвиток теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

Для досягнення мети дисертаційної роботи необхідно вирішити такі завдання:

1. Проаналізувати методи, моделі та засоби інтелектуального опрацювання біомедичних даних малих обсягів.
2. Розвинути метод кусково-лінійної апроксимації для моделювання багатовимірних поверхонь відгуку.
3. Удосконалити метод формування комбінованої поверхні відгуку меншої вимірності з застосуванням кластерного аналізу на вибірках даних малих обсягів.
4. Розробити метод нелінійного доповнення простору незалежних ознак біомедичної вибірки даних вихідними сигналами шару підсумовування ймовірнісної нейронної мережі.
5. Розробити методологію розв'язання задач регресії та класифікації на екстремально-коротких біомедичних вибірках з використанням аугментації даних.
6. Розробити метод синтезу ансамблів лінійних нейроподібних структур з неітеративним машинним навчанням з використанням різних підходів до агрегування кусково-лінійно апроксимованих поверхонь відгуку кожного його елемента.
7. Розробити метод синтезу ансамблю гомогенних нелінійних штучних нейронних мереж, що функціонують на незалежно сформованих вибірках даних з використанням симетричних випадкових зміщень та з різними способами агрегування вихідних результатів окремих моделей.
8. Розробити метод часткової компенсації систематичної похибки нейронної мережі узагальненої регресії.

9. Удосконалити метод глобально-локальної апроксимації шляхом незалежного виділення глобальної та локальної компонент сигналу ансамблю нейронних мереж.

10. Розробити прямий та зворотний методи компенсації похибок регресійних моделей з використанням каскадних ансамблів на основі раціональних дробів.

11. Розвинути теорію синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням на вибірках біомедичних даних малих обсягів.

12. Розробити програмне забезпечення для реалізації усіх методів та виконати оцінювання їх результатів на основі як модельних даних, так і даних біомедичних досліджень.

Об’єкт дослідження – процеси розроблення методів синтезу нейроподібних ансамблевих структур для інтелектуального аналізу структурованих вибірок даних малого обсягу.

Предмет дослідження – моделі, методи та засоби синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для розв’язання задач регресії та класифікації на біомедичних вибірках даних малого обсягу

Методи дослідження. У дисертаційній роботі використано методи математичного аналізу та чисельного моделювання для формалізації, побудови та теоретичного обґрунтування нових ансамблів штучних нейронних мереж, методи теорії ймовірностей і математичної статистики для оцінювання точності та здатності до узагальнення розроблених методів, а також для визначення критеріїв зупинки і перевірки їх ефективності в умовах обмежених вибірок даних, методи теорії штучних нейронних мереж і ансамблевого навчання для побудови неітеративних ансамблів ШНМ, розроблення методів їх навчання, а також реалізації механізмів компенсації похибок, методи структурного й компонентного аналізу даних, які використовувалися для розширення простору ознак з залученням полінома Колмогорова-Габора, декартового квадрата та ізоморфного

представлення вхідних компонент векторів, методи кластерного аналізу для побудови комбінованих поверхонь відгуку та реалізації аугментації даних на основі кластеризації. Методи комп'ютерного моделювання та експериментальної перевірки використовувалися для реалізації удосконалених та розроблених методів, проведення експериментів і виконання порівняльного аналізу ефективності моделей у задачах регресії та класифікації на біомедичних вибірках даних малих обсягів.

Наукова новизна роботи. На основі виконання теоретичних та експериментальних досліджень вирішено важливу теоретичну проблему розвитку основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів. При цьому отримано такі наукові результати:

уперше розроблено:

- методологію розв'язання задач регресії й бінарної класифікації на основі декартового квадрата множини вхідних компонент векторів, яка містить базовий, адитивний, каскадний методи, а також методи формування комбінованої поверхні відгуку та формування аугментованої вибірки даних необхідного обсягу, що забезпечує підвищення як точності роботи нелінійних моделей машинного навчання, так і їхні властивості узагальнення в умовах навчання на екстремально-коротких вибірках біомедичних даних;
- метод синтезу ансамблю лінійних нейроподібних структур, який завдяки агрегуванню кусково-лінійно апроксимованих поверхонь відгуку кожного його елемента забезпечує підвищення точності моделей регресії та класифікації;
- метод синтезу ансамблів гомогенних нелінійних штучних нейронних мереж, який шляхом агрегування апроксимованих поверхонь відгуку, сформованих на незалежних вибірках із використанням попарних симетричних випадкових зміщень, підвищує точність моделей регресії та класифікації;

- метод часткової компенсації систематичної похибки нейронної мережі узагальненої регресії, який за рахунок використання ансамблю з двох таких нейронних мереж забезпечує підвищення точності розв'язання задач регресії у випадку аналізу біомедичних даних малих обсягів;
- метод нелінійного розширення простору незалежних ознак біомедичної вибірки даних, який внаслідок доповнення їх вихідними сигналами шару підсумовування ймовірнісної нейронної мережі забезпечує підвищення точності класифікації даних на лінійних неітеративних нейроподібних структурах;
- прямий та зворотний методи компенсації похибки моделі регресії, у яких за рахунок перетворень на основі раціональних дробів досягнуто підвищення точності та робастності апроксимації в задачах аналізу біомедичних даних малих обсягів з можливістю додаткового ітеративного уточнення результатів;

удосконалено:

- метод глобально-локальної апроксимації біомедичних даних малих обсягів, який на відміну від наявного використовує незалежне виділення глобальної та локальної компонент сигналу з використанням ансамблю з двох нейронних мереж узагальненої регресії та надає змогу додатково ітеративно уточнити шукане значення;
- метод формування комбінованої поверхні відгуку з застосуванням кластерного аналізу, який на відміну від наявного доповнює простір ознак значенням залежного атрибута центру кластера, що забезпечує підвищення точності моделей регресії та класифікації на вибірках даних малих обсягів без втрати узагальнення;

набули подальшого розвитку:

- теорія синтезу нейроподібних ансамблевих структур із неітеративним машинним навчанням, яка ґрунтується на розроблених методах нелінійного збільшення вимірності простору ознак, аугментації даних, декореляції вибірок та

компенсації похибок для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів;

- метод кусково-лінійної апроксимації, який за рахунок ізоморфного подання кожного значення ознаки у вигляді її локальних компонент збільшує вимірність простору вхідних даних та забезпечує локальну корекцію параметрів багатовимірної поверхні відгуку лінійними неітеративними нейроподібними структурами.

Практичне значення одержаних результатів. Дисертаційне дослідження має значне наукове значення і спрямоване на розвиток основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням, орієнтованої на ефективний аналіз структурованих біомедичних даних малих обсягів. Розроблені моделі та методи створюють передумови для розв'язання актуальних практичних задач інтелектуального аналізу даних малих обсягів, більшої вимірності, нестабільності та гетерогенності вибірок даних, що характерно для сучасної біомедичної інженерії, біоматеріалознавства, персоналізованої медицини та суміжних галузей.

Отримані результати дослідження включено до окремих розділів лекційних курсів з дисципліни “Основи штучного інтелекту” для студентів освітнього ступеня “бакалавр” та дисципліни “Моделювання та аналіз даних у інформаційних системах” для студентів освітнього ступеня “магістр” за спеціальністю 122 *Комп'ютерні науки*. Зокрема, навчальні матеріали доповнено розділами, присвяченими методам нелінійного розширення простору вхідних даних для підвищення точності класифікації та регресії, методам інтелектуального аналізу структурованих вибірок даних малих обсягів, методам синтезу гомогенних ансамблів штучних нейронних мереж лінійного й нелінійного типу з різними підходами до агрегування результатів їхньої роботи, а також методам компенсації похибок нейронних мереж. Впровадження підтверджено відповідним актом про впровадження кафедри систем штучного інтелекту Національного університету

“Львівська політехніка”.

Практична цінність роботи полягає у тому, що розроблені та вдосконалені методи і програмні засоби є ефективними інструментами для інтелектуального аналізу біомедичних даних обмеженого обсягу. Ці методи відзначаються високими показниками точності, робастності до шумів та аномалій, а також обчислювальною ефективністю. Використання розроблених та удосконалених методів забезпечує:

- суттєве підвищення точності лінійних моделей машинного навчання, зберігаючи їх здатність до узагальнення за рахунок незначного збільшення вимірності простору незалежних ознак кожного вектора даних;
- підвищення точності нелінійних моделей машинного навчання до 59% та їх стійкості до спотворень у даних шляхом використання розробленої методології розв’язання задач регресії та бінарної класифікації на основі декартового квадрата множини вхідних компонент векторів;
- збільшення точності на 5 – 6% в умовах аналізу даних малих обсягів на основі розроблених ансамблів лінійних і нелінійних штучних нейронних мереж з застосуванням різних способів декореляції вхідних даних та агрегування результатів;
- зменшення середньоквадратичної похибки на 4 – 8% залежно від способу агрегування результатів роботи методу часткового усунення систематичної похибки нейронної мережі узагальненої регресії;
- зменшення середньоквадратичної похибки результатів розв’язання задачі регресії на 13% з використанням удосконаленого методу глобально локальної апроксимації, та на 9% та 11% відповідно з використанням прямого і зворотного методів компенсації похибок на основі раціональних дробів.

Розроблені програмні засоби, захищені авторським правом на комп’ютерні програми (додаток Д), забезпечили можливість експериментального оцінювання ефективності розроблених методів як на модельних, так і на реальних даних біомедичних досліджень. Отримані результати успішно апробовані при розв’язанні

низки прикладних задач у різних напрямках біомедичних досліджень, зокрема, в біоматеріалознавстві, ревматології, травматології, стоматології та трансплантології. Їх впровадження здійснювалося у співпраці з медичними установами та науково-дослідними організаціями, а також в межах виконання наукового-дослідних робіт в Україні та міжнародних грантових проєктів за кордоном, що засвідчено відповідними актами впровадження та використання. Зокрема, результати проведених досліджень впроваджено (додаток Г):

- у Військово-медичному клінічному центрі Західного регіону в період з січня 2024 по січень 2025 року, що підтверджено відповідним Актом від 22.04.2025 року;
- у Комунальному некомерційному підприємстві Харківської обласної ради “Обласний медичний клінічний центр урології і нефрології ім. В. І. Шаповала” у 2022 – 2024 роках, що підтверджено відповідним Актом впровадження;
- на кафедрі стоматології дитячого віку та імплантології Харківського національного медичного університету у 2022 – 2024 роках, що підтверджено відповідним Актом про впровадження від 04.03.2024 року;
- у Комунальному некомерційному підприємстві Харківської обласної ради “Обласна клінічна лікарня” у 2022 – 2024 роках, що підтверджено Актом від 02.12.2024 року;
- у межах виконання міжнародного наукового гранту EURIZON “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” (№ 871072, 2024 – 2025), що підтверджено Актом використання наукових результатів №17244/3 від 17.03.2025 року, виданим Українським науково-технологічним центром;
- у межах виконання НДР в Україні, які фінансувалися Міністерством освіти і науки України а також Національним Фондом Досліджень України, що підтверджено відповідними Актами.

Окрім цього, розроблені та удосконалені методи інтелектуального аналізу біомедичних даних малих обсягів можуть використовуватися для розв’язання задач, що виходять за межі опрацювання суто біомедичних даних малих обсягів. Такий розвиток вимагає попереднього вивчення особливостей постановки відповідних задач та характеристик вхідних даних. Проте, завдяки універсальності закладених підходів, розроблені у дисертації методи можуть бути успішно адаптовані та ефективно використані для аналізу даних у суміжних прикладних галузях, де також спостерігається дефіцит обсягу вибірки даних для ефективної реалізації методів навчання існуючих нелінійних ШНМ. Прикладами такого застосування є адаптації каскаду штучних нейронних мереж узагальненої регресії [1] та методу розв’язання задачі регресії на основі декартового квадрата впорядкованих множин входів з використанням принципів каскадування ШНМ [2] для розв’язання задачі неруйнівної діагностики стану мостової конструкції, а також ієрархічного методу розв’язання задач регресії в умовах даних малих обсягів [3] для оцінювання теплового та охолоджувального навантаження будівлі. Ці методи було адаптовано та використано здобувачем в процесі виконання ним міжнародного наукового проєкту за програмою Horizon Europe “Innovative Methodologies for the Design of Zero-Emission and Cost-Effective Buildings Enhanced by Artificial Intelligence - ZEBAI” (№ 101138678).

Особистий внесок здобувача. Усі наукові результати, подані в дисертації, є особистим досягненням здобувача. Усі положення, що виносяться на захист, належать автору особисто. Наукові праці [4], [5] виконані одноосібно. У працях, опублікованих у співавторстві, концептуальні ідеї, теоретичні підходи та методи, що лягли в основу дисертаційного дослідження, розроблені здобувачем самостійно. Відповідно до пунктів наукової новизни дисертаційної роботи здобувачем зроблено такі внески: розроблено методологію розв’язання задач регресії і бінарної класифікації на основі декартового квадрата множини вхідних компонент векторів [6], [7], яка включає розроблення базового [8], [9], [10], [11], адитивного [12], [13],

каскадного [14], [15] методів, а також розроблення методів формування комбінованої поверхні відгуку [16] та аугментованої вибірки даних обмеженого обсягу [17]; розвинуто метод кусково-лінійної апроксимації для задач регресії і класифікації [18], [19], [20] та розроблено метод синтезу ансамблю лінійних нейроподібних структур на його основі [21], [22]; розроблено методи синтезу ансамблів гомогенних нелінійних штучних нейронних мереж та нейроподібних структур з РБФ-розширенням входів з різними способами формування підвибірок даних та агрегування результатів [23], [24], [25], [26], [27], [28], [29], [30]; розроблено метод часткового усунення систематичної похибки нейронної мережі узагальненої регресії з використанням трьох різних способів агрегування результатів [31], [32], [33]; розроблено метод нелінійного доповнення простору незалежних ознак біомедичної вибірки даних вихідними сигналами шару підсумовування різними модифікаціями ймовірнісної нейронної мережі [34], [35], [36]; розроблено прямий та зворотний методи компенсації похибки моделі регресії з можливістю ітеративного уточнення результатів [37], [38]; удосконалено метод глобально локальної апроксимації біомедичних даних малих обсягів [39], [40]; удосконалено метод формування комбінованої поверхні відгуку з застосуванням кластерного аналізу [41], [42]; розвинуто теорію синтезу нейроподібних ансамблевих структур з неітеративним машинним навчання, а саме: розроблено метод нелінійного розширення входів поліномом Колмогорова – Габора [43], розроблено двокроковий метод нормалізації табличних вибірок даних [44], розроблено каскадний ансамбль неітеративних ШНМ на основі використання полінома Колмогорова – Габора та метод його поліноміального представлення у режимі застосування [45], розроблено метод зменшення вимірності простору вхідних даних для кожного члена каскадного ансамблю [46], розроблено метод зменшення тривалості процедури навчання каскадного ансамблю [47], розроблено метод каскадування алгоритмів машинного навчання [48], досліджено особливості функціонування ЙНМ в процесі аналізу даних малих обсягів [49], досліджено

ефективність використання різних методів нормалізації даних на точність інтелектуального аналізу даних малих обсягів [50], розроблено метод розширення входів НМУР [51], досліджено вплив ефективності існуючих методів аугментації табличних вибірок даних на точність роботи алгоритмів машинного навчання [52], розроблено каскадний ансамбль неітеративних ШНМ на основі принципу лінеаризації поверхні відгуку [53], запропоновано програмні засоби для реалізації розроблених методів [54], [55], [56].

Апробація результатів дисертації. Основні теоретичні та прикладні результати дисертаційної роботи широко представлені та обговорені на провідних науково-практичних конференціях, симпозіумах і семінарах, що відбувалися в 13 країнах світу, а саме: XVI Міжнародній науково-практичній конференції “Інформаційні технології і автоматизація – 2023” (жовтень 2023, Одеса, Україна), 15-тій Ювілейній Міжнародній науково-практичній конференції молодих вчених “Економічний і соціальний розвиток України в XXI столітті: національна візія та виклики глобалізації” (травень 2023, Тернопіль, Україна), XXI-й Міжнародній науковій конференції “Нейромережні технології та їх застосування” (НМТЗ–2023) (2023, Краматорськ–Вінниця–Тернопіль, Україна), 33th International Conference on Database and Expert Systems Applications (DEXA 2022) (22–24 серпня 2022, Відень, Австрія), International Conference on Informatics & Data-Driven Medicine (IDDM 2020) (19–21 листопада 2020, Векше, Швеція), IEEE International Conference on Electronics and Information Technologies (ELIT 2023) (26–28 вересня 2023, Львів, Україна), International Conference on Advanced Computer Information Technologies (ACIT 2022) (26–28 вересня 2022, Списька Капітула, Словаччина), International Conference on Decision Aid Sciences and Application (DASA 2020) (8–9 листопада 2020, Сахір, Королівство Бахрейн), International Work-Conference on Artificial Neural Networks (IWANN 2019) (12–14 червня 2019, Гран-Канарія, Іспанія), 6th International Symposium on Emerging Information, Communication and Networks (EICN 2019) (4–7 листопада 2019, Коїмбра, Португалія), 2-nd International

Conference on Computer Science, Engineering and Education Applications (ICCSEEA2019) (26–27 січня 2019, Київ, Україна), International Conference on Artificial Intelligence and Logistics Engineering (ICAILE2021) (22–24 січня 2021, Київ, Україна), 12th International Conference on Ambient Systems, Networks and Technologies (ANT 2021) (23–26 березня 2021, Варшава, Польща), 4th International Conference on Emerging Data and Industry 4.0 (EDI40 2021) (23–26 березня 2021, Варшава, Польща), International Conference on Informatics & Data-Driven Medicine (IDDM 2021) (19–21 листопада 2021, Валенсія, Іспанія), International Conference on Data Science and Applications (ICDSA 2021) (10–11 квітня 2021, Калькутта, Індія), International Conference on Future Networks and Communications / Mobile Systems and Pervasive Computing (FNC 2024) (5–7 серпня 2024, Гантінгтон, Західна Вірджинія, США), 2-nd International Workshop on Reliability Engineering and Computational Intelligence (RECI 2022) (14–15 листопада 2022, Жиліна, Словаччина), International Conference on Big Data Innovations and Applications (Innovate-Data 2019) (2019, Стамбул, Туреччина), International Conference on Mobile Web and Intelligent Information Systems (MobiWIS 2019) (2019, Стамбул, Туреччина), IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON 2019) (4–6 липня 2019, Львів, Україна), International Workshop on Informatics & Data-Driven Medicine (IDDM 2018) (листопад 2018, Львів, Україна), IEEE International Conference on Computer Science and Information Technologies (CSIT 2023) (18–20 жовтня 2023, Львів, Україна), Mediterranean Smart Cities Conference (MSCC 2024) (травень 2024, Мартіль–Тетуан, Марокко) та International Workshop on Small and Big Data Approaches in Healthcare (SBDaH 2021) (1–4 листопада 2021, Льовен, Бельгія).

Публікації. Основні результати дисертаційного дослідження опубліковано у 53 наукових публікаціях. Зокрема, опубліковано два розділи у колективних монографіях в Україні та за кордоном, 22 статті у наукових журналах, з яких 6 – у фахових виданнях України категорії Б, а 16 – у високореєтингових закордонних журналах, що індексуються у наукометричних базах Scopus та/або Web of Science;

серед них чотири статті опубліковано у журналах першого квартилю (Q1), а одинадцять – у журналах другого квартилю (Q2). Крім того, матеріали дисертації висвітлено у 26 публікаціях у збірниках матеріалів міжнародних наукових конференцій, з яких 22 індексуються у наукометричній базі Scopus. За результатами дослідження отримано три свідоцтва України про реєстрацію авторського права на комп'ютерні програми.

Структура та обсяг роботи. Дисертаційна робота складається з вступу, семи розділів, висновків, списку використаних джерел, що включає 310 найменувань, та п'яти додатків. Загальний обсяг становить 371 сторінку, з яких основний текст займає 310 сторінок. Робота містить 50 рисунків і 17 таблиць, що ілюструють основні результати дослідження.

РОЗДІЛ 1

АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ІНТЕЛЕКТУАЛЬНОГО ОПРАЦЮВАННЯ БІОМЕДИЧНИХ ДАНИХ МАЛИХ ОБСЯГІВ

У розділі розглянуто поняття біомедичних даних та проаналізовано ключові характеристики, які зумовлюють специфіку розроблення, впровадження та застосування методів інтелектуального аналізу в різних наукових напрямках та спеціалізаціях біомедичних досліджень. На основі аналізу сучасних наукових джерел наведено різні підходи до трактування поняття малих обсягів даних, сформульовані експертами різних країн. Подано узагальнене визначення цього поняття з урахуванням особливостей біомедичних задач. Систематизовано типові труднощі та обмеження, що виникають в процесі аналізу біомедичних вибірок даних малих обсягів з використанням моделей штучних нейронних мереж. На основі цього сформульовано наукову проблему дисертаційного дослідження. Проаналізовано ефективність використання існуючих лінійних та нелінійних моделей штучних нейронних мереж для аналізу таких даних, методи нелінійного розширення вхідного простору ознак та аугментації даних. Окреслено їх переваги, а також виявлено обмеження в застосуванні. На основі проведеного аналізу сформульовано узагальнене поняття ансамблю штучних нейронних мереж. Проаналізовано існуючі класи ансамблів, які використовуються в процесі розв'язання задач регресії й класифікації на основі біомедичних вибірок даних малих обсягів. Сформульовано мету та окреслено основні завдання дисертаційного дослідження.

Результати, описані в цьому розділі, висвітлено в періодичних виданнях України та у журналах інших держав, які включено до наукометричних баз Scopus та Web of Science: [12], [44], [45], [48], [57], а також обговорювалися та опубліковані у матеріалах українських та закордонних конференцій: [7], [19], [20], [22], [42], [50], [52].

1.1 Поняття біомедичних даних та їх особливості

У сучасних умовах розвитку науки та медицини зростає потреба у точних, ефективних і безпечних технологіях, здатних забезпечити нові підходи до діагностики, лікування та моніторингу стану здоров'я людини [58]. Ця потреба зумовила появу і стрімкий розвиток міждисциплінарних напрямів, що поєднують знання з різних галузей – технічних, біологічних і медичних. Однією з таких ключових галузей є біомедична інженерія.

Біомедична інженерія – це міждисциплінарна галузь, яка поєднує інженерно-технічні підходи, фізико-математичні методи та медико–біологічні принципи задля вирішення складних біологічних і медичних завдань. Серед них слід виділити розроблення інноваційних методів і засобів проведення медичних досліджень, діагностики, терапії, лікування та профілактики захворювань, а також реабілітації [59]. Вона охоплює більше 18 різних напрямків наукових досліджень і спеціалізацій, серед яких слід виділити біоматеріали [60], клінічну інженерію [61], біоінформатику [62], тканинну інженерію [63], нейроінженерію [64] тощо.

Усі вищевказані напрямки біомедичної інженерії продукують різноманітні дані, які прийнято називати біомедичні дані [59], [65]. Біомедичні дані охоплюють широкий спектр інформації, що генерується з різноманітних джерел, включаючи лікарняне та домашнє обладнання для моніторингу, переносні та імплантовані сенсори, мобільні застосунки, електронні медичні записи, медичні реєстри й клінічні випробування, пацієнто-орієнтовані опитування, соціально-економічні та екологічні показники, медичні зображення й сигнали, відкриті публічні бази даних та дані лабораторних експериментів [65], [66], [67], [68].

Біомедичні дані відзначаються низкою унікальних характеристик, які суттєво відрізняють їх від інших типів даних [65]. Врахування цих особливостей є ключовим фактором для ефективної розробки, впровадження та експлуатації

моделей інтелектуального аналізу таких даних в різних напрямках та спеціалізаціях біомедичних досліджень.

У таблиці 1.1 наведено узагальнення результатів аналізу наукової літератури, що охоплює основні характеристики біомедичних даних, які суттєвим чином впливають на подальший інтелектуальний аналіз таких даних.

Таблиця 1.1 – Особливості біомедичних даних, що впливають на ефективність інтелектуального аналізу

№	Особливість	Пояснення	Джерело
1.	Різні формати даних	Дані можуть бути структурованими (табличні записи), напівструктурованими (JSON/XML-журнали) або неструктурованими (клінічні нотатки, зображення, часові сигнали).	[69]
2.	Змінний обсяг даних	Біомедичні дані продукуються як у надзвичайно великих обсягах, так і у вигляді малих вибірок (десятки спостережень у пілотних дослідженнях).	[70]
3.	Складність структури даних	Необхідність інтегрувати інформацію з різних напрямків біомедичної інженерії призводить до складної мультидоменної структури даних.	[71]
4.	Висока вимірність даних	Безліч ознак при обмеженій кількості пацієнтів створюють “прокляття вимірності”.	[72]

№	Особливість	Пояснення	Джерело
5.	Гетерогенність форматів даних	Поєднання числових, номінальних та порядкових показників у межах однієї вибірки даних.	[73]
6.	Пропуски у даних	Частка відсутніх значень може сягати 10–30 % (інколи більше) через технічні обмеження, помилки введення або дизайн дослідження.	[74]
7.	Шум та аномалії у даних	Високий рівень випадкового та системного шуму (артефакти сенсорів, коливання вимірювань, помилки введення даних).	[75]

Усі ці особливості біомедичних даних зумовлюють ряд суттєвих обмежень для прямого застосування інструментарію машинного навчання в процесі аналізу таких даних [59], [76], [77], [78]. Зокрема, гетерогенність форматів даних потребує складних моделей попереднього опрацювання, в яких кожен тип даних опрацьовується окремими методами. Екстремальна варіативність обсягів даних – від петабайтних томографічних матриць до екстремально коротких вибірок у десятки спостережень табличного типу – суперечить припущенням статистичної ефективності відповідно до теорії Вапніка-Червоненкіса [79], [80] про обчислювальні межі узагальнення. Згідно з цією статистичною теорією для гарантованої високої властивості узагальнення потрібно, щоб розмір вибірки даних був щонайменше того ж порядку, що й складність самої моделі [81]. Висока вимірність і складність мультидоменної структури біомедичних даних породжують “прокляття вимірності”, коли кількість вимірюваних ознак перевищує число пацієнтів [82], [83], [84]. У таких умовах класичні методи машинного навчання за рядом винятків зменшують або втрачають властивість узагальнювати через надмірну розмірність даних. Це призводить до перенавчання моделі. Окрім цього,

наявність пропусків, викидів чи аномалей, а також високий рівень шуму у біомедичних даних ускладнюють застосування методів, які припускають незалежність та однаковий розподіл даних [85].

1.2 Визначення малих даних, їх різновиди

У сучасних біомедичних дослідженнях збір великих масивів даних часто утруднений через етичні, фінансові та часові обмеження [86]. Водночас саме дані, отримані в контрольованих умовах і подані чітко структурованими таблицями, становлять основу для пілотного моделювання, первинних висновків чи формування гіпотез [87], [88].

Зокрема, у клінічній інженерії пілотні випробування нових терапевтичних агентів, як правило, спрямовані на оцінку їх безпеки та початкової ефективності [89]. Для таких досліджень залучають не більше кількох десятків пацієнтів, кожен з яких проходить серію стандартизованих вимірювань (біомаркери, клінічні шкали тощо). В результаті формується структурована вибірка даних з невеликою кількістю рядків, але великою кількістю ознак.

У галузі біоматеріалознавства адитивне виготовлення біосумісних імплантів з титану [90] чи цирконію [91] передбачає значні витрати порошкових матеріалів. Готовий виріб повинен задовольняти суворі вимоги до біосумісності, механічної міцності, твердості й тріщиностійкості [92], [93]. Оцінка цих властивостей традиційно виконується з використанням дорогого лабораторного обладнання: випробувальні машини для розтягнення, фрактографічні системи тощо [90], [91]. Якщо хоча б одна характеристика не відповідає специфікації, то матеріал бракується й утилізується. Це призводить до значних фінансових витрат (порошок, електроенергія), трудових ресурсів (підготовка та тестування зразків) і часових затримок на повторне виготовлення [94]. Проте сукупність результатів таких випробувань формує невелику, але цінну структуровану вибірку даних. На її основі

можна розробити пілотні моделі, в тому числі з використанням ШНМ для оцінювання властивостей нових зразків.

Подібна ситуація виникає і в багатьох інших напрямках та спеціалізаціях біомедичної інженерії – від тканинної інженерії [95] до нейроінтерфейсів [96]. Всі вони зводяться до задачі аналізу вибірок даних обмежених обсягів (малі дані).

Для розуміння цієї задачі розглянемо поняття малих даних. У таблиці 1.2 на основі проведеного аналізу літературних джерел подано різні визначення поняття малих даних, сформульовані експертами галузі з різних країн світу.

Таблиця 1.2 – Найвагоміші визначення поняття малих даних у науковій літературі.

№	Визначення	Перший автор	Країна	Джерело
1.	Мала вибірка даних – це вибірка, що містить менше, ніж десять спостережень на кожну цільову змінну.	Torgyn Shaikhina	Великобританія	[86]
2.	Під “малими даними” розуміють суворе використання даних для окремої особи, клініки, системи охорони здоров’я тощо з метою покращення опису та контролю для цієї одиниці.	Eric B. Hekler;	США	[97]
3.	Малі дані – це дані, які характеризуються обмеженим обсягом, непостійним збором, вузькою різноманітністю та	Rob Kitchin	Ірландія	[98]

№	Визначення	Перший автор	Країна	Джерело
	зазвичай генеруються для відповіді на конкретні питання.			
4.	Малі дані – це вибірка персоналізованої інформації на основі аналізу власних цифрових слідів.	Deborah Estrin	США	[99]
5.	Малі дані – це вибірки, які часто використовуються в наукових та інженерних дослідженнях через обмеження часу, коштів, етики, конфіденційності, безпеки та технічні бар'єри для їх збору.	Jiayu Jiang	Китай	[100]
6.	Малі дані – це вибірка даних, в якій число спостережень значно менше за число ознак. В цьому випадку класичні умови вибіркової складності не виконуються, і гарантоване узагальнення стає неможливим без зменшення складності моделі або	Peter Kokol	Словенія	[101]

№	Визначення	Перший автор	Країна	Джерело
	суттєвого збільшення кількості спостережень.			
7.	Малі дані визначаються як сценарій, у якому відношення кількості навчальних зразків до складності моделі (наприклад, кількість параметрів) є меншим за 20.	Chuanjiang Li	Китай, Греція	[102]
8.	Малі дані – це дані обмеженого обсягу, отримані в результаті експериментів, проведених людьми, або зібрані суб’єктивно.	Pengcheng Xu	Китай	[103]

На основі узагальнення результатів, представлених у таблиці 1.2, у роботі формулюється і використовується наступне визначення поняття малих даних.

Малі дані – це вибірка даних малого обсягу та вузької спрямованості, яка призначена для розв’язання конкретних задач за умов ресурсних, часових, етичних або конфіденційних обмежень і характеризується високим співвідношенням числа ознак до кількості спостережень. Такі дані зазвичай:

- містять значно меншу кількість спостережень, ніж традиційні дані великих обсягів [98];
- мають обмежену різноманітність і нерегулярну структуру збору часто через експериментальні або суб’єктивні методи отримання [103];

- виникають під впливом часових, фінансових та технічних бар'єрів, а також з міркувань безпеки й конфіденційності [104];
- зазвичай не задовольняють класичних вимог теорії Вапніка–Червоненкіса [79], [80], що ускладнює гарантію узагальнення без спрощення моделей або значного розширення вибірки [97], [98], [101], [103], [104].

Слід зазначити, що, окрім поняття малих даних, у науковій літературі зустрічається термін “критично малі дані” або “екстремально-короткі вибірки даних”. Під екстремально-короткими вибірками даних розуміють підвибірки малих даних, що містять значно менше, ніж 100 спостережень [86], [97], [105]. Це поняття має важливе значення для аналізу та інтерпретації результатів статистичних досліджень, особливо в напрямку можливості інтелектуального аналізу даних.

В основі визначення поняття екстремально коротких вибірок даних покладено концепцію мінімального обсягу даних, необхідного для забезпечення надійності статистичного аналізу. У [86] екстремально-коротка вибірка визначається як така, що містить менше, ніж десять спостережень на одну цільову змінну (наприклад, один клас). За таких умов можливий розрахунок тільки базових статистичних показників, таких як середнє значення (математичне очікування), медіана та дисперсія. Проте навіть ці оцінки мають обмежену точність через високу чутливість до вибірових коливань. Використання складніших статистичних методів або побудова моделей на таких вибірках є проблематичними через високий ризик статистичної похибки та нестабільність результатів [106].

Інше визначення екстремально коротких вибірок, представлене у [105], яке вказує на те, що такі вибірки охоплюють обсяг даних до 50 спостережень. На відміну від [86], така класифікація дозволяє застосовувати більш складні статистичні методи, зокрема, обчислення коефіцієнта кореляції та проведення різних статистичних тестів значущості. Однак навіть за таких умов надійність одержаних результатів залишається обмеженою через малий обсяг вибірки, що знижує статистичну потужність аналізу [106].

Застосування теореми Чебишева [107] дозволяє оцінити ймовірність того, що відхилення випадкової величини від її математичного очікування не перевищує певного значення навіть за умов малих вибірок. Це забезпечує певний рівень надійності статистичних оцінок, однак такі оцінки залишаються обмеженими через широкий інтервал довірчої ймовірності [108]. Водночас, центральна гранична теорема втрачає свою ефективність при аналізі малих вибірок, оскільки розподіл вибірових середніх не наближається до нормального розподілу, як це відбувається для великих вибірок [109]. Це означає, що навіть, якщо дані мають випадковий характер, на малих вибірках їх розподіл може залишатися асиметричним або мати великі відхилення, що спотворює результати статистичного аналізу [109].

Окрім цього, аналіз великої кількості дисертацій на здобуття наукового ступеня кандидата медичних наук (доктора філософії), демонструє, що статистичний аналіз часто проводиться на вибірках, які не перевищують 100 пацієнтів [110], [111], [112]. Незважаючи на обмежений обсяг даних, дослідники активно застосовують різноманітні статистичні методи, включаючи кореляційний аналіз, регресійне моделювання та тестування гіпотез [110], [111], [112]. Це демонструє, що навіть за умов невеликої вибірки даних можливе застосування статистичних методів, що за належної інтерпретації результатів може забезпечити отримання важливих наукових висновків.

1.3 Проблема інтелектуального аналізу біомедичних даних малих обсягів

У сучасних біомедичних дослідженнях дослідники дедалі частіше стикаються з ситуаціями, коли обсяг доступних даних є вкрай обмеженим [97], [113]. Зокрема, у багатьох випадках накопичення великої навчальної вибірки є економічно недоцільним або технічно неможливим, наприклад, через тривалість і вартість кожного експерименту [90], [91], [112]. Саме тому часові, фінансові, етичні та інші чинники призводять до формування малих структурованих вибірок даних з

залученням десятків або сотень спостережень з великою кількістю параметрів. Зокрема, у напрямках тканинної інженерії та біоматеріалознавства експериментальні протоколи обмежують кількість зразків через високу вартість реагентів і тривалість культивування клітин [33], [34]. У нейроінженерії та клінічній інженерії дані часто залучають від кількох десятків пацієнтів або прототипу пристрою, що значно обмежує їхній обсяг [114]. Проте навіть такі, невеликі за обсягом, вибірки даних можуть бути цінними джерелами інформації. Вони дозволяють виконувати пілотне моделювання та формулювати гіпотези чи попередні висновки за умови належного опрацювання таких даних [106].

Для ефективного аналізу біомедичних даних обмежених обсягів доцільно використовувати інструменти штучного інтелекту, які здатні виявляти приховані, нелінійні та складні залежності у структурованих вибірках [104]. На відміну від традиційних статистичних підходів моделі МН, зокрема, ШНМ, дозволяють розпізнавати складні взаємозв'язки, які залишаються непомітними для людини. Однак застосування таких моделей на малих вибірках супроводжується низкою суттєвих обмежень [59], [114], [115]. Перш за все ці обмеження зумовлені характерними особливостями структурованих біомедичних даних, наведеними у таблиці 1.1. Такі особливості значно посилюються в умовах аналізу обмежених обсягів даних [116].

Розглянемо основні обмеження та труднощі, що виникають в процесі інтелектуального аналізу малих біомедичних вибірок даних з використанням моделей МН, зокрема, ШНМ:

1. Власне обсяг малої або екстремально-короткої вибірки біомедичних даних.

Основні труднощі в цьому випадку полягають в тому, що класичні методи МН чи ШНМ потребують тисяч або навіть десятків тисяч прикладів для ефективної реалізації процедури навчання. На малих вибірках даних такі моделі або перенавчаються, запам'ятовуючи навчальні дані без здатності до

узагальнення, або взагалі не здатні виділити значущі взаємозв'язки. Це підтверджується статистичною теорією Вапніка-Червоненкіса [79], [80]:

2. Складний характер даних, нелінійні взаємозв'язки в середині множини незалежних атрибутів досліджуваної вибірки даних. Це суттєве обмеження стосовно використання лінійних методів МН та лінійних ШНМ в процесі аналізу таких даних. Зазвичай, такі моделі не здатні ефективно моделювати подібні залежності [117]. Зокрема, у випадку використання моделей, що припускають лінійність, результати аналізу можуть бути некоректними або поверхневими. Це підтверджується теоремою універсальної апроксимації (теорема Цибенка) [118], згідно з якою лише нелінійні моделі здатні апроксимувати довільні функції з необхідною точністю.
3. Велика кількість незалежних атрибутів (ознак, параметрів) малої вибірки біомедичних даних. Така структура даних зазвичай передбачає використання моделей з невеликою кількістю параметрів, зокрема, лінійних моделей або простих ШНМ, щоб відповідати умовам теореми Вапніка-Червоненкіса [79], [80] щодо властивості узагальнення моделей. Окрім цього, застосування існуючих методів нелінійного розширення простору вхідних ознак у таких умовах є ризикованим, оскільки може призвести до виникнення “прокляття вимірності” [119]. Це явище характеризується тим, що зі збільшенням кількості ознак обсяг даних, необхідний для надійного навчання моделі, зростає експоненціально. Це робить навчання нестабільним або навіть неможливим на малих вибірках даних.
4. Наявність пропусків у біомедичних даних. Видалення спостережень з пропусками призводить до ще більшого зменшення обсягу вибірки, що критично для малих даних [120]. Замість видалення пропусків рекомендують застосовувати прості методи їх відновлення, такі як імпутація середнім, медіаною або спеціалізовані моделі імпутації на основі простих методів

машинного навчання [121]. Використання складніших методів у випадку аналізу малих даних повинно детально досліджуватися [122].

5. Наявність шумів та аномалій у біомедичних даних. Це може суттєвим чином спотворювати результати аналізу. Шуми та аномалії посилюються в умовах аналізу малих даних, де кожне спостереження має значний вплив на фінальні результати. Через обмежений обсяг даних видалення аномальних значень або їх приглушення за допомогою стандартних методів часто є неефективним [123]. У таких випадках необхідно застосовувати робастні методи інтелектуального аналізу даних, які здатні знижувати вплив аномалій без втрати основної інформації (наприклад, ансамблі моделей ШНМ [124]).
6. Обмежений обсяг валідаційної вибірки. Для коректного налаштування гіперпараметрів складних моделей машинного навчання, наприклад, з великою кількістю параметрів, необхідна надійна валідаційна вибірка [114]. Однак за умов малих даних обсяг валідаційної вибірки є обмеженим, що знижує надійність отриманих результатів [125]. Це підвищує ризик перенавчання та вимагає застосування методів перехресної валідації (крос-валідації) [114].
7. Ітераційний характер методів навчання більшості існуючих ШНМ. Більшість моделей ШНМ використовують ітеративний підхід до навчання, що потребує значних обчислювальних ресурсів. За умов малих вибірок це може призводити до перевитрати ресурсів та часу, оскільки кожна ітерація (єпоха) може значно змінювати ваги моделі, що погіршуватиме її стабільність [126].
8. Нетривіальна задача визначення розподілу даних у екстремально-короткій вибірці. Невизначеність у розподілі даних ускладнює застосування методів аналізу, що передбачають певний розподіл (наприклад, нормальний). Зокрема, це унеможлиблює застосування простих методів аугментації даних (збільшення обсягу вибірки даних), які ґрунтуються на припущенні про відомий розподіл в даних [127].

Усі зазначені труднощі призводять до низки взаємопов'язаних протиріч, що виникають в процесі побудови моделей машинного навчання для аналізу біомедичних даних малих обсягів. Насамперед спостерігається протиріччя між критичним дефіцитом даних для реалізації методів навчання і потребою у великому обсязі даних для забезпечення узагальнюваності моделей. Додатково біомедичні дані вирізняються високою вимірністю, гетерогенністю, нестабільністю та наявністю шумів і пропусків, що породжує протиріччя між складністю реальних даних і недостатньою адаптивністю класичних моделей машинного навчання, зокрема, нейронних мереж. Також гетерогенні ансамблі нейронних мереж мають високий потенціал, але через обмеженість обсягу даних їх ефективний синтез часто неможливий, що створює протиріччя між теоретичними перевагами та їх практичною реалізацією. Окрім того ітеративні методи навчання, хоча й забезпечують високу точність, є обчислювально затратними, що унеможлиблює їх ефективне використання в умовах обмежених обчислювальних ресурсів.

Сукупність цих протиріч породжує актуальну теоретичну проблему розвитку основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

1.4 Аналіз моделей та методів для розв'язання задач регресії/класифікації

У задачах інтелектуального аналізу даних, зокрема, регресії та класифікації, при обмеженій кількості спостережень досить часто використовуються лінійні моделі, зокрема, класична лінійна регресія, а також її регуляризовані варіанти Ridge, LASSO, ElasticNet тощо [128]. Такі моделі характеризуються високим рівнем інтерпретованості, обчислювальною ефективністю та стабільністю при мінімальній кількості параметрів, що робить їх особливо привабливими в умовах аналізу малих вибірок даних [129].

Серед неітеративних апроксимаційних методів поширеними є підходи, засновані на методі найменших квадратів [130]. Проте їх ефективність значно знижується у випадках майже виродженої задачі, зокрема, при сильній мультиколінеарності або нестачі даних. Застосування методів сингулярного розкладу дозволяє частково компенсувати ці проблеми [131]. Проте використання цього методу супроводжується істотним зростанням обчислювальних витрат та накладає обмеження на характеристики вхідних даних. Це зменшує їх придатність для розв'язування широкого класу біомедичних задач.

Враховуючи обмеження вищевказаних методів, значно кращим варіантом можна вважати використання лінійної нейроподібної структури на основі Моделі Послідовних Геометричних Перетворень [19], [132], [133], [134], яка реалізується за допомогою покрокового, неітеративного методу навчання. Цей тип ШНМ забезпечує високу стійкість до мультиколінеарності, добре масштабується при невеликому обсязі даних та дозволяє здійснювати ефективну апроксимацію без необхідності складної процедури оптимізації вагових коефіцієнтів [19], [134]. Враховуючи це, ЛНПС МПГП може бути логічним та технологічно-ефективним розширенням класичних лінійних методів в процесі розв'язання задач аналізу малих вибірок біомедичних даних.

Розглянемо засади методу МПГП для випадку автоасоціативного режиму, тобто, режиму навчання без вчителя [132].

Нехай вибірка даних для навчання задана сукупністю n векторів реалізацій $x_i \in \mathbb{R}^m$, де $i = \overline{1, n}$. Компоненти вектора $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})$ – це атрибути (ознаки) спостереження. Навчання моделі здійснюється шляхом покрокових геометричних перетворень у m -вимірному просторі реалізацій, де m – кількість ознак [133].

У результаті таких перетворень формується проміжна система координат головних компонент, напрямки якої відповідають найдовшим осям еліпсоїда

розсіювання, заданого точками x_i . Таким чином, базою для реалізації методу навчання є матриця вхідних векторів:

$$A = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \dots \\ x_n^\top \end{bmatrix} \in \mathbb{R}^{n \times m}. \quad (1.1)$$

Метод навчання ЛНПС МПГП відомий [19], [132], [133], [134] і складається з послідовності кроків:

1. Встановити початковий номер кроку ортогоналізації, $St := 1$.
2. Для кожного вектора $x_i \in \mathbb{R}^m$ з матриці (1.1), обчислити норму:

$$\|x_i\| = \sqrt{\sum_{j=1}^m x_{i,j}^2}, i = \overline{1, n} \quad (1.2)$$

та обрати базовий рядок $x_b^{(St)}$, для якого норма максимальна:

$$\|x_b^{(St)}\| = \max_i \|x_i\|. \quad (1.3)$$

3. Для кожного x_i обчислити коефіцієнт проєкції $v_i^{(St)}$ (вихід St -го нейрона прихованого шару):

$$v_i^{(St)} = \frac{x_i^\top x_b^{(St)}}{\|x_b^{(St)}\|^2}, i = \overline{1, n}. \quad (1.4)$$

4. Оновити кожен вектор x_i шляхом проєкції на нормальну гіперплощину, ортогональну до $x_b^{(St)} \in \mathbb{R}^m$:

$$x_i := x_i - v_i^{(St)} \times x_b^{(St)}, i = \overline{1, n}. \quad (1.5)$$

5. Якщо $St < m$, інкрементувати $St := St + 1$ та повернутися до кроку 2. В іншому випадку – перейти до кроку 6.
6. У випадку, коли всі m ортогональних компонент побудовано, отримаємо матрицю залишків нульового порядку:

$$A^{(m)} = 0. \quad (1.6)$$

Це можливо, якщо початкові вектори x_i лінійно незалежні. Слід зазначити, що у випадку лінійної залежності обнулення матриці наступає за меншу кількість кроків.

7. Для кожного i отримуємо вектор ознак у новій системі координат (вихід прихованого шару):

$$v_i = \begin{bmatrix} v_i^{(1)} \\ v_i^{(2)} \\ \dots \\ v_i^{(m)} \end{bmatrix} \in \mathbb{R}^m. \quad (1.7)$$

У режимі реконструкції (автоасоціації) вхідного вектора за збереженими базовими напрямками $x_b^{(st)}$ виконуємо зворотне перетворення:

$$\dot{x}_i = \sum_{j=1}^m v_i^{(j)} \times x_b^{(j)}, \quad (1.8)$$

де $\dot{x}_i \in \mathbb{R}^m$ – реконструйований вектор (вихід моделі).

Граф функціонування цієї моделі можна подати у вигляді нейроподібної структури (рис. 1.1).

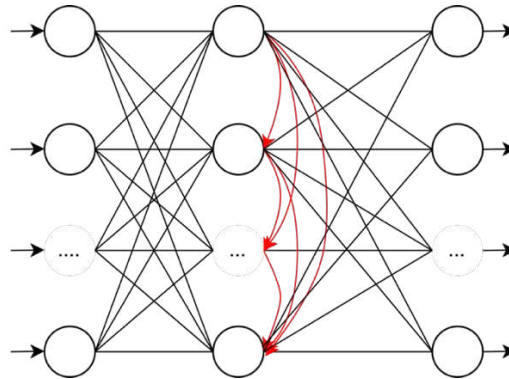


Рисунок 1.1 – Топологія нейроподібної структури МПГП в автоасоціативному режимі

Якщо розглядати функціонування ЛНПС МПГП у режимі з вчителем, то він особливий тим, що передбачає наявність не лише вхідних векторів $x_i \in \mathbb{R}^m$, де $i =$

$\overline{1, n}$, а й відповідних вихідних значень $y_i \in \mathbb{R}$. Останні відомі лише для навчальної вибірки [19], [132], [133]. Це означає, що для тестових (нових) векторів на етапі пошуку значення y_i невідоме, а, отже, точні значення проєкцій (коефіцієнтів v_i^j), які залежать від структури x_i , можливо знаходити лише для навчальних даних.

Отже, можна застосувати стандартні кроки навчання ЛНПС МПГП лише до навчальної вибірки даних, для якої відомі y_i . В результаті цього буде отримано набір проєкцій (1.7), які утворюють новий простір ознак. У цьому просторі можна задати у табличній формі залежність між ними (1.7) та вихідною змінною y_i :

$$y_i = f(v_i^{(1)}, v_i^{(2)}, \dots, v_i^{(m)}), i = \overline{1, n}. \quad (1.9)$$

Оскільки (1.9) задана лише для навчальних векторів, її апроксимують на базі структури штучної нейронної мережі (ШНМ), топологія якої зображена на рис. 1.2 [19], [132], [133].

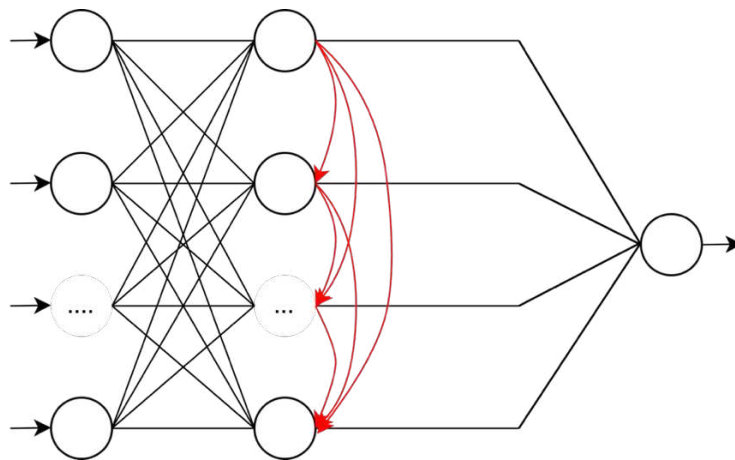


Рисунок 1.2 – Топологія лінійної нейроподібної структури МПГП в режимі з вчителем

Лінійна нейроподібна структура МПГП має низку суттєвих переваг в процесі аналізу зашумлених, високовимірних та обмежених за обсягом біомедичних даних порівняно з іншими класичними лінійними методами [19], [132], [133].

Перш за все модель ЛНПС МПГП є концептуально універсальною [19]. Вона може бути використана як в автоасоціативному режимі (для відновлення сигналу чи виявлення аномалій, (1.8), так і в режимі з вчителем (для побудови регресійної чи класифікаційної залежності (1.9)). При цьому одна й та сама структура ЛНПС МПГП (див. рис. 1.1 і 1.2) адаптується до задачі за рахунок зміни функції активації або структури вихідного шару. У випадку лінійної апроксимації функції $f(\cdot)$, вона реалізується у вигляді зваженої суми проєкцій (аналог лінійної регресії у декорельованому просторі даних), де параметри оцінюються за критерієм найменших квадратів (аналогічно класичній задачі Гауса–Маркова [135]).

ЛНПС МПГП реалізує послідовне головне перетворення, в основі якого покладено інкрементне ортогоналізування вхідних векторів у багатовимірному просторі [19]. На кожному кроці ортогоналізації (1.2) – (1.5) відбувається вибір напрямку з максимальною евклідовою нормою та проєкцій всіх векторів на нормальну гіперплощину, ортогональну до цього напрямку. Такий підхід є конструктивною реалізацією ортогоналізації Грама–Шмідта, що забезпечує виключення колінеарних або надлишкових ознак без потреби у явній регуляризації. Це особливо важливо для біомедичних вибірок з високою корельованістю між ознаками.

Оскільки формування базису у ЛНПС МПГП відбувається не глобально, а поетапно (покроковий вибір найбільш інформативних напрямків у розсіюванні), метод зменшує вплив слабких шумових компонент [134]. Вектори з малими нормами не можуть бути обрані як базові (1.3), що забезпечує низьку чутливість до флуктуацій шуму. Зокрема, класичний метод головних компонент, базується на спектральному розкладі коваріаційної матриці і враховує всі вектори одночасно [136]. На відміну від нього МПГП виконує локальну адаптивну декомпозицію [132]. Такий підхід узгоджується з гіпотезою локальної енергії сигналу [137] та принципом збереження високої варіативності за малих вкладень шуму.

Завдяки тому, що модель не потребує ітеративної оптимізації параметрів чи побудови глобальної моделі з великої кількості зразків, ЛНПС МПГП є ефективною в процесі аналізу обмежених обсягів даних [132], [133]. Зокрема, послідовне перетворення, реалізоване у (1.4) – (1.5), дозволяє замінити оригінальний простір ознак на ортогональний простір проєкцій (1.7), у якому подальша апроксимація (1.9) є більш стійкою.

Останньою, проте не менш важливою перевагою ЛНПС МПГП є те, що у разі необхідності підвищення точності результату модель може бути розширена шляхом заміни лінійного вихідного шару на нелінійний апроксиматор, наприклад, РБФ-мережу, що реалізує універсальну апроксимацію (теорема Колмогорова–Арнольда–Габора про представлення довільної неперервної функції у вигляді суперпозицій). Слід зазначити, що у випадку, коли поверхня відгуку є лінійною, ЛНПС МПГП навчається з нульовою систематичною помилкою (похибкою методу) [19]. Однак ця властивість не зберігається при моделюванні нелінійних поверхонь у режимі застосування [43], [134], [138].

Попри численні переваги в процесі аналізу складних та структурно нелінійних взаємозв'язків у біомедичних вибірках даних модель ЛНПС МПГП, як і більшість лінійних методів, може виявитися недостатньо гнучкою для точного відтворення функціональних залежностей між вхідними параметрами та цільовим атрибутом. Це обмеження зумовлене тим, що всі моделі, які лінійно обробляють дані у вихідному або лінійно ортогоналізованому просторі, апіорі передбачають глобальну лінійність шуканої функції [129], [135]. Проте, справжні залежності між змінними біомедичної вибірки даних часто мають багатовимірний нелінійний характер, який не можна точно апроксимувати жодною лінійною комбінацією, навіть після ортогоналізації [139], [140], [141].

У таких випадках доцільним є використання методів нелінійного розширення вхідного простору ознак, які дають змогу підвищити точність результату, не змінюючи принципи їх навчання [142]. Такий підхід вперше формалізовано у [142].

Він є попередником подальших теорем, зокрема, таких як теорема Ковера [143]. Серед найпоширеніших методів розширення входів слід відзначити поліноміальне розширення (включення до набору ознак усіх степеневих і комбінаційних членів до певного порядку) [142], [144], розширення на основі радіально-базисних функцій (РБФ) [145], [146], [147], функціональні або аналітичні розширення (застосування до кожної ознаки елементарних функцій: логарифм, експонента, синус, гіперболічний тангенс) [148] тощо.

Поліноміальні й РБФ-перетворення широко використовуються в задачах аналізу біомедичних даних [30]. Вони дозволяють апроксимувати складні варіації сигналів [86], [116]. Втім, при прямому застосуванні до високорозмірних біомедичних даних вони можуть призводити до ефекту “прокляття вимірності” [72], [149], що погіршує обчислювальну ефективність та властивість узагальнення [150]. У випадку аналізу малих біомедичних вибірок даних застосування традиційного поліноміального розширення не завжди є доцільним через велику кількість параметрів, які потрібно оцінювати [151].

Радіально-базисні розширення, навпаки, характеризуються вищою гнучкістю, що дозволяє ефективніше відображати складну локальну структуру даних [145], [146]. Зокрема, ці функції дають змогу локалізувати вплив кожного центру у просторово обмеженій області, що відповідає принципам теорії локальної апроксимації функцій [152] і забезпечує кращу адаптацію моделі до нелінійних залежностей у даних. Крім того, поєднання РБФ-розширення з лінійним поданням початкових вхідних ознак дозволяє зберігати здатність до узагальнення за межами локальних областей впливу, що формує екстраполятивні властивості моделі [145], [146]. Сукупно ці характеристики сприяють підвищенню робастності моделі ШНМ до шуму та зниженню ризику перенавчання. Це є важливою перевагою в процесі роботи з малими та зашумленими біомедичними вибірками даних.

Особливої уваги заслуговує можливість розширення моделі ЛНПС МПГП за рахунок додавання нелінійного шару, зокрема, на основі РБФ-функцій [138]. Така

топология та метод її навчання існує (рис. 1.3) і забезпечує суттєве підвищення точності аналізу даних [145], [146].

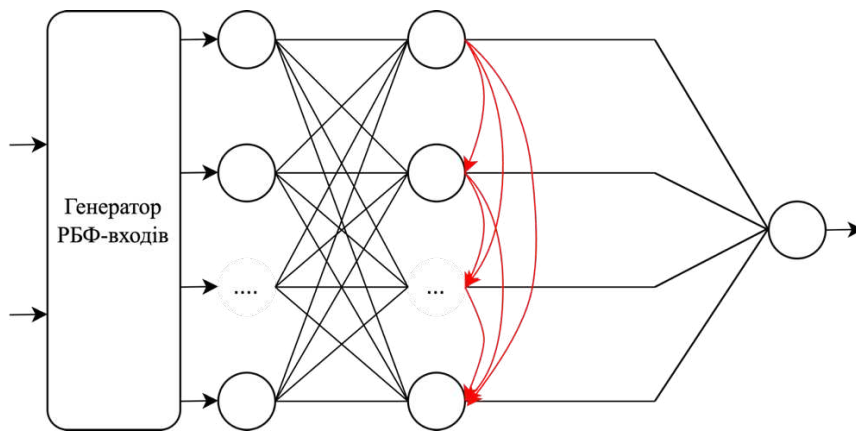


Рисунок 1.3 – Топологія лінійної нейроподібної структури МПГП з РБФ-розширенням входів

Зокрема, використання ЛНПС МПГП з попереднім радіально-базисним розширенням входів [145], [146] дозволяє реалізувати універсальну апроксимацію неперервної функції відповідно до теореми Колмогорова–Арнольда–Габора [153]. Згідно з цією теоремою будь-яку неперервну функцію багатьох змінних можна подати скінченною суперпозицією неперервних функцій однієї змінної та операцій додавання [153]. Такий підхід відкриває можливість розширити апроксимувальну здатність лінійної нейроподібної структури без радикального ускладнення її архітектури.

Таким чином, у ситуаціях, коли базова лінійна модель ЛНПС МПГП не забезпечує належної точності через обмеження, пов’язані з апіорною лінійністю, застосування нелінійного розширення вхідного простору виступає доцільним та обґрунтованим кроком. Однак варто зазначити, що більшість відомих методів розширення – такі як поліноміальні чи РБФ-перетворення – суттєво збільшують вимірність простору ознак. У задачах, пов’язаних з аналізом біомедичних вибірок обмеженого обсягу, це є критичним недоліком, що може призвести до “прокляття вимірності” та погіршення властивості узагальнення моделі ШНМ [72].

Зважаючи на це, ефективний інтелектуальний аналіз у галузі біомедичної інженерії, особливо в умовах малих даних, потребує не лише застосування існуючих, але й удосконалення або розроблення нових методів нелінійного розширення вхідного простору для лінійних ШНМ. Такі методи повинні, з одного боку, забезпечувати високу точність розв'язання поставленої задачі, а з іншого, – обмежувати зростання вимірності простору ознак, який і без того є високорозмірним у більшості біомедичних задач. Це критично важливо для можливості використання подібних методів при збереженні властивості узагальнення моделі відповідно до умов теорії Вапніка–Червоненкіса [79], [80].

Іншим перспективним напрямом підвищення точності аналізу біомедичних даних малого обсягу або екстремально коротких вибірок даних є застосування нелінійних ШНМ. Серед найбільш ефективних моделей у таких умовах слід виокремити нейронну мережу узагальненої регресії та ймовірнісну нейронну мережу. Обидва ці типи ШНМ поєднують низку властивостей [1], [15], які роблять їх особливо придатними до задач аналізу обмежених і зашумлених даних, характерних для біомедичної інженерії.

Обидві ШНМ належать до класу одношарових радіально-базисних мереж інтерполяційного типу, для яких відсутня класична процедура навчання у звичному розумінні, тобто процес ітеративного пошуку та коригування ваг, характерний для багатшарових нейронних мереж [133]. У таких моделях навчання полягає у фіксації навчальної вибірки, а всі обчислення виконуються онлайн, тобто в процесі подачі нового вхідного вектора, вихідне значення якого невідоме. Результат обчислюється шляхом інтерполяції на основі всієї або частини навчальної вибірки [85].

Обидві архітектури потребують налаштування лише одного параметра – ширини (розмаху) гаусівської функції σ , яка визначає ступінь згладжування функції активації [1], [154]. Це значення може бути підібране як емпірично, так і за

допомогою оптимізаційних процедур, що істотно спрощує процес використання ШНМ і зменшує потребу в глибоких технічних знаннях з боку користувача [154].

ЙМН і НМУР, будучи реалізацією радіально-базисного підходу, відповідають умовам теореми Колмогорова–Арнольда [153], [155]. Ці моделі використовують ядрові функції, центровані в навчальних точках, що дозволяє формувати адаптивні локальні апроксимації у вхідному просторі даних, які чутливі до структурних властивостей вибірки [156]. Особливою перевагою обох ШНМ є їх здатність забезпечувати гладку інтерполяцію між навчальними точками, що гарантує неперервність вихідної функції [157], [158]. Це критично важлива властивість для обробки біомедичних даних, які схильні до локальних флуктуацій. Завдяки локальній природі обчислень та відсутності глобального навчання з мінімізацією втрат [157], [158] ці мережі демонструють високу робастність до викидів, шумів і часткової неповноти даних. Подібні ситуації часто трапляються в процесі розв’язання різноманітних прикладних задач і галузі біомедичної інженерії. Зокрема, формування відповіді через згладжене агрегування множини локальних реакцій дозволяє уникати надмірного впливу окремих аномальних точок, зменшуючи чутливість моделі до неконтрольованих збурень у навчальній множині [159].

Завдяки використанню радіально-базисних функцій, центрованих у точках навчальної вибірки, НМУР адаптивно моделює складні зв’язки між вхідними та вихідними змінними. Це забезпечує високу точність при малій кількості спостережень [157]. Зі свого боку, ЙНМ формує ймовірнісну інтерпретацію вихідного результату, тобто оцінює вірогідність того, що вхідні дані належать до певного класу [49], [154]. Це є особливо цінним при наявності невизначеності або неповноти даних, що властиво багатьом біомедичним задачам, у яких дані збираються вручну або з використанням гетерогенних джерел [65]. ЙНМ використовує ядрові функції, зазвичай Гаусові, для побудови оцінки щільності

ймовірності. Це дозволяє побудувати гнучку модель розподілу класів, яка краще враховує локальні особливості вибірки порівняно з класичними підходами [36].

Обидві архітектури демонструють одні з найвищих серед існуючих ШНМ властивості узагальнення, що є критично важливо для забезпечення коректної роботи моделі на нових, раніше не бачених зразках [1], [160], [161]. Важливо зазначити, що НМУР демонструє стійкість до шуму та коливань у вибірці, тоді як ЙНМ є менш схильною до перенавчання завдяки імовірнісній природі своїх оцінок.

Враховуючи зазначене, як НМУР, так і ЙНМ можуть забезпечити ефективні підходи до аналізу малих та зашумлених біомедичних вибірок. Їх здатність до непараметричного навчання, адаптація до локальної структури даних та обчислювальна ефективність роблять ці архітектури привабливими для задач, де обсяг спостережень є обмеженим, а якість вхідних ознак нерівномірна [14], [162].

Разом з тим, слід враховувати, що точність роботи таких моделей значною мірою залежить від властивостей самої вибірки, зокрема, від кількості спостережень, узгодженості ознак, рівня шуму та внутрішньої структурованості даних [10]. У практиці біомедичних досліджень нерідко трапляються ситуації, коли ознаки мають різну фізичну природу, масштаб або ступінь достовірності, що може суттєво впливати на стабільність результатів моделі [150], [163]. Особливо це актуально у випадках екстремально коротких вибірок даних, де обсяг даних настільки малий, що не забезпечує навіть базового статистичного подання цільових залежностей [106]. У таких умовах продуктивність навіть найстійкіших моделей, зокрема, НМУР і ЙНМ, може виявитися недостатньою для досягнення задовільної точності або робастності.

Саме тому, коли обсяг доступних даних обмежує якість аналітичного висновку, доцільним є розгляд додаткових стратегій доповнення або структурного розширення вибірки. Одним з найбільш ефективних і теоретично обґрунтованих підходів у цьому контексті є застосування методів аугментації даних [164], [165]. Такий підхід дозволяє не лише збільшити розмір вибірки, але й покращити її

різноманітність, що, своєю чергою, може забезпечити підвищення точності, робастності та властивостей узагальнення моделей ШНМ в задачах біомедичного аналізу даних [164], [166].

Згідно зі словником термінів у сфері штучного інтелекту [167], підготовленим Міністерством цифрової трансформації України, аугментація даних – це *“техніка, за якої розмір та якість тренувального набору даних (датасету) збільшується шляхом зміни оригінальних тренувальних даних для створення нових тренувальних прикладів, щоб покращити моделі машинного навчання”* [167].

Серйозне теоретичне підґрунтя для використання методів аугментації даних, які дозволяють штучно розширювати вихідну вибірку без залучення додаткових реальних спостережень, подано у [164]. Перш за все, аугментація даних дозволяє зменшити дисперсії оцінок параметрів моделі, що зменшує ймовірність перенавчання [168]. Нові спостереження, згенеровані на основі вихідних даних, дозволяють усереднити випадкові флуктуації у навчальній вибірці, що формалізується як зменшення варіативності емпіричного ризику [169]. Це є особливо критичним для моделей, навчених на малих обсягах даних, де статистична стабільність оцінки суттєво обмежена [164], [170]. Також аугментація даних забезпечує рівномірніше покриття простору вхідних даних, що узгоджується з класичною теоремою Колмогорова–Арнольда про апроксимацію. У практиці нейронних мереж це означає, що додаткові дані забезпечують кращу апроксимацію функціональних залежностей, особливо в регіонах, які недостатньо представлені у вихідній вибірці [118]. Окрім цього, багато методів аугментації базуються на інваріантності цільової функції до певних перетворень вхідних даних, наприклад, масштабування, перестановка ознак, інверсія ознак, генерація синтетичних векторів на основі статистичних властивостей вибірки тощо [168]. Такі перетворення можуть бути описані в термінах теорії груп [164]. Зокрема, якщо модель навчена на множині прикладів, інваріантній щодо певної групи, вона автоматично набуває

робастності до широкого класу збурень вхідних даних, що підвищує її стійкість та властивість узагальнення [164].

Саме тому такий підхід досить широко використовується для підвищення точності та робастності моделей штучних нейронних мереж в процесі аналізу біомедичних даних, зокрема, у випадках обмежених за обсягом вибірок.

На сьогодні існує значна кількість методів аугментації, спеціально розроблених чи адаптованих для роботи з табличними наборами даних. У роботі [171] подано одну з найбільш повних класифікацій простих методів аугментації для табличних вибірок даних. Автором [171] виконано систематизацію існуючих підходів, які згруповано за типами впливу (рядкові, стовпцеві, комбіновані) та реалізовано у вигляді Python-бібліотеки під назвою DeltaPy. Ця бібліотека надає користувачам гнучкий інструментарій для побудови та оцінювання моделей ШНМ з використанням аугментованих вибірок табличних даних. Застосування DeltaPy дозволяє суттєво підвищити ефективність моделей, зокрема, в умовах недостатнього представлення вибірки (задачі регресії або апроксимації) або її структурної дисбалансованості (задачі класифікації).

У сучасних дослідженнях з аугментації табличних даних виокремлюють два основні класи методів: синтез спостережень та екстракція нових спостережень [171]. Подібні методи спрямовані на покращення якості моделей машинного навчання, особливо в умовах обмежених або незбалансованих вибірок.

Методи з першого класу (синтез спостережень) передбачають генерацію нових даних шляхом створення синтетичних прикладів, які імітують властивості наявних спостережень. Одним з відомих методів у цьому класі є SMOTE [52], який генерує нові зразки для менш представлених класів шляхом інтерполяції між існуючими спостереженнями цього класу. Цей підхід дозволяє зменшити дисбаланс класів та покращити здатність моделей до узагальнення, особливо в задачах класифікації. Зокрема, у дослідженні [172] продемонстровано ефективність

поєднання SMOTE з генеративними змагальними мережами для покращення якості синтетичних даних у клінічних випробуваннях.

Перевагою методів цього класу є те, що вони дозволяють збільшити обсяг даних без необхідності в зборі нових спостережень. Це є важливою перевагою у випадках обмежених ресурсів або складності збору даних. Крім того, синтетичні дані, згенеровані за допомогою цих методів, можуть покращити ефективність моделей ШНМ, підвищити їх стійкість до перенавчання та забезпечити кращу здатність до узагальнення [168].

Методи з класу екстракції нових спостережень реалізуються шляхом модифікації наявних даних для створення нових прикладів. Це може включати додавання шуму, перестановку значень або інші трансформації, що зберігають основні характеристики даних [26], [52], [57]. Такі методи дозволяють збільшити обсяг даних без необхідності в зборі нових спостережень, що є особливо корисним у випадках обмежених ресурсів або складності збору даних. Зокрема, у дослідженні [165] показано важливість та ефективність генеративних підходів у аугментації табличних даних. Це пояснюється тим, що систематичне застосування методів аугментації може суттєво покращити здатність моделей до узагальнення, особливо при роботі з обмеженими або незбалансованими вибірками [168].

Перевагами методів цього класу є те, що вони дозволяють моделювати варіативність даних, зберігаючи при цьому їх статистичні властивості. Це забезпечує підвищення стійкості моделей до перенавчання та покращує їх властивість узагальнення, особливо в умовах аналізу обмежених або незбалансованих вибірок даних.

Попри численні переваги, прості методи аугментації табличних даних мають низку обмежень, що ускладнює їх застосування у задачах аналізу складних високовимірних біомедичних вибірок. Зокрема, методи на основі інтерполяції, ефективні лише в умовах локальної щільності даних і можуть генерувати штучні спостереження, які не мають фізіологічного чи клінічного сенсу [52], [173]. У

випадках багатовимірних структур з сильно корельованими або латентно пов'язаними ознаками це може призводити до створення некоректних або навіть суперечливих спостережень.

Методи, що базуються на випадкових модифікаціях існуючих рядків, зокрема, додавання шуму або перестановка значень, не враховують внутрішню структуру даних та взаємозв'язки між змінними [52], [174]. У біомедичних задачах, де навіть незначна зміна вхідних параметрів може суттєво вплинути на результат, наприклад, діагностичне рішення або ймовірність летальності, така неконтрольована аугментація може порушити біологічну достовірність вибірки.

Якщо говорити про більш складні методи аугментації даних, то доцільно виокремити підходи, що ґрунтуються на генеративно-змагальних мережах та автоенкодерах, зокрема, варіаційних автоенкодерах. Методи на основі ГЗМ, такі як CTGAN [175], CTabGAN [176] та їх модифікації, продемонстрували здатність ефективно моделювати складні багатовимірні розподіли табличних даних [177]. Проте вони мають низку обмежень. Серед них – потреба у великій кількості даних, нестабільність навчання, а також низька інтерпретованість згенерованих результатів.

Варіаційні автоенкодери, такі як TVAE [178], DPTVAE [179] та CTVAE [180], забезпечують стабільніше навчання, є менш чутливими до розміру вибірки та дозволяють інтерпретувати латентні простори. Це робить їх ефективними в задачах біомедичного аналізу. Однак, вони також мають певні недоліки. Синтетичні дані, згенеровані ВА, іноді мають розмиту структуру через апроксимацію розподілу латентного простору нормальним розподілом [181]. Це може знижувати точність моделювання дискретних або багатомодальних даних. Також ВА є менш гнучкими у відтворенні складних залежностей між категоріальними та числовими ознаками. Окрім цього, вони характеризуються перевизначенням латентного простору, коли модель захоплює лише основні варіації даних, ігноруючи дрібні, але важливі взаємозв'язки [182].

Незважаючи на бурхливий розвиток генеративних підходів, їх ефективне застосування у біомедичних задачах з обмеженим обсягом даних залишається проблематичним. У випадках екстремально коротких вибірок даних, зокрема, в клінічній практиці або експериментальній медицині, ці методи можуть генерувати приклади, які не відповідають ані внутрішній структурі даних, ані медичному контексту [183]. Це зумовлює необхідність удосконалення існуючих та розроблення нових методів аугментації табличних вибірок даних та адаптації їх до високовимірних задач з урахуванням вимог до збереження не лише статистичної подібності, але й біологічної інтерпретованості та клінічної валідності згенерованих даних.

1.5 Особливості синтезу ансамблів ШНМ для розв’язання задач регресії/класифікації

У випадку, коли одна модель штучної нейронної мережі (ШНМ) не забезпечує достатню точність розв’язання задачі регресії або класифікації, одним з можливих підходів до покращення якості моделювання є використання кількох таких моделей одночасно. Цей метод машинного навчання відомий як ансамблеве навчання [184] і призначений для підвищення точності розв’язання поставленої задачі.

Одним з базових теоретичних обґрунтувань ефективності використання ансамблів є теорема Кондорсе про журі присяжних 1785 року, згідно з якою, якщо кожен член ансамблю має ймовірність правильного рішення, що перевищує 0.5, і ці рішення незалежні, то ймовірність правильності колективного рішення зростає зі збільшенням числа моделей і прямує до одиниці [185]. Принцип “мудрості натовпу”, який вперше опубліковано Френсісом Галтоном [186], підтвердив це. Зокрема, результати його експерименту щодо оцінки ваги бика на сільському ярмарку показали, що середнє значення всіх оцінок, взятих анонімно й незалежно, виявилось ближчим до істинної ваги тварини, ніж більшість індивідуальних оцінок,

включаючи експертні. Згодом ця ідея отримала широку популяризацію завдяки книзі [187], у якій автор розглянув численні приклади з різних галузей науки та культури, де колективні рішення натовпу (за певних умов) виявлялися ефективнішими за дії окремих експертів чи лідерів.

Іншим підходом до обґрунтування ефективності ансамблевого навчання є класичний розклад середньої квадратичної похибки на зміщення та дисперсію [188], в якому підкреслюється, що усереднення незалежних моделей у складі ансамблю дозволяє суттєво зменшити дисперсію, зберігаючи при цьому низький рівень зміщення.

Проте, сама нова теорія, яка розширює попереднє обґрунтування, представлена у роботі [189]. В її основі покладено введення розширеного розкладу похибки, що включає не лише класичні компоненти, зміщення та дисперсію, але й різноманітність між моделями як окремий фактор. Автори у [189] доводять, що належна диверсифікація між базовими моделями ансамблю може суттєво зменшити загальну похибку методу, і формалізують умови, за яких це зменшення є статистично обґрунтованим. Розроблена теорія є універсальною для широкого класу функцій втрат і комбінаторів. Вона створює підґрунтя для об'єктивної оцінки ефективності ансамблевих методів, зокрема, в задачах регресії та класифікації з використанням ШНМ.

Загалом, ансамблеве навчання, згідно з результатами численних досліджень, демонструє перевагу над використанням окремих (індивідуальних) моделей [190]. Це обґрунтовується підвищеною властивістю узагальнення, стабільністю результатів та зменшенням ризику перенавчання у випадку використання ансамблів ШНМ.

Для подальшого розгляду методів цього класу, сформалізуємо поняття ансамблю.

У роботі використовуємо узагальнене визначення ансамблю ШНМ, сформоване на основі аналізу наукової літератури [184], [189], [191], [192], [193].

Під ансамблем штучних нейронних мереж розуміється система, що складається з двох або більше окремих ШНМ, яка включає такі компоненти:

1. Вибір чи генерація набору базових моделей ШНМ, що можуть бути однорідними/гомогенними (однакової архітектури) або неоднорідними/гетерогенними (різної архітектури).
2. Метод формування навчальних вибірок/підвибірок даних для усіх базових моделей ансамблю.
3. Спосіб агрегування виходів окремих моделей, наприклад, усереднення, зважене голосування, мета-модель.
4. Метод навчання моделей ШНМ ансамблю, який може бути незалежним для кожної моделі або скоординованим у випадку складніших ансамблів.
5. Метод застосування ансамблю ШНМ на нових даних.

Ансамблі моделей ШНМ можуть бути класифіковані за різними критеріями [190], [194], [195], [196]: способом агрегування виходів, режимом навчання, особливостями формування вхідних даних, а також за типом базових моделей. Кожен тип ансамблю має свої переваги та обмеження, особливо в процесі розв'язання задач регресії/класифікації у галузі біомедичної інженерії, де вибірки даних є обмеженими, нерівномірними та багатовимірними. Розглянемо кожен з них детальніше.

Перший критерій класифікації – це спосіб агрегування виходів базових моделей ансамблю. Для задач регресії найпоширенішим є ансамбль ШНМ шляхом усереднення їх роботи [21]. Його можна реалізувати з використанням середнього арифметичного або зваженого усереднення результатів базових ШНМ. Такий ансамбль ефективно зменшує дисперсію, що є корисним у випадках з сильним шумом у вхідних даних. Однак, за наявності систематичного зсуву в усіх моделях, агреговане передбачення також буде зміщеним [197]. Для задач класифікації ефективним є ансамбль шляхом голосування як простого, так і зваженого [198].

Його перевага полягає у високій стійкості до локальних помилок окремих моделей, однак в ньому не враховується рівень впевненості кожної моделі [21].

Другий критерій – режим навчання базових моделей у складі ансамблю. У разі паралельного навчання, наприклад, у ансамблів з класу бегінгових, усі моделі навчаються незалежно одна від одної [199]. Такий процес може відбуватися на різних підвибірках або з застосуванням випадкових трансформацій [21]. Це дозволяє зменшити дисперсію без ризику перенавчання, що робить такий підхід відносно ефективним в процесі аналізу біомедичних вибірок даних обмежених обсягів.

Окрему групу становлять ансамблі з навчанням на ознаках другого рівня (мета-ознаках), до яких належить стекінг [57], [200]. У цьому підході базові моделі навчаються паралельно, але їх виходи використовуються як вхідні ознаки для мета-моделі, яка додатково навчається на окремій вибірці [201]. Це дозволяє побудувати більш гнучку модель, що враховує взаємозв'язки між результатами окремих моделей [26]. У випадку біомедичних задач це може бути корисно для інтеграції результатів моделей, що працюють на різних рівнях біологічної організації (молекулярний, фізіологічний, клінічний) [202]. Проте головним обмеженням стекінгу є потреба в додатковій валідаційній підвибірці для навчання мета-моделі, що зменшує ефективний обсяг даних у задачах з малою вибіркою [25]. Для подолання цього обмеження застосовуються методи крос-валідаційного стекінгу або регуляризовані мета-моделі.

На противагу обома попереднім підходам з паралельним режимом навчання стоїть послідовне навчання, притаманне ансамблям з класу бустингових [203]. В таких ансамблях кожна наступна модель коригує помилки попередньої. Зазвичай, вихід кожної моделі додається до сумарного результату з певною вагою [204]. Цей підхід значно підвищує точність, однак є більш чутливим до шуму, що обмежує його застосування при роботі з біомедичними даними низької якості [160].

Компромісом між двома підходами є каскадні ансамблі, у яких виходи моделей попереднього рівня використовуються як вхідні ознаки на наступному рівні [193], [205]. Така архітектура дозволяє реалізувати поетапне уточнення результатів, проте вимагає особливої уваги до накопичення похибок на ранніх рівнях [5], [206].

Третім важливим критерієм класифікації ансамблевих моделей є спосіб формування вхідних даних для кожної базової моделі [21], [207]. Цей аспект визначає рівень декореляції між моделями в ансамблі, що безпосередньо впливає на його здатність зменшувати дисперсію й підвищувати властивість узагальнення. Найпоширенішим підходом тут є бутстрепінг – випадкове вибірконе витягування прикладів з навчальної вибірки з поверненням, що формує унікальні підмножини даних для кожної моделі [208]. Це дозволяє зменшити ризик перенавчання, оскільки кожна модель бачить дещо інший розподіл даних. Особливо важливим є те, що бутстрепінг працює навіть на відносно малих вибірках, зберігаючи основні статистичні властивості набору [209]. Проте в біомедичних задачах з екстремально-короткою вибіркою даних його ефективність знижується через надмірне дублювання елементів.

Ще однією ефективною стратегією є використання випадкових підвибірок ознак, де кожна модель працює лише з частиною вхідних ознак. Такий підхід реалізований, наприклад, у методі Випадкового Лісу і дозволяє моделі зосередитися на різних аспектах вхідного простору [94]. Для задач з великою кількістю біомедичних ознак, наприклад, геноміка, спектрометрія, такий підхід може забезпечити підвищення різноманітності в ансамблі та зменшити ризик перенавчання.

Додатковим джерелом варіативності може стати використання різних ініціалізацій параметрів моделей, зокрема, випадкових ваг на старті навчання ШНМ [23]. Цей підхід не змінює вхідні дані, проте впливає на результати навчання, особливо в нестабільних моделях з сильною залежністю від початкових умов,

наприклад, багатошарові перцептрони або автоенкодери. У випадку аналізу малих вибірок даних, такий підхід дозволяє створити ансамблі без потреби у додаткових даних, що є суттєвою перевагою. Однак надмірна залежність від початкової ініціалізації може призводити до нерепрезентативних моделей при слабкому контролі регуляризації [210].

У практиці аналізу біомедичних даних доцільно комбінувати вказані стратегії. Зокрема, можна застосовувати бутстрепінг для вибірок, підвибірки ознак для зменшення вимірності, та варіативність ініціалізацій для підвищення рівня декореляції. Комплексне використання цих методів дозволяє одержати різноманітність для ансамблю, що забезпечить можливість компенсації недоліків кожної окремої моделі в процесі аналізу обмежених та гетерогенних вибірок біомедичних даних [195].

Ще одним ключовим критерієм класифікації ансамблів ШНМ є тип базових моделей, що використовуються у їхньому складі. Залежно від архітектурної однорідності виділяють гомогенні (однорідні) та гетерогенні (неоднорідні) ансамблі [21], [207]. Цей параметр визначає не лише структурну організацію ансамблю, але й рівень внутрішньої різноманітності моделей, що суттєво впливає на його властивість узагальнення.

Гомогенні ансамблі складаються з однакових за типом базових моделей. Усі вони навчаються на незалежно створених (декорельованих) вибірках/підвибірках даних або з різними параметрами для кожного члена ансамблю. Проте архітектурно всі члени ансамблю залишаються однаковими. Основна перевага гомогенних ансамблів полягає у простоті реалізації, зменшенні дисперсії та високій масштабованості. У задачах з малими біомедичними вибірками гомогенні ансамблі дають змогу уникнути перенавчання, зокрема, шляхом багатократного усереднення нестабільних оцінок [211].

Гетерогенні ансамблі формуються з моделей різної природи. Такий підхід дозволяє кожній моделі спеціалізуватися на різних аспектах даних або типах

структурних залежностей [207]. У біомедичних задачах це особливо цінно, оскільки дані можуть мати ієрархічну природу, наприклад, поєднувати молекулярні, фізіологічні та клінічні показники. Основна перевага гетерогенних ансамблів – висока різноманітність рішень, що призводить до підвищеного рівня властивості узагальнення та стійкості до похибок окремих моделей [200]. Водночас їх недоліком є складність налаштування та потреба у спеціалізованому механізмі агрегування, наприклад, стекінг або зважене голосування, що може бути критичним обмеженням при роботі з малим обсягом навчальних даних. У таких випадках доцільним є використання слабко параметризованих моделей на мета-рівні або застосування крос-валідаційних стратегій [212].

Загалом, вибір між гомогенним і гетерогенним ансамблем має базуватись на співвідношенні інформаційного обсягу вибірки, складності задачі та потреби в інтерпретованості. У випадку аналізу екстремально коротких вибірок даних доцільно розпочинати з гомогенних ансамблів на базі стабільних, неітеративних моделей. У випадку доступності структурно різнорідних даних або наявності попередньо навчених моделей потрібно поступово переходити до гетерогенних конфігурацій. Обидва підходи можуть бути поєднані в ієрархічну або каскадну архітектуру, що забезпечує гнучке балансування між стабільністю та варіативністю.

Попри численні переваги всі типи ансамблів мають специфічні обмеження, які особливо проявляються при роботі з біомедичними вибірками обмеженого обсягу. Зокрема, методи усереднення та голосування не враховують взаємозв'язки між моделями, стекінгові ансамблі потребують додаткової валідаційної вибірки, а бустингові ансамблі чутливі до шуму. Ансамблі з класу бегінг не компенсують систематичних похибок, а каскадні ансамблі схильні до накопичення похибок. Методи декореляції такі як бутстрепінг, підвибірki ознак, випадкові ініціалізації підвищують варіативність. Проте, при обмеженій кількості даних зростає ризик дублювання або втрати інформативності. Гомогенні ансамблі є простими, але менш

виразними, тоді як гетерогенні – гнучкі, проте складні у налаштуванні, що ускладнює їх застосування на вибірках даних обмеженого обсягу.

З огляду на вищевикладене, виникає необхідність удосконалення існуючих та розроблення нових ансамблів ШНМ, здатних ефективно працювати в умовах обмеженого обсягу біомедичних даних.

1.6 Формулювання мети і постановка задач дослідження

Метою роботи є розвиток основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

Для досягнення мети дисертаційної роботи необхідно вирішити такі завдання:

1. Проаналізувати методи, моделі та засоби інтелектуального опрацювання біомедичних даних малих обсягів.
2. Розвинути метод кусково-лінійної апроксимації для моделювання багатовимірних поверхонь відгуку.
3. Удосконалити метод формування комбінованої поверхні відгуку меншої вимірності з застосуванням кластерного аналізу на вибірках даних малих обсягів.
4. Розробити метод нелінійного доповнення простору незалежних ознак біомедичної вибірки даних вихідними сигналами шару підсумовування ймовірнісної нейронної мережі.
5. Розробити методологію розв'язання задач регресії та класифікації на екстремально-коротких біомедичних вибірках з використанням аугментації даних.
6. Розробити метод синтезу ансамблів лінійних нейроподібних структур з неітеративним машинним навчанням з використанням різних підходів до агрегування кусково-лінійно апроксимованих поверхонь відгуку кожного його

елемента.

7. Розробити метод синтезу ансамблю гомогенних нелінійних штучних нейронних мереж, що функціонують на незалежно сформованих вибірках даних з використанням симетричних випадкових зміщень та з різними способами агрегування вихідних результатів окремих моделей.

8. Розробити метод часткової компенсації систематичної похибки нейронної мережі узагальненої регресії.

9. Удосконалити метод глобально-локальної апроксимації шляхом незалежного виділення глобальної та локальної компонент сигналу ансамблю нейронних мереж.

10. Розробити прямий та зворотний методи компенсації похибок регресійних моделей з використанням каскадних ансамблів на основі раціональних дробів.

11. Розвинути теорію синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням на вибірках біомедичних даних малих обсягів.

12. Розробити програмне забезпечення для реалізації усіх методів та виконати оцінювання їх результатів на основі як модельних даних, так і даних біомедичних досліджень.

1.7. Висновки до розділу

1. Аналіз сутності біомедичних даних та їх характеристик дозволив встановити, що такі дані мають низку специфічних властивостей. Це, зокрема, неоднорідність, висока вимірність, змінний обсяг, наявність пропусків, шумів та аномалій тощо. Встановлено, що в умовах екстремально коротких обсягів біомедичних вибірок даних порушуються класичні припущення статистичної теорії узагальнення. Це вимагає використання моделей ШНМ, стійких до обмеженого

обсягу даних, неоднорідності та високої варіативності інформації в процесі аналізу таких даних.

2. Узагальнений аналіз типових умов формування біомедичних вибірок даних дозволив сформулювати науково обґрунтоване визначення малих даних як вибірок обмеженого обсягу, високої вимірності вхідних даних задачі та нерегулярної структури збору. Вони виникають під впливом етичних, фінансових, часових або технічних обмежень. Особливої уваги потребують екстремально-короткі вибірки даних, які не задовольняють класичних статистичних припущень і, зазвичай, унеможливають гарантії узагальнення з використанням існуючих методів машинного навчання.

3. Встановлено, що біомедичні дані малих обсягів характеризуються рядом критичних обмежень, які ускладнюють або й взагалі унеможливають застосування класичних методів машинного навчання та штучних нейронних мереж. До ключових проблем належать: недостатній обсяг вибірки, висока вимірність простору вхідних даних задачі, наявність пропусків, шумів та аномалій, а також невизначеність у розподілах і обмежені можливості валідації моделей. Ці фактори призводять до порушення умов статистичної узагальнюваності й суттєво підвищують ризик перенавчання.

4. Узагальнення зазначених труднощів дозволило сформулювати науково-прикладну проблему цього дослідження, що полягає у необхідності розвитку основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

5. Встановлено, що лінійні методи, зокрема, лінійні нейроподібні структури МПГП, є ефективними для опрацювання високовимірних біомедичних вибірок малого обсягу завдяки своїй інтерпретованості, неітеративному навчанню

та стійкості до мультиколінеарності. Проте обмеження, пов'язані з апіорною лінійністю моделі, потребують розроблення і застосування методів нелінійного розширення вхідного простору даних для підвищення точності аналізу.

6. Показано, що НМУР та ЙНМ демонструють високу точність і робастність у задачах, де доступна лише обмежена вибірка спостережень. Завдяки своїй інтерполяційній природі та здатності до локального навчання ці моделі адаптивно описують складні залежності навіть за високого рівня шуму. Проте, у випадку критично малих вибірок даних, де обсяг даних не забезпечує навіть мінімальної статистичної стабільності, точність таких моделей може бути недостатньою.

7. Показано необхідність використання методів аугментації табличних даних (розширення обсягу вибірки даних), що дозволяє підвищити щільність покриття вхідного простору даних та знизити ризик перенавчання. Аналіз сучасних підходів (від простих перетворень до генеративних моделей) засвідчив, що аугментація даних може суттєво покращити ефективність аналізу за умови збереження структурної, статистичної та клінічної валідності даних. Водночас ефективне використання цих підходів у біомедичних задачах з обмеженим обсягом даних потребує розроблення нових методів з урахуванням специфіки високорозмірних та гетерогенних біомедичних вибірок даних.

8. Аналіз показав, що ансамблі моделей штучних нейронних мереж здатні суттєво підвищити точність, стійкість та властивості узагальнення в процесі розв'язання задач регресії/класифікації на біомедичних вибірках обмеженого обсягу. Позитивний ефект ансамблевого навчання обґрунтовується як класичними теоріями, зокрема, розкладом середньої квадратичної похибки та теоремою Кондорсе, так і сучасними підходами до врахування диверсифікації моделей у складі ансамблю.

9. Показано, що ефективність ансамблів в умовах аналізу малих обсягів даних значною мірою залежить від типу базових моделей, методу декореляції вибірок даних, механізму агрегування результатів та конфігурації ансамблю (гомогенної або гетерогенної). Встановлено, що для забезпечення високої точності, робастності та властивостей узагальнення в умовах аналізу малих біомедичних вибірок даних необхідним є розроблення ансамблів моделей неітеративного машинного навчання, які враховуватимуть обмеженість даних, варіативність ознак та специфіку задач аналізу біомедичних даних.

РОЗДІЛ 2

МЕТОДИ НЕЛІНІЙНИХ ПЕРЕТВОРЕНЬ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ ДЛЯ ФОРМУВАННЯ ПРОСТОРІВ ОЗНАК ВИЩОЇ ВИМІРНОСТІ

У розділі 2 викладено метод нелінійного доповнення входів лінійної неітеративної нейроподібної структури МПГП членами полінома Колмогорова-Габора, а також метод визначення коефіцієнтів цього полінома. Розроблено метод розв'язання задач класифікації, який ґрунтується на розширенні входів ШНМ вихідними сигналами шару підсумовування ймовірнісної нейронної мережі. Удосконалено метод формування комбінованої поверхні відгуку з застосуванням кластерного аналізу для підвищення точності розв'язання задач регресії та класифікації у випадку аналізу даних малих обсягів. Розвинуто одновимірний метод кусково-лінійної апроксимації для аналізу багатовимірних поверхонь відгуку, який збільшує вимірність простору вхідних даних на основі ізоморфного подання кожного одиничного значення ознаки як сукупності локальних ознак і дозволяє моделі локально коригувати параметри апроксимації досліджуваної залежності.

Наукові та практичні результати, описані в цьому розділі, висвітлено в періодичних виданнях України та у високореєтингових журналах інших держав, які включено до наукометричних баз Scopus та Web of Science: [4], [34], [35], [43], [44], а також апробовано і опубліковано у матеріалах українських та закордонних конференцій: [19], [20], [36], [41], [42], [49], [50], [51]

2.1 Розширення простору вхідних компонент векторів даних для нейроподібних структур МПГП

У багатьох задачах машинного навчання одним з важливих етапів опрацювання даних розглядається перетворення вхідних даних у простір вищої вимірності, де можна застосовувати лінійні моделі для розв'язання задач, які в

оригінальному просторі мають складну нелінійну природу. Зокрема, метод нелінійних розширень для підвищення точності розв'язання задач регресії з використанням ШНМ прямого поширення достатньо загально подається Йок-Кхан Пао [142]. Для підвищення точності розв'язання задач класифікації такий підхід обґрунтовується теоремою Ковера [143]. Згідно з цією теоремою, якщо деякі вхідні дані, що спочатку не є лінійно роздільними в їх первинному (низьковимірному) просторі, підлягають нелінійному перетворенню у простір достатньо високої вимірності, то з дуже високою ймовірністю вони стають лінійно роздільними [143].

Завдання цього розділу полягає у формуванні таких нелінійних перетворень, які проєктують вхідні дані в простір вищої вимірності, де залежності між розширеними змінними можуть бути лінійними. Це дозволить застосовувати лінійні неітеративні нейроподібні структури МПГП [134], а також синтезовані на їх основі аналітичні вирази [20], [213] для подальшого аналізу або класифікації даних.

Формальна постановка задачі виглядає так. Нехай маємо навчальну вибірку малого обсягу даних:

$$A \in \mathbb{R}^{n \times m}, \quad (2.1)$$

де n – кількість векторів, а m – вимірність кожного вектора. При цьому кожен рядок A позначимо як $x_i \in \mathbb{R}^m$ для $i = \overline{1, n}$, а відповідні значення залежного атрибута y_i для кожного $x_i \rightarrow y_i$ подані одним з варіантів:

$$\begin{cases} y_i \in \{1, 2, \dots, K\}, & \text{для задач класифікації;} \\ y_i \in \mathbb{R}, & \text{для задач регресійного аналізу.} \end{cases} \quad (2.2)$$

Проблема полягає у тому, щоб знайти нелінійне відображення:

$$\phi: \mathbb{R}^m \rightarrow \mathbb{R}^H \quad (2.3)$$

таке, яке відображатиме кожен вектор x_i у новий простір вищої вимірності H ($H > m$), таким чином, щоб залежність між x_i та y_i у цьому просторі була максимально лінійною.

Проекція x_i у простір вищої вимірності, зокрема, через відображення (2.3) значно збільшить ймовірність лінійного розділення даних, які в початковому просторі демонструють складні нелінійні залежності. Окрім цього, само відображення ϕ наповнює вектор ознак новими композиціями початкових координат – поліноміальними, радіальними тощо [214]. Завдяки цьому ШНМ отримує додаткові ступені свободи, які дозволяють йому краще апроксимувати складні функції залежності між вхідними даними та вихідним (залежним) атрибутом. Таке збагачення ознак може підвищити гнучкість моделі і дати змогу знайти і врахувати самі тонкі взаємозв'язки, недоступні у початковому просторі вхідних даних.

Водночас збільшення вимірності простору несе з собою й ряд ризиків. Зокрема, без достатнього обсягу даних для реалізації процедур навчання обраних засобів машинного навчання, висока кількість ознак може призвести до “прокляття вимірності” та значного підвищення варіації оцінок моделі [82]. Окрім цього, такий підхід може суттєво зменшити властивості узагальнення обраним засобом машинного навчання і навіть спричинити перенавчання. Саме тому виникає завдання розроблення нових та удосконалення існуючих методів нелінійного перетворення вхідних даних для формування простору ознак вищої вимірності в умовах аналізу даних малих обсягів. Вибір конкретного перетворення ϕ і контроль складності моделі мають базуватися на балансі між зменшенням систематичної похибки й обмеженням надлишкового ускладнення.

2.2 Доповнення входів нейроподібних структур МПГП нелінійними членами поліному Колмогорова-Габора

Підвищення точності розв'язання задач апроксимації нелінійних залежностей відносять до однієї з важливих задач обчислювального інтелекту. Існуючі, зокрема, лінійні методи машинного навчання не завжди забезпечують можливість їх використання для отримання достатньо точних результатів розв'язання цієї задачі.

Одним з підходів щодо підвищення точності роботи таких моделей є нелінійне розширення вхідних даних. Воно дозволяє перетворити простір вхідних даних так, щоб побудувати більш інформативні характеристики для подальшого інтелектуального аналізу.

Серед ефективних методів нелінійного розширення є використання поліному Колмогорова-Габора (або розклад Іто, або поліном Вінера), який застосовується для розширення простору вхідних даних з подальшим його використанням штучними нейронними мережами (ШНМ) чи методами машинного навчання [215], [43], [216]. Такий підхід використовується вже доволі давно для розроблення нелінійних моделей апроксимації [149]. Його використання обґрунтовується високими апроксимаційними властивостями цього полінома, який, по суті, є дискретним аналогом ряду Вольтерри [43], [213].

В Україні цей поліном почав використовуватися Івахненком О.Г. у його відомому Методі Групового Врахування Аргументу (МГУА) [217], а також продовжує активно використовуватися сьогодні, в процесі розв'язання різноманітних задач апроксимації, у працях відомих вчених Зайченка О.П., Снитюка Є.В.

Якщо говорити про задачу апроксимації, то теоретичним обґрунтуванням використання цього підходу є теорема Вєрштраса. Вона стверджує, що будь-яка неперервна функція на заданому інтервалі може бути як завгодно точно описана рядом поліномів [218]. Якщо говорити про задачу класифікації, то використання такого підходу також знаходить своє місце і теоретично обґрунтовується теоремою Ковера, яка стверджує що нелінійне розширення простору вхідних даних задачі збільшує ймовірність правильної класифікації лінійними методами [143].

Загальний вигляд поліному Колмогорова-Габора можна подати наступним чином [216]:

$$Y(x_1, \dots, x_n) = \theta_0 + \sum_{i=1}^n \theta_i x_i + \sum_{i=1}^n \sum_{j=1}^n \theta_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \theta_{ijk} x_i x_j x_k + \dots, \quad (2.4)$$

де n – число вхідних ознак, k – максимальний ступінь полінома, θ – коефіцієнти при членах полінома.

Як видно з (2.4), поліном Вінера виражається через члени, що є ступенями вхідних змінних та їх добутками. Такі перетворення забезпечують значне розширення інформації. Нейронна мережа може використовувати його для підвищення точності апроксимації за рахунок синтезу нових ознак, які виходять за межі лінійної комбінації вхідних даних. Саме тому у реальних задачах, де залежності між біомедичними даними є складними і нелінійними, подібна модель може бути ефективною.

Зайченко Ю.П. [219], а згодом і Снитюк В.Є. [220] пропонують використовувати дещо інший вигляд цього полінома:

$$Y(x_1, \dots, x_n) = \theta_0 + \sum_{i=1}^n \theta_i x_i + \sum_{i=1}^n \sum_{j=i}^n \theta_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=i}^n \sum_{k=j}^n \theta_{ijk} x_i x_j x_k + \dots. \quad (2.5)$$

Загалом, використання (2.5) замість (2.4) обумовлено необхідністю уникнення симетричних опорних функцій в поліномі, що може змінити результати апроксимації відповідно до обраних критеріїв якості. Зокрема, у таблиці 2.1 показано, як зведене упорядковане сумування в поліномі (2.5) суттєво зменшує кількість параметрів і обчислювальну складність порівняно з повним класичним поліномом (2.4), при цьому знижуючи ризик перенавчання.

Однак, не зважаючи на те, який з варіантів полінома використовувати (2.4) чи (2.5), проблемою на практиці залишається визначення значень коефіцієнтів θ для кількості коефіцієнтів q цього полінома. Використання методів лінійної регресії різних варіантів [221] через велику обчислювальну складність і належність завдання до класу майже вироджених задач ускладнює точне їх розв'язання.

Таблиця 2.1 – Порівняння характеристик форм полінома Колмогорова-Габора (2.4)-(2.5).

Показник	Форма полінома (2.4)	Форма полінома (2.5)
Кількість коефіцієнтів	$\sum_{q=0}^k n^q$	$\sum_{q=0}^k \binom{n+q-1}{q}$
Обчислювальна складність	$O(n^q)$	$O\left(\sum_{q=0}^k \binom{n+q-1}{q}\right)$
Перенавчання	Дуже високий ризик	Менший ризик, але залежить від k

У цьому розділі описано метод розширення входів ШНМ поліномом Колмогорова-Габора, а також розроблено метод визначення коефіцієнтів при членах цього полінома, побудованого на основі (2.5), з використанням лінійної, неітеративної нейроподібної структури Моделі Послідовних Геометричних Перетворень (МППГ), розробленої Ткаченком Р.О. [19].

Такий підхід забезпечує суттєве підвищення точності розв'язання задач класифікації чи регресії, а також можливість переходу від нейромережевої форми застосування моделі до поліноміальної на основі можливості синтезу коефіцієнтів при членах полінома Колмогорова-Габора [43], [45].

Структурно-функційну схему роботи цього методу подано на рис. 2.1. Схема містить два блоки. Перший з них (генератор членів полінома Колмогорова-Габора) призначений для формування членів цього полінома. Це відбувається згідно (2.5) для заданих значень кількості вхідних змінних n і степеня полінома k . Члени полінома Вінера формуються послідовно для $k = 1, 2, 3, \dots$

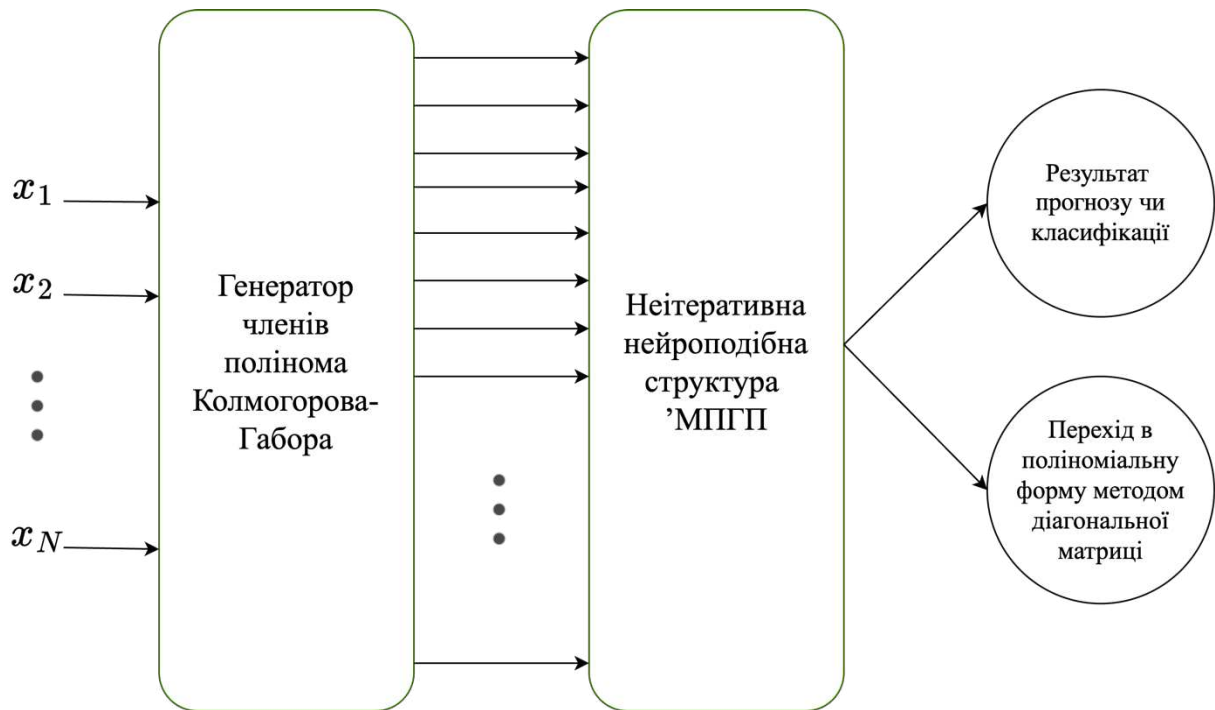


Рисунок 2.1 – Структурно-функційна схема нелінійного розширення входів нейроподібної структури МПГП членами полінома Колмогорова-Габора

Наступний блок схеми – це лінійний варіант неітеративної нейроподібної структури МПГП. На її входи подаємо розширений набір даних, який містить члени полінома Колмогорова-Габора, сформовані для первинних входів (значень незалежних змінних конкретної задачі) та заданого степені поліному.

Таке поліноміальне розширення входів лінійного варіанту МПГП забезпечить можливість підвищення точності МПГП без алгоритмічного ускладнення методів її навчання та функціонування [43], [216]. Це відбувається за рахунок подолання обмежень лінійної апроксимації. Включення до вхідного простору додаткових змінних, що відповідають добуткам, степеням та поліноміальним перетинам початкових ознак, трансформує задачу у простір вищої вимірності, у якому лінійна модель здатна апроксимувати складні нелінійні залежності. Це особливо важливо в умовах аналізу біомедичних вибірок, де взаємозв'язки між параметрами часто є нелійними. Завдяки такому розширенню модель зберігає простоту методу навчання, одночасно покращуючи здатність до

узагальнення, стійкість до шуму та адаптивність до прихованих взаємодій між змінними. Окрім цього, враховуючи неітеративну природу методу навчання МПГП, такий підхід забезпечує високу швидкодію опрацювання даних, зокрема, в порівнянні з методом найменших квадратів чи сингулярною декомпозицією [216].

Окрім цього, розроблена схема надає можливість представити результат в компактному вигляді полінома Колмогорова-Габора [213]. Це можливо, зважаючи на повторюваність розв'язку, яка забезпечується обраним інструментом навчання. Для цього необхідно синтезувати коефіцієнти полінома Колмогорова-Габора [20]. Представимо основні кроки цього методу, беручи до уваги попередньо навчену нейроподібну структуру МПГП, яка використовує розширені вектори, отримані з використанням (2.4).

Нехай для кожного вхідного вектора $x_i \in \mathbb{R}^m, i = \overline{1, n}$ отримано розширений вектор:

$$x_i^+ = \phi(x_i) \in \mathbb{R}^g \quad (2.6)$$

за допомогою полінома Колмогорова-Габора другого степеня (2.4).

Метод визначення коефіцієнтів при членах полінома Колмогорова-Габора складається з таких кроків:

1. Визначити кількість незалежних атрибутів g кожного розширеного вектора даних. Зокрема, g для x^+ при використанні другого степеня полінома дорівнює числу всіх лінійних і квадратичних мономів [33], [43]:

$$g(m) = m + \frac{m(m+1)}{2}, \quad (2.7)$$

де m – кількість початкових входів задачі.

2. Сформувати матрицю тестових сигналів T :

$$T = \begin{pmatrix} 0_{1 \times g} \\ I_g \end{pmatrix} \in \mathbb{R}^{(g+1) \times g}, \quad (2.8)$$

де $0_{1 \times g}$ - нульовий вектор довжини g (перший рядок матриці T), I_g - одинична матриця розміром $g \times g$.

3. Отримати вектор коефіцієнтів поверхні відгуку шляхом застосування матриці T як вхідного сигналу попередньо навченої лінійної неітеративної нейроподібної структури МПГП:

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_{g+1} \end{pmatrix} \in \mathbb{R}^{g+1}, \quad (2.9)$$

де $\alpha = SGTM(T)$ – результат подачі T на вхід попередньо натренованої нейроподібної структури МПГП (SGTM).

4. Сформувати коефіцієнти лінійного полінома згідно такого виразу:

$$\theta_t = \begin{cases} \alpha_1, & \text{якщо } t = 1, \\ \alpha_t - \alpha_1, & \text{якщо } t = 2, g+1. \end{cases} \quad (2.10)$$

5. Сформувати лінійний поліном для розв'язання поставленої задачі лінійної нейроподібною структурою МПГП з нелінійним розширенням входів поліномом Колмогорова-Габора другого степеня

$$y^{pred} = \theta_1 + \sum_{t=2}^g \theta_t x_{t-1}^+. \quad (2.11)$$

Поліноміальне представлення результатів задач регресії чи класифікації, зокрема, в області біомедичної інженерії дозволяє отримати чітку аналітичну структуру моделі, яка чітко відображає взаємозв'язки між численними медико-біологічними та інженерно-технічними показниками. Завдяки тому, що кінцева функція y^{pred} задається лінійним поліномом над розширеними ознаками, отриманими через поліном Колмогорова – Габора, можна відразу побачити, які саме взаємодії між вхідними змінними враховано, та інтерпретувати значення кожного коефіцієнта θ_t . Такий підхід значно підвищує інтерпретованість моделі, що робить процес прийняття рішень прозорішим, зокрема, в порівнянні з “чорними скриньками” глибоких нейромереж. Як наслідок, така форма представлення

результату викликає більшу довіру з боку клініцистів, навіть якщо вони не є експертами в області штучного інтелекту [45].

Обчислювальна ефективність є ще одною значною перевагою поліноміального представлення результатів роботи нейроподібної структури МПГП з нелінійним розширення входів поліномом Колмогорова-Габора другого ступеня. Зокрема, зменшення обчислювальної складності з $O(p^2 \times g)$ у випадку використання нейромережевого представлення методу [19], [134] до $O(p^2)$ у випадку використання поліноміального представлення методу (2.11) [216], [222] дозволяє розробляти високоефективні програмні та апаратні засоби для інтелектуального аналізу даних у реальних біомедичних застосуваннях.

2.3 Метод розширення входів нейроподібних засобів на основі вихідних сигналів шару сумування ймовірнісної нейронної мережі

Штучні нейронні мережі (ШНМ) є потужними інструментами для вирішення практичних задач в області біомедичних досліджень. Однією з важливих складових ефективності нейронної мережі є правильне формулювання вхідних даних і способу їх обробки. Як вже зазначалося у підрозділі 2.1, нелінійне розширення входів, зокрема, членами полінома Колмогорова-Габора забезпечує підвищення точності регресії (теорема Веєрштраса) чи класифікації (теорема Ковера) даних. Проте серед особливостей біомедичних даних можна виділити велику кількість незалежних атрибутів, які слід враховувати в процесі аналізу, та можливий обмежений обсяг даних для реалізації процедури навчання. Враховуючи ці дві обставини, використання полінома Колмогорова-Габора може бути недоречним в процесі розв'язання ряду задач. Це пояснюється збільшенням вимірності простору вхідних даних задачі та супроводжуватиметься значною тривалістю процедури навчання у випадку вибору високих ступенів цього полінома [213], [223].

Іншим можливим варіантом розширення простору вхідних даних задачі є вихідні сигнали НМУР чи ЙНМ, які ефективно працюють з короткими вибірками

даних, забезпечують високі властивості узагальнення та можуть використовуватися як засоби попереднього опрацювання даних.

У випадку використання ймовірнісної нейронної мережі можливі два варіанти механізму виводу результатів: у вигляді мітки (маркеру) класу, до якого належить спостереження, або у вигляді набору ймовірностей приналежності до кожного класу, визначених задачею [36], [49], [154]. Останній підхід можна використовувати для синтезу нових ознак для вхідних даних, що відображають більш складні ймовірнісні взаємозв'язки між змінними. Наприклад, у задачу можна вводити нові ознаки - ймовірності належності до певних класів або варіації цих ймовірностей через їх сумування, середнє чи використання інших функцій [34], [35], [224]. Таке розширення може бути особливо корисним у випадках, коли вихідні дані мають високий рівень невизначеності, або коли прості лінійні чи навіть нелінійні зв'язки між вхідними ознаками не можуть адекватно охопити всі складні взаємозалежності у короткій вибірці біомедичних даних.

Загалом подібні методи можуть значно підвищити здатність нейронної мережі до навчання на складних даних, що мають високу варіативність або невизначеність. Проте перш, ніж аналізувати їх детальніше, розглянемо функціонування та можливі варіанти формування вихідного сигналу ймовірнісної нейронної мережі.

Топологія ймовірнісної нейронної мережі (ЙНМ) має чотири шари: вхідний, шар образів, шар сумування і вихідний шар [49], [225]. Її наведено на рис. 2.2.

Проаналізуємо основні кроки роботи ЙНМ, а особливо можливі варіанти формування ймовірностей приналежності до кожного класу задачі у шарі сумування топології ЙНМ.

Нехай кожен вектор $x_i \in \mathbb{R}^m$ для $i = \overline{1, n}$, заданої опорної вибірки даних A (2.1) подано у вигляді $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})$. Відповідний залежний атрибут визначається як у (2.2).

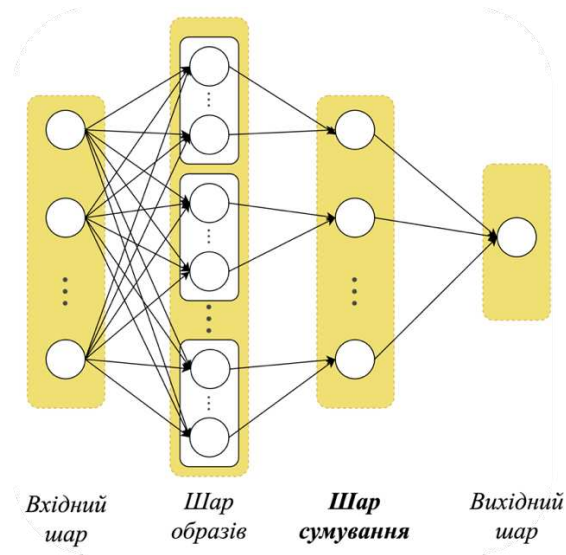


Рисунок 2.2 – Топологія ймовірнісної нейронної мережі

Позначимо тестовий вектор $u_i = (u_{e,1}, u_{e,2}, \dots, u_{e,m})$, $e = \overline{1, U}$, для якого невідома належність до одного з визначених класів задачі, тобто відсутній вихідний елемент y_{u_e} .

Нехай існує K класів і N_k – кількість навчальних векторів у класі k , причому $\sum_{k=1}^K N_k = n$.

Реалізація ЙМН передбачає виконання наступних кроків:

1. Визначити кількість вхідних атрибутів та заданих класів задачі.
2. Побудувати топологію ЙМН відповідно до кроку 1.
3. Виконати нормування усіх векторів x_i :

$$\acute{x}_i = \mathcal{N}(x_i), i = \overline{1, n}, \quad (2.12)$$

де $\mathcal{N}(\cdot)$ - обрана функція нормалізації.

4. Виконати нормування поточного вектора u_i з невідомим виходом на основі $\mathcal{N}(\cdot)$, отриманої у п. 3: $\acute{u}_e = \mathcal{N}(u_e)$, $e = \overline{1, U}$.
5. Обчислити Евклідові відстані D_{e,i_k} між $\acute{u}_i$ та усіх $\acute{x}_i$ для кожного з представлених класів задачі k , $i_k = 1$, k згідно виразу:

$$D_{e,i_k} = \sqrt{\sum_{j=1}^m (\dot{u}_{e,j} - \dot{x}_{i_k,j})^2}, e = \overline{1 \ U}, i_k = \overline{1 \ N_k}, k = \overline{1 \ K}. \quad (2.13)$$

6. Обчислити гаусівські функції G_{e,i_k} від (2.13) згідно виразу:

$$G_{e,i_k} = \exp\left(-\frac{(D_{e,i_k})^2}{\sigma^2}\right), \sigma > 0, \quad (2.14)$$

де σ – розмах гаусівської функції, єдиний параметр ЙНМ, який треба налаштовувати.

7. Обчислити ймовірності $P_{raw}(e, k)$ приналежності поточного вектора $\bar{x}_{m,j}$ до кожного з заданих класів задачі [49]:

$$P_{raw}(e, k) = \frac{\sum_{i_k=1}^{N_k} G_{e,i_k}}{N_k}. \quad (2.15)$$

Вираз (2.15) відповідає процедурі формування ймовірностей при використанні класичного варіанту ЙМН [49], [154]. Проте цей вираз не відповідає умові (2.16):

$$\sum_{k=1}^K P_{raw}(e, k) \neq 1, \quad (2.16)$$

тобто не забезпечує формування повної системи подій.

Така форма не є ефективною з точки зору прикладного застосування у різноманітних областях біомедичної інженерії. Це пояснюється тим, що у цій галузі важливо знати саме величину ймовірності настання кожної з подій при виконанні (2.16). Використання (2.15) може зменшувати інформативність результату для приймання на його основі відповідного рішення та, відповідно, зменшує довіру до використання подібного засобу для не фахівця в області штучного інтелекту.

У (2.17) цей недолік виправлено [34]:

$$P_{sum}(e, k) = \frac{\sum_{i_k=1}^{N_k} G_{e,i_k}}{\sum_{r=1}^K \sum_{j_r=1}^{N_r} G_{i,j_r}}, \quad (2.17)$$

де k – індекс того класу, для якого обчислюється ймовірність, r – додатковий індекс класу, що використовується у зовнішніх сумах у знаменнику.

Вираз (2.17) описує повну систему подій, що є перевагою в процесі аналізу даних, зокрема, в різних галузях біомедичної інженерії. Проте цей вираз не враховує нерівномірність подання кількості векторів з різних класів, бо це виконується у (2.5). Таким чином, менший клас подавляється більшим класом. У випадку необхідності розв'язання задач бінарної класифікації, які дуже часто виникають у біомедичній інженерії, при необхідності опрацювання обмежених обсягів даних та ще й їх незбалансованості це є суттєвим недоліком. Саме тому, пропонуємо третій підхід до формування вихідних сигналів шару сумування ЙНМ для визначення ймовірності $P(e, k)$ належності поточного вектора до одного з класів [36]:

$$P(e, k) = \frac{\frac{\sum_{i_k=1}^{N_h} G_{e, i_k}}{N_k}}{\sum_{r=1}^K \left(\frac{\sum_{j_r=1}^{N_r} G_{i, j_r}}{N_r} \right)}. \quad (2.18)$$

Таке, досить просте, проте ефективне рішення задовольняє обом вимогам: тут зберігається повна система подій та враховується нерівномірність подання усіх класів задачі у заданій вибірці даних [36].

8. Для кожного вектора з невідомим виходом знайти:

$$y_{u_e} = \arg \max_{k=1, K} P(e, k) \quad (2.19)$$

на основі (2.15) чи (2.17) чи (2.18). Максимальне значення y_{u_e} є вихідним сигналом роботи ЙМН, який визначатиме приналежність поточного вектора до одного з заданих класів задачі.

На основі вихідних сигналів шару підсумовування ЙНМ розроблено метод та ряд його варіацій для розширення простору вхідних даних щодо підвищення точності розв'язання задач класифікації лінійними ШНМ. В основі методу

покладено використання вихідних сигналів шару сумування ЙНМ, який формує набір ймовірностей належності кожного спостереження (вектора) до кожного з визначених задачею класів [34], [35], [36]. Слід зазначити, що тут можна використовувати також сигнали вихідного шару цієї ШНМ.

Розроблений метод складається з двох етапів: етап навчання та етап використання.

Нехай $x_i \in \mathbb{R}^m, i = \overline{1, n}$, вектор ознак i -того спостереження з опорної вибірки даних A (2.1). Задача передбачає існування K класів, причому в опорній вибірці клас k містить N_k векторів при умові, що $\sum_{k=1}^K N_k = n$.

Етап навчання передбачає послідовне виконання таких кроків:

1. Визначити набір ймовірностей належності кожного поточного спостереження $\acute{x}_i \in \mathbb{R}^m$ для $i = \overline{1, n}$, нормалізованого згідно (2.12) до кожного з k класів, $k = \overline{1, K}$:

$$p_i = \begin{pmatrix} P(i, 1) \\ P(i, 2) \\ \dots \\ P(i, K) \end{pmatrix}, \quad (2.20)$$

де кожне $P(i, k)$ обчислюється за однією з формул (2.15), (2.17) або (2.18). Слід зазначити, що в процесі визначення p_i для усіх x_i , поточні спостереження $\acute{x}_i$ тимчасово вилучаються з опорної вибірки.

2. Сформувати нові вектори виду:

- шляхом заміни початкових входів задачі $\acute{x}_i$ на набір ймовірностей, отриманих на кроці 1. Такий варіант може бути корисним в процесі розв'язання задач багатокласової класифікації у випадку аналізу високовимірних вибірок даних малих обсягів. Позначимо це як *Варіант 1*:

$$x_i^+ = p_i \in \mathbb{R}^K. \quad (2.21)$$

- шляхом розширення початкових входів задачі $\acute{x}_i$ набором ймовірностей, отриманих на кроці 1. Такий варіант забезпечить незначне розширення простору вхідних даних задачі і може забезпечити суттєве збільшення точності лінійних засобів машинного навчання. Позначимо це як *Варіант 2*:

$$x_i^+ = (\acute{x}_i, p_i) \in \mathbb{R}^{m+K}. \quad (2.22)$$

- шляхом нелінійного розширення векторів виду (2.22) членами полінома Колмогорова-Габора $Poly(\cdot)$. Такий варіант забезпечить додаткове моделювання взаємозв'язків між початковими входами задачі $\acute{x}_i$ та ймовірностями, визначеними ЙМН. Такий підхід може забезпечити підвищення точності аналізу низьковимірних біомедичних вибірок даних. Позначимо це як *Варіант 3*:

$$x_i^+ = Poly(\acute{x}_i, p_i) \in \mathbb{R}^g, \quad (2.23)$$

де $g = \dim(Poly(\mathbb{R}^{m+K}))$ – обчислюється за формулою (2.7).

3. Сформувані навчальну вибірку даних A^+ на основі (2.21), (2.22) чи (2.23):

$$A^+ = \begin{bmatrix} x_1^+ \\ x_2^+ \\ \dots \\ x_n^+ \end{bmatrix} \in \begin{cases} \mathbb{R}^{n \times K}, \text{ якщо } x_i^+ \text{ сформовано на основі (2.21),} \\ \mathbb{R}^{n \times (m+K)}, \text{ якщо } x_i^+ \text{ сформовано на основі (2.22),} \\ \mathbb{R}^{n \times g}, \text{ якщо } x_i^+ \text{ сформовано на основі (2.23).} \end{cases} \quad (2.24)$$

4. Виконати навчання лінійної нейроподібної структури $SGTM(\cdot)$ з використанням нової вибірки даних A^+ :

$$y_i^{pred} = SGTM(A^+), i = \overline{1, n}. \quad (2.25)$$

Слід зазначити, що y_i^{pred} використовується лише для оцінки похибок роботи досліджуваного методу на етапі навчання.

Розширення простору вхідних даних задачі на основі (2.24) для аналізу лінійними ШНМ повинно збільшити ймовірність правильної класифікації даних, зокрема, лінійними методами (теорема Ковера).

На рис. 2.3, рис. 2.4 та рис. 2.5 відповідно подано візуалізації схем етапу навчання розробленого методу з різними варіантами розширення входів лінійних ШНМ на основі вихідних сигналів шару сумування ЙНМ.

Етап навчання передбачає послідовне виконання таких кроків:

1. Визначити набір ймовірностей належності кожного поточного спостереження $\hat{u}_e \in \mathbb{R}^m$ для $e = \overline{1, U}$, нормалізованого згідно (2.12), до кожного з k класів, $k = \overline{1, K}$:

$$p_e = \begin{pmatrix} P(e, 1) \\ P(e, 2) \\ \dots \\ P(e, K) \end{pmatrix}, \quad (2.26)$$

де кожне $P(e, K)$ обчислюється за однією з формул (2.15), (2.17) або (2.18).

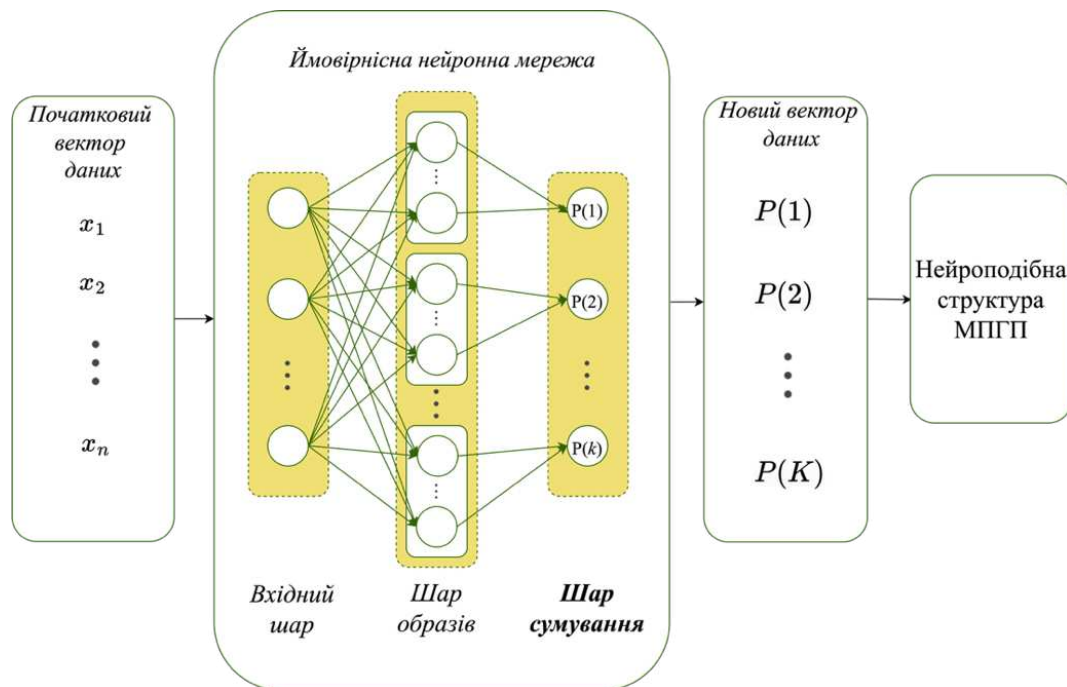


Рисунок 2.3 – Структурно-функційна схема розробленого методу з першим варіантом розширення входів на основі ЙНМ

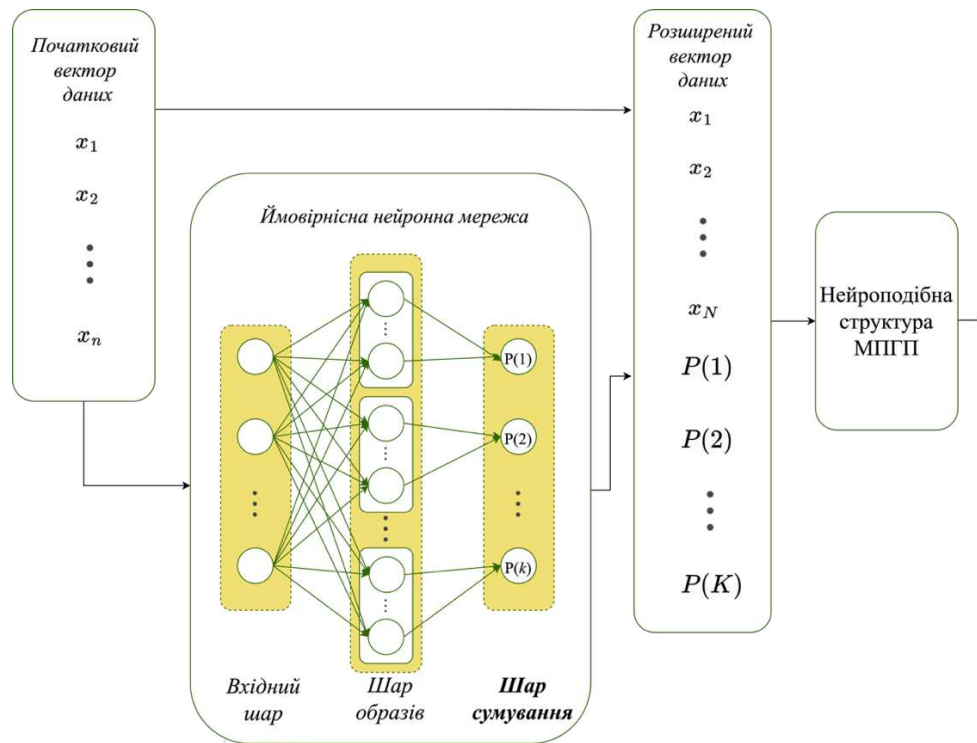


Рисунок 2.4 – Структурно-функційна схема розробленого методу з другим варіантом розширення входів на основі ЙНМ

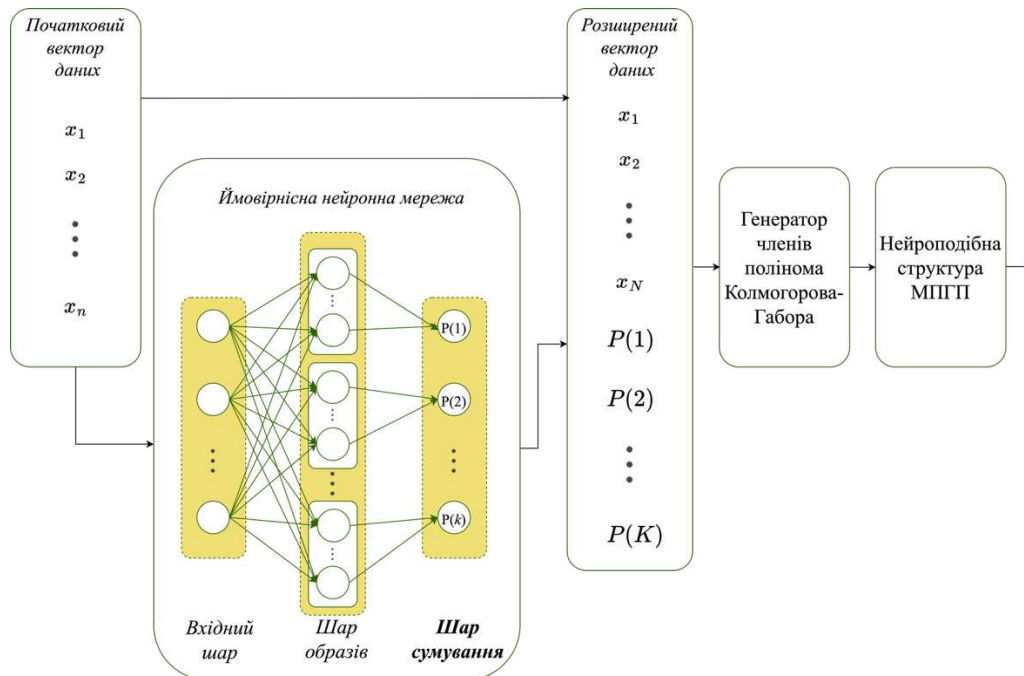


Рисунок 2.5 – Структурно-функційна схема розробленого методу з третім варіантом розширення входів на основі ЙНМ

2. Сформуувати розширені вектори ознак:

- у випадку використання Варіанту 1:

$$u_e^+ = p_e \in \mathbb{R}^K. \quad (2.27)$$

- у випадку використання Варіанту 2:

$$u_e^+ = (\acute{u}_e, p_e) \in \mathbb{R}^{m+K}. \quad (2.28)$$

- у випадку використання Варіанту 3:

$$u_e^+ = Poly(\acute{u}_e, p_e) \in \mathbb{R}^g, \quad (2.29)$$

де g – обчислюється за формулою (2.7).

3. Застосувати (2.27), (2.28) чи (2.29) на попередньо навченій лінійній нейроподібній структурі $SGTM(\cdot)$ для отримання вихідного сигналу:

$$y_e^{pred} = SGTM(u_e^+), i = \overline{1 \ n}. \quad (2.30)$$

Проаналізуємо переваги кожної з варіацій розробленого методу.

Використання розробленого методу з першим варіантом розширення входів згідно (2.21), який замінює початкові ознаки виключно на вектор ймовірностей $p_i \in \mathbb{R}^K$, забезпечує різке пониження вимірності вхідного простору до числа класів. Це корисно, коли вибірка високовимірна, первинні ознаки мають високу кореляцію або шум, або сама задача класифікації вже добре апроксимується моделлю ЙНМ. Використання лише p_i підвищує стійкість до перенавчання та значно зменшує обчислювальні витрати для лінійної ШНМ. Це пояснюється тим, що методу не доводиться опрацьовувати великомасштабні вектори m -вимірного простору. Такий підхід особливо доцільний у випадках, де важливо швидке прийняття рішення на обмеженому апаратному ресурсі, а точність попередньої оцінки класової належності все ще є недостатньою.

Розроблений метод з другим варіантом розширення входів згідно (2.22) поєднує вихідний вектор нормованих ознак $\acute{x}_i \in \mathbb{R}^m$ з вектором ймовірностей

$p_i \in \mathbb{R}^K$, створюючи розширений простір вимірності $m + K$. Такий підхід дозволяє зберегти детальну інформацію про початкові атрибути, одночасно підсилюючи модель даними про ймовірні класи. Це забезпечує врахування взаємозв'язків між ознаками та ймовірнісними сигналами шарів підсумовування ЙМН, що підвищує лінійну роздільність даних та, відповідно, точність класифікації. За рахунок помірною зростання вимірності вхідних даних, використання (2.22) зберігає обчислювальну ефективність, залишаючи об'єм параметрів у прийнятних межах для лінійної ШНМ, що забезпечує виконання умов теореми Вапніка-Червоненкіса [79], [80].

Розроблений метод з третім варіантом розширення входів згідно (2.23) передбачає нелінійне перетворення входів $x_e^+ \in \mathbb{R}^g$ поліномом Колмогорова-Габора, які вже містять як початкові ознаки, так і ймовірності належності до усіх класів задачі. Такий підхід дозволяє враховувати складні взаємозв'язки не лише серед початкових ознак, але й між останніми ознаками та ймовірностями, а також самими ймовірнісними сигналами. Завдяки теоремі Ковера лінійна роздільність класів у цьому багатовимірному просторі зростає, що особливо важливо при аналізі малих або нерівномірних біомедичних вибірок. Хоча вимірність $g > (m + K)$, підбір оптимального степеня полінома разом з регуляризацією може забезпечити уникнення перенавчання, отримуючи при цьому найвищу точність на складних нелінійних задачах.

2.4 Метод формування комбінованої поверхні відгуку на основі виділених кластерів

Важливою складовою компонентою сучасних задач МН є ефективне використання вхідних даних, оскільки від цього залежить здатність моделі до узагальнення та її точність. У випадках аналізу складних, високовимірних, неоднорідних, зашумлених даних глобальна модель ШНМ в процесі аналізу досліджуваної вибірки даних може не враховувати усіх можливих варіантів

залежностей між даними. Це впливатиме на точність її роботи, недостатній рівень якої нівелюватиме можливість використання такої моделі апроксимації чи класифікації даних на практиці.

Одним з ефективних способів підвищення якості вхідних даних, а, отже, і точності регресійних або класифікаційних моделей, є їх попереднє опрацювання за допомогою кластерного аналізу. Цей підхід базується на гіпотезі компактності, яка стверджує, що елементи одного класу в задачі класифікації мають бути розташовані у вхідному просторі ближче один до одного, ніж до елементів інших класів [226]. Тобто, дані одного класу утворюють компактну (щільну) групу в ознаковому просторі.

Якщо це припущення виконується, методи класифікації можуть легше провести чітке розділення між класами, оскільки міжкласові межі стають виразнішими [226]. Однак на практиці один клас може містити декілька компактних підмножин точок, тобто окремі замкнені, щільні групи в ознаковому просторі. Наприклад, якщо клас має складну топологію, то навіть після нелінійних перетворень залишаються декілька компактних регіонів. Власне використання такої інформації та інтеграція її у методи класифікації чи регресії може сприяти покращенню точності розв'язання поставленої задачі. Крім того, поєднання кластерного аналізу з методами регресії чи класифікації може зменшити час навчання моделей. Наприклад, використання кластеризації перед регресією чи класифікацією дозволяє розділити дані на більш однорідні підмножини, що спрощує побудову локальних моделей для кожного окремого кластера [227].

Відомо чимало методів сумісного використання кластеризатора-регресора чи кластеризатора-класифікатора (ієрархічні, або двокрокові методи), які передбачають поділ даних на групи (кластери) з подібними властивостями. Далі відбувається побудова окремої моделі класифікації чи регресії в кожному окремому кластері [227].

Такий підхід забезпечує отримання суттєво кращих результатів інтелектуального аналізу в межах кожного окремого кластеру. Більше того, він дозволяє обробляти складні дані з різними характеристиками і забезпечує кращу адаптацію регресора чи класифікатора до специфіки кожного кластера. Проте він демонструє свою ефективність в процесі аналізу великої кількості даних. У випадку обмеженості вибірки для навчання кластеризація може формувати кластери з критично малою кількістю спостережень. Це зменшуватиме або й взагалі унеможливлуватиме використання класифікаторів чи регресорів для подальшого аналізу таких даних [228].

У [229] запропоновано інший підхід. В його основі покладено використання результатів кластеризації для розширення простору вхідних даних задачі, зокрема, маркерами належності до усіх виділених кластерів (формування комбінованої поверхні відгуку). Такий підхід забезпечує підвищення точності роботи класифікатора чи регресора і дозволяє включати етап кластеризації для підвищення ефективності опрацювання обмежених вибірок даних (екстремально-короткі чи малі дані). Проте у випадку аналізу малих вибірок біомедичних даних високої вимірності, що дуже характерно для цієї галузі, а також виділення великої кількості кластерів простір вхідних даних може суттєво збільшитися. Це може провокувати перенавчання і низьку точність подальшого інтелектуального аналізу отриманих даних на основі обраного регресора чи класифікатора.

У цьому підрозділі удосконалено цей метод для підвищення ефективності розв'язання задач класифікації чи регресії у випадку аналізу високовимірних вибірок даних малих обсягів [4], [41]. Ідея удосконаленого методу полягає в реалізації двокрокової стратегії навчання, де на початковому етапі виконується паралельне застосування лінійного методу і кластеризації. Лінійний метод використовується для отримання попереднього результату апроксимації або класифікації. Після цього всі початкові входи замінюються отриманим значенням. Паралельно виконується кластеризація для виділення кластерів (компактних

множин точок) з подальшим формуванням комбінованої поверхні відгуку різними варіантами, які розглянемо нижче. Завершальний етап передбачає виконання другого кроку навчання на нових даних з використанням другого лінійного методу, який формує фінальний результат.

Такий підхід дозволяє досягти двох ключових переваг в процесі аналізу високовимірних даних малих обсягів. По-перше, він значно зменшує кількість незалежних змінних, необхідних для подальшого аналізу, що спрощує модель і робить її більш ефективною. По-друге, інтеграція інформації, отриманої через кластеризацію, забезпечує додаткову інформативність для регресора чи класифікатора за рахунок формування таким чином, комбінованої поверхні відгуку. Обидві переваги допомагають підвищити точність роботи лінійних методів і їх здатність до узагальнення, що є важливим для застосування на нових, не бачених раніше даних.

Удосконалений метод складається з двох етапів: етап навчання та етап застосування.

Нехай маємо навчальну вибірку даних у вигляді матриці (2.1).

Етап навчання методу розв'язання задач регресії/класифікації на основі формування комбінованої поверхні відгуку передбачає виконання таких кроків:

1. Виконати нормування усіх векторів x_i матриці A :

$$\acute{x}_i = \mathcal{N}(x_i), i = \overline{1 \ n}, \quad (2.31)$$

де $\mathcal{N}(\cdot)$ - обрана функція нормалізації.

2. Виконати навчання першого лінійного класифікатора/регресора з використанням лінійної неітеративної нейроподібної структури МПГП і застосувати його для отримання результату для кожного вектора навчальної вибірки даних:

$$y_i^{(1)pred} = SGTM_1(\acute{x}_i), \quad (2.32)$$

де $y_i^{(1)pred}$ – вихідний сигнал для кожного i -того вектора, $SGTM_1(\cdot)$ – функція класифікації/регресії на основі лінійної неітеративної нейроподібної структури МПГП.

3. Визначити оптимальну кількість кластерів C_{opt} в процесі кластеризації одним з обраних методів (метод силуету [230], метод ліктя [231], Gap- статистику [232], індекс Калінського–Харабаша [233], індекс Девіса–Боулдіна [234] тощо). Слід зазначити, що для задачі класифікації кількість кластерів повинна бути більшою за кількість класів $C_{opt} > K$, що забезпечить можливість врахування додаткової інформації про кластери даних у досліджуваній вибірці даних (один клас може мати декілька кластерів даних або компактних множин даних).
4. Виконати кластеризацію навчальної вибірки даних методом *k-means* з використанням C_{opt} двома способами:

- на основі множини всіх векторів $\acute{x}_i, i = \overline{1, n}$ навчальної вибірки даних A , тобто без урахування вихідного атрибута:

$$\mu_c = \frac{1}{N_c} \sum_{i=1}^n cl_{i,c} \acute{x}_i, N_c = \sum_{i=1}^n cl_{i,c}, c = \overline{1, C_{opt}}, \quad (2.33)$$

$$cl_{i,c} = \begin{cases} 1, \text{ якщо } a_i = c, \\ 0, \text{ інакше} \end{cases}, c = \overline{1, C_{opt}}, \quad (2.34)$$

$$a_i = \arg \min_{c=\overline{1, C_{opt}}} \|\acute{x}_i - \mu_c\|^2,$$

де a_i – значення кластера, яке мінімізує відстань до $\acute{x}_i$, N_c - кількість навчальних векторів, віднесених до кластеру c , μ_c – центр кластера, $cl_{i,c}$ – бінарні мітки належності до кожного з C_{opt} ;

- на основі множини розширених векторів виду $\ddot{x}_i = (\acute{x}_i; y_i) \in \mathbb{R}^{m+1}$ навчальної вибірки даних A , тобто з урахуванням вихідного (залежного) атрибута:

$$\mu_c^+ = \frac{1}{N_c} \sum_{i=1}^n cl_{i,c} \ddot{x}_i, N_c = \sum_{i=1}^n cl_{i,c}, c = \overline{1, C_{opt}}. \quad (2.35)$$

5. Сформувати комбіновану поверхню відгуку шляхом утворення нової множини атрибутів відповідно до одного з методів:

- кожен новий вектор комбінованої поверхні відгуку включає результуюче значення $y_i^{(1)pred}$ для поточного вектора, отримане за допомогою лінійного класифікатора чи регресора (крок 2) та міток приналежності цього вектора $cl_{i,c}$ до кожного з C_{opt} кластерів:

$$x_i^+ = \left(y_i^{(1)pred}, cl_{i,1}, cl_{i,2}, \dots, cl_{i,C_{opt}} \right) \in \mathbb{R}^{1+C_{opt}}. \quad (2.36)$$

Позначимо це як *Варіант 1*.

- кожен новий вектор комбінованої поверхні відгуку включає усі нормалізовані компоненти для поточного вектора $\acute{x}_i$ та середнє значення залежного атрибуту $y_{a_i}^{centr}$ (у-координата) відповідного центра кластера, до якого належить $\acute{x}_i$:

$$x_i^+ = (\acute{x}_i, y_{a_i}^{centr}) \in \mathbb{R}^{m+1}, \quad (2.37)$$

де $y_{a_i}^{centr} = \frac{1}{N_{a_i}} \sum_{j=1}^n cl_{j,a_i} y_j$. Позначимо це як *Варіант 2*.

- кожен новий вектор комбінованої поверхні відгуку включає результуюче значення $y_i^{(1)pred}$ для поточного вектора, отримане за допомогою лінійного класифікатора чи регресора (крок 2) та значення вихідного атрибуту (у-координата) відповідного центра кластера, до якого належить поточний вектор:

$$x_i^+ = \left(y_{a_i}^{(1)pred}, y_c^{centr} \right) \in \mathbb{R}^2. \quad (2.38)$$

Позначимо це як *Варіант 3*.

6. Виконати навчання другого (фінального) лінійного класифікатора чи регресора на основі розширеної вибірки векторів x_i^+ , який буде використовувати ці атрибути для отримання остаточного результату $y_i^{(2)pred}$:

$$y_i^{(2)pred} = SGT M_2(x_i^+), i = \overline{1, n}. \quad (2.39)$$

На рис. 2.6, рис. 2.7 та рис. 2.8 відповідно подано візуалізації етапу навчання для різних варіантів удосконаленого методу.

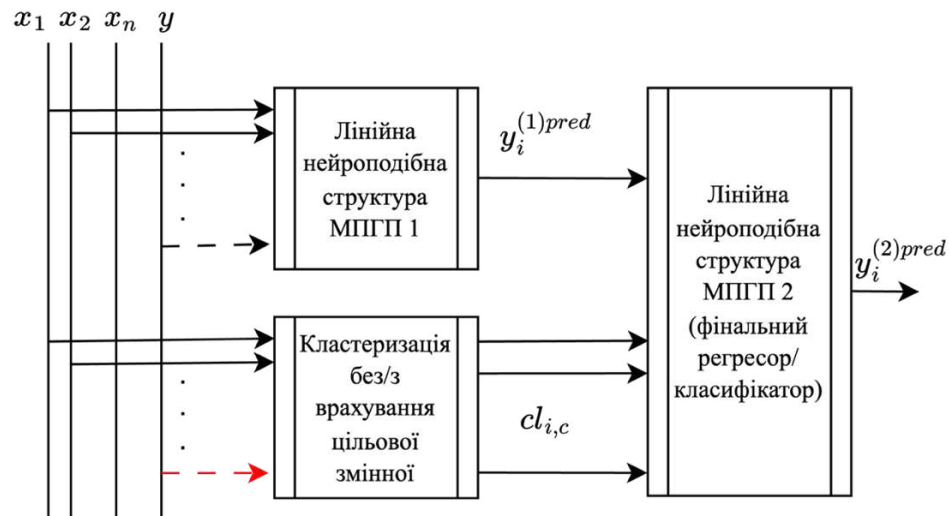


Рисунок 2.6 – Структурно-функційна схема удосконаленого методу розширення входів на основі формування комбінованої поверхні відгуку згідно (2.36)

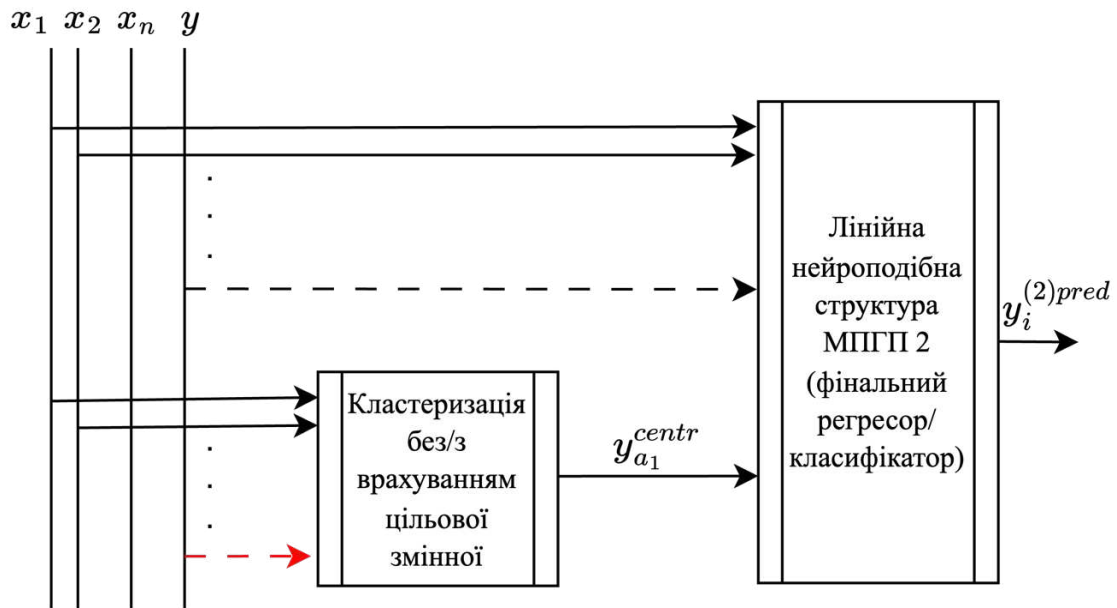


Рисунок 2.7 – Структурно-функційна схема удосконаленого методу розширення входів на основі формування комбінованої поверхні відгуку згідно (2.37)

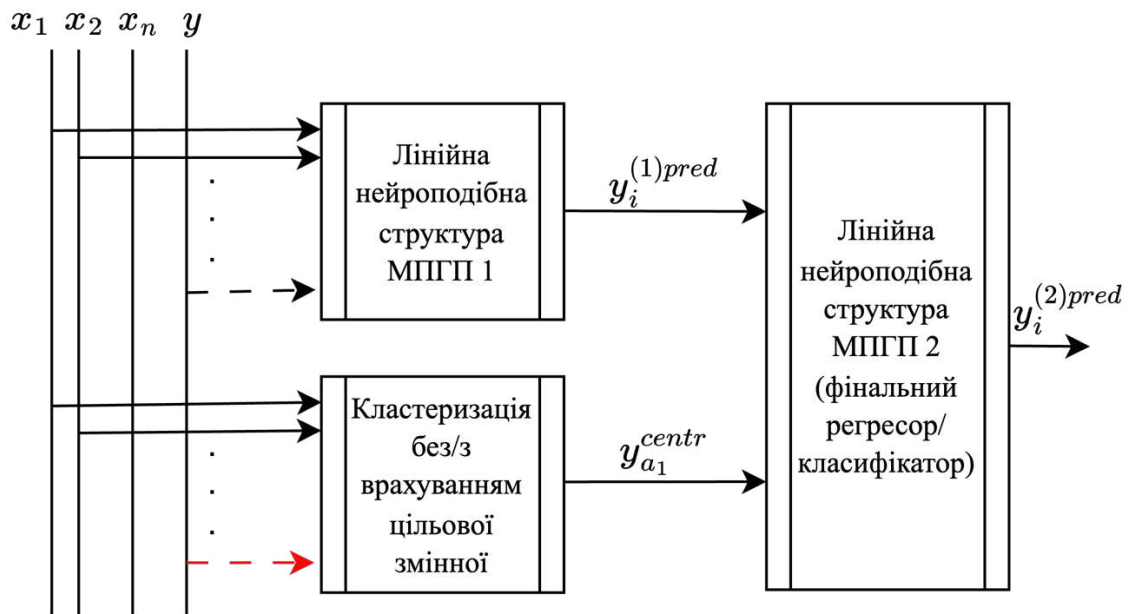


Рисунок 2.8 – Структурно-функційна схема удосконаленого методу розширення входів на основі формування комбінованої поверхні відгуку згідно (2.38)

Етап застосування методу розв'язання задач регресії/класифікації на основі формування комбінованої поверхні відгуку передбачає виконання таких кроків:

1. Пронормувати тестовий вектор виду $u \in \mathbb{R}^m$ згідно (2.31): $\acute{u} = \mathcal{N}(u)$.
2. Отримати перше значення вихідного сигналу для $\acute{u}$ з використанням попередньо навченої лінійної неітеративної нейроподібної структури МПГП (2.32):

$$y_u^{(1)pred} = SGTM_1(\acute{u}). \quad (2.40)$$

3. Визначити належність a_u тестового вектора u до одного з C_{opt} виділених кластерів, зокрема, шляхом пошуку найменшого значення Евклідової відстані a_u між тестовим вектором та усіма центрами кластерів:

$$a_u = \arg \min_{c=1, \dots, C_{opt}} \|\acute{u} - \mu_c\|^2. \quad (2.41)$$

Тут виконується пошук індексів кластеру c , центр якого μ_c , отриманий у (2.33) або (2.34), наближає $\acute{u}$ найкраще.

4. Встановити для вектора $\acute{u}$ індикатори приналежності $cl_{u,c}$ до кожного a_u :

$$cl_{u,c} = \begin{cases} 1, & \text{якщо } a_u = c, \\ 0, & \text{інакше} \end{cases} \quad c = 1, \dots, C_{opt}. \quad (2.42)$$

5. Сформувати комбіновану поверхню відгуку шляхом розширення вектора $\acute{u}$ одним з обраних методів:

- розширений вектор u^+ згідно з (2.36) матиме такий вигляд:

$$u^+ = (y_u^{(1)pred}, cl_{u,1}, cl_{u,2}, \dots, cl_{u,C_{opt}}) \in \mathbb{R}^{1+C_{opt}}. \quad (2.43)$$

- розширений вектор u^+ згідно з (2.37) матиме такий вигляд:

$$u^+ = (\acute{u}_i, y_{a_u}^{centr}) \in \mathbb{R}^{m+1}, \quad (2.44)$$

де $y_{a_u}^{centr}$ – середнє значення цільової змінної в кластері, до якого належить $\acute{u}_i$.

- розширений вектор u^+ згідно з (2.38) матиме такий вигляд:

$$u^+ = (\acute{u}_i, y_{a_u}^{centr}) \in \mathbb{R}^2. \quad (2.45)$$

6. Застосувати другу, попередньо навчену згідно (2.39) лінійну ШНМ на одному з варіантів (2.43) або (2.45) для отримання шуканого результату:

$$y_u^{(2)pred} = SGT M_2(u^+). \quad (2.46)$$

Удосконалений метод поєднує двоетапну процедуру навчання для формування єдиної комбінованої поверхні відгуку. Це може дозволити значно підвищити точність розв’язання задач класифікації чи регресії, використовуючи прості лінійні моделі на другому кроці навчання [225], [235]. Крім того, за рахунок відбору лише найінформативніших ознак, наприклад, міток кластерів або локальних середніх цільової змінної вдається утримувати вимірність фінального вектора атрибутів на мінімально необхідному рівні. Така економія вимірності критична при роботі з високовимірними біомедичними вибірками малого обсягу, де важливі як висока точність, так і здатність моделі коректно узагальнювати на нові, раніше не відомі дані.

Удосконалений метод формування комбінованої поверхні відгуку згідно (2.36) передбачає доповнення початкового результату роботи лінійної моделі ШНМ набором індикаторів приналежності до кожного з C_{opt} кластерів. Такий підхід забезпечує високу гнучкість. В цьому випадку кожна мітка $cl_{i,c}$ є окремою ознакою, що дозволяє фінальній моделі коригувати вплив приналежності до конкретного кластера незалежно від інших. Для малих вибірок біомедичних даних, де початкова вибірка даних може бути високовимірною, кількість індикаторів C_{opt} зазвичай значно менша за вимірність m , тому суттєво зменшується ризик “прокляття вимірності”.

Удосконалений метод формування комбінованої поверхні відгуку згідно (2.37) доповнює нормований вектор $\acute{x}_i$ єдиним локальним атрибутом – середнім значенням цільової змінної в кластері a_i . Цей підхід дозволяє зберегти низьку вимірність простору вхідних даних, лише $m + 1$ ознак незалежно від числа

кластерів, і водночас врахувати локальну структуру цільової змінної. Для обмежених біомедичних вибірок це забезпечить додаткову інформативність та підвищить властивість узагальнення моделі, оскільки одна додаткова ознака, що узагальнює поведінку у сусідстві кожного сусіднього спостереження, значно зменшує вплив шуму.

Удосконалений метод формування комбінованої поверхні відгуку згідно (2.38) поєднує глобальний та локальний сигнали в двовимірному векторі на основі конкатенацій (2.38). Зокрема, $y_i^{(1)pred}$ з (2.38) несе глобальну інформацію першого результату, а $y_{a_i}^{centr}$ – локальну інформацію кластеру. Така компактна двовимірна репрезентація дозволяє навчати фінальну лінійну модель ШНМ надзвичайно швидко, мінімізуючи ризик перенавчання та забезпечуючи достатню точність результатів навіть на дуже малих вибірках даних.

2.5 Метод кусково-лінійної апроксимації багатовимірних поверхонь відгуку

Лінійні методи за своєю природою апроксимують функції площинами або гіперплощинами у випадку багатовимірних даних, що є особливо актуальним для галузі біомедичної інженерії. У класичній лінійній нейронній мережі або іншій лінійній моделі після процедур навчання встановлюються ваги, які діють як коефіцієнти полінома, що апроксимує функцію відгуку. Цей підхід передбачає, що кожна компонента вхідного вектора множиться на відповідний коефіцієнт, а модель апроксимує функцію відгуку як одну гіперплощину.

Однак багато залежностей мають нелінійний характер. Щоб врахувати локальні нелінійності, можливо розбити кожну компоненту $x_j, j = \overline{1, m}$ вхідного вектора з (2.2) на S частин $x_{j,s}, s = \overline{1, S}$ таким чином, щоб:

$$x_j = \sum_{s=1}^S x_{j,s}, \quad (2.47)$$

де $x_{j,s}$ відповідає значенню компоненти x_j в певному інтервалі.

Використовуючи ці розділені частини як окремі входи лінійної ШНМ, замість стандартної форми апроксимації:

$$y_i^{pred} = \sum_{j=1}^m w_j x_j + b, i = \overline{1, n} \quad (2.48)$$

отримуємо нову, яку можна подати у такому вигляді:

$$y_i^{pred} = \sum_{j=1}^m \sum_{s=1}^S w_{j,s} x_{j,s} + b,$$

де $x_{j,s}$ має свій коефіцієнт $w_{j,s}$.

Таким чином, замість одного коефіцієнта w_j для цілої ознаки x_j отримуємо S коефіцієнтів $w_{j,s}$, які відповідають різним ділянкам значення поточної ознаки. Це дозволить моделі локально коригувати апроксимацію залежності, адже для різних інтервалів значення поточної ознаки встановлюються різні коефіцієнти [21].

Такий підхід відповідає ідеї розширення простору ознак або використання базисних функцій, що є класичним прийомом для підвищення здатності лінійних моделей апроксимувати нелінійні залежності. Окрім цього він знаходить підтвердження в теоремі Стоуна–Вейєрштраса, яка гарантує, що будь-яку неперервну функцію на компактній множині можна рівномірно апроксимувати сумою базисних функцій [236]. Зокрема, розбиття ознак можна інтерпретувати як використання простих ступінчастих базисних функцій, де кожна функція-індикатор відповідає певному інтервалу значень. Таким чином, розбиття вхідних ознак дає змогу лінійній моделі здійснювати кусково-лінійну апроксимацію тільки для багатовимірної поверхні відгуку, що дозволяє краще адаптуватися до локальних нелінійностей у високівимірних даних різноманітних задач у галузі біомедичної інженерії.

Принцип суперпозиції лінійних операторів забезпечує те, що сума окремих лінійних функцій є також лінійною [237]. Отже, при подачі окремих функцій на входи лінійної ШНМ кожна з них множиться на окремий коефіцієнт, оптимізований

за критерієм мінімізації обраної похибки апроксимації. Такий підхід дозволяє отримати більш гнучку модель, здатну точніше відображати локальні варіації даних, зберігаючи при цьому переваги лінійних методів: простоту та інтерпретованість.

У цьому підрозділі розвинуто метод кусково-лінійної апроксимації, який використовувався в аналоговій техніці, на випадок формування багатовимірної поверхні відгуку. В його основу покладено ідею розбиття кожної компоненти вектора (незалежного атрибуту) на однакову наперед визначену кількість частин, сума значень яких відображатиме початкове значення поточного компонента вектора.

Нехай початкова вибірка даних представлена матрицею (2.1). Окрім цього, кожному рядку $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})$, $i = \overline{1, n}$ поставлено у відповідність значення виходу y_i згідно (2.2).

Завдання полягає у пошуку такого перетворення (2.3), яке трансформує кожен вектор x_i у новий простір вищої вимірності згідно (2.8), (2.9) і дозволяє застосовувати лінійні або слабо-нелінійні методи навіть у випадку сильно нелінійної залежності між ознаками.

Метод кусково-лінійної апроксимації багатовимірної поверхні відгуку передбачає послідовне виконання таких кроків:

1. Виконати нормалізацію усіх векторів x_i , $i = \overline{1, n}$ матриці A :

$$\acute{x}_i = \mathcal{N}(x_i), \acute{x}_i = (\acute{x}_{i,1}, \acute{x}_{i,2}, \dots, \acute{x}_{i,m}), \quad (2.49)$$

де $\mathcal{N}(\cdot)$ – обрана функція нормалізації.

2. Визначити крайні межі $[a, b]$ діапазону розбиття кожної компоненти (незалежного атрибуту) заданого вектора (2.49). У випадку використання розробленого методу нижня a та верхня b межі діапазону ставляться у відповідність до значень, отриманих в результаті застосування обраної функції нормалізації $\mathcal{N}(\cdot)$.

3. Визначити кількість частин S , $s = \overline{1, S}$, на які буде розділений кожен компонент $\dot{x}_{i,j}$ вектора $\dot{x}_i$. Число S задає користувач або воно є гіперпараметром, який налаштовується в процесі оптимізації методу.
4. Задати межі для S частин межах діапазону $[a, b]$, що може відбуватися як випадковим чином, так і за допомогою вказівки користувача. Важливо, щоб значення межі кожної частини задовольняли умову:

$$a = a_0 < a_1 < a_2 < \dots < a_{S-1} < a_S = b. \quad (2.50)$$

Ця умова забезпечує коректне розбиття діапазону $[a, b]$ на S частин з інтервалами $[a_0, a_1], [a_1, a_2], \dots, [a_{S-1}, a_S]$.

5. Виконати процедуру розбиття значення кожної компоненти $\dot{x}_{i,j}$ кожного i -го вектора, $i = \overline{1, n}$ на S частин, використовуючи задані межі:

$$\dot{x}_{i,j,s} = \begin{cases} 0, & \text{якщо } \dot{x}_{i,j} \leq a_{s-1}; \\ \dot{x}_{i,j} - a_{s-1}, & \text{якщо } a_{s-1} < \dot{x}_{i,j} \leq a_s, \quad s = \overline{1, S}; \\ a_s - a_{s-1}, & \text{якщо } \dot{x}_{i,j} > a_s, \end{cases} \quad (2.51)$$

за умови, що сума усіх частин згідно перетворення (2.51) зберігає початкове значення $\dot{x}_{i,j}$:

$$\dot{x}_{i,j} = \sum_{s=1}^S \dot{x}_{i,j,s}. \quad (2.52)$$

6. Після виконання перетворень (2.51) для кожної компоненти i -го вектора $\dot{x}_{i,j}$ побудувати вектор її частин:

$$x_{i,j}^+ = (\dot{x}_{i,j,1}, \dot{x}_{i,j,2}, \dots, \dot{x}_{i,j,s}, \dots, \dot{x}_{i,j,S}). \quad (2.53)$$

7. Сформувати розширену матрицю для реалізації процедури навчання обраної ШНМ. Для цього об'єднати всі $x_{i,j}^+$ з (2.53) у єдиний вектор:

$$x_i^+ = (x_{i,1}^+, x_{i,2}^+, \dots, x_{i,m}^+) \in \mathbb{R}^{mS}, \quad (2.54)$$

з якого сформувати матрицю:

$$A^+ = \begin{bmatrix} x_1^+ \\ x_2^+ \\ \vdots \\ x_n^+ \end{bmatrix} \in \mathbb{R}^{n \times mS}. \quad (2.55)$$

8. Виконати навчання лінійної ШНМ - ($SGTM(\cdot)$):

$$y_i^{pred} = SGTM(A^+). \quad (2.56)$$

На етапі застосування розвинутого методу необхідно виконати усі кроки (2.49)-(2.54) лише для вхідного вектора даних з невідомим значенням вихідного сигналу. Кроки 2 та 3 методу, тобто кількість частин, на які розбиватиметься кожна компонента вхідного вектора, а також межі для цих частин мають бути такими ж, як і для навчальної вибірки даних. Розширений вектор даних згідно (2.51) подається на попередньо навчений лінійний метод машинного навчання (2.56) для отримання фінального результату розв'язання задачі регресії/класифікації.

Метод кусково-лінійної апроксимації багатовимірної поверхні відгуку має ряд важливих переваг. Він дозволяє розбивати складну нелінійну залежність на сукупність локальних лінійних моделей, які легко оптимізувати, аналізувати та інтерпретувати. Крім того, використання лінійних методів МН забезпечує більш стабільний та швидкий процес навчання, що є особливо актуальним при розв'язанні прикладних задач у галузі біомедичної інженерії, де важлива не лише точність, але й зрозумілість моделі та її здатність до узагальнення. Саме тому застосування розробленого методу в поєднанні з лінійними моделями є практичним рішенням в процесі аналізу біомедичних вибірок даних різних обсягів.

2.6. Висновки до розділу

1. Метод нелінійного розширення входів лінійних неітеративних нейроподібних структур МПГП поліномом Колмогорова-Габора дозволяє значно підвищити точність розв'язання нелінійних задач регресії та класифікації лінійними методами за рахунок подолання обмежень лінійної апроксимації. Подання

результатів роботи нейроподібної структури з нелінійним розширенням входів у поліноміальній формі дозволяє отримати інтерпретовану аналітичну модель, що робить процес прийняття рішень прозорішим і викликає більшу довіру з боку клініцистів, навіть якщо вони не є експертами в області штучного інтелекту. Окрім цього, така форма застосування нейромережевої моделі зменшує обчислювальну складність в процесі її використання, що дозволяє розробляти високоефективні програмні та апаратні засоби для інтелектуального аналізу даних у реальних біомедичних застосуваннях.

2. Розроблений метод розширення входів лінійних ШНМ на основі набору ймовірностей з шару підсумовування ймовірнісної нейронної мережі може забезпечити суттєве підвищення точності розв'язання задач класифікації та регресії лінійними ШНМ особливо у випадках аналізу екстремально коротких вибірок даних. Використання різних підходів до формування вихідних сигналів шару підсумовування ймовірнісної нейронної мережі може бути основою для розроблення різних каскадних методів аналізу даних малих обсягів залежно від їх вимірності та рівня незбалансованості.

3. Удосконалений метод формування комбінованої поверхні відгуку на основі кластерного аналізу значно знижує виміри простору вхідних даних задачі порівняно з базовим методом. Це забезпечує його ефективне використання в процесі аналізу біомедичних даних малих обсягів. Крім того, застосування різних варіацій цього методу може забезпечити підвищення точності інтелектуального аналізу даних малих обсягів як високої, так і низької вимірності.

4. Розвиток методу кусково-лінійної апроксимації для формування багатовимірної поверхні відгуку дозволяє розширювати простір вхідних даних задачі через ізоморфне подання кожного одиничного значення ознаки як сукупності локальних ознак. Такий підхід, зокрема, дозволяє більш точно коригувати локальні

області функції відгуку. Це, в свою чергу, може суттєво підвищити точність роботи лінійних методів. Крім цього, цей метод може використовуватися для створення унікальних, декорельованих вибірок даних як для розроблення однорідних (гомогенних) ансамблів штучних нейронних мереж, так і для попереднього опрацювання даних перед застосуванням мета-моделей у гомогенних/гетерогенних стекінгових та блендінгових ансамблях. Це може забезпечити отримання результатів з підвищеною точністю при використанні лінійних мета-моделей для агрегування результатів.

РОЗДІЛ 3

МЕТОДОЛОГІЯ АУГМЕНТАЦІЇ ДАНИХ НА ОСНОВІ ДЕКАРТОВОГО КВАДРАТА МНОЖИНИ ВХІДНИХ КОМПОНЕНТ ВЕКТОРІВ

У розділі 3 подано новий підхід до аугментації табличних вибірок даних. Він не передбачає введення синтетичних даних до вибірки, а ґрунтується на доступній інформації з вибірки даних. Такий підхід забезпечує можливість збільшення обсягу вибірки даних з використанням базових структурних операцій. Розроблено і обґрунтовано методологію розв’язання задач регресії та класифікації на основі декартового квадрата множини вхідних компонент векторів в умовах аналізу екстремально коротких вибірок даних. Показано, що використання базового та адитивного методів дозволяє обґрунтовано використовувати існуючі нелінійні моделі ШНМ для аналізу екстремально коротких вибірок даних. Подано модифікації цих методів з використанням принципів каскадування штучних нейронних мереж і сформованої шляхом кластеризації комбінованої поверхні відгуку. Додатково розглянуто метод формування аугментованої вибірки даних необхідного обсягу за рахунок використання кластеризації для оптимізації часу навчання та точності роботи розроблених методів.

Наукові і практичні результати, одержані у третьому розділі, опубліковано у розділі колективної монографії за кордоном [6], у високореєтингових журналах інших держав, які включено до наукометричних баз Scopus та Web of Science [12], [13], [14], [15], [16], а також апробовано і опубліковано у матеріалах українських та закордонних конференцій [7], [8], [9], [10], [11], [17], [52].

3.1 Аугментація даних на основі декартового квадрата множини вхідних компонент векторів

Сучасний розвиток біомедичної інженерії, яка поєднує інженерно-технічні та медико-біологічні знання, вимагає активного використання інструментів машинного навчання для вирішення ряду прикладних завдань. Ефективність цих інструментів значною мірою залежить від наявності достатньої кількості зразків для реалізації процедур навчання. Однак у різних спеціалізаціях та наукових напрямках цієї галузі все частіше виникають ситуації з використанням обмеженого обсягу доступних даних для ефективного навчання моделей машинного навчання. Це зумовлено низкою факторів, зокрема, значними часовими витратами на отримання вибірок даних необхідного обсягу, значними ресурсними, в тому числі фінансовими, витратами або просто відсутністю достатньої кількості даних, які на визначений момент часу відібрати проблематично.

У випадку необхідності аналізу екстремально-короткої вибірки даних, подібні методи не забезпечують задовільних результатів щодо точності роботи. Це пояснюється рядом потенційних проблем в процесі опрацювання таких даних, серед яких: проблеми перенавчання, що дуже характерно в процесі опрацювання обмеженого обсягу даних; величезний вплив шумових компонент, а також незбалансованості вибірки даних у випадку розв'язання задачі класифікації; дуже низька здатність до узагальнення на нових даних; проблеми підбору оптимальних параметрів для ітераційних методів машинного навчання через малий обсяг валідаційної вибірки тощо. У зв'язку з цим постає важлива задача збільшення обсягу даних, або аугментація, що дозволяє забезпечити більш ефективне навчання моделей машинного навчання в умовах обмежених даних.

Нехай екстремально-коротка вибірка незалежних змінних для реалізації методу навчання ШНМ представлена матрицею

$$A \in \mathbb{R}^{n \times m}, \quad (3.1)$$

де n – кількість векторів, а m – вимірність кожного вектора.

У загальному випадку завдання аугментації даних полягає в тому, щоб побудувати розширену вибірку даних:

$$A^+ \in \mathbb{R}^{N \times m}, \quad (3.2)$$

де $N > n$. Вона містить як оригінальні вектори, так і/або нові, синтезовані на їх основі з можливістю збільшення вимірності m кожного початкового вектора даних, щодо підвищення різноманітності векторів даних для навчання при збереженні статистичних властивостей початкової вибірки.

Існує широкий спектр методів аугментації числових вибірок даних по рядках і по стовпцях. Вони використовують різні види перетворень (3.2), кожен з яких має свої переваги та обмеження. До основних класів таких методів можна віднести методи додавання шуму, масштабування, зсуву, інтерполяції, генерації синтетичних даних, перетворення ознак тощо. Проте кожен з цих підходів має свої недоліки. Наприклад, додавання шуму може призвести до суттєвого спотворення даних, що підвищує чутливість моделі до навіть незначних коливань. Масштабування та зсуви можуть порушити важливі взаємозв'язки в даних, особливо в контексті нелінійних моделей. Інтерполяція може викликати аномалії або артефакти, якщо вихідні дані є рідкісними або нелінійними. Методи генерації синтетичних даних можуть вводити нереалістичні зразки, що спотворює реальність і, як наслідок, знижує ефективність моделі. Крім того, усі ці методи виконують такі перетворення (3.2), які модифікують або синтезують нові, штучні дані, що не завжди є виправданим, може бути ризикованим і зменшує довіру фахівця в галузі біомедичної інженерії до розробленого на основі таких даних рішення.

У роботі вперше запропоновано принципово інший підхід до аугментації табличних вибірок даних, який не використовує штучних методів введення додаткової інформації в дані шляхом інтерполяції, додавання шуму, генерації синтетичних даних, тощо, а базується на використанні інформації, наявної тільки у

заданій вибірці даних, де ефект забезпечується виключно за допомогою розробленого методу декартового квадрата множини вхідних компонент векторів.

Декартовий квадрат вибірки даних $A \in \mathbb{R}^{n \times m}$ подається як розширена матриця $A^+ \in \mathbb{R}^{n^2 \times 2m}$, утворена множиною всіх можливих впорядкованих пар векторів ознак $x_q^+ = (x_i, x_k) \in \mathbb{R}^{2m}$, $q = (i-1)n + k$, $i, k = \overline{1, n}$:

$$A^+ = \begin{bmatrix} (x_1^+)^T \\ (x_2^+)^T \\ \dots \\ (x_{n^2}^+)^T \end{bmatrix} \in \mathbb{R}^{n^2 \times 2m}, \quad (3.3)$$

де кожен рядок містить конкатенацію двох m -вимірних векторів (x_i, x_k) . Таким чином, аугментована матриця A^+ має n^2 рядків та $2m$ стовпців.

Залежний атрибут $z \in \mathbb{R}^{n^2 \times 1}$ для кожної такої пари формується як різниця залежних атрибутів обох векторів пари:

$$z_{(i-1)n+k} = y_k - y_i, \quad i, k = \overline{1, n}. \quad (3.4)$$

По суті (3.3) – це квадратичне збільшення обсягу вибірки даних для реалізації методів навчання ШНМ порівняно з початковою вибіркою даних. Окрім цього, вимірність кожного нового вектора є подвоєною. Завданням такої непрямої постановки задачі регресії є апроксимація не значень цільових атрибутів досліджуваної вибірки даних, а різниць між ними з подальшим використанням нелінійних ШНМ для реалізації процесу навчання. Саме через непряму постановку задачі регресії з використанням (3.3), (3.4) вона має ряд особливостей в режимі застосування ШНМ, які будуть детально розглянуті в процесі опису методів застосування кожного з розроблених у цьому розділі методів. Цей напрям, згодом, після перших англомовних публікацій автора дисертації на цю тему [3], [7], почав активно досліджуватися і використовуватися за кордоном під різними назвами. Зокрема, група науковців Університету Ватерлоо (Канада) запропонувала модифікацію методу для розв'язання задачі регресії [238]. Дещо модифіковану

процедуру аугментації використано групою науковців Лос-Аламоської національної лабораторії та Массачусетського технологічного інституту (США) [239]. Ефективну адаптацію розробленого у цій дисертації методу для розв'язання задачі класифікації запропоновано групою науковців Мюнхенського університету Людвіга-Максиміліана та компанії IDIADA Fahrzeugtechnik GmbH (Німеччина) [240].

Загалом, декартів квадрат множини A , як метод аугментації екстремально-короткої вибірки даних, має ряд важливих переваг в процесі застосування інструментарію машинного навчання. По-перше, розроблений автором метод є простою та ефективною формою штучної аугментації даних без введення зовнішнього шуму, яка дозволяє суттєво збільшити обсяг вибірки даних для навчання. Це відкриває можливості для використання ядрових методів машинного навчання як ітеративного, так і неітеративного типу, зокрема, методів з РБФ-ядром для розв'язання різноманітних задач регресії та класифікації в умовах екстремально коротких вибірок даних. По-друге, розроблений підхід дозволяє значно покращити властивості узагальнення нелінійних штучних нейронних мереж, оскільки модель отримає можливість працювати з більш різноманітними варіаціями вхідних даних, що допомагає запобігти перенавчанню. По-третє, використання (3.3) може істотно підвищити точність роботи нелінійних методів машинного навчання і значно покращити робастність обраної моделі навіть за умов дуже обмеженого обсягу вихідних даних. Варто підкреслити, що цей підхід покращуватиме результати навчання лише при використанні нелінійних методів машинного навчання. У протилежному випадку, коли застосовуватимуться лінійні методи машинного навчання, позитивного ефекту не спостерігатиметься внаслідок ефекту лінійної суперпозиції сигналів.

Розглянемо методи розв'язання задач регресії та класифікації екстремально коротких вибірок даних на основі декартового квадрата множини вхідних компонент векторів, а також кроки етапів їх навчання і застосування.

3.2 Метод розв'язання задач регресії та класифікації на основі декартового квадрата множини вхідних компонент векторів

Кожен рядок екстремально-короткої вибірки незалежних змінних (3.1) позначимо як $x_i \in \mathbb{R}^m$ для $i = \overline{1, n}$, де n – кількість векторів спостережень, m – кількість вхідних ознак (атрибутів) кожного вектора спостережень. Його нормалізовано обраним методом нормалізації. При цьому відповідні значення цільової змінної подані таким чином:

$$\begin{cases} y_i \in \{1, 2, \dots, K\}, & \text{для задач класифікації;} \\ y_i \in \mathbb{R}, & \text{для задач регресійного аналізу.} \end{cases} \quad (3.5)$$

Метод розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів включає ряд етапів, зокрема,

- аугментації даних;
- навчання обраної нелінійної ШНМ на аугментованій вибірці даних;
- застосування методу.

Розглянемо кожен з вищенаведених компонент методу (Метод 3.1) детальніше.

Етап аугментації даних для цього методу (Методу 3.1) передбачає послідовне виконання таких кроків:

1. Сформувані аугментовану вибірку даних A^+ на основі A з використанням декартового квадрата множини вхідних компонент векторів (3.3):

$$A^+ = \begin{bmatrix} (x_1^+)^T \\ (x_q^+)^T \\ \dots \\ (x_{n^2}^+)^T \end{bmatrix} \in \mathbb{R}^{n^2 \times 2m}, \quad (3.6)$$

де кожен розширений вектор x_a^+ формується наступним чином:

$$x_q^+ = (x_i, x_k) \in \mathbb{R}^{2m}, \quad q = (i-1)n + k, \quad i, k = \overline{1, n}, \quad q = \overline{1, n^2}. \quad (3.7)$$

У (3.5) кожен елемент $x_q^+ = (x_i, x_k)$ розглядається як вектор у $\mathbb{R}^{2m}$. Загальна кількість векторів x_q^+ для A^+ становить n^2 .

2. Обчислити нове значення залежного атрибуту для кожної пари (i, k) як різницю цільових атрибутів кожного поточного k -того вектора і усіма іншими векторами y_i вибірки даних A , включаючи його самого:

$$z_{(i-1)n+k} = y_k - y_i, \quad i, k = \overline{1, n}. \quad (3.8)$$

На рис. 3.1. наведено структурно-функційну схему аугментації даних згідно описів, поданих вище.

3. Виконати навчання $f_1(A^+)$ обраної нелінійної ШНМ $f_1(\cdot)$ на основі використання аугментованої вибірки даних (3.6) та відповідних значень нової цільової змінної (3.8).

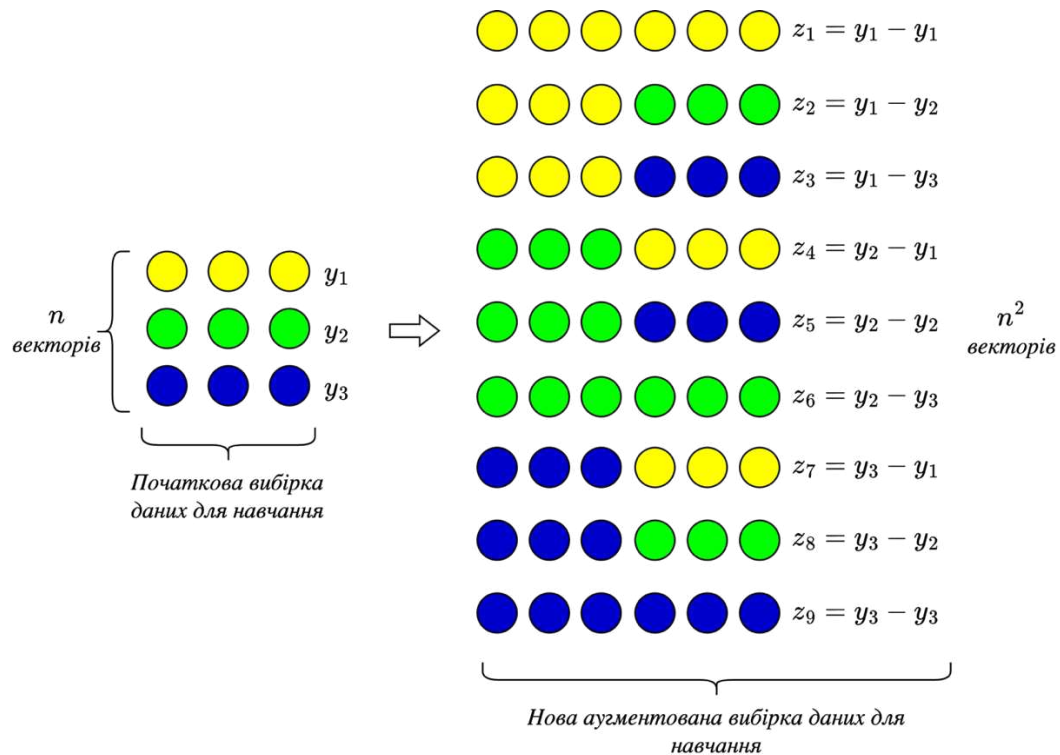


Рисунок 3.1 – Структурно-функційна схема аугментації даних для методу розв’язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів

Введемо до розгляду вектор $u \in \mathbb{R}^m$, для якого невідомо значення вихідного сигналу. Тоді етап застосування Методу 3.1 передбачає послідовного виконання таких кроків:

1. Сформувати тимчасову вибірку даних $B^+ \in \mathbb{R}^{n \times 2m}$ шляхом конкатенації вектора u з кожним i -тим вектором x_i початкової вибірки даних A з (3.1),

$$B^+ = \begin{bmatrix} (u_1^+)^T \\ (u_2^+)^T \\ \vdots \\ (u_n^+)^T \end{bmatrix} \in \mathbb{R}^{n \times 2m}, \quad (3.9)$$

де кожен розширений вектор u_i^+ формується наступним чином:

$$u_i^+ = (u, x_i) \in \mathbb{R}^{2m}, \quad i = \overline{1, n}, \quad (3.10)$$

2. Для кожного рядка u_i^+ матриці B^+ отримати шукане значення $z_i^{pred} = f_1(u_i^+)$, $i = \overline{1, n}$, шляхом застосування попередньо навченого нелінійного засобу машинного навчання $f_1(\cdot)$.
3. У випадку розв'язання задачі регресії остаточний шуканий сигнал y_u^{pred} для одного вектора u з використанням методу розв'язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів, обчислити згідно виразу:

$$y_u^{pred} = \frac{1}{n} \sum_{i=1}^n z_i^{pred} + Q, \quad (3.11)$$

де $Q = const$ – середнє значення суми відомих вихідних сигналів для усіх векторів з (3.1), яке для кожного вхідного вектора тестової вибірки буде однаковим:

$$Q = \frac{1}{n} \sum_{i=1}^n y_i. \quad (3.12)$$

Слід зазначити, що у (3.11) за рахунок $\frac{1}{n} \sum_{i=1}^n y_i$ імплементується принцип усереднення і часткової взаємної компенсації випадкових похибок з різними

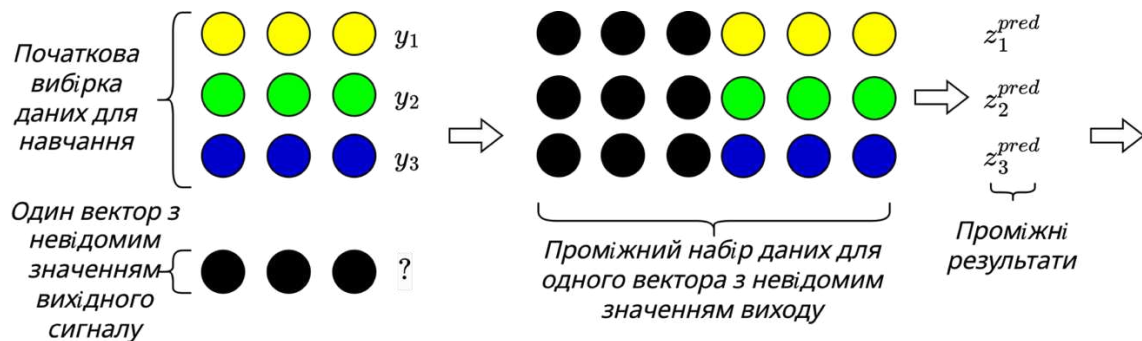
знака мивиходів векторів спостережень, що є характерним для роботи ансамблевих методів.

4. У випадку розв'язання задачі класифікації шуканий сигнал методу класифікації на основі декартового квадрата множини вхідних компонент векторів обчислити згідно наступного виразу:

$$y_u^{pred} = \arg \max_{i=1}^n (y_i + z_i^{pred}). \quad (3.13)$$

Слід зазначити, що завдяки (3.13) реалізується принцип голосування більшістю, що також є характерним для роботи ансамблевих методів.

На рис. 3.2 наведено структурно-функційну схему етапу застосування методів інтелектуального аналізу для розв'язання задач регресії і класифікації на основі декартового квадрата множини вхідних компонент векторів, згідно кроків, описаних вище.



$$\Rightarrow y^{pred} = \frac{\sum_{i=1}^n y_i + \sum_{i=1}^n z_i^{pred}}{n} \quad (1)$$

$$\Rightarrow y^{pred} = \arg \max \left\{ (y_1 + z_1^{pred}), (y_2 + z_2^{pred}), (y_3 + z_3^{pred}) \right\} \quad (2)$$

(1) використовується у випадку розв'язання задачі регресії

(2) використовується у випадку розв'язання задачі класифікації

Рисунок 3.2 – Структурно-функційна схема етапу застосування методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів

Загалом, розроблений метод дозволяє квадратично збільшити обсяг вибірки даних, що є формою штучної аугментації даних, заснованої виключно на наявних даних, без додавання штучного шуму або синтетичних точок. Додатковою перевагою такого методу є зменшення впливу випадкових похибок у даних, які є типовими для різноманітних вимірювань у медичних дослідженнях. Розроблений метод дозволяє моделі навчатися на відносних відмінностях між прикладами. Агрегація результатів, отриманих на етапі застосування методу, дозволяє реконструювати шукану відповідь за допомогою усереднення (3.11) чи голосування більшістю (3.13). Таким чином реалізуються принципи ансамблювання ШНМ, які дозволяють частково компенсувати похибки при опрацюванні малих обсягів даних та підвищують робастність методу в цілому. Проте найбільшою перевагою цього методу є те, що він фокусується на різницях між векторами, а не на прямому відображенні ознак у мітки. Такий підхід дає змогу виявляти приховані залежності між векторами, наприклад, між пацієнтами, які могли б залишитися прихованими в класичних підходах регресії. Наприклад, незначна зміна в рівні біомаркера між двома схожими пацієнтами може призвести до суттєвої різниці у результатах, що виявляється лише через навчання ШНМ на основі (3.6). Це робить метод придатним до адаптації під різні структури даних та особливості клінічного сценарію.

3.3 Адитивний метод розв'язання задачі регресії з застосуванням декартового квадрата множини вхідних компонент векторів

В основі адитивного методу розв'язання задачі на основі декартового квадрата множини вхідних компонент векторів покладено можливість часткової взаємної компенсації випадкових похибок різних знаків. Він реалізується на основі симетричного формування значення залежного атрибуту (3.8) для кожної пари (i, k) в режимі застосування методу.

На відміну від (3.8) введемо два варіанти формування значення залежного атрибуту, які будуть використовуватися в режимі застосування методу:

$$z_i^{(1)} = y_k - y_i, \quad z_i^{(2)} = y_i - y_k, \quad (3.14)$$

де значення k фіксоване й відповідає позиції вектора з невідомим вихідним сигналом u , а індекс усіх векторів навчальної вибірки даних $i = 1, n$.

Виконавши алгебраїчні перетворення над рівностями (3.14), зокрема, підсумовуючи обидві частини кожної з них за всіма $i = \overline{1, n}$, отримаємо:

$$\sum_{i=1}^n z_i^{(1)} = ny_k - \sum_{i=1}^n y_i, \quad (3.15)$$

$$\sum_{i=1}^n z_i^{(2)} = \sum_{i=1}^n y_i - ny_k. \quad (3.16)$$

Віднявши рівність (3.15) від рівності (3.16):

$$\begin{aligned} \sum_{i=1}^n z_i^{(1)} - \sum_{i=1}^n z_i^{(2)} &= \\ &= (ny_k - \sum_{i=1}^n y_i) - (\sum_{i=1}^n y_i - ny_k) = 2ny_k - 2 \sum_{i=1}^n y_i, \end{aligned} \quad (3.17)$$

отримаємо вираз, який дозволяє оцінити значення y_k з урахуванням взаємних різниць між спостереженнями:

$$y_k = \frac{\sum_{i=1}^n z_i^{(1)} - \sum_{i=1}^n z_i^{(2)}}{2n} + \frac{\sum_{i=1}^n y_i}{n}, \quad (3.18)$$

де $z_i^{(1)}$ та $z_i^{(2)}$ визначаються з використанням обраного методу машинного навчання, y_i – відоме значення залежного атрибуту для усієї початкової вибірки даних.

Такий підхід забезпечить можливість підвищення точності аналізу екстремально коротких вибірок даних в порівнянні з Методом 3.1.

Розглянемо покрокову реалізацію адитивного методу (Метод 3.2). Як і у попередньому випадку, задано екстремально-коротку вибірку даних у вигляді матриці (3.1). Етап аугментації даних згідно цього методу відбувається на основі виразів (3.6), (3.7). Етап навчання на основі аугментованої вибірки даних виконується з використанням обраної нелінійної ШНМ $f_2(\cdot)$.

Етап застосування для цього методу (Методу 3.2.) дещо відрізняється від (3.9)-(3.11) і вимагає послідовного виконання таких кроків.

Оперуємо вектором $u \in \mathbb{R}^m$, для якого невідомо значення вихідного сигналу.

1. Сформувати дві тимчасові вибірки даних $B^{(1)+} \in \mathbb{R}^{n \times 2m}$ та $B^{(2)+} \in \mathbb{R}^{n \times 2m}$ шляхом конкатенації вектора u з кожним i -тим вектором x_i з A . Основна різниця – це позиція обох векторів u та x_i у розширеному векторі даних. Зокрема, матрицю $B^{(1)+}$ можна записати так:

$$B^{(1)+} = \begin{bmatrix} \left(u_1^{(1)+}\right)^T \\ \left(u_2^{(1)+}\right)^T \\ \dots \\ \left(u_n^{(1)+}\right)^T \end{bmatrix} \in \mathbb{R}^{n \times 2m}, \quad (3.19)$$

де кожен розширений вектор $u_i^{(1)+}$ формується наступним чином:

$$u_i^{(1)+} = (u, x_i) \in \mathbb{R}^{2m}, \quad i = \overline{1 \ n}. \quad (3.20)$$

2. Матрицю $B^{(2)+}$ можна представити наступним чином:

$$B^{(2)+} = \begin{bmatrix} \left(u_1^{(2)+}\right)^T \\ \left(u_2^{(2)+}\right)^T \\ \dots \\ \left(u_n^{(2)+}\right)^T \end{bmatrix} \in \mathbb{R}^{n \times 2m}, \quad (3.21)$$

де кожен розширений вектор $u_i^{(2)+}$ формується наступним чином:

$$u_i^{(2)+} = (x_i, u) \in \mathbb{R}^{2m}, \quad i = \overline{1 \ n}. \quad (3.22)$$

3. Отримати $z_i^{(1)pred} = f_1(u_i^{(1)+})$, $i = \overline{1 \ n}$, для кожного рядка $u_i^{(1)+}$ матриці $B^{(1)+}$ шляхом застосування попередньо навченої нелінійної ШНМ $f_2(\cdot)$.

4. Отримати $z_i^{(2)pred} = f_1(u_i^{(2)+})$, $i = \overline{1, n}$, для кожного рядка $u_i^{(2)+}$ матриці $B^{(2)+}$, шляхом застосування того ж самого рядка $u_i^{(2)+}$ на попередньо навчений нелінійний ШНМ $f_2(\cdot)$.
5. Сформуванати остаточний шуканий сигнал y_u^{pred} роботи адитивного методу розв'язання задачі регресії/класифікації на основі декартового квадрата множини вхідних компонент векторів. На відміну від (3.11)-(3.12), y_u^{pred} обчислюється згідно виразу:

$$y_u^{pred} = \frac{\sum_{i=1}^n y_i}{n} + \frac{\sum_{i=1}^n z_i^{(1)pred}}{2n} - \frac{\sum_{i=1}^n z_i^{(2)pred}}{2n}. \quad (3.23)$$

На рис. 3.3 наведено структурно-функційну схему етапу застосування адитивного методу розв'язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів згідно кроків, описаних вище.

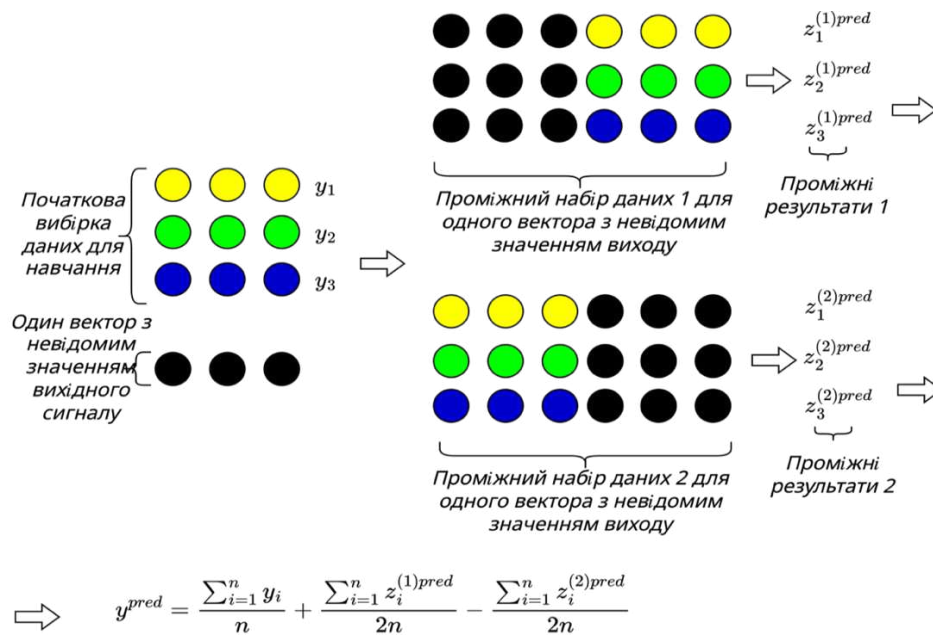


Рисунок 3.3 – Структурно-функційна схема етапу застосування адитивного методу розв'язання задачі регресії на основі декартового квадрата множини вхідних

КОМПОНЕНТ

Слід зазначити, що у (3.23), як і у (3.11), також імплементується принцип усереднення результату, що характерний для роботи ансамблевих методів. Окрім цього, використання обох доданків $\left(\frac{\sum_{i=1}^n z_u^{pred(1)}}{2n} - \frac{\sum_{i=1}^n z_u^{pred(1)}}{2n} \right)$ у (3.23) на відміну від (3.8), який використовує лише $\frac{\sum_{i=1}^n z_u^{pred}}{n}$, забезпечує часткову взаємну компенсацію випадкових похибок з різними знаками у рівності (3.20). Оскільки біомедичні вимірювання часто супроводжуються значним рівнем випадкових флуктуацій, остання перевага дозволяє зменшити вплив таких коливань за рахунок побудови парних взаємодій між векторами і підвищити робастність цього методу.

Такий підхід особливо ефективний у випадках, коли кількість доступних зразків недостатня для стійкого навчання класичних моделей машинного навчання. Крім того, розроблений метод дозволяє ефективно інтегрувати будь-які базові нелінійні штучні нейронні мережі або ядрові методи. Це забезпечує гнучкість адаптації до конкретного типу даних або задачі, зокрема, для багатовимірних просторів ознак з обмеженим числом спостережень, що є типовим для задач медичної діагностики, оцінки ризиків тощо.

3.4 Метод розв'язання задач регресії та класифікації з використанням принципів каскадування

У багатьох випадках інтелектуального аналізу даних малих обсягів, використання відомих штучних нейронних мереж, таких як ймовірнісна нейронна мережа, нейронна мережа узагальненої регресії, різновидів неітеративної нейроподібної структури з РБФ-розширенням входів, тощо, в процесі інтелектуального аналізу екстремально-короткої вибірки даних може демонструвати порівняно високі результати. Це залежить від обсягу, якості, вимірності тощо заданої вибірки даних. Проте точність розв'язання конкретної задачі все ще може не відповідати вимогам користувачів в певній області застосування. Для розв'язання цієї задачі можна використати принцип

каскадування методів машинного навчання [15], описаний у другому розділі, і підвищити точність роботи методу розв’язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів.

У цьому підрозділі запропонована реалізація такого принципу через використання вихідних сигналів першої нейронної мережі (ШНМ-1) як додаткових атрибутів для заданої екстремально-короткої вибірки даних з подальшим опрацюванням розширеної вибірки даних другим методом (ШНМ-2). Такий підхід, зокрема, для задачі класифікації обґрунтовується теоремою Т. Ковера [143]. Вона стверджує, що перехід до вищої вимірності може значно полегшити завдання класифікації, перетворюючи дані, які не є лінійно роздільними, на лінійно роздільні і, відповідно, демонструє суттєве підвищення точності розв’язання поставленої задачі.

У другому розділі досліджено ефективність використання цього принципу за рахунок розширення простору вхідних даних задачі з використанням вихідних сигналів шару сумування ймовірнісної нейронної мережі. Експериментально доведено (розділ 2), що такий підхід забезпечує суттєве підвищення точності роботи методів машинного навчання на розширеній вибірці даних. Використаємо цей підхід для модифікації методу класифікації на основі декартового квадрата множини вхідних компонент векторів, який описано у підрозділі 3.1 (Метод 3.3).

Як і у попередньому випадку оперуємо матрицею (3.1) для подання заданої екстремально-короткої вибірки даних.

Етап аугментації даних для методу розв’язання задач регресії та класифікації з використанням принципів каскадування (Методу 3.3) передбачає виконання таких кроків:

1. Виконати навчання першої штучної нейронної мережі $f_3^{(1)}(A) \in \mathbb{R}^n$, яка забезпечує достатньо високу точність на заданій початковій вибірці даних з (3.1):

$$A = \begin{bmatrix} x_1^T \\ x_2^T \\ \dots \\ x_n^T \end{bmatrix} \in \mathbb{R}^{n \times m}. \quad (3.24)$$

2. Застосувати навчену нейронну мережу до усіх n рядків навчальної вибірки даних для отримання вихідних сигналів $y_i^{tempPred}$:

$$y_i^{tempPred} = f_3^{(1)}(x_i), \quad i = 1, n. \quad (3.25)$$

Слід зазначити, що у випадку використання ймовірнісної нейронної мережі чи нейронної мережі узагальненої регресії в процесі реалізації цього кроку методу, поточне спостереження, для якого виконується аналіз, тимчасово видаляється з опорної вибірки даних.

3. Для кожного i -того вектора додати $y_i^{tempPred}$ до x_i як новий атрибут, утворивши розширену матрицю $\hat{A} \in \mathbb{R}^{n \times (m+1)}$.

$$\hat{A} = \begin{bmatrix} \hat{x}_1^T \\ \hat{x}_2^T \\ \dots \\ \hat{x}_n^T \end{bmatrix} \in \mathbb{R}^{n \times (m+1)}, \quad \hat{x}_i = (x_i, y_i^{tempPred}) \in \mathbb{R}^{m+1}. \quad (3.26)$$

4. Виконати процедуру аугментації даних згідно (3.6), (3.7) на основі $\hat{A}$:

$$\hat{A}^+ = \begin{bmatrix} (\hat{x}_1^+)^T \\ (\hat{x}_2^+)^T \\ \dots \\ (\hat{x}_{n^2}^+)^T \end{bmatrix} \in \mathbb{R}^{n^2 \times 2(m+1)}, \quad (3.27)$$

де для $a = (i - 1)n + k$, $i, k = \overline{1 \ n}$:

$$\hat{x}_a^+ = (\hat{x}_i, \hat{x}_k) \in \mathbb{R}^{2(m+1)}. \quad (3.28)$$

5. Обчислити нове значення залежного атрибуту z_i для кожного a так само як у (3.8).

6. Виконати навчання другої ШНМ $f_3^{(2)}(\hat{x}_a^+) \in \mathbb{R}^n$ (можливо тієї ж архітектури, що і перша ШНМ) на основі розширеної матриці $\hat{A}^+$.

На рис. 3.4 наведено структурно-функційну схему етапу аугментації даних згідно кроків, поданих вище.

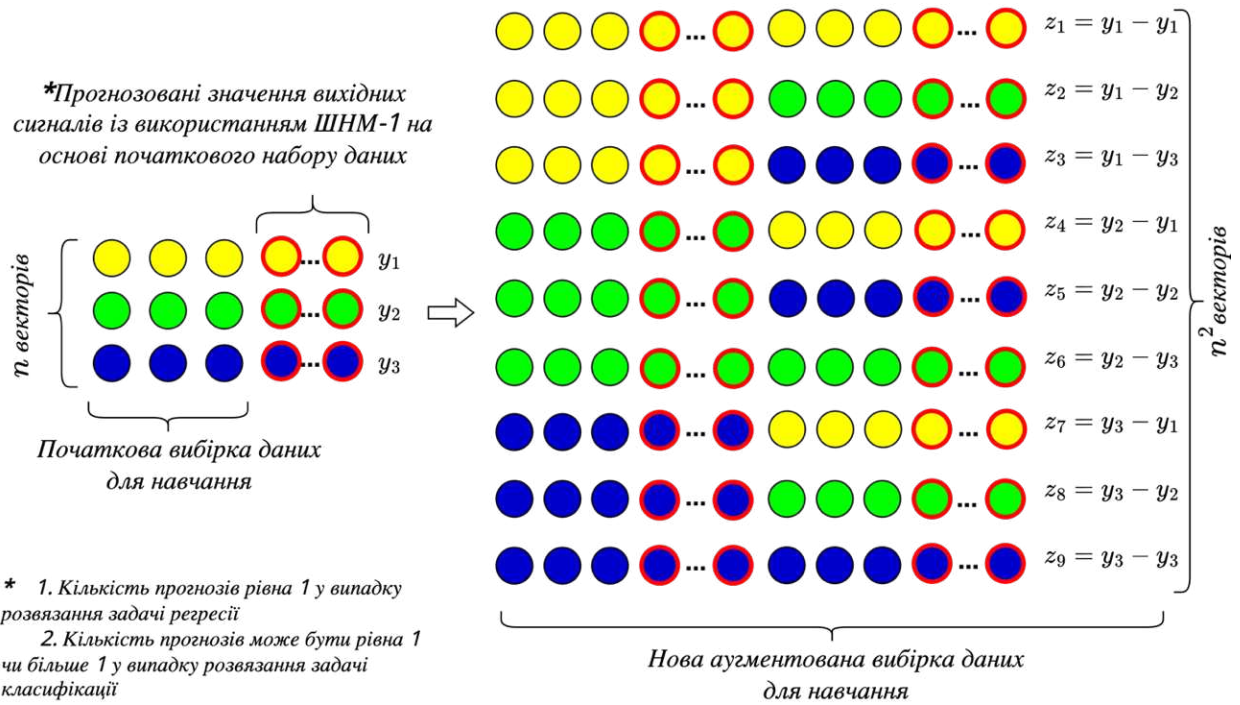


Рисунок 3.4 – Структурно-функційна схема етапу аугментації даних для методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використання принципів каскадування ШНМ

Етап застосування методу розв'язання задач регресії та класифікації з використанням принципів каскадування (Методу 3.3) передбачає виконання таких кроків.

Як і у попередніх випадках оперуємо вектором $u \in \mathbb{R}^m$, для якого невідомо значення вихідного сигналу.

1. Застосувати вектор $u \in \mathbb{R}^m$ на попередньо навченій першій ШНМ:

$$y_u^{tempPred} = f_3^{(1)}(u) \in \mathbb{R}. \quad (3.29)$$

2. Додати $y_i^{tempPred}$ до u як новий атрибут, утворивши розширений вектор:

$$\hat{u} = (u, y_u^{tempPred}) \in \mathbb{R}^{m+1}. \quad (3.30)$$

3. Сформувати тимчасову вибірку даних $\hat{B}^+ \in \mathbb{R}^{n \times 2m}$ шляхом конкатенації вектора $\hat{u}$ з кожним i -тим вектором $\hat{x}_i$ вибірки даних $\hat{A}$ з (3.26)

$$\hat{B}^+ = \begin{bmatrix} (\hat{u}_1^+)^T \\ (\hat{u}_2^+)^T \\ \vdots \\ (\hat{u}_n^+)^T \end{bmatrix} \in \mathbb{R}^{n \times 2(m+1)}, \quad (3.31)$$

де кожен розширений вектор $\hat{u}_i^+$ формується таким чином:

$$\hat{u}_i^+ = (\hat{u}, \hat{x}_i) \in \mathbb{R}^{2(m+1)}, \quad i = \overline{1 \ n}. \quad (3.32)$$

4. Для кожного рядка $\hat{u}_i^+$ матриці $\hat{B}^+$ отримати шукане значення $\hat{z}_i^{pred}$ шляхом застосування попередньо навченої ШНМ:

$$\hat{z}_i^{pred} = f_3^{(2)}(\hat{u}_i^+), \quad i = \overline{1 \ n}. \quad (3.33)$$

5. Сформувати остаточний шуканий сигнал y_u^{pred} для одного вектора u у випадку розв'язання задачі регресії згідно виразу:

$$y_u^{pred} = \frac{1}{n} \sum_{i=1}^n \hat{z}_i^{pred} + \frac{1}{n} \sum_{i=1}^n y_i. \quad (3.34)$$

6. У випадку розв'язання задачі класифікації, зокрема, двокласової, y_u^{pred} обчислюється згідно наступного виразу:

$$y_u^{pred} = \arg \max_{i=\overline{1 \ n}} (y_i + \hat{z}_i^{pred}). \quad (3.35)$$

На рис. 3.5 наведено структурно-функційну схему етапу застосування методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини

вхідних компонент векторів з використанням принципу каскадування ШНМ, згідно кроків, описаних вище.

Розроблений метод забезпечує можливість підвищення точності апроксимації при незначному збільшенні тривалості процедури навчання обраної нелінійної ШНМ.

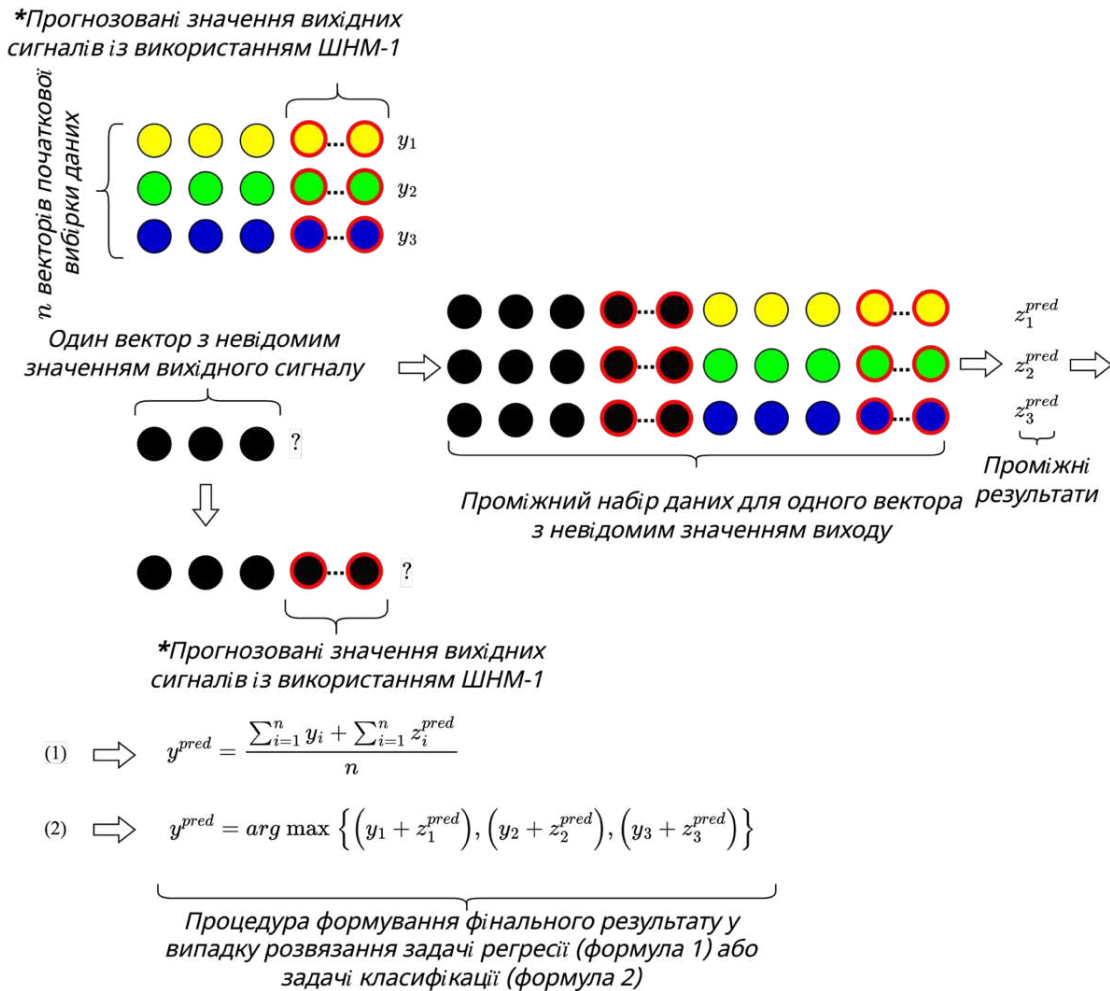


Рисунок 3.5 – Структурно-функційна схема етапу застосування методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використанням принципів каскадування ШНМ

Порівняно з базовим методом на основі декартового квадрата множини вхідних компонент векторів (підрозділ 3.2), розроблений каскадний підхід дозволяє

суттєво підвищити точність моделювання за рахунок розширення простору ознак [2]. Включення вихідних сигналів першої моделі (ШНМ-1) як нових атрибутів підвищує інформативність кожного вектора, що, у свою чергу, покращує здатність моделі другого рівня (ШНМ-2) до навчання навіть на екстремально коротких вибірках. Це також підвищує робастність методу до зашумлених даних. Реалізацію цього підходу можна виконати також для методу з підрозділу 3.3. Окрім цього, суттєвою перевагою запропонованого методу є можливість використання різних моделей на першому та другому етапах. Така гнучкість архітектури дозволяє, наприклад, поєднувати швидку ЙНМ для отримання попереднього результату (груба оцінка) з більш складною глибокою моделлю на другому рівні. У результаті досягається компроміс між точністю і швидкодією. Принцип каскадування, який покладено в основу роботи цього методу, також підвищує стійкість до перенавчання. Це відбувається тому, що додаткові ознаки не вводяться довільно, а генеруються на основі результатів первинної моделі, яка вже враховує основні закономірності у даних. Це особливо важливо при роботі з біомедичними даними високої вимірності в умовах обмеженої кількості спостережень.

3.5 Метод розв'язання задач регресії та класифікації на основі формування комбінованої поверхні відгуку

У другому розділі описано удосконалені автором методи підвищення точності роботи регресорів і класифікаторів на основі машинного навчання з використанням кластерного аналізу. В їх основу покладено можливість розширення простору вхідних даних задачі інформацією, отриманою після кластеризації. Використання різновидів таких методів дозволяє як суттєвим чином розширити простір вхідних даних, наприклад, мітками належності кожного спостереження до усіх центрів кластерів отриманих внаслідок кластеризації, так і несуттєво збільшити вимірність простору задачі, наприклад, вихідним сигналом центру відповідного кластеру. Утворена таким чином комбінована поверхня відгуку

забезпечує можливість підвищення точності розв'язання задач регресії чи класифікації, зокрема, як показано у розділі 2, – з використанням лінійних ШНМ. Такий підхід можна використати для підвищення точності роботи методів розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів, особливо у випадку, коли кластер описано декількома компактними множинами точок.

Враховуючи те, що вищевказані методи подвоюють кількість незалежних атрибутів задачі, доречно буде застосувати метод розширення початкової вибірки даних лише одним атрибутом – виходом центра відповідного кластера (підрозділ 2.2). Це зменшить тривалість процедури навчання методу у порівнянні з використанням інших методів, поданих у другому розділі. Слід зазначити, що застосування інших методів до розширення початкової вибірки даних результатами кластеризації, наприклад, мітками приналежності до усіх визначених кластерів задачі, також можливе для випадку невеликої кількості незалежних атрибутів у початковій вибірці біомедичних даних.

Розглянемо етап аугментації та етап застосування методу розв'язання задачі регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використанням комбінованої поверхні відгуку.

Як і у попередньому випадку оперуємо матрицями (3.1) для представлення заданої екстремально-короткої вибірки даних.

Основні кроки етапу аугментації даних для методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використанням комбінованої поверхні відгуку (Методу 3.4) такі.

1. Виконати кластеризацію заданої вибірки даних обраним методом, наприклад, залученням [241] на основі (2.33) чи (2.35) і визначенням оптимальної кількості кластерів C_{opt} , $c(i) \in \{1, \dots, C_{opt}\}$. Для цього можна використати методи з підрозділу 2.4.

2. Обчислити координати центрів включно з координатою залежного атрибуту ($y_{c(i)}^{clusterOutput}$, $i = \overline{1, n}$ для усіх виділених кластерів.

$$(\mu_c, y_c^{clusterOutput}) \in \mathbb{R}^m \times \mathbb{R}, c = \overline{1, C_{opt}}. \quad (3.36)$$

3. Виконати розширення кожного i -того вектора x_i початкової вибірки даних (3.1) за рахунок додавання $y_{c(i)}^{clusterOutput}$, до якого належить i -тий поточний вектор. Нова розширена матриця $\check{A} \in \mathbb{R}^{n \times (m+1)}$ матиме вигляд:

$$\check{A} = \begin{bmatrix} \check{x}_1^T \\ \check{x}_2^T \\ \dots \\ \check{x}_n^T \end{bmatrix} \in \mathbb{R}^{n \times (m+1)}, \quad \check{x}_i = (x_i, y_{c(i)}^{clusterOutput}) \in \mathbb{R}^{m+1}, \quad (3.37)$$

де $c(i)$ – індекс кластеру для x_i .

4. Виконати процедуру аугментації даних згідно (3.6), (3.7) на основі $\check{A}$:

$$\check{A}^+ = \begin{bmatrix} (\check{x}_1^+)^T \\ (\check{x}_2^+)^T \\ \dots \\ (\check{x}_{n^2}^+)^T \end{bmatrix} \in \mathbb{R}^{n^2 \times 2(m+1)}, \quad (3.38)$$

де для $a = (i-1)n + k$, $i, k = \overline{1, n}$:

$$\check{x}_a^+ = (\check{x}_i, \check{x}_k) \in \mathbb{R}^{2(m+1)}. \quad (3.39)$$

5. Обчислити нове значення залежного атрибуту z_i для кожного a так само як у (3.8).
6. Виконати навчання обраної ШНМ $f_4(\check{x}_a^+)$ на основі розширеної матриці $\check{A}^+$.

На рис. 3.6 наведено структурно-функційну схему етапу аугментації даних згідно кроків, поданих вище.

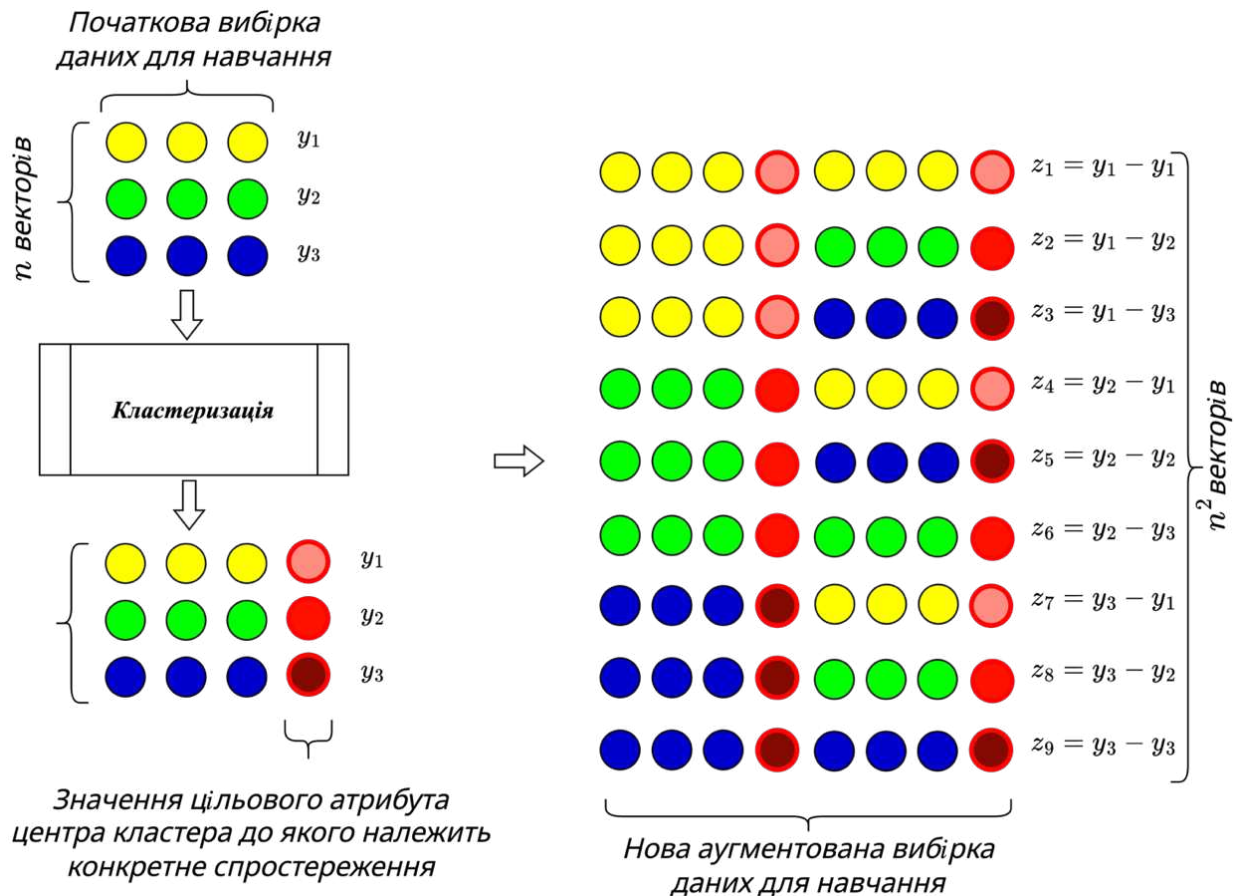


Рисунок 3.6 – Структурно-функційна схема етапу аугментації даних для методу розв’язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використанням комбінованої поверхні відгуку

На етапі застосування, як і у попередніх випадках, оперуємо вектором $u \in \mathbb{R}^m$, для якого невідомо значення вихідного сигналу.

Етап застосування Методу 3.4 передбачає виконання таких кроків.

1. Визначити належність тестового вектора u до одного з C_{opt} виділених кластерів, зокрема, шляхом пошуку найменшого значення Евклідової відстані між тестовим вектором та усіма центрами кластерів на основі 2.41.

2. Виконати розширення кожного u – того вектора за рахунок додавання $y_{c(u)}^{clusterOutput}$, до якого він належить, утворивши розширений вектор

$$\check{u} = (u, y_{c(u)}^{clusterOutput}) \in \mathbb{R}^{m+1}, \quad (3.40)$$

де $c(u)$ – індекс кластеру, обраного для u на основі 2.41.

3. Сформувати тимчасову вибірку даних $\check{B}^+ \in \mathbb{R}^{n \times 2m}$ шляхом конкатенації вектора $\hat{u}$ з кожним i -тим вектором $\hat{x}_i$ вибірки даних $\hat{A}$ з (3.26):

$$\check{B}^+ = \begin{bmatrix} (\check{u}_1^+)^T \\ (\check{u}_2^+)^T \\ \vdots \\ (\check{u}_n^+)^T \end{bmatrix} \in \mathbb{R}^{n \times 2(m+1)}, \quad (3.41)$$

де кожен розширений вектор $\hat{u}_i^+$ формується таким чином:

$$\check{u}_i^+ = (\check{u}, \hat{x}_i) \in \mathbb{R}^{2(m+1)}, \quad i = \overline{1 \ n}. \quad (3.42)$$

4. Для кожного рядка $\check{u}_i^+$ матриці $\hat{B}^+$ отримати шукане значення $\check{z}_i^{pred}$ шляхом застосування попередньо навченої ШНМ:

$$\check{z}_i^{pred} = f_4(\check{u}_i^+), \quad i = \overline{1 \ n}. \quad (3.43)$$

5. Сформувати остаточний шуканий сигнал y_u^{pred} для одного вектора u у випадку розв'язання задачі регресії згідно виразу:

$$y_u^{pred} = \frac{1}{n} \sum_{i=1}^n \check{z}_i^{pred} + \frac{1}{n} \sum_{i=1}^n y_i. \quad (3.44)$$

6. У випадку розв'язання задачі класифікації значення y_u^{pred} обчислити згідно наступного виразу:

$$y_u^{pred} = \arg \max_{i=\overline{1 \ n}} (y_i + \check{z}_i^{pred}). \quad (3.45)$$

На рис. 3.7 наведено структурно-функційну схему етапу застосування методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини

вхідних компонент векторів з використанням комбінованої поверхні відгуку, згідно з кроками, описаними вище.

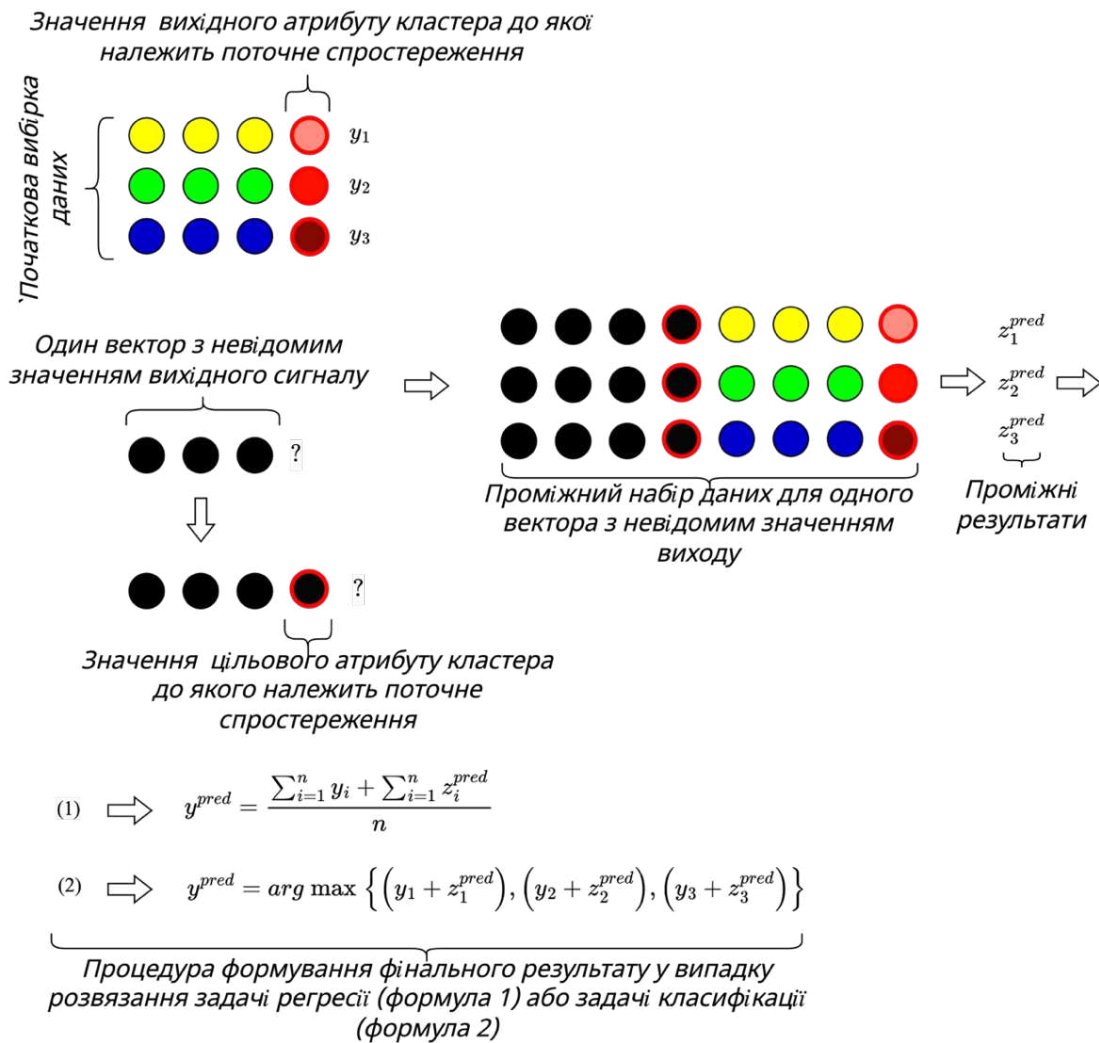


Рисунок 3.7 – Структурно-функційна схема етапу застосування методу розв'язання задач регресії/класифікації на основі декартового квадрата множини вхідних компонент векторів з використання комбінованої поверхні відгуку

Розроблений метод забезпечує можливість підвищення точності розв'язання задач регресії чи класифікації за рахунок врахування додаткової інформації, отриманої після виконання процедури кластеризації.

На відміну від базового методу (підрозділ 3.2), який оперує тільки ознаками вхідного простору, цей метод дозволяє додатково враховувати внутрішню

структуру даних, виявлену в процесі кластеризації. Це забезпечує додаткову інформацію про локальні закономірності та дозволяє уточнити результати навіть у випадках високого рівня перекриття між класами чи значною варіативністю регресійної залежності.

Метод також демонструє переваги в порівнянні з каскадним підходом (підрозділ 3.3.). Зокрема, якщо каскадування дозволяє розширювати вибірку вихідними сигналами попередньої моделі, то комбінована поверхня відгуку виконує розширення за рахунок геометричних властивостей простору, фактично реалізуючи адаптивну локалізацію ознак. Додатковою перевагою є те, що вихідна кластерна ознака, наприклад, середнє значення цільової змінної в межах локальної множини, забезпечує згладжування шумів та компенсує одиничні аномалії у даних. Це суттєво підвищує робастність моделі до флуктуацій, які часто трапляються у біомедичних вимірюваннях з обмеженою кількістю спостережень.

3.6 Метод формування аугментованої вибірки даних необхідного обсягу за рахунок кластеризації

Усі описані у попередніх пунктах цього розділу методи дозволяють підвищити точність розв'язання задач регресії чи класифікації у випадку аналізу екстремально коротких наборів даних. Проте, не зважаючи на обрану модифікацію методу, в їх основу покладено авторську процедуру аугментації даних. Згідно з нею обсяг аугментованої вибірки даних зростає квадратично по відношенню до обсягу початкової вибірки даних. У випадку наявності невеликої кількості векторів у вибірці для навчання (або опорній вибірці), наприклад 30, такий підхід має суттєві переваги, оскільки аугментована вибірка містить вже 900 векторів, які можливо ефективно аналізувати відомим методами. Проте, коли кількість векторів у початковій вибірці значно більша, наприклад 100, то аугментована вибірка міститиме 10000 векторів для реалізації процедури навчання. Це може суттєвим чином збільшити час навчання обраного ядрового методу навіть з неітеративним

машинним навчанням або інтерполяційних штучних нейронних мереж, враховуючи необхідність підбору оптимальних його чи її параметрів. Особливо гостро ця проблема постає в процесі використання методів з пункту 3.4. Вони передбачають використання двох ШНМ в процесі реалізації процедур навчання методів розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з використанням принципу каскадування методів машинного навчання.

У цьому підрозділі запропоновано підхід до підвищення ефективності роботи усіх, представлених у попередніх підрозділах, методів розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів. Розроблений метод суттєвим чином скорочуватиме обсяг аугментованої вибірки даних при збереженні, а в деяких випадках і підвищенні точності його роботи.

В основу розробленого методу покладено використання кластеризації як в процесі реалізації процедури аугментації даних, так і в процесі виконання етапу застосування довільного методу розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів.

Розглянемо їх детальніше. Як у попередніх випадках, для позначення навчальної вибірки даних використаємо (3.1).

Основні кроки етапу аугментації даних для методу формування аугментованої вибірки даних зменшеного обсягу (Методу 3.5.) такі.

1. Виконати кластеризацію заданої вибірки даних обраним методом на основі (2.33) чи (2.35) з визначенням оптимальної кількості кластерів $C_{opt}, c = \overline{1, C_{opt}}$. Для цього використати методи з підрозділу 2.4, розділу 2.
2. Обчислити координати центрів включно з координатою залежного атрибуту $(y_{c(i)}^{clusterOutput}, i = \overline{1, n})$ для кожного з $c = \overline{1, C_{opt}}$ кластерів:

$$(\mu_c, y_c^{clusterOutput}), \in \mathbb{R}^m \times \mathbb{R} \quad (3.46)$$

та представити їх у вигляді матриці:

$$C = \begin{bmatrix} \mu_1^T & y_1^{clusterOutput} \\ \mu_2^T & y_2^{clusterOutput} \\ \dots & \dots \\ \mu_{C_{opt}}^T & y_{C_{opt}}^{clusterOutput} \end{bmatrix} \in \mathbb{R}^{C_{opt} \times (m+1)}. \quad (3.47)$$

3. Сформувати матрицю аугментованої вибірки даних шляхом конкатенації усіх векторів з початкової вибірки даних у (3.1) та усіх центрів кластерів у (3.47):

$$\check{A}^+ = \begin{bmatrix} (\check{x}_1^+)^T \\ (\check{x}_2^+)^T \\ \dots \\ (\check{x}_{nC_{opt}}^+)^T \end{bmatrix} \in \mathbb{R}^{nC_{opt} \times 2m}, \quad (3.48)$$

де для кожної пари (i, c) , $i = \overline{1, n}$, $c = \overline{1, C_{opt}}$, $b = \overline{1, nC_{opt}}$:

$$\check{x}_b^+ = (x_i, \mu_c) \in \mathbb{R}^{2m}, b = (i - 1)C_{opt} + c. \quad (3.49)$$

4. Обчислити нове значення залежного атрибуту z_i для кожного a так:

$$\check{z}_b = y_c^{clusterOutput} - y_i, \quad i = \overline{1, n}, \quad c = \overline{1, C_{opt}}, \quad b = (i - 1)C_{opt} + c. \quad (3.50)$$

5. Виконати навчання нелінійної ШНМ $f_5(\check{x}_a^+)$, $a = \overline{1, nC_{opt}}$ на основі розширеної матриці $\check{A}^+$.

В результаті виконання усіх вищевказаних дій обсяг аугментованої вибірки даних зростає не квадратично (n^2), а пропорційно до оптимальної кількості компактних множин точок, виділених з досліджуваної вибірки даних ($n \times C_{opt}$).

На рис. 3.8 наведено структурно-функційну схему етапу аугментації даних згідно описів, поданих вище.

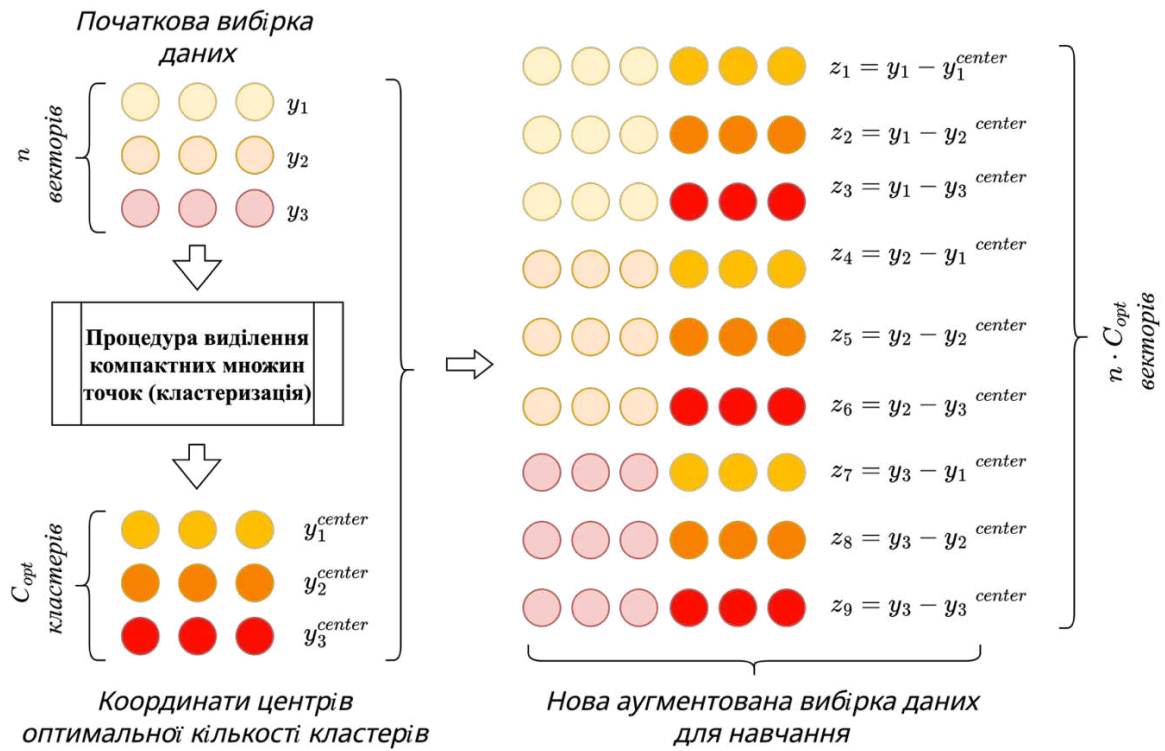


Рисунок 3.8 – Структурно-функційна схема формування аугментованої вибірки даних зменшеного обсягу у методах розв’язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів

На основі процедури (3.6), побудованої з використанням (3.26), виконується процедура навчання обраної нелінійної ШНМ.

Як і у попередніх випадках для опису етапу застосування розробленого методу, оперуємо вектором $u \in \mathbb{R}^m$, для якого невідоме значення вихідного сигналу.

Основні кроки етапу застосування для Методу 3.5:

1. Сформувати тимчасову вибірку даних $\check{B}^+ \in \mathbb{R}^{n \times 2m}$ шляхом конкатенації вектора $\check{y}$ з кожним s -тим центром кластера μ_c :

$$\check{B}^+ = \begin{bmatrix} (\check{y}_1^+)^T \\ (\check{y}_2^+)^T \\ \dots \\ (\check{y}_{C_{opt}}^+)^T \end{bmatrix} \in \mathbb{R}^{C_{opt} \times 2m}, \quad (3.51)$$

де кожен розширений вектор $\check{u}_i^+$ формується таким чином:

$$\check{u}_c^+ = (u, \mu_c) \in \mathbb{R}^{2m}, \quad c = \overline{1, C_{opt}}. \quad (3.52)$$

2. Для кожного рядка $\check{u}_i^+$ матриці $\check{B}^+$ отримати шукане значення $\check{z}_i^{pred}$ шляхом застосування попередньо навченої ШНМ:

$$\check{z}_c^{pred} = f_5(\check{u}_c^+), \quad c = \overline{1, C_{opt}}. \quad (3.53)$$

3. Сформувати остаточний шуканий сигнал y_u^{pred} для заданого вектора u у випадку розв'язання задачі регресії згідно виразу:

$$y_u^{pred} = \frac{1}{C_{opt}} \sum_{c=1}^{C_{opt}} \check{z}_c^{pred} + \frac{1}{C_{opt}} \sum_{c=1}^{C_{opt}} y_c^{clusterOutput}. \quad (3.54)$$

4. У випадку розв'язання задачі класифікації значення y_u^{pred} слід обчислювати згідно наступного виразу:

$$y_u^{pred} = \arg \max_{c=1}^{C_{opt}} (y_c^{clusterOutput} + \check{z}_c^{pred}). \quad (3.55)$$

На рис. 3.9 наведено структурно-функційну схему етапу застосування методів розв'язання задач регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів з залученням сформованої аугментованої вибірки даних зменшеного обсягу.

На відміну від базового методу (підрозділ 3.2), який оперує повним декартовим добутком вхідних векторів і має квадратичну складність $O(n^2)$ щодо обсягу навчальної вибірки даних, розроблений метод використовує компактне представлення даних через центри кластерів. Це дозволяє лінійно скоротити розмір аугментованої вибірки до $O(nC_{opt})$, $C_{opt} \ll n$ без втрати репрезентативності ознак.

Окрім цього, цей метод забезпечує кращу здатність до узагальнення, оскільки формує навчальні приклади на основі представників кластерів. Це зменшує надмірність парних зв'язків та ризик перенавчання.

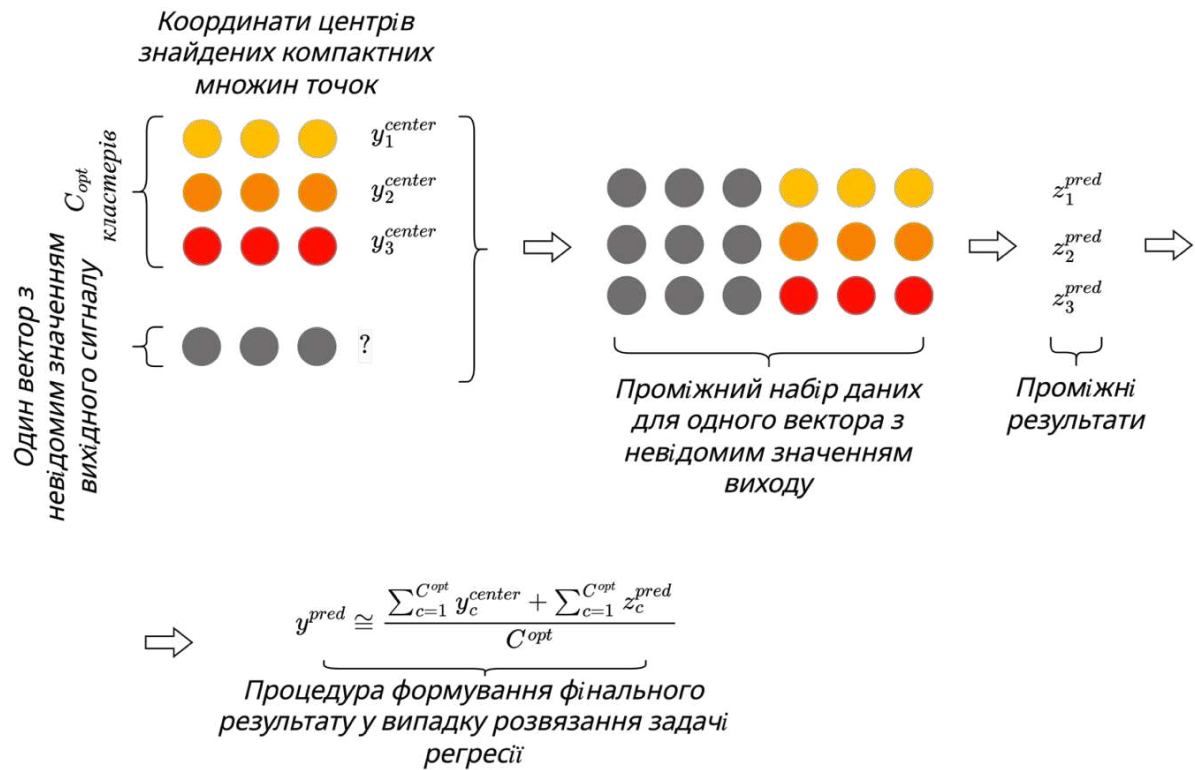


Рисунок 3.9 – Структурно-функційна схема етапу застосування методів розв'язання задач регресії або класифікації на основі аугментованої вибірки даних зменшеного обсягу

Найбільшою перевагою цього методу є те, що він є універсальним у поєднанні з будь-яким іншим методом розв'язання задачі на основі декартового добутку множини вхідних компонент векторів. Його можна застосувати до будь-якого з попередніх методів як механізм попереднього скорочення обсягу аугментованої вибірки. Це значно підвищує практичну гнучкість та масштабованість методів, а також зменшує час їх навчання при збереженні високої точності роботи.

3.7 Висновки до розділу

1. У розділі запропоновано принципово новий підхід до аугментації табличних вибірок даних та розроблено методологію розв'язання задач регресії та класифікації на його основі. Основою цього підходу є використання декартового

квадрата множини вхідних компонент векторів, що забезпечує квадратичне збільшення обсягу вибірки для ефективного виконання етапів навчання. Такий підхід дозволяє використовувати існуючі ШНМ нелінійного типу для аналізу екстремально коротких вибірок даних. Більше того, він дозволяє охопити ширший спектр нелінійних моделей ШНМ, включаючи ітеративні, неітеративні та інтерполяційні ШНМ, які, зазвичай, демонструють недостатню ефективність в таких умовах. Водночас, важливо підкреслити, що застосування розробленого підходу до лінійних ШНМ не призводить до покращення результатів через ефект лінійної суперпозиції сигналів.

2. Розроблений метод розв'язання задач регресії та класифікації завдяки авторській процедурі аугментації, неітеративному навчанню з використанням нейроподібної структури МППП з РБФ-розширенням входів, а також принципом усереднення чи голосування на етапі застосування методу може значно підвищити точність порівняно з рядом існуючих методів, що використовуються для аналізу вибірок даних подібних обсягів.

3. Використання адитивного методу розв'язання задач регресії/класифікації забезпечує часткову взаємну компенсацію випадкових похибок різних знаків. Це пояснюється використанням удосконаленої процедури отримання фінального результату, яка ґрунтується на принципах ансамблевого навчання. Цей метод може значно підвищити точність результатів, зокрема, в процесі аналізу екстремально коротких вибірок біомедичних даних. Окрім цього, така перевага забезпечує робастність до аномалій та викидів у даних, ідентифікація яких є складною задачею при обробці обмежених вибірок даних.

4. Використання принципів каскадування штучних нейронних мереж, детально розглянутих у другому розділі, дозволило модифікувати базовий та адитивний методи розв'язання задач регресії чи класифікації. В основу модифікації покладено можливість інтеграції результатів роботи ШНМ першого рівня на другий рівень за умови, що вони забезпечують задовільну точність результату

апроксимації на початковій вибірці даних. У таких випадках ефективним може бути використання нейронної мережі узагальненої регресії або ймовірнісної нейронної мережі як ШНМ першого рівня методу. Подальший аналіз розширеної таким чином, аугментованої вибірки даних з використанням операції декартового квадрата та неітеративної нейроподібної структури МППП з РБФ-розширенням входів повинний значно підвищити точність методу порівняно з точністю за базовим варіантом.

5. Модифікація розроблених методів шляхом формування комбінованої поверхні відгуку забезпечує можливість врахування геометричних властивостей досліджуваної вибірки даних. Використання різних варіантів побудови комбінованої поверхні відгуку, детально розглянутих у другому розділі, дозволяє ефективно розв'язувати широкий спектр прикладних задач з урахуванням таких параметрів, як обсяг вибірки, кількість незалежних атрибутів тощо. Це забезпечить підвищену точність, а особливо адаптивність методів цього класу до особливостей аналізованих даних.

6. Для скорочення тривалості навчання методів, описаних у цьому розділі, особливо в умовах аналізу вибірок даних малих обсягів (від 100 до 1000 спостережень), розроблено метод формування аугментованої вибірки даних зменшеного обсягу. В його основу покладено використання кластерного аналізу на початковій вибірці даних. Операція декартового квадрата в цьому випадку виконується шляхом конкатенації кожного вектора початкової вибірки даних з кожним визначеним центром кластера. Суттєве скорочення обсягу аугментованої вибірки даних в цьому випадку дозволить суттєво прискорити навчання нелінійних моделей при збереженні чи незначному погіршенні точності їх роботи.

РОЗДІЛ 4

МЕТОДИ СИНТЕЗУ АНСАМБЛІВ ШНМ НА ОСНОВІ КОМБІНАЦІЙ АПРОКСИМОВАНИХ ПОВЕРХОНЬ ВІДГУКУ

У розділі запропоновано підхід до розв’язання задач регресії та класифікації для малих вибірок даних, що базуються на комбінуванні глобально апроксимованих поверхонь відгуку. На цій основі розроблено методи синтезу ансамблів гомогенних штучних нейронних мереж, які дозволяють комбінувати як зміщені функції, так і зони кусково-лінійної апроксимації. В основу запропонованого підходу покладено та обґрунтовано можливість формування композиції ансамблів з однорідних нейронних мереж лінійного або нелінійного типу шляхом створення декорельованих вибірок даних, унікальних для кожного елемента ансамблю. Для цього використано кілька методів декореляції. Зокрема, ансамбль лінійних неітеративних нейроподібних структур МПГП синтезується з використанням кусково-лінійної апроксимації багатовимірної поверхні відгуку. Ансамбль інтерполятивних нейронних мереж синтезується з використанням методу випадкових зміщень, а ансамбль неітеративних нейроподібних структур МПГП з РБФ-розширенням входів – з використанням попарних симетричних випадкових зміщень для декореляції вхідної вибірки даних. Показано, що усі синтезовані ансамблі можуть бути реалізовані на основі усереднення чи голосування більшістю (клас бегінгових методів), а також блендінгу (спрощеної схеми стекінгу) чи класичної схеми стекінгу.

Наукові і практичні результати, описані у цьому розділі, опубліковано у періодичних виданнях України [25], [23], [26], [24], а також апробовано і опубліковано у матеріалах українських та закордонних конференцій: [18], [22], [27], [28], [29], [30].

4.1 Комбінування апроксимованих поверхонь відгуку для підвищення точності інтелектуального аналізу даних малих обсягів

Проаналізуємо альтернативні варіанти апроксимації поверхонь відгуку при вирішенні завдань регресії та класифікації багат шаровим перцептроном та нейронними мережами з радіально-базисними функціями. Хоча обидва методи належать до класу нейронних мереж, вони принципово відрізняються за способом формування апроксимувальних функцій і геометричною інтерпретацією просторових форм цих функцій [147].

Багат шаровий перцептрон використовує комбінацію поверхонь схилів. Це означає, що нейрони прихованих шарів БШП реалізують лінійні комбінації вхідних сигналів, які трансформуються за допомогою нелінійних активаційних функцій (зазвичай сигмоїдальної, гіперболічного тангенса або ReLU) [242]. Таким чином, кожен нейрон БШП формує своєрідну поверхню-схил у просторі вхідних даних. Збільшуючи кількість нейронів, створюється композиція поверхонь різних напрямів і крутизни, що дозволяє апроксимувати складні функції з різкими перепадами та складною геометрією [243]. Теоретичне обґрунтування такої здатності ґрунтується на універсальних апроксимаційних властивостях БШП [117], які свідчать про можливість наближення практично будь-якої неперервної функції за умови достатньої кількості нейронів у прихованих шарах.

На противагу цьому мережі з радіально-базисними функціями (РБФ-мережі) формують поверхню відгуку як комбінацію радіально-симетричних базисних функцій [84]. Основним елементом таких мереж є нейрони, виходи яких визначаються відстанню між вхідним вектором і центром нейрона [244]. Типовим представником таких функцій є гаусівська функція, поверхня якої має форму симетричного склепіння. Поверхня відгуку РБФ-мережі формується як лінійна комбінація таких гіркоподібних функцій з різними ширинами і розташуванням центрів [245]. За результатами досліджень у [246] саме така композиція радіально-

базисних функцій дозволяє якісно апроксимувати складні нелінійні залежності, забезпечуючи хорошу інтерполяцію локальних особливостей цільової функції, проте погіршуючи екстраполятивну здатність мережі.

Незважаючи на це, використання одної ШНМ у випадку аналізу малих даних не завжди забезпечує такий рівень точності, який необхідний користувачу. Це обумовлено рядом причин, серед яких слід виділити похибки моделі ШНМ, висока дисперсія, сигнали шумів у даних тощо.

Розглянемо окремий регресор (класифікатор) $f(x)$. Середньоквадратична похибка результатів його роботи в процесі розв'язання задач регресії/класифікації може бути розкладена за формулою [247]:

$$E[(f(x) - y)^2] = E_s^2(f(x)) + B(f(x)) + \eta^2, \quad (4.1)$$

де η – шумова компонента в даних; E_s – систематична похибка моделі, яка характеризує відхилення середнього результату роботи моделі від істинного значення; B – дисперсія моделі, яка відображає ступінь варіації результатів апроксимації при зміні навчальної вибірки; $f(x)$ – результат апроксимації моделі (результуючого значення) в точці x ; y – істинне (спостережуване) значення цільової змінної у точці x .

Такий підхід можна реалізувати з використанням концепції ансамблевого навчання, що є одним з потужних інструментів підвищення точності розв'язання задач класифікації та регресії ШНМ/методами машинного навчання [227], [248], [249]. В його основі покладено ідею використання декількох моделей для отримання фінального результату, що, як правило, демонструє суттєво вищу точність результату, в тому числі й за рахунок різних способів агрегування результатів роботи різних моделей. Такий підхід дозволяє зменшити як η , так і систематичні помилки E_s окремих моделей. Зокрема, при комбінуванні L базових моделей $f_l(x), l = \overline{1, L}$, кожна з яких має однакову дисперсію $B(f(x))$, середнє арифметичне результатів апроксимації визначається так:

$$y^{pred}(x) = \frac{1}{L} \sum_{l=1}^L f_l(x), \quad (4.2)$$

За припущенням незалежності кожної з базових моделей ШНМ $f_1(x), f_2(x), \dots, f_L(x)$ (слабкі регресори, класифікатори) дисперсія комбінованого результату апроксимації:

$$\mathcal{B}(y^{pred}(x)) = \frac{1}{L^2} \sum_{l=1}^L \mathcal{B}(f_l(x)) = \frac{\mathcal{B}(f_l(x))}{L}. \quad (4.3)$$

Таким чином, збільшення кількості L базових моделей призводить до зменшення дисперсії, що є однією з основних переваг ансамблевих методів. Окрім цього, ефективність такого підходу також базується і на принципі “мудрості натовпу”, що підтверджується теоремою Кондорсе [250] і розвинено у першому розділі цієї роботи. Також у [249] розглянуто граничні випадки агрегування рішень елементів ансамблю у випадку розв’язання задачі класифікації. Тут показано, що ймовірність правильної класифікації суттєво зростає і наближається до одиниці при збільшенні кількості елементів ансамблю, зокрема, за умови декорельованості характеристик членів ансамблю.

Одним з ефективних способів формування результату розв’язання задач регресії чи класифікації підвищеної точності на основі зваженого агрегування результатів роботи різних моделей є стекінг або його простіша форма – блендінг. Такий підхід запропоновано Вольпартом ще у 1992 році [251]. В його основу покладено ідею паралельного використання L базових моделей $f_1(x), f_2(x), \dots, f_L(x)$ машинного навчання (слабкі регресори, класифікатори), вихідні сигнали яких формують нову вибірку даних для навчання мета-моделі $g(\cdot)$, тобто мета-регресор у випадку розв’язання задачі регресії чи мета-класифікатор у випадку розв’язання задачі класифікації. Результати роботи власне мета-моделі будуть результатами роботи усього стекінгового ансамблю:

$$y^{pred}(x) = g(f_1(x), f_2(x), \dots, f_L(x)). \quad (4.4)$$

Якщо кожна з базових моделей $f_1(x), f_2(x), \dots, f_L(x)$ робить різні помилки, тобто їх похибки є некорельованими або слабо корельованими [249], мета-модель здатна навчитися коректно їх “зважувати” та компенсувати слабкі сторони окремих елементів ансамблю.

Зокрема, якщо мета-модель використовує лінійну комбінацію результатів апроксимації, то (4.2) можна записати так:

$$y^{pred}(x) = \sum_{l=1}^L w_l f_l(x), \quad (4.5)$$

де $\sum_{l=1}^L w_l = 1$.

Тоді (4.3) набуде наступного вигляду:

$$\mathcal{B}(y^{pred}(x)) = \sum_{l=1}^L w_l^2 \mathcal{B}(f_l(x)). \quad (4.6)$$

При однакових вагових коефіцієнтах $w_l = \frac{1}{L}$ (4.6) співпадає з виразом (4.3), який відповідає способу формування результатів роботи ансамблю на основі середнього арифметичного (розділ 3). Однак, мета-модель може навчитися оптимізувати ваги w_l таким чином, щоб забезпечити більше зниження дисперсії, ніж підхід на основі середнього арифметичного, особливо коли якість окремих моделей суттєво відрізняється.

Ефективність стекінгового ансамблю значною мірою залежить від розмаїття його компонент [252]. Припустимо, що похибки окремих моделей мають кореляцію ρ . Якщо усі базові моделі мають однакову дисперсію $\mathcal{B}(f(x))$ тоді дисперсія середнього результату апроксимації може бути записана так [253]:

$$\mathcal{B}(y^{pred}(x)) = \mathcal{B}(f(x)) \left(\frac{1}{L} + \frac{L-1}{L} \rho \right). \quad (4.7.)$$

Якщо ρ є малим або близьким до нуля, зменшення дисперсії буде суттєве, що наочно демонструє важливість розмаїття базових моделей. Так само можна показати необхідність незалежності вибірок даних для ансамблів з класу бегінг.

Отож, серед основних способів досягнення високої ефективності роботи такого підходу слід виділити два [27]. Перший з них передбачає використання декількох різних ШНМ (гетерогенні ансамблі), результати яких агрегуються обраною мета-моделлю. В основу другого підходу покладено можливість використання однакових ШНМ (однорідні ансамблі), проте з різним підходом до формування навчальної або опорної для випадку використання інтерполятивних нейронних мереж вибірки даних.

Обидва підходи відповідають принципам побудови бегінгових чи стекінгових ансамблів, проте у випадку аналізу даних малих обсягів вони мають свої особливості. Композиція ансамблів гетерогенних методів машинного навчання має ряд суттєвих недоліків у випадку аналізу даних малих обсягів. Серед них слід виділити такі [27]:

- необхідність ретельного вибору різних ШНМ як компонент першого рівня ансамблю, які забезпечують достатню точність аналізу або, принаймні, здатні генерувати адекватні результати розв'язанням задач регресії чи класифікації;
- необхідність поділу вибірки даних для навчання на дві частини, де перша використовується для навчання елементів першого рівня ансамблю, а друга – для навчання мета-моделі. Це є критичною проблемою у випадку аналізу малих даних і може призводити до втрати важливої інформації;
- висока тривалість процедури навчання всіх членів ансамблю у випадку використання методів з ітеративними методами навчання, що ускладнює застосування таких методів у реальному часі;
- необхідність підбору великої кількості оптимальних параметрів, які можуть суттєво відрізнятися для кожного члена гетерогенного ансамблю, що робить процес налаштування такого ансамблю складним та ресурсоємким;
- інші недоліки.

Саме тому подібні ансамблі обмежуються певним вузьким класом задач в галузі біомедичної інженерії та не завжди знаходять практичне застосування в процесі розв'язання ширшого спектру задач [27].

На відміну від описаних класичних підходів у цьому розділі запропоновано інший підхід. Він полягає в тому, що вибірка даних апроксимується задану кількість разів глобальними на області визначення моделями ШНМ, а потім результати агрегуються, наприклад, шляхом усереднення чи за схемою стекінгу. У випадку використання однорідних ШНМ для кожного з них формується своя унікальна дещо зміщена вибірка даних.

Ансамблі однорідних ШНМ (гомогенні ансамблі) передбачають використання різних підвибірок даних для навчання кожного окремого елемента ансамблю. Причому кількість таких підвибірок залежить від кількості елементів ансамблю, яких, на основі (4.3), повинно бути багато. Традиційний підхід до формування таких вибірок, який використовується у випадку аналізу даних великих обсягів, зокрема, за рахунок поділу загальної вибірки на частини, не належить до раціональних варіантів [27]. Це пояснюється обмеженістю обсягу даних у випадку аналізу малих вибірок. Саме тому у цьому розділі пропонуються нові ансамблі з включенням однорідних ШНМ за рахунок формування декорельованих вибірок даних для кожного окремого елемента ансамблю. Їх можна використовувати для розв'язання різноманітних задач регресії чи класифікації в умовах дефіциту навчальних даних.

Такий підхід також знижує ризик перенавчання. У порівнянні з попереднім класом ансамблів (гетерогенні ансамблі) зниження ризику перенавчання забезпечується тим, що базові регресори/класифікатори згідно розробленого підходу навчаються на декорельованих вибірках даних. Це дозволяє мінімізувати випадкові помилки та покращити властивості узагальнення. Агрегування вихідних сигналів однорідних ШНМ забезпечує отримання стабільних результатів і значно спрощує налаштування параметрів ансамблю, оскільки гіперпараметри його

елементів можна оптимізувати уніфіковано [27]. Це дозволяє ефективно використовувати обмежений обсяг даних і знизити вплив варіативності.

Крім того, використання нейронних мереж узагальненої регресії чи ймовірнісних нейронних мереж як базових елементів ансамблю усуває необхідність поділу вибірки на дві частини для ансамблю з однорідних елементів. Це дозволяє як базовим моделям, так і мета-моделі проводити аналіз усієї доступної вибірки даних, що є критично важливо в процесі опрацювання даних обмежених обсягів [21], [22]. А композиції ансамблів на основі лінійних чи нелінійних неітеративних штучних нейронних мереж або ШНМ, що не потребують окремої процедури навчання для кожного члена ансамблю, додатково скорочують час роботи таких методів.

Нехай маємо навчальну вибірку даних малого обсягу:

$$A \in \mathbb{R}^{n \times m}, \quad (4.8)$$

де n – кількість векторів, m – вимірність кожного вектора. При цьому кожен рядок A позначимо як $x_i \in \mathbb{R}^m$ для $i = \overline{1, n}$, а відповідні значення залежного атрибута y_i для кожного $x_i \rightarrow y_i$ подані одним з варіантів:

$$\begin{cases} y_i \in \{1, 2, \dots, K\}, & \text{для задач класифікації;} \\ y_i \in \mathbb{R}, & \text{для задач регресійного аналізу.} \end{cases} \quad (4.9)$$

Проблема полягає у виборі базових моделей (слабкі регресори/класифікатори), які будуть мати однакову архітектуру з спільною функціональною формою $h(\tilde{A}^{(l)}; \theta)$, де θ – оптимальні параметри моделі; але кожна з них буде навчатися на власній незалежно утвореній вибірці даних $\tilde{A}^{(l)} \in \mathbb{R}^{n \times m}$ для $l = \overline{1, L}$.

Кожна базова модель ШНМ генерує результат, який можна подати у вигляді $y_{basic}^{(l)} = h(\tilde{A}^{(l)}; \theta_l)$.

Далі потрібно виконати агрегування результатів усіх базових моделей:

$$y^{pred} = g(h(\tilde{A}^{(1)}; \theta_1), \dots, h(\tilde{A}^{(L)}; \theta_L); \phi), \text{ або } y^{pred} = \frac{1}{L} \sum_{l=1}^L y_{basic}^{(l)}, \quad (4.10)$$

де ϕ – оптимальні параметри обраної мета-моделі.

Цей підхід дозволяє, використовуючи однакову модельну архітектуру, застосовувати базові моделі ШНМ на різних декорельованих варіантах вибірки (матриці) A . Це забезпечує зменшення кореляції результатів їх роботи. Подальше агрегування отриманих результатів через використання обраної мета-моделі $g(\cdot; \phi)$ забезпечує підвищення точності розв'язання задач регресії або класифікації на основі таких методів.

4.2 Ансамбль лінійних однорідних ШНМ з використанням методу кусково-лінійної апроксимації багатовимірної поверхні відгуку

Чимало задач інтелектуального аналізу малих обсягів даних можна розв'язувати за допомогою лінійних засобів машинного навчання. Ці методи вирізняються високою швидкістю, простотою реалізації та мінімальними затратами часу й ресурсів на налаштування параметрів. Одним з перспективних підходів до розв'язання задач регресії або класифікації в цьому напрямку є використання лінійної неітеративної нейроподібної структури МПГП [134]. Принципи функціонування її та основні кроки реалізації детально описано у першому розділі цієї роботи. Серед переваг такого рішення є однотипність методів навчання та функціонування, що визначає простоту реалізації нейроподібної структури на основі використання єдиного набору базових арифметичних операцій [134]. Також лінійна нейроподібна структура МПГП практично не вимагає налаштування параметрів, що робить її простим та автоматичним рішенням для користувачів з обмеженим досвідом у галузі штучного інтелекту [20]. Окрім цього тривалість неітеративної процедури навчання цієї ШНМ визначається заздалегідь відомою кількістю кроків трансформації і залежить тільки від апаратного забезпечення [19].

Незважаючи на усі вказані переваги, цей тип неітеративних ШНМ та інші лінійні методи не завжди демонструють задану точність [43]. У випадку, коли точність роботи лінійної нейроподібної структури МПГП є задовільною, проте все ще недостатньою для розв'язання конкретної задачі в галузі біомедичної інженерії, варто використати ансамбль таких моделей. В його основі можна покласти стекінговий/бегінговий підхід до агрегування результатів роботи ряду незалежних моделей, що дозволяє зменшити як систематичні похибки, так і дисперсію результатів апроксимації. З одного боку такий підхід компенсує індивідуальні недоліки окремих моделей ШНМ, підвищуючи загальну точність методу розв'язання задач регресії чи класифікації. З іншого, – зберігає простоту реалізації подібних моделей [22]. Окрім цього, неітеративний метод навчання у випадку використання лінійних нейроподібних структур МПГП забезпечить високу швидкодію роботи в процесі практичної реалізації такої моделі.

У цьому підрозділі описано новий ансамбль лінійних неітеративних нейроподібних структур МПГП для розв'язання задач регресії чи класифікації в умовах малих вибірок даних. В основу композиції такого ансамблю покладено метод декореляції малої вибірки даних із залученням кусково-лінійної апроксимації багатовимірної поверхні відгуку [21], який детально описано у другому розділі цієї роботи. Він дозволяє створювати унікальні вибірки даних для кожного окремого елемента ансамблю лінійних ШНМ.

На рис. 4.1 подано структурно-функційну схему роботи ансамблю лінійних неітеративних нейроподібних структур МПГП, побудованих на декорельованих вибірках даних методом кусково-лінійної апроксимації.

Розглянемо методи навчання та застосування розробленого ансамблю.

Неітеративний метод навчання розробленого ансамблю лінійних неітеративних нейроподібних структур МПГП, побудованих на декорельованих вибірках даних з використанням методу кусково-лінійної апроксимації багатовимірної поверхні відгуку, передбачає послідовне виконання таких кроків.

1. Нормалізувати незалежні атрибути навчальної вибірки даних A згідно обраної функції $\mathcal{N}(\cdot)$. В результаті нормалізації отримаємо нову матрицю $\hat{A}$ виду:

$$\hat{A} = \mathcal{N}(A) = \mathcal{N}[\hat{x}_{i,j}], i = \overline{1, n}, j = \overline{1, m}. \quad (4.11)$$

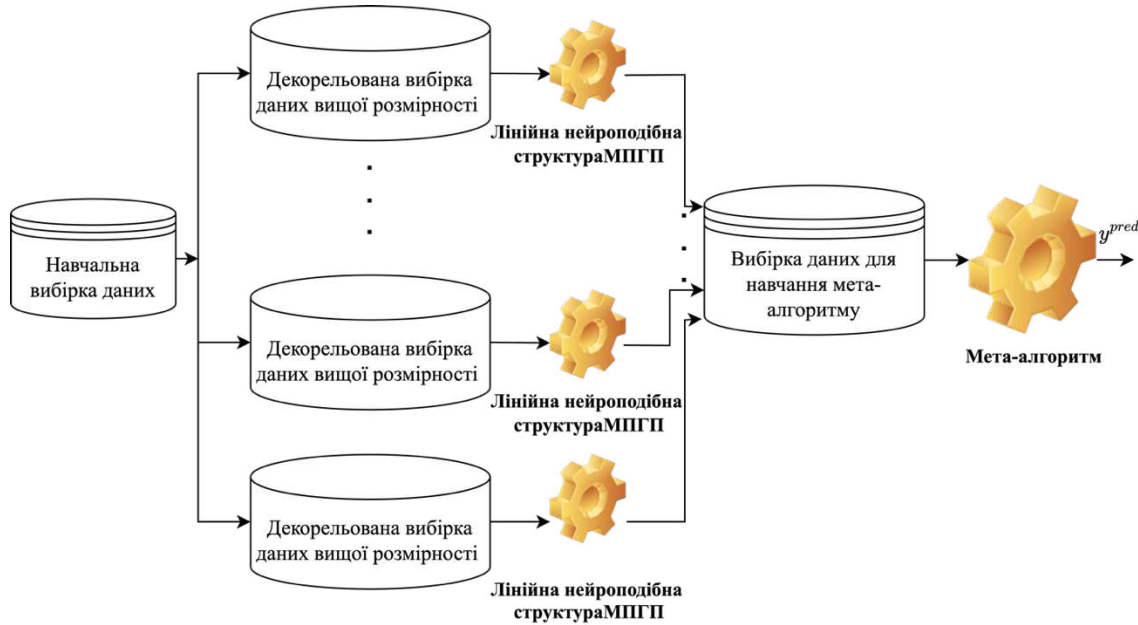


Рисунок 4.1 – Структурно-функційна схема роботи ансамблю лінійних неітеративних нейроподібних структур МПГП, побудованих на незалежно сформованих вибірках даних методом кусково-лінійної апроксимації багатовимірної поверхні відгуку.

2. Встановити загальну кількість L елементів ансамблю $l = \overline{1, L}$. Її може задати або визначити як параметр процедури оптимізації користувач. Відповідно до методу кусково-лінійної апроксимації багатовимірної поверхні відгуку встановити кількість частин s , $s = \overline{1, S}$, на які буде розділений кожен незалежний атрибут $\hat{x}_{i,j}$. Значення кожної з s частин буде знаходитися в межах діапазону $[a, b]$, де значення a, b визначаються обраною функцією нормалізації $\mathcal{N}(\cdot)$. В загальному випадку S задає

користувач. Проте S може бути гіперпараметром, який налаштовується в процесі оптимізації методу.

3. Побудувати ансамбль з L однорідних лінійних нейроподібних структур МПГП SGTM одночасно для кожного елемента ансамблю з індексом $l = \overline{1, L}$. Побудувати унікальну декорельовану вибірку даних $\tilde{A}^{(l)}$, кожний елемент якої формується методом кусково-лінійної апроксимації багатовимірної поверхні відгуку, який описано у другому розділі. Зокрема, у випадку утворення L нових векторів, вирази (2.29), (2.30) набувають такого вигляду:

$$\dot{x}_{i,j,s}^{(l)} = \begin{cases} 0, & \text{якщо } \dot{x}_{i,j} \leq a_{s-1}^{(l)} \\ \dot{x}_{i,j} - a_{s-1}^{(l)}, & \text{якщо } a_{s-1}^{(l)} < \dot{x}_{i,j} \leq a_s^{(l)}, s = \overline{1, S}, \\ a_s^{(l)} - a_{s-1}^{(l)}, & \text{якщо } \dot{x}_{i,j} > a_s^{(l)} \end{cases} \quad (4.12)$$

за умови, що сума усіх частин згідно перетворення (4.12) зберігає початкове значення $\dot{x}_{i,j}$:

$$\dot{x}_{i,j} = \sum_{s=1}^S \dot{x}_{i,j,s}^{(l)}, i = \overline{1, n}, j = \overline{1, m}. \quad (4.13)$$

У випадку побудови ансамблю з кількістю частин $s, s = \overline{1, S}$, на які буде поділятися кожна незалежна ознака кожного вектора даних, s буде однаковою для усіх L елементів ансамблю. Проте границі:

$$a = a_0^{(l)} < a_1^{(l)} < a_2^{(l)} < \dots < a_{S-1}^{(l)} < a_S^{(l)} = b \quad (4.14)$$

для S частин кожного окремого l -того елемента ансамблю $l = \overline{1, L}$ будуть різними. Вони генеруватимуться випадковим чином в межах діапазону $[a, b]$. Саме тому вираз (4.12) забезпечить утворення унікальних декорельованих компонент $\tilde{x}_{i,j,s}^{(l)} = \dot{x}_{i,j,s}^{(l)}, i = \overline{1, n}, j = \overline{1, m}, s = \overline{1, S}$ для розширених векторів виду:

$$\tilde{x}_{i,j}^{(l)} = (\dot{x}_{i,j,1}^{(l)}, \dot{x}_{i,j,2}^{(l)}, \dots, \dot{x}_{i,j,S}^{(l)}), \quad (4.15)$$

а для вхідного вектора:

$$\tilde{x}_i^{(l)} = (\tilde{x}_{i,1}^{(l)}, \tilde{x}_{i,2}^{(l)}, \dots, \tilde{x}_{i,m}^{(l)}). \quad (4.16)$$

4. Зберегти кількість частин s , $s = \overline{1, S}$, на які буде розділений кожен незалежний атрибут, а також границі областей цих частин (4.14) для кожного l -того елемента ансамблю $l = \overline{1, L}$.
5. На основі (4.16) утворити l декорельованих вибірок навчальних даних $\tilde{A}^{(l)} \in \mathbb{R}^{n \times (m \times S)}$ вищої вимірності, кожна з яких використовується виключно для побудови відповідної моделі регресії (класифікації) $SGTM_l(\tilde{A}^{(l)})$:

$$\tilde{A}^{(l)} = \begin{bmatrix} \tilde{x}_1^{(l)} \\ \tilde{x}_2^{(l)} \\ \dots \\ \tilde{x}_n^{(l)} \end{bmatrix}, \in \mathbb{R}^{n \times (m \cdot S)}, l = \overline{1, L}, \quad (4.16)$$

$$\tilde{x}_i^{(l)} \in \mathbb{R}^{m \cdot S}, i = \overline{1, n}.$$

Слід зазначити, що залежний атрибут y_i , $i = \overline{1, n}$ для кожного вектора $\tilde{x}_i^{(l)}$ залишається таким самим як і для початкової вибірки даних A відповідно до виразу (4.9).

6. Виконати неітеративний метод навчання для усіх L однорідних лінійних нейроподібних структур МПГП - $SGTM_l(\tilde{A}^{(l)})$.
7. Отримати вихідні сигнали від усіх L елементів ансамблю $y_i^{pred(l)} = SGTM_l(\tilde{A}^{(l)})$ для кожного i -го вектора з $\tilde{A}^{(l)}$, $i = \overline{1, n}$. Для цього слід застосувати відповідну декорельовану навчальну вибірку даних $\tilde{A}^{(l)}$ на відповідній попередньо навченій $SGTM_l(\tilde{A}^{(l)})$ для l -того елемента ансамблю.

8. Сформуувати розширену вибірку даних для навчання мета-моделі. Для цього з отриманих вихідних сигналів $y_i^{pred(l)}$ кожної ШНМ ансамблю, які можна вважати мета-ознаками, сформуувати нову вибірку даних (мета-вибірку) $M \in \mathbb{R}^{n \times L}$ для навчання мета-моделі (мета-регресора або мета-класифікатора):

$$M = \begin{bmatrix} y_1^{pred(1)} & y_1^{pred(2)} & \dots & y_1^{pred(L)} \\ y_2^{pred(1)} & y_2^{pred(2)} & \dots & y_2^{pred(L)} \\ \dots & \dots & \dots & \dots \\ y_n^{pred(1)} & y_n^{pred(2)} & \dots & y_n^{pred(L)} \end{bmatrix} \in \mathbb{R}^{n \times L}, \quad (4.17)$$

де $y_i^{pred(l)}$ – вихідний сигнал, отриманий від l -того елемента ансамблю для i -го зразка.

Слід зазначити, що залежний атрибут y_i , $i = \overline{1, n}$ для кожного вектора з вибірки M залишається таким, як і для початкової вибірки даних A , тобто (4.9).

9. Навчити обрану мета-модель на основі (4.17):

$$g(M; \phi), \quad (4.18)$$

де ϕ – параметр обраної мета-моделі.

Метод застосування розробленого ансамблю лінійних неітеративних нейроподібних структур МППП, побудованих на декорельованих вибірках даних з використанням методу кусково-лінійної апроксимації багатовимірної поверхні відгуку, передбачає послідовне виконання таких кроків:

1. Нормалізувати поточний вхідний вектор даних $u = (u_1, u_2, \dots, u_m)$ з невідомим значенням вихідного сигналу з використанням коефіцієнтів, отриманих в процесі нормалізації в режимі навчання ансамблю:

$$\acute{u} = \mathcal{N}(u), \quad (4.19)$$

де $\acute{u}$ – нормалізований вектор ознак для поточного спостереження, $\mathcal{N}(\cdot)$ – функція обраного методу нормалізації, що використовує визначені у (4.11) коефіцієнти.

2. Створити L декорельованих векторів $\tilde{u}^{(l)}, l = \overline{1, L}$, використовуючи збережені в процесі навчання кількості частин $s, s = \overline{1, S}$, на які буде поділений кожен незалежний атрибут, а також межі цих частин $a_0^{(l)} < a_1^{(l)} < a_2^{(l)} < \dots < a_S^{(l)}$. В результаті застосування (4.19) на (4.12) отримаємо:

$$\dot{u}_{j,s}^{(l)} = \begin{cases} 0, & \text{якщо } \dot{u}_j \leq a_{s-1}^{(l)}, \\ \dot{u}_j - a_{s-1}^{(l)}, & \text{якщо } a_{s-1}^{(l)} < \dot{u}_j \leq a_s^{(l)}, \quad s = \overline{1, S}, \\ a_s^{(l)} - a_{s-1}^{(l)}, & \text{якщо } \dot{u}_j > a_s^{(l)}. \end{cases} \quad (4.20)$$

Це надасть можливість створити розширений вектор для кожної компоненти:

$$\tilde{u}_j^{(l)} = (\dot{u}_{j,1}^{(l)}, \dot{u}_{j,2}^{(l)}, \dots, \dot{u}_{j,S}^{(l)}). \quad (4.21)$$

Після цього потрібно сформулювати повний розширений вектор ознак для l -того елемента ансамблю $l = \overline{1, L}$:

$$\tilde{u}^{(l)} = (\tilde{u}_1^{(l)}, \tilde{u}_2^{(l)}, \dots, \tilde{u}_m^{(l)}). \quad (4.22)$$

3. Утворити L декорельованих векторів $\tilde{u}^{(l)}$, кожен з яких використовуватиметься відповідним l -тим елементом ансамблю з лінійних нейроподібних структур МПГП.
4. Отримати вихідний сигнал від усіх L елементів ансамблю для відповідних $\tilde{u}^{(l)}$ з використанням попередньо навчених лінійних нейроподібних структур (*SGTM*), утворених в режимі підготовки методу:

$$y_{pred}^{(l)} = SGTM_l(\tilde{u}^{(l)}). \quad (4.23)$$

5. Сформулювати новий розширений вектор даних (мета-вектор) $\hat{u} \in \mathbb{R}^{1 \times L}$ для використання попередньо навченого мета-регресора або мета-

класифікатора з отриманих вихідних сигналів $y_{pred}^{(l)}$, які можна вважати мета-ознаками, кожного елементу ансамблю:

$$\hat{u} = (y_{pred}^{(1)}, y_{pred}^{(2)}, \dots, y_{pred}^{(L)}), \quad (4.24)$$

де $y^{(l)pred}$, $l = \overline{1, L}$ - вихідний сигнал, отриманий у (4.23) після застосування l -того елемента ансамблю до $\tilde{u}^{(l)}$ -го зразка.

Слід зазначити, що на цьому кроці, використовуючи (4.24), можна сформулювати остаточний вихідний сигнал роботи ансамблю на засадах принципів усереднення для задачі регресії чи голосування більшістю для задачі класифікації. В такому випадку ансамбль можна віднести до класу бегінгових.

6. Застосувати обрану мета-модель з виразу (4.18) у виразі (4.24), щоб отримати шуканий вихідний сигнал y^{final} для поточного вхідного вектора даних

$$y^{final} = g(\hat{u}; \phi), \quad (4.25)$$

де ϕ – оптимальний параметр обраної мета-моделі, підібраний у режимі підготовки ансамблю.

В такому випадку ансамбль можна віднести до класу стекінгових.

Розроблений неітеративний ансамбль лінійних нейроподібних-структур з використанням декорельованих вибірок даних, сформованих методом кусково-лінійної апроксимації, є особливо ефективним для аналізу біомедичних даних малих обсягів. Він забезпечує можливість формування нових представлень вхідних даних шляхом детермінованого поділу ознак на випадкові частини діапазону, що дозволяє суттєво розширити простір ознак при збереженні статистичної цілісності вихідної вибірки. Це дозволяє підвищити точність ансамблю навіть за умов обмеженого обсягу навчальних даних.

На відміну від класичних ансамблів розроблений ансамбль не передбачає багаторазового коригування ваг або гіперпараметрів. Це мінімізує ризик

перенавчання, що є критично важливим аспектом у біомедичних застосуваннях. Декореляція даних у межах ансамблю підвищує рівень різноманіття вибірки та здатності ансамблю до узагальнення, тоді як особливості функціонування нейроподібних структур МПГП забезпечують інтерпретованість результатів та їх робастність, що необхідно для медичних рішень. Гнучкість у виборі мета-моделі дозволяє адаптувати ансамбль до широкого спектру задач від діагностики до оцінки ризиків, зберігаючи при цьому високу обчислювальну ефективність.

4.3 Ансамбль однорідних штучних нейронних мереж на основі методу випадкових зміщень

Нейронна мережа узагальненої регресії часто використовується для розв’язання різноманітних нелінійних задач регресії в умовах аналізу даних обмеженого обсягу [85]. Це пояснюється її найвищими властивостями узагальнення серед існуючих типів ШНМ, простотою реалізації, необхідністю підбору лише одного параметра та відсутністю процедури навчання [159]. Подібні властивості демонструє також ймовірнісна нейронна мережа, яка часто використовується в процесі розв’язання задач класифікації малих обсягів даних [36]. Незважаючи на це, ці типи ШНМ не завжди демонструють достатню точність роботи [10]. Саме тому у цьому підрозділі розроблено ансамбль таких штучних нейронних мереж для розв’язання задач регресії чи класифікації [10], [254].

В основі ансамблю з однорідним ШНМ покладено метод декореляції малої вибірки даних [197]. Він дозволяє створювати унікальну вибірку даних для кожного окремого елемента ансамблю. Формування декорельованих вибірок даних для кожного елемента ансамблю однорідних ШНМ базується на генерації на заданому інтервалі та додаванні випадкових зміщень до незалежних ознак початкової вибірки даних [197], [254].

Нехай вхідний вектор ознак $x_i \in \mathbb{R}^m$, $i = \overline{1, n}$, вибірки даних A з (4.8) нормалізовано за допомогою обраної функції нормалізації $\mathcal{N}: \mathbb{R}^m \rightarrow \mathbb{R}^m$, тобто:

$$x'_i = \mathcal{N}(x_i), i = \overline{1, n}, \quad (4.26)$$

де x'_i – нормалізований вектор ознак для i -го зразка.

Основні кроки реалізації методу випадкових зміщень такі.

1. Згенерувати вектор випадкових зміщень:

$$\Delta = (\Delta_j) \in \mathbb{R}^m, j = \overline{1, m}, \quad (4.27)$$

де кожен елемент Δ_j генерується випадковим чином на інтервалі $[-\delta, \delta]$:

$$\Delta_j \in [-\delta, \delta], j = \overline{1, m}. \quad (4.28)$$

2. Сформувати нові вектори $\tilde{x}_i, i = \overline{1, n}$ шляхом покомпонентного додавання елементів вектора зміщень Δ_j та нормалізованого вектора x'_i :

$$\tilde{x}_{ij} = x'_{ij} + \Delta_j, j = \overline{1, m}. \quad (4.29)$$

3. Сформувати нову декорельовану вибірку $\tilde{A}$ шляхом виконання операції (4.29) для кожного вектора x'_i :

$$\tilde{x}_i = (\tilde{x}_{i1}, \tilde{x}_{i2}, \dots, \tilde{x}_{im})^T, i = \overline{1, n}. \quad (4.30)$$

Таким чином, нова вибірка даних $\tilde{A}$, декорельована по відношенню до початкової вибірки даних A , набуває вигляду:

$$\tilde{A} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \dots \\ \tilde{x}_n \end{bmatrix}, \tilde{x}_i \in \mathbb{R}^m, i = \overline{1, n}. \quad (4.31)$$

Описаний метод створення декорельованої вибірки даних дозволяє генерувати довільну кількість унікальних вибірок даних. Це дозволяє створювати ансамблі однорідних ШНМ для підвищення ефективності розв'язання задач регресії та класифікації в процесі аналізу малих обсягів даних [197]. Такі ансамблі можуть бути реалізовані за принципами усереднення результату (бегінг), блендінгу (спрощеної схеми стекінгу), а також класичної схеми стекінгу [254]. Крім того

використання ймовірнісної нейронної мережі замість нейронної мережі узагальненої регресії дозволяє будувати ансамблі однорідних ШНМ для розв'язання задач класифікації. У цьому випадку замість агрегування результатів апроксимації через усереднення використовується ансамбль, який приймає рішення на основі голосування більшістю. При цьому схеми блендингу та стекінгу також будуть ефективними варіантами.

Розглянемо композицію ансамблю на основі набору однорідних нейронних мереж (нейронних мереж узагальненої регресії – GRNN або ймовірнісних нейронних мереж – PNN), який дозволяє розв'язувати задачі регресії чи класифікації у випадку обмежених обсягів даних [25].

На рис. 4.2 наведено структурно-функційну схему ансамблю, побудованого на основі однорідних ШНМ, з застосуванням описаного методу декореляції вибірки даних.

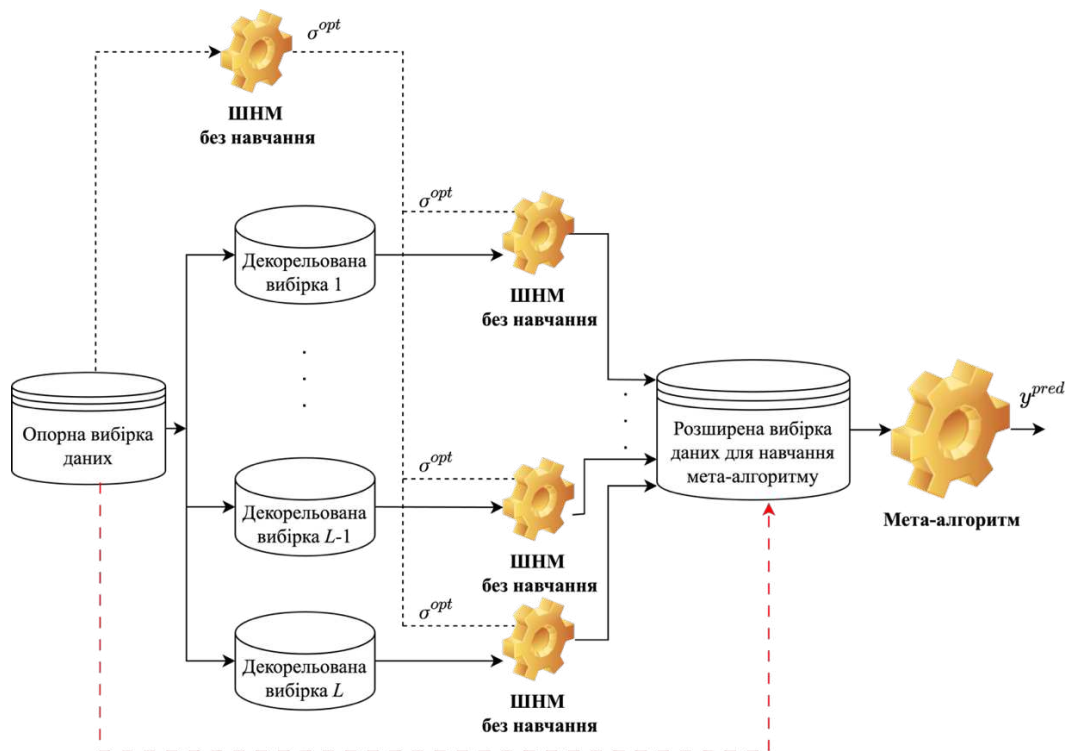


Рисунок 4.2 – Структурно-функційна схема роботи ансамблю однорідних ШНМ, побудованих на незалежно сформованих вибірках даних методом випадкових зміщень

Розглянемо методи підготовки та застосування розробленого ансамблю.

Метод підготовки (з можливим подальшим навчанням мета-моделі) ансамблю однорідних інтерполяційних ШНМ (НМУР та ЙМН), побудованих на декорельованих вибірках даних, з використанням випадкових зміщень передбачає послідовне виконання таких кроків.

1. Нормалізувати елементи опорної вибірки даних A відповідно до обраної функції $\mathcal{N}(\cdot)$, оскільки процедура навчання відсутня. Отримаємо матрицю $\acute{A}$:

$$\acute{A} = \mathcal{N}(A). \quad (4.32)$$

2. Знайти оптимальне значення розмаху гаусівської функції σ^{opt} для досліджуваної ШНМ з використанням кросвалідації за обраним критерієм. На цьому кроці слід реалізувати нейронну мережу узагальненої регресії $GRNN(\acute{A})$ для розв'язання задачі регресії або ймовірнісну нейронну мережу $PNN(\acute{A})$ для розв'язання задачі класифікації. Слід зазначити, що в процесі аналізу кожного поточного вектора даних цими ШНМ поточний вектор слід тимчасово вилучати з опорної вибірки.
3. Встановити загальну кількість L елементів ансамблю $l = \overline{1, L}$ (задати користувачем або визначити як параметр процедури оптимізації).
4. Побудувати ансамбль з L однорідних ШНМ, передавши кожній мережі значення оптимального параметру σ^{opt} , визначене на кроці 2. Одночасно для кожного елемента ансамблю з індексом $l = \overline{1, L}$ за допомогою методу генерації випадкових зміщень (4.27)-(4.31) створити унікальні декорельовані вибірки даних $\tilde{A}^{(l)}$, кожна з яких використовується виключно для побудови відповідної моделі регресії $GRNN_l(\tilde{A}^{(l)})$ чи класифікації $PNN_l(\tilde{A}^{(l)})$. Зберегти l векторів зміщень $\Delta^{(l)}$.

5. Отримати вихідний сигнал від усіх L елементів ансамблю $y_i^{pred(l)} = GRNN_l(\acute{A})$, чи $y_i^{pred(l)} = PNN_l(\acute{A})$ для кожного i -го вектора з $\acute{A}, i = \overline{1, n}$ на основі використання відповідних $\tilde{A}^{(l)}$ як опорних вибірок для обчислення $y_i^{pred(l)}$. Слід зазначити, що в процесі такого аналізу з використанням ШНМ поточний вектор треба тимчасово вилучати з опорної вибірки $\tilde{A}^{(l)}$.
6. Сформувати розширену вибірку даних одним з двох способів для навчання мета-моделі:
 - з отриманих вихідних сигналів $y_i^{pred(l)}$ кожної ШНМ ансамблю, які можна вважати мета-ознаками, сформувати нову вибірку даних (мета-вибірку) $M \in \mathbb{R}^{n \times L}$ для навчання мета-моделі (мета-регресора або мета-класифікатора):

$$M = \begin{bmatrix} y_1^{pred(1)} & y_1^{pred(2)} & \dots & y_1^{pred(L)} \\ y_2^{pred(1)} & y_2^{pred(2)} & \dots & y_2^{pred(L)} \\ \dots & \dots & \dots & \dots \\ y_n^{pred(1)} & y_n^{pred(2)} & \dots & y_n^{pred(L)} \end{bmatrix} \in \mathbb{R}^{n \times L}, \quad (4.33)$$

де $y_i^{pred(l)}$ – вихідний сигнал, отриманий від l -го елемента ансамблю для i -го зразка.

Слід зазначити, що залежний атрибут $y_i, i = \overline{1, n}$ для кожного вектора M залишається таким же, як і для початкової вибірки даних, тобто (4.9). Такий ансамбль можна віднести до класу стекінгових.

На цьому кроці можна сформувати остаточний вихідний сигнал роботи ансамблю, використовуючи ансамблеві принципи усереднення (для задачі регресії) чи голосування більшістю (для задачі класифікації) [21]. Такий ансамбль можна віднести до класу бегінгових;

- інший спосіб формування вибірки даних для мета-регресора полягає у додаванні початкової нормалізованої матриці $\hat{A}$ до матриці M , отримуючи розширену матрицю $\ddot{M}$:

$$\ddot{M} = [M, \hat{A}] \in \mathbb{R}^{n \times (L+m)}. \quad (4.34)$$

Такий підхід може зменшити кількість ШНМ ансамблю і забезпечити підвищення точності розв'язання задачі через спрощення поверхні відгуку, що пояснюється принципом лінеаризації поверхні відгуку [25].

7. Навчити обрану мета-модель $\mathcal{G}(M; \phi)$ на основі (4.33) чи $\mathcal{G}(\ddot{M}; \phi)$ на основі (4.34), де ϕ – параметр обраної мета-моделі.

Метод реалізації розробленого ансамблю однорідних ШНМ, побудованих на декорельованих вибірках даних з використанням випадкових зміщень, передбачає послідовне виконання таких кроків:

1. Нормалізувати поточний вхідний вектор даних $u = (u_1, u_2, \dots, u_m)$ при невідомих значеннях вихідного сигналу з використанням коефіцієнтів, отриманих в процесі нормалізації в режимі підготовки ансамблю:

$$\acute{u} = \mathcal{N}(u), \quad (4.35)$$

де $\acute{u}$ – нормалізований вектор ознак для поточного спостереження, $\mathcal{N}(\cdot)$ – функція обраного методу нормалізації, що використовує визначені у (4.32) коефіцієнти.

2. Отримати вихідний сигнал від усіх L елементів ансамблю для вектора $\acute{u}^{(l)}$ з використанням відповідних l -тих опорних вибірок $\tilde{A}^{(l)}$, утворених в режимі підготовки методу:

$$y_{pred}^{(l)} = \begin{cases} GRNN_l(\acute{u}^{(l)}), & \text{у випадку розв'язання регресії,} \\ PNN_l(\acute{u}^{(l)}), & \text{у випадку розв'язання класифікації.} \end{cases} \quad (4.36)$$

3. Сформувані новий розширений вектор даних для подальшого застосування мета-моделі:

- з отриманих вихідних сигналів $y_{pred}^{(l)}$ кожної ШНМ ансамблю, які можна вважати мета-ознаками, сформуванати новий вектор даних (мета-вектор):

$$\hat{u} \in \mathbb{R}^{1 \times L} \quad (4.37)$$

для застосування попередньо навченого мета-регресора або мета-класифікатора, який матиме наступний вигляд:

$$\hat{u} = (y_{pred}^{(1)}, y_{pred}^{(2)}, \dots, y_{pred}^{(L)}), \quad (4.38)$$

де $y^{(l)pred}, l = \overline{1, L}$ – вихідний сигнал, отриманий у (4.36) після застосування l -го елемента ансамблю до $\hat{u}^{(l)}$ -го зразка.

Додатково на цьому кроці можна використати ансамблеві принципи усереднення (для задачі регресії) чи голосування більшістю (для задачі класифікації) і сформуванати остаточний вихідний сигнал роботи ансамблю [21].

- для другого способу формування мета-вибірки до вектора $\hat{u}$ з (4.38) треба додати усі ознаки з початкового нормалізованого вектора $\hat{u}$, отримуючи розширений вектор $\check{u}$ виду:

$$\check{u} = (\hat{u}, \hat{u}) \in \mathbb{R}^{1 \times (L+m)}. \quad (4.39)$$

4. Застосувати попередньо навчену мета-модель на (4.38):

$$y^{final} = g(\hat{u}; \phi), \quad (4.40)$$

чи на (4.39):

$$y^{final} = g(\check{u}; \phi), \quad (4.41)$$

щоб отримати шуканий вихідний сигнал y^{final} для поточного вхідного вектора даних. Слід зазначити, що ϕ – оптимальний параметр мета-моделі, підібраний у режимі підготовки ансамблю.

У порівнянні з попереднім лінійним ансамблем на основі методу кусково-лінійної апроксимації багатовимірної поверхні відгуку запропонований ансамбль однорідних нелінійних штучних нейронних мереж (нейронна мережа узагальненої регресії чи ймовірнісна нейронна мережа), побудований на випадкових зміщеннях, має низку суттєвих переваг для аналізу біомедичних даних малих обсягів. Насамперед він повністю усуває потребу в навчанні базових моделей, що істотно спрощує реалізацію та зменшує обчислювальні витрати. Також розроблений ансамбль не потребує поділу вибірки на частини як у класичних стекінгових ансамблях. Це забезпечує повноцінне використання всієї доступної інформації в кожному елементі ансамблю. Окрім цього процедура формування декорельованих вибірок шляхом додавання випадкових зміщень до нормалізованих ознак є обчислювально простою і дозволяє уникнути складних процедур сегментації чи побудови підпросторів.

Завдяки імовірнісній природі ймовірнісної нейронної мережі, що ґрунтується на згладженні, розроблений ансамбль демонструє високу стійкість до шуму та здатність до узагальнення навіть у випадках з обмеженим числом зразків. Крім того ансамбль вимагає мінімального налаштування (тільки розмах гаусівської функції), що значно полегшує його адаптацію до нових задач. Різноманітні варіанти агрегування результатів (усереднення, голосування, блендінг або стекінг) роботи усіх елементів ансамблю розширюють можливості застосування цього ансамблю для розв'язання різноманітних задач регресії/класифікації на основі малих біомедичних вибірок даних.

4.4 Ансамбль неітеративних нейроподібних структур з РБФ-розширенням входів на основі випадкових симетричних зміщень

Розмаїття завдань у спеціалізованих напрямках біомедичної інженерії спричиняє ситуації, коли лінійні методи машинного навчання та деякі інтерполяційні штучні нейронні мережі демонструють низьку ефективність в

процесі розв'язання задач регресії чи класифікації на основі обмежених обсягів даних [255]. Побудова однорідних ансамблів з таких моделей не завжди дозволить досягти необхідної точності.

З огляду на широкий спектр прикладних завдань, специфіку обсягів і вимірності даних, а також складні взаємозв'язки між медико-біологічними та інженерно-технічними параметрами при роботі з малою вибіркою біомедичних даних доцільним може бути використання ядрових методів машинного навчання [84], [256], [256]. Зокрема, у випадку, коли мала вибірка даних містить близько 1000 спостережень і має невисоку вимірність, використання ядрового підходу може суттєво підвищити ефективність розв'язання різноманітних прикладних задач регресії та класифікації.

Одним з ефективних ядрових ШНМ з неітеративним методом навчання для розв'язання подібних задач є нейроподібна структура МПГП з РБФ-розширенням входів [138], яку описано у першому розділі. Нелінійне розширення входів з використання РБФ-функцій має ряд переваг для підвищення точності інтелектуального аналізу даних [28], [29], [30]. Зокрема, РБФ-функції перетворюють вихідний простір так, що навіть тонкі нелінійні зв'язки між ознаками стають більш вираженими. Це дозволяє моделі точніше розпізнавати закономірності в даних [24]. Окрім цього, як зазначено та обґрунтовано у другому розділі роботи, перетворення даних у простір вищої вимірності забезпечує можливість більш ефективного лінійного розділення класів. Особливу увагу також заслуговує локальна природа РБФ-функцій, завдяки якій визначена область впливу кожної функції дозволяє зосередитися на релевантних регіонах простору ознак, мінімізуючи вплив шуму та неінформативних даних [84].

У випадку використання нейроподібної структури МПГП з РБФ-розширенням входів до вищевказаних переваг додається неітеративний, а, отже, швидкий процес навчання і можливість отримання задовільних розв'язків для

навчальних вибірок різного обсягу за єдиною методологічною основою розв'язання задач [23], [24].

Для ще більшого підвищення точності розв'язання задач регресії та класифікації у цьому підрозділі розроблено ансамбль зазначених ядрових ШНМ неітеративного типу [23], [24]. В основі роботи такого ансамблю покладено генерацію ряду нейроподібних структур МПГП з РБФ-розширенням входів, кожна з яких працює з власною, декорельованою вибіркою даних. Останні у випадку аналізу малих даних формуються з використанням методу попарних симетричних випадкових зміщень. Розглянемо його детальніше.

Як і у попередньому випадку (4.11), вхідний вектор ознак $x_i \in \mathbb{R}^m$, $i = \overline{1, n}$ вибірки даних $A \in \mathbb{R}^{n \times m}$ повинен бути нормалізований за допомогою обраної функції нормалізації $\mathcal{N}: \mathbb{R}^m \rightarrow \mathbb{R}^m$.

Метод попарних симетричних випадкових зміщень передбачає послідовне виконання таких кроків:

1. Згенерувати перший вектор Δ випадкових зміщень відповідно до (4.28):

$$\Delta = (\Delta_1, \Delta_2, \dots, \Delta_m) \in \mathbb{R}^m. \quad (4.43)$$

2. Згенерувати другий вектор випадкових зміщень шляхом дзеркального відображення значень компонент першого вектору Δ випадкових зміщень (4.43):

$$-\Delta = (-\Delta_1, -\Delta_2, \dots, -\Delta_m) \in \mathbb{R}^m. \quad (4.44)$$

3. Сформувані два нові вектори $\tilde{x}_i$ шляхом покомпонентного додавання векторів зміщень (4.43) та (4.44) до нормалізованого вектора x'_i :

$$\tilde{x}_{i,j}^{(+)} = x'_{ij} + \Delta_j, \quad \tilde{x}_{i,j}^{(-)} = x'_{ij} - \Delta_j, \quad j = \overline{1, m}, \quad (4.45)$$

де x'_i — нормалізований вектор ознак для i -го зразка, $i = \overline{1, n}$.

4. Сформувані дві нові декорельовані вибірки даних шляхом виконання операцій (4.45) для кожного вектора x'_i :

$$\tilde{A}^{(+)} = \begin{bmatrix} \tilde{x}_1^{(+)} \\ \tilde{x}_2^{(+)} \\ \dots \\ \tilde{x}_n^{(+)} \end{bmatrix} \in \mathbb{R}^{n \times m}, \quad \tilde{A}^{(-)} = \begin{bmatrix} \tilde{x}_1^{(-)} \\ \tilde{x}_2^{(-)} \\ \dots \\ \tilde{x}_n^{(-)} \end{bmatrix} \in \mathbb{R}^{n \times m}. \quad (4.46)$$

5. Сформувати фінальну декорельовану вибірку даних подвоєного обсягу, удвічі аугментовану, шляхом вертикального об'єднання (конкатенації) рядків матриць $\tilde{A}^{(+)}$ та $\tilde{A}^{(-)}$ з (4.46):

$$\tilde{A}_{aug} = \begin{bmatrix} \tilde{A}^{(+)} \\ \tilde{A}^{(-)} \end{bmatrix} = \begin{bmatrix} \tilde{x}_1^{(+)} \\ \tilde{x}_2^{(+)} \\ \dots \\ \tilde{x}_n^{(+)} \\ \tilde{x}_1^{(-)} \\ \tilde{x}_2^{(-)} \\ \dots \\ \tilde{x}_n^{(-)} \end{bmatrix} \in \mathbb{R}^{2n \times m}. \quad (4.47)$$

Метод симетричних випадкових зміщень як і класичний метод випадкових зміщень дозволяє генерувати довільну кількість унікальних декорельованих вибірок даних. Це є критично важливо при побудові ансамблів однорідних елементів для розв'язання задач регресії та класифікації в умовах обмежених обсягів даних. Основною перевагою цього методу є можливість формування векторів зміщень з різними знаками. Завдяки тому, що зміщення можуть бути як додатними, так і від'ємними, метод дозволяє ефективно компенсувати похибки різних знаків, що виникають в процесі навчання і використання нейроподібних структур МППП з РБФ-розширенням входів. Така компенсація сприяє підвищенню робастності отриманих результатів за рахунок зниження впливу аномалей, викидів та інших відхилень у досліджуваній вибірці даних і, відповідно, покращує точність розв'язання практичних задач [25].

Крім того застосування симетричних випадкових зміщень забезпечує можливість подвоєння обсягу навчальної вибірки порівняно з її початковим

розміром. З одного боку це дозволяє використовувати ширший спектр нелінійних ШНМ як базових моделей ансамблю, що є особливо важливим в умовах обмежених даних. З іншого боку, враховуючи необхідність поділу початкової вибірки на дві рівні частини для навчання слабких моделей першого рівня та мета-моделі другого рівня можливість аугментації даних у два рази дозволяє частково компенсувати зменшення обсягу вибірки для базових моделей. Це, у свою чергу, дозволяє підвищити властивості узагальнення та робастність усього ансамблю.

Усі ці переваги є передумовою для створення ефективних композицій ансамблів, які можуть використовувати різні стратегії агрегування вихідних сигналів такі як голосування, усереднення, стекінг або блендінг [21], [24], [25], [26]. Застосування цього підходу дозволяє підвищити загальну ефективність (точність та узагальнення) розв'язання різноманітних задач регресії та класифікації у галузі біомедичної інженерії в умовах аналізу малих обсягів даних.

У цьому підрозділі розроблено новий ансамбль неітеративних нейроподібних структур МПГП з РБФ-розширенням входів, що формуються на декорельованих вибірках даних з використанням методу симетричних випадкових зміщень. Його призначено для розв'язання задач регресії чи класифікації у випадку аналізу обмежених обсягів даних.

На рис. 4.3 наведено структурно-функційну схему ансамблю, побудованого на основі однорідних нелінійних нейроподібних структур МПГП з використанням розробленого методу декореляції вибірки даних.

Розглянемо методи навчання та застосування розробленого ансамблю.

Неітеративний метод навчання розробленого ансамблю нейроподібних структур МПГП з РБФ-ядром, побудованих на декорельованих вибірках даних, з використанням симетричних випадкових зміщень, передбачає послідовне виконання таких кроків:

1. Нормалізувати елементи навчальної вибірки даних A згідно обраної функції $\mathcal{N}(\cdot)$. Отримаємо вибірку $\hat{A}$:

слабких моделей, отриманих на тих же прикладах, що й в процесі їх навчання.

3. Використовуючи першу частину вибірки $\hat{A}_1$, знайти оптимальне значення кількості центрів РБФ-функцій та значення розмаху гаусівської функції σ^{opt} для досліджуваної ШНМ з використанням кросвалідації за обраним критерієм. На цьому кроці слід реалізувати неітеративну нейроподібну структуру з РБФ-розширенням входів $SGTM\_RBF(\hat{A})$ для розв'язання задач регресії (класифікації).
4. Встановити загальну кількість L елементів ансамблю $l = \overline{1, L}$ (задати користувачем або визначити як параметр процедури оптимізації).
5. Побудувати ансамбль з L неітеративних нейроподібних структур з РБФ-розширенням входів, передавши кожній мережі значення оптимальних параметрів, визначене на кроці 3. Одночасно для кожного елемента ансамблю з індексом $l = \overline{1, L}$ на основі першої частини вибірки $\hat{A}_1$ за допомогою методу генерації випадкових зміщень (4.43)–(4.47) створити унікальні декорельовані вибірки даних $\tilde{A}_1^{(l)}$. Кожна з них використовується виключно для побудови відповідної моделі регресії чи класифікації $SGTM\_RBF_l(\tilde{A}_1^{(l)})$. Слід зазначити, що відповідно до процедури (4.43)–(4.47) обсяг $\tilde{A}_1^{(l)}$ буде подвоєно, тобто $\tilde{A}_1^{(l)} \in \mathbb{R}^{2n \times m}$.
6. Зберегти l векторів зміщень $\Delta^{(l)}$ та $-\Delta^{(l)}, l = \overline{1, L}$. Це є суттєвою перевагою в процесі застосування ядрових методів для аналізу малих обсягів даних.
7. Для кожного елемента ансамблю з індексом $l = \overline{1, L}$ на основі другої частини вибірки $\hat{A}_2$ отримати вихідний сигнал від усіх L елементів ансамблю $y_i^{pred(l)} = SGTM\_RBF_l(\hat{A}_2^{(l)})$ для кожного i -го вектора з $\hat{A}_2^{(l)}, i = \overline{1, 2n}$.

8. Сформуувати розширену вибірку даних для навчання мета-моделі. Для цього з отриманих вихідних сигналів $y_i^{pred(l)}$ кожної ШНМ ансамблю, які можна вважати мета-ознаками, сформуувати нову вибірку даних (мета-вибірку) $M \in \mathbb{R}^{2n \times L}$ для навчання мета-моделі (мета-регресора або мета-класифікатора):

$$M = \begin{bmatrix} y_1^{pred(1)} & y_1^{pred(2)} & \dots & y_1^{pred(L)} \\ y_2^{pred(1)} & y_2^{pred(2)} & \dots & y_2^{pred(L)} \\ \dots & \dots & \dots & \dots \\ y_{2n}^{pred(1)} & y_{2n}^{pred(2)} & \dots & y_{2n}^{pred(L)} \end{bmatrix} \in \mathbb{R}^{2n \times L}, \quad (4.49)$$

де $y_i^{pred(l)}$ – вихідний сигнал, отриманий від l -того елемента ансамблю для i -го зразка.

Слід зазначити, що залежний атрибут $y_{i_1} = y_{i_2}$, $i_1 = \overline{1, n}$, $i_2 = \overline{n+1, 2n}$ для кожного вектора з вибірки M залишається таким же, як і для початкової вибірки даних, тобто (4.9).

На цьому кроці можна сформуувати остаточний вихідний сигнал роботи ансамблю, використовуючи ансамблеві принципи усереднення (для задачі регресії) чи голосування більшістю (для задачі класифікації) [21]. Такий ансамбль можна віднести до класу бегінгових.

9. Навчити обрану мета-модель $\mathcal{G}(M; \phi)$ на основі (4.49), де ϕ – параметр обраної мета-моделі.

Метод застосування ансамблю нейроподібних структур МППП з РБФ-ядром, побудованих на декорельованих вибірках даних з використанням симетричних випадкових зміщень, передбачає послідовне виконання таких кроків:

1. Нормалізувати поточний вхідний вектор даних $u = (u_1, u_2, \dots, u_m)$ з невідомим значенням вихідного сигналу з використанням коефіцієнтів, отриманих в процесі нормалізації в режимі підготовки ансамблю:

$$\hat{u} = \mathcal{N}(u), \quad (4.51)$$

де $\hat{u}$ – нормалізований вектор ознак для поточного спостереження, $\mathcal{N}(\cdot)$ – функція обраного методу нормалізації, що використовує визначені у (4.48) коефіцієнти.

2. Отримати вихідні сигнали $y_{pred}^{(l)}$ від усіх L елементів ансамблю шляхом застосування нормалізованого вектора $\hat{u}$ (без зміщень) на попередньо навчених $SGTM_{RBF_l}(\tilde{A}_1^{(l)})$, утворених в режимі навчання методу з використанням формули:

$$y_{pred}^{(l)} = SGTM_{RBF_l}(\hat{u}^{(l)}). \quad (4.53)$$

3. Сформувати новий розширений вектор даних для подальшого застосування мета-моделі:

- з отриманих вихідних сигналів $y_{pred}^{(l)}$ кожної ШНМ ансамблю, які можна вважати мета-ознаками, сформувати новий вектор даних (мета-вектор):

$$\hat{u} \in \mathbb{R}^{1 \times L} \quad (4.54)$$

для застосування попередньо навченого мета-регресора або мета-класифікатора, який має такий вигляд:

$$\hat{u} = (y_{pred}^{(1)}, y_{pred}^{(2)}, \dots, y_{pred}^{(L)}), \quad (4.55)$$

де $y^{(l)pred}, l = \overline{1, L}$ – вихідний сигнал, отриманий у (4.53) після застосування l -го елемента ансамблю до $\hat{u}^{(l)}$ -го зразка.

Окрім цього на поточному кроці можна використати ансамблеві принципи усереднення (для задачі регресії) чи голосування більшістю (для задачі класифікації) і сформувати остаточний вихідний сигнал роботи ансамблю [21];

- для другого способу формування мета-вибірки в режимі підготовки методу до вектора $\hat{u}$ з (4.54) слід додати усі ознаки з

початкового нормалізованого вектора $\hat{u}$, отримуючи розширений вектор $\check{u}$ виду:

$$\check{u} = (\hat{u}, \hat{u}) \in \mathbb{R}^{1 \times (L+m)}. \quad (4.56)$$

Такий спосіб агрегації результатів дозволяє віднести розроблений ансамбль до класу стекінгових.

4. Застосувати попередньо навчену мета-модель на (4.55):

$$y^{final} = \varphi(\hat{u}; \phi) \quad (4.57)$$

чи на (4.56):

$$y^{final} = \varphi(\check{u}; \phi) \quad (4.58)$$

для отримання шуканого вихідного сигналу y^{final} для поточного вхідного вектора даних. Слід зазначити, що ϕ – оптимальний параметр мета-моделі, підібраний у режимі підготовки ансамблю.

У порівнянні з попередніми ансамблями як лінійним ансамблем нейроподібних структур з РБФ-розширенням входів на основі використання кусково-лінійної апроксимації поверхні відгуку, так і ансамблем однорідних інтерполяційних нейронних мереж на основі випадкових зміщень розроблений ансамбль неітеративних нейроподібних структур МПГП з РБФ-розширенням входів, сформованих на симетричних випадкових зміщеннях має низку ключових переваг. Перш за все, на відміну від ансамблю з пункту 3.3 цей ансамбль використовує ядрову апроксимацію з нелінійним розширенням ознак, що дозволяє йому ефективно виявляти складні локальні взаємозв'язки у даних і покращувати здатність до узагальнення. Завдяки РБФ-функціям простір ознак розширюється таким чином, що навіть слабо виражені нелінійні закономірності стають краще роздільними для лінійних моделей.

У порівнянні з лінійним ансамблем з підрозділу 3.2, який оперує в лінійному просторі даних, цей ансамбль забезпечує вищу адаптивність до складної геометрії

вхідних даних, що є характерним для біомедичних задач. Крім того метод симетричних випадкових зміщень дозволяє не лише формувати унікальні декорельовані вибірки, але й подвоювати обсяг навчальних даних без втрати їх інформаційної цінності. Це не лише компенсує необхідність поділу вибірки на підмножини як для класичного стекінгу, але й дозволяє досягнути вищої точності та робастності завдяки частковій компенсації похибок протилежних знаків. Це особливо корисно в умовах наявності шуму чи викидів у досліджуваній вибірці даних. Окрему перевагу становить гнучкість параметризації цього ансамблю. Композиція ансамблю дозволяє незалежно налаштовувати кількість центрів радіальних базисних функцій (РБФ) та ширини гаусівських ядер для усіх елементів ансамблю, що забезпечує баланс між його складністю і здатністю до узагальнення. Таким чином, розроблений ансамбль поєднує в собі високу точність, робастність до шуму, здатність працювати з малою кількістю даних і ефективність реалізації у вигляді неітеративного ансамблю, що сукупно робить його найпотужнішим серед розглянутих у цьому розділі.

4.5 Ієрархічні методи розв'язання задач регресії і класифікації з додатковою кластеризацією даних

У біомедичній інженерії задачі інтелектуального аналізу даних характеризуються підвищеною складністю через високу вимірність, значну неоднорідність, наявність шумів, а також складну, часто невідому, топологію розподілу вхідних ознак. У випадку аналізу даних середнього та великого обсягу дані можуть надходити з різноманітних біомедичних сенсорів, діагностичних систем або результатів клінічних спостережень [257]. Вони можуть містити приховані структури, що не є очевидними при глобальному аналізі, але відіграють ключову роль у визначенні функціональних залежностей. У таких умовах застосування єдиної глобальної моделі, в тому числі ансамблевої, для розв'язання задач регресії або класифікації виявляється неефективним. Це пояснюється

неможливістю врахування локальних закономірностей, характерних для окремих підмножин даних або компактних множин точок.

Одним з перспективних підходів до подолання цієї проблеми є використання ієрархічних методів, заснованих на попередньому кластерному поділі вхідної вибірки [258]. Проблемою кластеризації даних, а також розроблення нових методів її розв'язання, зокрема, у випадку перетинання класів займається Бодянський Є.В., Камінський Р.М та Шафроненко А. Ю. Результати їх досліджень показують, що кластеризація дозволяє виявити умовно однорідні підмножини даних або компактні множини точок, кожна з яких характеризується власними спрощеними залежностями між вхідними та вихідними змінними [257], [259], [260]. Відповідно подальша побудова локальних моделей регресії чи класифікації в межах таких кластерів забезпечує адаптацію до специфіки даних, зменшує складність апроксимації та підвищує її точність. Крім того ієрархічний підхід дозволяє враховувати різну природу джерел даних, що особливо актуально для біомедичних систем, де об'єднуються результати лабораторних аналізів, зображень, електрофізіологічних сигналів тощо.

Теоретичну доцільність використання ієрархічних методів розв'язання задач регресії та класифікації на основі кластеризації можна обґрунтувати положеннями гіпотези компактності [226], [253]. Згідно з нею об'єкти, що мають подібні значення цільової змінної (задача регресії) або належать до одного класу (задача класифікації) зазвичай розташовуються поруч в ознаковому просторі, формуючи щільні підмножини або компактні множини точок. Це твердження відповідає природі багатьох реальних задач, зокрема, в біомедичній інженерії, де дані дуже часто утворюють кластери з внутрішньо подібною структурою.

Відповідно поділ вхідної вибірки на такі компактні регіони дозволяє будувати локальні моделі, які ефективніше відображають внутрішню закономірність у межах кожного підпростору. Тобто функціональна залежність між ознаками та цільовою змінною у межах компактної множини точок часто має

спрощену форму. Вона може бути лінійною, монотонною або гладкою. Таким чином, локальне моделювання забезпечує зменшення складності побудови регресійної або класифікаційної функції, що, у свою чергу, дозволяє уникнути надлишкової складності в процесі реалізації усієї моделі на глобальному рівні.

Подібний підхід узгоджується з концепцією моделі оптимальної складності, запропонованої академіком Івахненком О. Г. [149], [261]. Згідно з цією концепцією ефективна модель має бути побудована так, щоб забезпечити найкращу узагальнюючу здатність без перенавчання. У разі аналізу складних структурно неоднорідних біомедичних даних доцільно не створювати єдину надскладну глобальну модель, а натомість реалізовувати композицію локальних моделей оптимальної складності, кожна з яких формується у межах окремої компактної множини точок. Таким чином, структура кожної локальної моделі визначається внутрішньою складністю відповідного кластера і не є фіксованою або наперед заданою. Такий підхід відповідає не тільки принципів оптимальної складності, але й підвищує точність результатів, дозволяючи адаптувати структуру моделі кожного окремого кластера до локальних особливостей даних.

У [227], [258] розроблено ієрархічний метод розв'язання задач регресії чи класифікації на основі сумісного використання кластеризації та засобів ансамблевого навчання у галузі медицини. Він демонструє суттєве підвищення точності роботи. Проте у випадку, коли утворені кластери або компактні множини точок після виконання процедури кластеризації описано обмеженим числом спостережень, застосування класичних регресорів чи класифікаторів для його аналізу суттєво обмежене або й взагалі не можливе.

У цьому розділі удосконалено ієрархічний метод розв'язання задач регресії чи класифікації у випадку аналізу обмеженого обсягу даних.

Розглянемо детальніше методи його навчання і застосування.

Навчання ієрархічного методу розв'язання задач регресії і класифікації з використанням кластеризації передбачає послідовне виконання таких кроків.

1. На першому етапі за допомогою обраного методу кластерного аналізу виконати поділ вхідної вибірки даних на локальні області, кластери з подібними характеристиками [227]. Такі кластери даних розглядаються як умовно однорідні підпростори, у межах яких спостерігаються спрощені або специфічні залежності між ознаками та цільовими змінними. Слід зазначити, що у випадку кластеризації малих обсягів даних утворені компактні множини точок можуть містити обмежену кількість зразків, необхідну для навчання існуючих регресорів чи класифікаторів на основі машинного навчання [228].
2. На наступному етапі для кожного кластера побудувати локальну модель, окрему модель регресії/класифікації (або ансамбль), що узгоджується з внутрішньою структурою даних у межах цього кластеру.

У випадку аналізу кластеру з достатнім обсягом даних для реалізації подальших процедур навчання в якості базових моделей можуть використовуватись класичні методи машинного навчання та штучні нейронні мережі [258]. Проте, як показано у цьому розділі, більш ефективним підходом такого аналізу буде використання ансамблевого навчання, зокрема, ансамблів з однорідних ШНМ лінійного та нелінійного типів.

У випадку, коли утворені кластери або компактні множини точок описано обмеженим числом спостережень (екстремально короткі вибірки даних), для їх аналізу варто використовувати методи регресії чи класифікації на основі декартового квадрата множини вхідних компонент векторів, які розглянуто у розділі 3 цієї роботи [228]. Такий підхід забезпечить як можливість їх ефективного аналізу за рахунок аугментації даних, так і підвищення точності результатів роботи з використанням принципів ансамблевого навчання, на основі яких будуються методи з розділу 3.

Застосування ієрархічного методу розв'язання задач регресії і класифікації з використанням кластеризації передбачає послідовне виконання таких кроків.

1. Віднести нове спостереження до одного з попередньо сформованих кластерів, наприклад, на основі мінімальної відстані до центроїда (розділ 2).
2. Застосувати поточний вектор даних на попередньо навченій локальній моделі регресії/класифікації на основі декартового квадрата множини вхідних компонент векторів, яку побудовано для аналізу відповідного кластеру даних. Вихід такої локальної моделі буде результатом роботи усього ієрархічного методу.

Такий підхід дозволяє враховувати нелінійність і неоднорідність даних, знижує загальну похибку моделі за рахунок адаптації до локальних особливостей даних кожної компактної множини точок та покращує властивості узагальнення особливо у випадках складної топології даних або багатокластерної структури класів.

4.6 Висновки до розділу

1. У розділі подано, обґрунтовано, розроблено та показано ефективність комбінування глобально апроксимованих поверхонь відгуку, яке реалізовано розробленими методами синтезу ансамблів однорідних ШНМ лінійного або нелінійного типу. Такий підхід забезпечує можливість підвищення точності розв'язання задач регресії та класифікації в галузі біомедичної інженерії. Розроблено нові методи синтезу ансамблів однорідних ШНМ на основі різних підходів до декореляції початкової вибірки даних та з різними варіантами агрегування результатів їх роботи (усереднення, голосування більшістю, блендінг, стекінг).

2. Розроблений метод синтезу ансамблю лінійних неітеративних нейроподібних структур МПГП базується на формуванні унікальних декорельованих вибірок даних для кожного елемента з використанням методу кусково-лінійної апроксимації багатовимірної поверхні відгуку. Такий підхід

дозволить кожному елементові ансамблю краще відображати локальні варіації даних, зберігаючи при цьому основні переваги лінійних моделей: швидкодію та мінімальну потребу в їх налаштуванні. Окрім цього, розроблений ансамбль забезпечить простоту та ефективність процесу реалізації ансамблю, а також підвищену стійкість до шуму завдяки використанню особливостей єдиної архітектури його базових елементів та неітеративного навчання. Це істотно знизить необхідні обчислювальні витрати.

3. Ансамбль інтерполяційних нейронних мереж (нейронна мережа узагальненої регресії, ймовірнісна нейронна мережа) синтезовано з використанням методу випадкових зміщень для створення унікальної декорельованої вибірки даних щодо кожного елемента такого ансамблю. Метод підготовки цього ансамблю не вимагає поділу вибірки на дві частини, що забезпечує використання усього обмеженого обсягу даних як базовими елементами ансамблю, так і мета-моделлю. Це сукупно з використанням нелінійних ШНМ забезпечить підвищення точності розв'язання задач аналізу малих даних при збереженні високої швидкодії роботи.

4. Метод синтезу ансамблю неітеративних нейроподібних структур з РБФ-розширенням входів реалізовано з використанням випадкових попарних симетричних зміщень. Такий підхід до декореляції вибірки даних дозволяє формувати для кожного елемента ансамблю унікальні та водночас удвічі аугментовані вибірки даних. Незважаючи на те, що синтез цього ансамблю, зокрема, з використанням блендінгу, вимагає поділу навчальної вибірки на дві частини, зазначений недолік ефективно компенсується за рахунок застосування обраного методу декореляції вибірки даних. Такий підхід дозволяє застосовувати ядрові методи аналізу даних до вже збільшеного обсягу інформації. Він також підвищить робастність ансамблю і зменшить сумарну похибку роботи за рахунок компенсації похибок протилежних знаків у декорельованих підвибірках, сформованих методом симетричних випадкових зміщень.

5. Ієрархічні методи розв'язання задач класифікації та регресії дозволять зводити аналіз відносно великих нерегулярних та неоднорідних вибірок даних до опрацювання окремих малих підвибірок, що, у свою чергу, суттєво підвищить точність аналізу. Такий підхід ґрунтується на гіпотезі компактності та реалізований у цій роботі через використання кластеризації. Зокрема, виділення компактних регіонів з подальшим їх опрацюванням обраним методом дозволить розділити загальну задачу на низку локальних підзадач, адаптуючи кожен модель до унікальних властивостей кожного окремого кластера. У випадку, коли виділений компактний регіон подано у вигляді екстремально-короткої підвибірки даних, доцільно використовувати методи, описані у третьому розділі цієї роботи. Натомість, якщо кластер міститиме достатню кількість даних для реалізації методу навчання, можна застосовувати ансамблі однорідних штучних нейронних мереж (ШНМ), розроблені та описані у цьому розділі.

РОЗДІЛ 5

МЕТОДИ ЧАСТКОВОЇ КОМПЕНСАЦІЇ ПОХИБОК ШНМ НА ОСНОВІ КАСКАДНИХ АНСАМБЛІВ В ЗАДАЧАХ РЕГРЕСІЇ

У цьому розділі розроблено нові та вдосконалено існуючі методи компенсації похибок штучної нейронної мережі на основі каскадних ансамблів нейромережевих моделей. Аналітично обґрунтовано можливість покращення результатів апроксимації та часткового усунення систематичної похибки нейронної мережі узагальненої регресії із використанням каскаду з двох таких нейронних мереж. Розроблено метод компенсації похибки, який включає можливість формування вихідного сигналу одним з трьох способів: підсумовуванням виходів двох нейромережевих моделей, зваженим підсумовуванням їх результатів з використанням коректуючої лінійної неітеративної нейроподібної структури, нелінійним розширенням вихідних сигналів обох нейронних мереж за допомогою полінома Колмогорова-Габора з подальшим коригуванням результатів додатковою нейронною мережею. Вдосконалено метод глобально локальної апроксимації, який забезпечує незалежне виділення компонент шуканого сигналу за допомогою ансамблю з двох нейронних мереж узагальненої регресії та дозволяє поетапно уточнювати шукані значення. Розроблено прямий і зворотний методи компенсації похибки моделі регресії на основі раціональних дробів. Вони забезпечують також можливість ітераційного підвищення точності шуканого значення.

Наукові і практичні результати, одержані у цьому розділі, опубліковано у розділі колективної монографії [32], періодичних виданнях України та високореєтингових журналах інших держав, які включено до наукометричних баз Scopus та Web of Science: [5], [31], [33], [39], [45], [46], [47], [48], а також апробовано і опубліковано у матеріалах закордонних конференцій: [37], [38], [40], [53].

5.1 Компенсація похибок штучної нейронної мережі для задач регресії

У сучасних прикладних задачах біомедичної інженерії моделі машинного навчання та штучних нейронних мереж (ШНМ) часто використовуються для аналізу складних нелінійних залежностей. Проте через обмеженість обсягу доступних клінічних або експериментальних даних, наявність шуму та нерівномірного розподілу вибірки, складність самої апаратно-програмної реалізації мережі отримати високоточні та робастні результати далеко не завжди вдається.

Зменшення точності ШНМ нерідко призводить до хибної інтерпретації результатів, що в біомедичних задачах може мати критичні наслідки для планування, лікування або раннього виявлення патологій. Крім того, нестабільність результатів унеможливорює впровадження подібних методів у реальні клінічні чи інженерні процеси без додаткової перевірки та верифікації. У таблиці 5.1 подано перелік основних чинників, які обмежують точність результатів апроксимації моделями штучних нейронних мереж (ШНМ) у таких задачах.

Таблиця 5.1 – Основні причини зменшення точності роботи ШНМ

№	Чинник	Пояснення	Джерело
1	Якість та кількість даних	Вибірки малого обсягу або незбалансовані вибірки даних ускладнюють навчання моделей, а наявність шуму зменшує їх здатність виділяти значущі взаємозв'язки.	[262], [263]
2	Складність моделі	Складні моделі можуть захоплювати шум з даних і перенавчатися, тоді як надто прості – не вловлювати ключові залежності тобто недонавчатися.	[264], [265]

3	Алгоритмічні обмеження	Випадкова ініціалізація ваг і невдалий вибір активаційних функцій або числа шарів ШНМ можуть уповільнювати збіжність та знижувати точність ітеративних засобів машинного навчання.	[263], [266]
4	Обчислювальна складність	Високі обчислювальні витрати можуть обмежувати масштабування моделі в той час, коли оптимізація суттєво уповільнює процес апроксимації.	[267]
5	Представлення даних і відбір ознак	Використання великої кількості ознак призводить до «прокляття вимірності», а недостатня селекція – до втрати інформативності.	[268], [269]
6	Зміна розподілу даних	Якщо в реальних умовах розподіл вхідних даних відрізняється від того, на якому навчали ШНМ, її точність погіршується. Таке явище може виникати через зміну способу збору даних, оновлення сенсорів або еволюцію характеристик об'єктів, що апроксимуються.	[270]

Враховуючи вищезгадані обмеження, виникає проблема підвищення точності роботи моделей машинного навчання та штучних нейронних мереж у випадку їх значних похибок в процесі розв'язування різноманітних біомедичних завдань. Розглянемо її детальніше.

Нехай маємо вибірку даних малого обсягу:

$$A = [x_{i,j}] \in \mathbb{R}^{n \times m}, i = \overline{1, n}; j = \overline{1, m}, \quad (5.1)$$

де n – кількість векторів, m – вимірність кожного вектора. Кожний вхідний вектор позначимо як:

$$x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})^T \in \mathbb{R}^m, i = \overline{1, n}. \quad (5.2)$$

Відповідні значення цільової змінної подані таким чином:

$$\begin{cases} y_i \in \{1, 2, \dots, K\}, \text{ для задач класифікації;} \\ y_i \in \mathbb{R}, \text{ для задач регресійного аналізу;} \end{cases}, i = \overline{1, n}, \quad (5.3)$$

де K – кількість класів у задачі класифікації.

Нехай, базова ШНМ-1 дає початковий результат:

$$y_i^{(0)pred} = F_1(x_i; \theta_1), i = \overline{1, n}, F_1: \mathbb{R}^m \times \Theta_1 \rightarrow \mathbb{R}, \quad (5.4)$$

де Θ_1 – простір можливих значень параметрів θ_1 , які визначають структуру та властивості $F_1(\cdot)$.

Проте в багатьох випадках (див. табл. 5.1) обрана похибка, наприклад, середньоквадратична похибка E_{base} :

$$E_{base}(\theta_1) = \frac{1}{n} \sum_{i=1}^n \left(r_i^{(0)pred} \right)^2, r_i^{(0)pred} = y_i - y_i^{(0)pred} \quad (5.5)$$

може виявитися суттєвою.

Проблема полягає у тому, щоб знайти таку функцію уточнення F_2 :

$$F_2: \mathbb{R}^m \times \mathbb{R} \times \Theta_2 \rightarrow \mathbb{R}, \quad (5.6)$$

де Θ_2 – простір можливих значень параметрів θ_2 , які визначають структуру та властивості $F_2(\cdot)$ і спосіб поєднання його результатів з виходом $F_1(\cdot)$, щоб кінцевий результат апроксимації:

$$y_i^{pred} = F_2(x_i, r_i^{(0)pred}; \theta_2), i = \overline{1, n}$$

дозволяв отримати похибку:

$$E_{cascade}(\theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^{pred})^2, \quad (5.7)$$

для якої чинна наступна нерівність:

$$E_{cascade}(\theta_1, \theta_1) < E_{base}(\theta_1). \quad (5.8)$$

5.2 Метод часткової компенсації похибки роботи нейронної мережі узагальненої регресії

Нейронну мережу узагальненої регресії (НМУР) розроблено Donald F. Specht у 1991 році [157]. Ця нейронна мережа використовується для моделювання нерегулярних суттєво нелінійних поверхонь відгуку. З часу свого створення вона або її гібриди набули широкого застосування для розв'язання різноманітних практичних задач [271], [272].

Проаналізуємо особливості функціонування НМУР. Для усіх векторів x_i (5.2) опорної вибірки даних A , які нормалізовано обраною функцією $\mathcal{N}(\cdot)$:

$$\acute{x}_i = \mathcal{N}(x_i), \quad (5.9)$$

можна знайти значення вихідного сигналу:

$$y_i^{pred} = f_1(\acute{x}_i), i = \overline{1, n}. \quad (5.10)$$

Для пошуку y_k^{pred} на основі k -го вхідного вектора та усіх векторів опорної вибірки даних (5.9) можна використати НМУР. Реалізація НМУР передбачає виконання таких кроків [51], [273]:

1. Знайти евклідові відстані від вхідного вектора $\acute{x}_{k,j}$ до усіх векторів $\acute{x}_{i,j}$ з відомими значеннями виходів y_i , які вважаємо опорними:

$$D_{k,i} = \sqrt{\sum_{j=1}^m (\acute{x}_{k,j} - \acute{x}_{i,j})^2}, i, k = \overline{1, n}, j = \overline{1, m}. \quad (5.11)$$

2. Обчислити гаусівські функції від евклідових відстаней:

$$G_{k,i} = \exp\left(-\frac{(D_{k,i})^2}{2\sigma^2}\right), \sigma > 0, \quad (5.12)$$

де σ – розмах гаусівської функції, єдиний параметр НМУР, який потрібно налаштувати.

3. Обчислити значення шуканої величини y_k^{pred} :

$$y_k^{pred} = \frac{\sum_{i=1}^n y_i G_{k,i}}{\sum_{i=1}^n G_{k,i}}, k = \overline{1, n}. \quad (5.13)$$

Проаналізуємо складову систематичної похибки формування вихідного сигналу мережі НМУР на прикладі поточного вектора k :

$$\varepsilon_k = y_k - y_k^{pred}. \quad (5.14)$$

Для апроксимації (5.14) можна використати другу НМУР:

$$\varepsilon_k^{pred} = \frac{\sum_{i=1}^n \varepsilon_i \hat{G}_{k,i}}{\sum_{i=1}^n \hat{G}_{k,i}}, \quad (5.15)$$

але тільки для кожного вектора з відомим значенням вихідного сигналу з (5.2), де $\varepsilon_k = y_k - y_k^{pred}$ – похибка на k -му опорному векторі, $\hat{G}_{k,i}$ – гаусівська функція від (5.11) з суттєво меншим значенням розмаху гаусівської функції σ_{error} , яка використовується для обчислення гаусівських функцій для (5.15).

Ця формула показує, що поверхня відгуку похибки є достатньо гладкою і, отже, може бути певним чином промодельована в локальній області простору вхідних змінних. Застосування ще однієї НМУР зі зменшеною величиною розмаху гаусівської функції $\sigma_{error} \ll \sigma$ повинно забезпечити задовільне наближення похибки НМУР. Прийmemo до уваги, що для підвищення точності обчислень величини похибки згідно формули (5.15) значення коефіцієнта розмаху σ_{error} необхідно обирати значно меншим, ніж при їх використанні для формули (5.13). Це пояснюється відмінностями рельєфів поверхні відгуку багатовимірної функції НМУР і поверхні відгуку похибки цієї функції.

Фінальну формулу для пошуку y_u^{res} запишемо так:

$$y_k^{res} = \frac{\sum_{i=1}^n y_i G_{k,i}}{\sum_{i=1}^n G_{k,i}} + \frac{\sum_{i=1}^n (y_i - y_i^{pred}) \dot{G}_{k,i}}{\sum_{i=1}^n \dot{G}_{k,i}}. \quad (5.16)$$

Для пошуку значень y_k^{res} на основі двох доданків (5.16), де перший доданок $- y_k^{pred}$, а другий $- \varepsilon_i$, можна використати послідовний каскадний ансамбль з двох нейронних мереж узагальненої регресії.

Розглянемо етапи підготовки/навчання і застосування методу для реалізації такого каскадного ансамблю.

Етап підготовки/навчання каскадного ансамблю для часткової компенсації похибки роботи нейронної мережі узагальненої регресії передбачає виконання таких кроків:

1. Знайти вихідний сигнал y_k^{pred} для кожного вектора опорної вибірки даних $\dot{x}_i, i = \overline{1, n}$. Потрібно виконати обчислення для кожного поточного вектора $\dot{x}_k, k = \overline{1, n}$ на основі опорної вибірки з $n - 1$ векторів при $k \neq i$:

$$y_k^{pred} = \frac{\sum_{i \neq k}^n y_i \exp\left(-\frac{\sum_{j=1}^m (\dot{x}_{k,j} - \dot{x}_{i,j})^2}{2\sigma^2}\right)}{\sum_{i \neq k}^n \exp\left(-\frac{\sum_{j=1}^m (\dot{x}_{k,j} - \dot{x}_{i,j})^2}{2\sigma^2}\right)}. \quad (5.17)$$

2. Обчислити похибки НМУР-1 для кожного i -го опорного вектора:

$$\varepsilon_i = y_i - y_i^{pred}. \quad (5.18)$$

3. Сформувати опорну вибірку даних для отримання похибки НМУР-1 першого рівня каскаду, де набір незалежних атрибутів $\dot{x}_i, i = \overline{1, n}$ залишається таким самим, а відповідний i -тий залежний атрибут обчислено з використанням (5.18).

4. Виконати навчання коректуючої лінійної нейроподібної структури МПГП, на входи якої подати усі i -ті y_i^{pred} з (5.17) та ε_i з (5.18), а на вихід – відповідні значення залежного атрибуту y_i :

$$y_i^{res} = SGT M_1(y_i^{pred}, \varepsilon_i), i = \overline{1, n}. \quad (5.19)$$

Позначимо це як Варіант 2.

5. Виконати навчання коректуючої лінійної нейроподібної структури МПГП, входи якої розширити поліномом Колмогорова-Габора другого степеня $Poly(y_i^{pred} \varepsilon_i)$:

$$y_i^{res} = SGT M_2(Poly(y_i^{pred}, \varepsilon_i)), i = \overline{1, n}. \quad (5.20)$$

Позначимо це як Варіант 3.

У режимі застосування каскадного ансамблю маємо нормалізований на основі (5.9) тестовий вектор $\dot{u} = (\dot{u}_1, \dot{u}_2, \dots, \dot{u}_m)$, для якого невідоме значення вихідного сигналу, тобто відсутній вихідний елемент y_u .

Етап застосування каскадного ансамблю для часткової компенсації похибки роботи нейронної мережі узагальненої регресії передбачає виконання таких кроків:

1. Виконати застосування НМУР-1 каскадного ансамблю з використанням нормалізованої опорної вибірки даних (5.9) та відомих значень залежного атрибуту y_i з (5.3):

$$y_u^{pred} = \frac{\sum_{i=1}^n y_i \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma^2}\right)}. \quad (5.21)$$

2. Виконати застосування НМУР-2 каскаду для отримання похибки роботи НМУР-1 першого рівня каскадного ансамблю:

$$\varepsilon_u^{pred} = \frac{\sum_{i=1}^n \varepsilon_i \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma_{error}^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma_{error}^2}\right)}, \quad (5.22)$$

де ε_i визначається з (5.13).

3. Сформувати фінальний результат роботи каскадного ансамблю (Варіант 1) [161]:

$$y_u^{res} = y_u^{pred} + \varepsilon_u^{pred}. \quad (5.23)$$

4. Виконати застосування першої попередньо навченої лінійної нейроподібної структури МПГП (Варіант 2) [160]:

$$y_u^{res} = SGT M_1(y_u^{pred}, \varepsilon_u^{pred}). \quad (5.24)$$

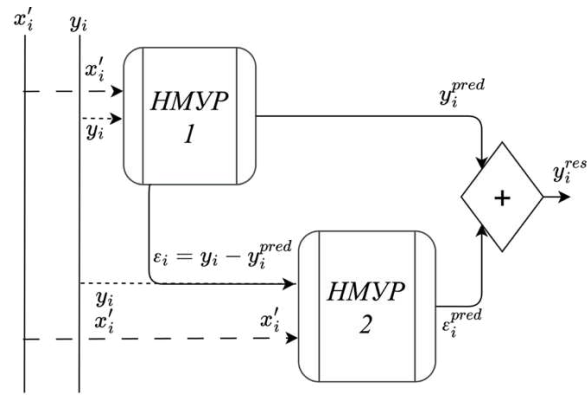
5. Виконати застосування другої попередньо навченої лінійної нейроподібної структури МПГП, входи якої розширити поліномом Колмогорова-Габора другого степеня (Варіант 3) [31]:

$$y_u^{res} = SGT M_2(Poly(y_u^{pred}, \varepsilon_u^{pred})). \quad (5.25)$$

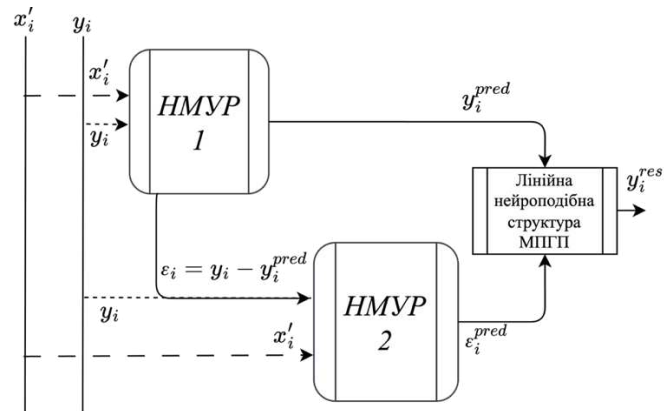
Залежно від обраного варіанту (5.23, 5.24 чи 5.25) можна отримати фінальний результат роботи каскаду.

Структурно-функційну схему методу з різними варіантами формування вихідного сигналу наведено на рис. 5.1.

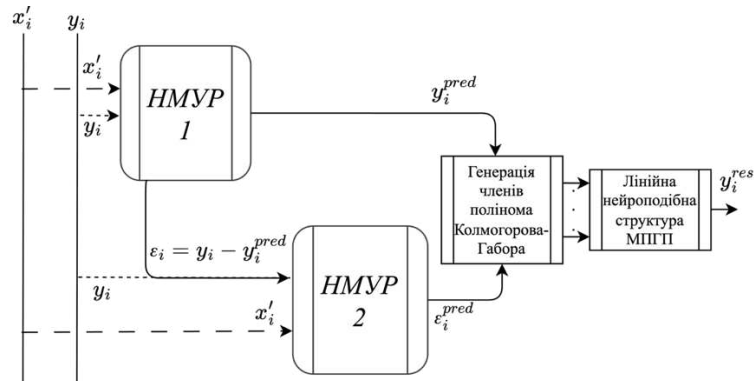
Метод часткової компенсації методичної похибки роботи нейронної мережі узагальненої регресії на основі каскадного ансамблю двох НМУР може забезпечити суттєве підвищення точності апроксимації в процесі аналізу біомедичних вибірок обмеженого обсягу. Це досягається завдяки поєднанню нечутливості до розподілу даних з цілеспрямованою компенсацією систематичної похибки. Уже сама НМУР відома своєю здатністю моделювати складні нелінійні залежності без жорстких припущень щодо форми поверхні відгуку [157].



а)



б)



в)

Рисунок 5.1 – Структурно-функційна схема методу часткової компенсації похибки роботи нейронної мережі узагальненої регресії на основі а) підсумовування (5.23); б) використання коректуючої лінійної нейроподібної структури МППП (5.24) ; в) нелінійного розширення вихідних сигналів обох НМУР поліномом Колмогорова-Габора з подальшим використанням коректуючої ШНМ (5.25).

Це важливо для біомедичних даних, де часто маємо справу з гетерогенними шумами [274]. Формула первинного результату (5.17) дозволяє отримувати гладку апроксимацію навіть на екстремально коротких вибірках даних. Це пояснюється тим, що кожен новий вектор оцінюється через ядрове зважування всіх наявних точок. Другу НМУР у каскаді призначено для компенсації систематичної похибки першого рівня, апроксимуючи залишкову різницю (5.18) як функцію вхідних ознак. Використовуючи невеликі значення σ_{error} у (5.22), формула (5.21) утворює високоточний локальний коректор, який компенсує систематичний зсув первинного результату апроксимації.

Агрегування результату на основі простого підсумовування первинного результату НМУР та похибки її роботи забезпечує зрозумілу корекцію систематичної похибки без додаткового навчання. У задачах з малою кількістю спостережень та складними нелінійностями, характерними для біомедичних даних, таке рішення дозволяє знизити упередження моделі. Це відбувається за рахунок точного обчислення похибки наявної моделі та її компенсації, проте не впливає на дисперсію, тобто сумарний розкид результатів лишається в межах первинної НМУР.

Використання коректуючої, лінійної, неітеративної нейроподібної структури МПГП, яка навчається лише на двох незалежних атрибутах, дозволяє оптимально зважити внесок первинного результату та похибки (5.19). За класичним розкладом на зсув і дисперсію [275] лінійний коректор одночасно знижує залишковий зсув і незначно збільшує дисперсію, якщо навчальна вибірка даних має обмежений обсяг.

Розширення входів лінійної нейроподібної структури МПГП поліномом Колмогорова–Габора другого степеня надає моделі здатність апроксимувати довільні непарні, нелінійні взаємозв'язки між y_u^{pred} та ε_u . Формула (5.20), де $Poly(\cdot)$ створює базисні мономи до другого степеня, базується на теоремі Колмогорова–Арнольда про універсальну апроксимацію [117]. Завдяки цьому лінійна неітеративна нейроподібна структура МПГП може коригувати складнішу

поверхню відгуку, зменшуючи як упередження, так і дисперсію в умовах малих вибірок даних, зокрема, за рахунок використаного нелінійного перетворення [43], [51], [216].

5.3 Удосконалений метод розв’язання задач регресії/класифікації на основі виділення глобальної та локальної компонент поверхні відгуку

Обмеженість обсягів біомедичних вибірок часто призводить до нестабільності в процесі використання існуючих ШНМ чи методів машинного навчання. У таких умовах подібні методи демонструють низьку здатність до узагальнення, а подекуди і перенавчання. Це зменшує можливість використання існуючих методів у клінічних і дослідницьких застосуваннях.

Глобально-локальна апроксимація вирішує цю проблему через розділення сигналу на дві доповнювані компоненти. Подібні підходи обґрунтовуються аналізом за способом “зсув–дисперсія” [275], знижуючи загальну дисперсію апроксимації без помітного зростання зсуву. Окрім цього глобально-локальний підхід підвищує рівень інтерпретованості методу, дозволяючи окремо оцінювати вклад кожного рівня апроксимації. Це може суттєво підсилити довіру клініцистів до інтелектуалізованих систем на основі таких методів в процесі розв’язання різноманітних задач біомедичної інженерії.

У цьому підрозділі описано удосконалений метод глобально локальної апроксимації для підвищення точності розв’язання задачі регресії у випадку аналізу малих даних [276]. В основу удосконалення покладено можливість використання ансамблю з двох нейронних мереж узагальненої регресії, які функціонуватимуть паралельно в режимі застосування методу для незалежного виділення локальної та глобальної компоненти поверхні відгуку. Удосконалений метод передбачає також можливість використання ітеративного режиму для отримання результатів підвищеної точності.

Нехай оперуємо вибіркою даних малого обсягу A з (5.1) та вектором $x_i \in \mathbb{R}^m, i = \overline{1, n}$ вибірки даних A . Відповідне значення залежного атрибуту для кожного i -го вектора опорної вибірки даних позначено у (5.3).

Удосконалений метод розв'язання задач регресії/класифікації на основі виділення глобальної та локальної компонент поверхні відгуку складається з двох етапів:

- навчання,
- застосування.

Етап навчання удосконаленого методу передбачає виконання таких основних кроків:

1. Сформувані розширену вибірку даних A^+ шляхом додавання значення залежного атрибуту як додаткову незалежну ознаку:

$$A^+ = \begin{bmatrix} (x_1^+)^T \\ (x_2^+)^T \\ \vdots \\ (x_n^+)^T \end{bmatrix} \in \mathbb{R}^{n \times (m+1)}. \quad (5.26)$$

2. Кожен її розширений вектор x_i^+ формувати таким чином:

$$x_i^+ = (x_i, y_i) \in \mathbb{R}^{m+1}, i = \overline{1, n}. \quad (5.27)$$

У (5.26) кожен елемент $x_i^+ = (x_i, y_i)$, розглядається як вектор у $\mathbb{R}^{m+1}$.

3. Нормалізувати усі розширені вектори $\acute{x}_i^+$ вибірки даних A^+ обраною функцією $\mathcal{N}(\cdot)$:

$$\acute{x}_i^+ = \mathcal{N}(x_i^+), i = \overline{1, n}. \quad (5.28)$$

4. Виконати навчання лінійної нейроподібної структури МППП “звуженого горла” $f_{SGTM}(\cdot)$ з використанням $\acute{A}^+$, де кількість нейронів вхідного шару становить $m + 1$, кількість нейронів прихованого шару – m , кількість нейронів вихідного шару – 1.

5. Застосувати кожен i -тий вектор $\dot{x}_i^+$ з $\dot{A}^+ \in \mathbb{R}^{n \times (m+1)}$ на попередньо навченій $f_{SGTM}(\cdot)$ для отримання глобальної компоненти поверхні відгуку:

$$y_i^{global} = f_{SGTM}(\dot{x}_i^+). \quad (5.29)$$

6. Для кожного вхідного i -го вектора, $i = \overline{1, n}$ з (5.2) обчислити локальну компоненту поверхні відгуку для (5.28) за формулою:

$$y_i^{local} = y_i - y_i^{global}. \quad (5.30)$$

Структурно-функційну схему роботи удосконаленого методу глобально локальної апроксимації подано на рис. 5.2.

На етапі застосування удосконаленого методу, на відміну від [276], використовується ансамбль з двох штучних нейронних мереж узагальненої регресії. Перша з них використовується для визначення глобальної компоненти відгуку, тоді як друга – для локальної. Для опису основних кроків процедури застосування оперуватимемо нормалізованим на основі (5.28) тестовим вектором $\dot{u} = (\dot{u}_1, \dot{u}_2, \dots, \dot{u}_m)$, для якого невідоме значення вихідного сигналу, тобто відсутній вихідний елемент y_u .

Етап застосування удосконаленого методу передбачає виконання таких основних кроків:

1. Виконати застосування НМУР-1 для отримання попереднього результату $y_u^{(0)}$ (груба оцінка) для вектора $\dot{u}$, який засновано на глобальних закономірностях опорної вибірки даних. Зокрема, якщо звести основні кроки роботи НМУР (5.11-5.12) у (5.13) то отримаємо загальну формулу для $y_u^{(0)}$:

$$y_u^{(0)} = \frac{\sum_{i=1}^n y_i \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma_{glob}^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{\sum_{j=1}^m (\dot{u}_j - \dot{x}_{i,j})^2}{2\sigma_{glob}^2}\right)}, \quad (5.31)$$

де σ_{glob} – параметр розмаху гаусівської функції, який використовується для НМУР-1.

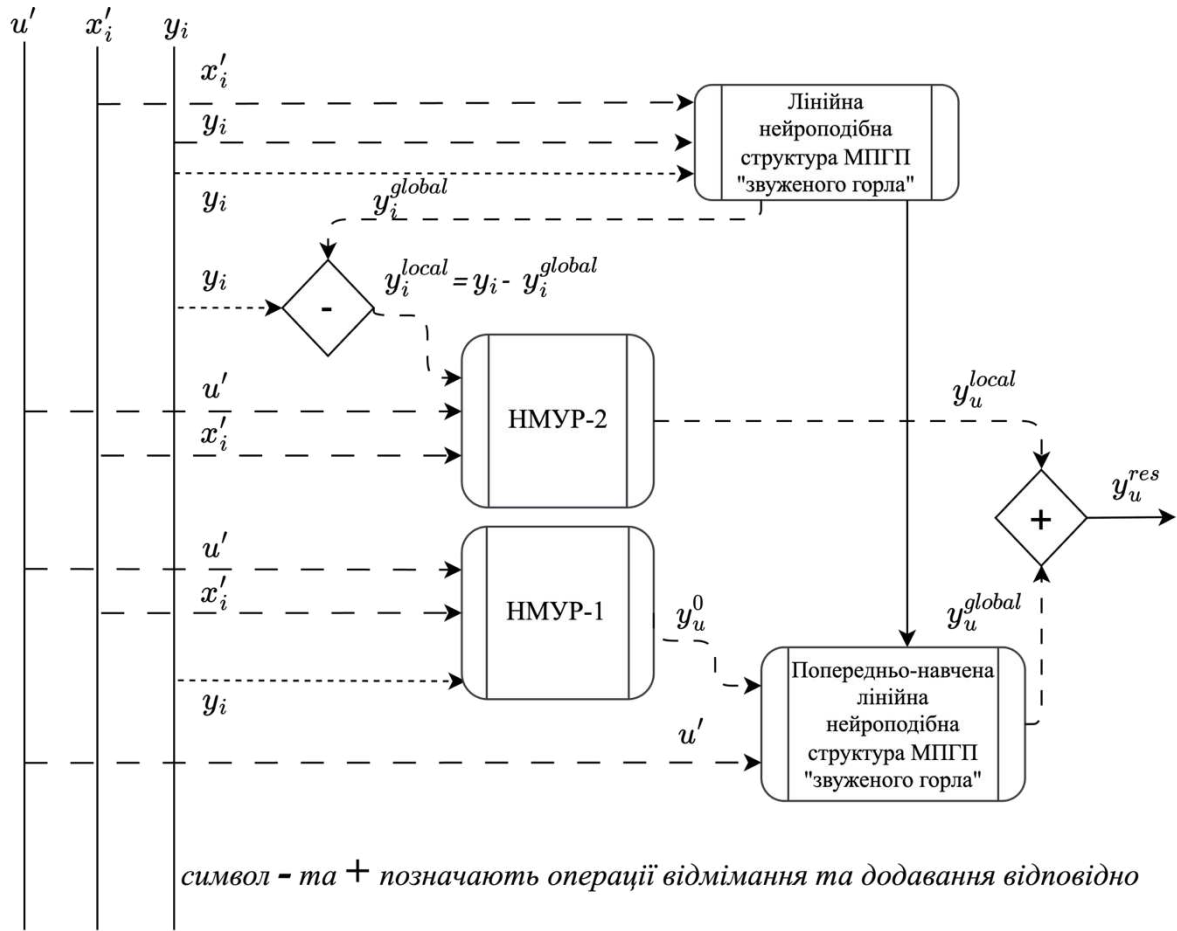


Рисунок 5.2 – Структурно-функційна схема удосконаленого методу глобально локальної апроксимації

2. Виділити глобальну частину невідомого значення вихідного сигналу для $\dot{u}$ шляхом побудови розширеного вектора виду:

$$u^+ = (\dot{u}, \dot{y}_u^{(0)}) \in \mathbb{R}^{m+1}, \quad (5.32)$$

де $\dot{y}_u^{(0)}$ – нормалізоване значення для (5.31) з використанням (5.30).

3. Застосувати u^+ на попередньо навченій $f_{SGTM}(\cdot)$:

$$y_u^{global} = f_{SGTM}(u^+). \quad (5.33)$$

4. Виділити локальну складову невідомого значення вихідного сигналу для $\hat{u}$ шляхом використання другої НМУР-2 з меншим значенням розмаху гаусівської функції $\sigma_{loc} \ll \sigma_{glob}$, яка апроксимує залишковий сигнал $y_i - y_i^{global}$ з (5.30) на основі опорної вибірки даних:

$$y_u^{local} = \frac{\sum_{i=1}^n (y_i - y_i^{global}) \exp\left(-\frac{\sum_{j=1}^m (\hat{u}_j - \hat{x}_{i,j})^2}{2\sigma_{loc}^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{\sum_{j=1}^m (\hat{u}_j - \hat{x}_{i,j})^2}{2\sigma_{loc}^2}\right)}. \quad (5.34)$$

5. Отримати фінальне значення вихідного сигналу для вектора $\hat{u}$ шляхом сумування глобальної та локальної складових, отриманих у (5.33-5.34):

$$y_u^{res} = y_u^{global} + y_u^{local}. \quad (5.35)$$

Для підвищення точності фінального результату отримане значення, y_u^{res} можна підставити у (5.32) замість $\hat{y}_u^{(0)}$ і повторити (5.33)-(5.35) до досягнення заданого критерію зупинки.

Удосконалений метод глобально локальної апроксимації для біомедичних вибірок даних малих обсягів має такі ключові переваги:

- використання лінійної нейроподібної структури МПГП з додатковою ознакою – цільовим атрибутом або її аналогу – авто-асоціативної нейроподібної структури в базовому [276] та удосконаленому методах дозволяє відтворити поверхню відгуку з нульовою похибкою роботи методу [132]. В той же час використання топології звуженого горла цієї ШНМ виконує нелінійну РСА-подібну проєкцію, зберігаючи головні напрямки варіації даних у “звуженому горлі” [276]. Це дозволяє досить точно відтворити глобальну компоненту поверхні відгуку;
- використання ансамблю з двох незалежних НМУР замість єдиної РБФ-мережі дозволяє розділити завдання пошуку глобальної та локальної

частин сигналу незалежно у порівнянні з базовим методом з [276]. НМУР значно ефективніша для роботи з обмеженими вибірками даних, в той час коли РБФ-мережі особливо ефективні в умовах аналізу достатніх об'ємів даних для реалізації процедури навчання;

- використання НМУР-1 на основі (5.31) з великим σ_{glob} виділяє глобальні закономірності в даних, тоді як НМУР-2 (5.34) з малим σ_{loc} спрямовано на визначення виключно локальних залишкових похибок. Така архітектура відповідає класичному аналізу “зсув–дисперсія” і дозволяє знижувати загальну дисперсію апроксимації без суттєвого збільшення зсуву [275];
- використання НМУР-1 на оригінальних y_i та НМУР-2 - на залишках $(y_i - y_i^{global})$ з (5.30) дозволяє кожній ШНМ зосередитися на своїй підзадачі. Завдяки цьому кожна модель витрачає обчислювальні ресурси на свою частину сигналу, а не на повторне відтворення уже добре апроксимованих залежностей;
- можливість ітеративного уточнення апроксимації у кілька кроків дає змогу ітераційно зменшити залишкову похибку, що є також перевагою в порівнянні з базовим методом [132]. Це реалізує механізм багатоетапного уточнення, подібний до ідеї багат шарового спуску у глибоких автоенкодерах [277]. Таким чином, можливість ітеративного уточнення апроксимації у кілька кроків підвищує стабільність апроксимації біомедичних вибірок даних.

5.4 Методи компенсації похибки моделі регресії з використанням раціональних дробів

Апроксимація раціональними дробами - це метод наближення функцій за допомогою відношення двох поліномів, що забезпечує високу точність

моделювання навіть у випадках з розривами або нерівномірною поведінкою цільових функцій [278]. Вона залишається одним з ключових інструментів у чисельному аналізі та наукових обчисленнях [37], [279].

На відміну від класичної поліноміальної апроксимації раціональні дробы дозволяють ефективно відтворювати асимптотичні властивості та локальні особливості функції при відносно невеликих степенях чисельника й знаменника. Особливо це помітно, коли слід зберегти точність на великому діапазоні аргументів, що дуже характерно для задач біомедичної інженерії або відтворити асимптотичні властивості функції [37].

Апроксимація раціональними дробами має ряд переваг над іншими видами апроксимації [280], [281]. Зокрема, раціональні дробы можуть забезпечувати [5], [38], [53], [279], [282], [283], [284]:

- високоточну апроксимацію функцій з сингулярностями, де поліноміальні підходи виявляються неефективними;
- можливість апроксимації складних функцій з високою точністю на великому діапазоні значень параметрів;
- достатню збіжність за рахунок того, що вони забезпечують збіжність не тільки в точках, але і в проміжках між точками, де значення функції можуть бути великими за модулем;
- можливість ефективної апроксимації функцій, які швидко змінюються в деяких точках, завдяки врахуванню особливостей поведінки функції в цих точках;
- більш гладку залежність між значеннями функції в окремих точках;
- збереження стабільності та високої точності за рахунок використання оптимізаційних методів, що є критично важливим для динамічних моделей;
- уникнення викидів на краях діапазону.

У цьому підрозділі розглянуто розроблені методи компенсації похибки моделі регресії з використанням раціональних дробів на основі прямого та зворотного підходів.

5.4.1. Прямий метод компенсації похибки моделі регресії

Розглянемо прямий метод компенсації похибки моделі регресії на основі використання раціональних дробів. Його реалізовано у вигляді дворівневого каскадного ансамблю з додатковою можливістю ітеративного уточнення результатів апроксимації для підвищення точності аналізу даних малих обсягів.

Нехай усі вектори x_i нормалізованої вибірки даних A нормалізовано обраною функцією $\mathcal{N}(\cdot)$ і апроксимовано деякою функцією $f_1(\cdot)$:

$$\acute{x}_i = \mathcal{N}(x_i), \quad y_i^{pred} = f_1(\acute{x}_i), \quad i = \overline{1, n}, \quad (5.36)$$

де $\acute{x}_i$ – нормалізований вектор, отриманий з x_i за допомогою $\mathcal{N}(\cdot)$, y_i^{pred} – апроксимоване значення для i -го вектора матриці $\acute{A}$, $f_1: \mathbb{R}^m \rightarrow \mathbb{R}$ – перша ШНМ.

Слід зазначити, що для виконання подальших кроків не потрібно використовувати метод розв’язання задачі регресії $f_1(\cdot)$, деталі його імплементації, оптимальні параметри тощо. Достатньо мати вхідні дані, на основі яких виконувалася апроксимація та отримані вихідні сигнали [5].

Завдання полягає у тому, щоб:

$$y_i^{pred} \approx y_i, \quad i = \overline{1, n}. \quad (5.37)$$

Реалізація прямого підходу компенсації похибки моделі регресії (5.37) полягає у введенні та використанні додаткової коригувальної функції виду [5]:

$$f_2: \mathbb{R}^{m+1} \rightarrow \mathbb{R}, \quad (5.38)$$

яка представлена другим методом машинного навчання/штучною нейронною мережею. Для цього використовується раціональний дріб виду:

$$y_i \approx \frac{y_i^{pred}}{1+f_2(\acute{x}_i)}. \quad (5.39)$$

На підставі формули (5.39) можна отримати таке наближення:

$$1 + f_2(\acute{x}_i) \approx \frac{y_i^{pred}}{y_i},$$

звідки:

$$f_2(\acute{x}_i) \approx \frac{y_i^{pred}}{y_i} - 1. \quad (5.40)$$

Праві частини (5.40) є відносними похибками апроксимації для кожного i -го вектора вибірки $\acute{A}$, отриманих з використанням першої ШНМ. Їх подамо у вигляді:

$$t_i = \frac{y_i^{pred}}{y_i} - 1, i = \overline{1, n}. \quad (5.41)$$

Етап навчання для прямого методу компенсації похибки моделі регресії полягає у послідовному виконанні таких кроків [5]:

1. Формування вибірки даних A^+ для реалізації процедури навчання другого засобу машинного навчання (коректувальної функції):

$$A^+ = (\acute{x}_i, t_i) \in \mathbb{R}^{n \times (m+1)}. \quad (5.42)$$

2. Навчання коректуючої ШНМ $f_2(A^+)$ і отримання апроксимованого значення t_i^{pred} для кожного вектора матриці A^+ :

$$t_i^{pred} = f_2(\acute{x}_i), i = \overline{1, n}. \quad (5.43)$$

3. Застосування навчальної вибірки даних на другій попередньо навченій ШНМ. Виконати уточнення значення i -го вихідного сигналу за формулою:

$$y_i^{(1)} = \frac{y_i^{pred}}{t_i^{pred} + 1}. \quad (5.44)$$

Слід зазначити, що крок 3 виконується для оцінки похибки на навчальній вибірці даних.

Структурно-функційну схему прямого методу компенсації похибки моделі регресії на основі використання раціональних дробів можна подати у вигляді дворівневого каскадного ансамблю, який подано на рис. 5.3.

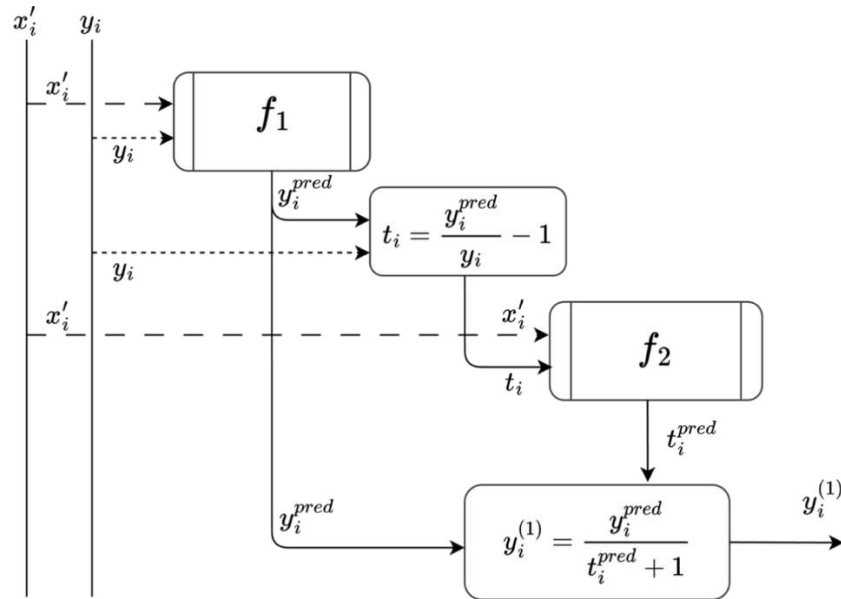


Рисунок 5.3 – Структурно-функційна схема прямого методу компенсації похибки моделі регресії на основі використання раціональних дробів

Етап застосування для прямого методу компенсації похибки моделі регресії полягає у послідовному виконанні таких кроків [5]:

1. Нормалізація поточного вектора $u = (u_1, u_2, \dots, u_m,)$ з невідомим значенням вихідного сигналу функцією нормалізації $\mathcal{N}(\cdot)$, яка використовувалася для навчальної вибірки даних.
2. Апроксимація значення $y_u^{pred} = f_1(u)$ для поточного вектора $u = (u_1, u_2, \dots, u_m,)$ шляхом його застосування на першій попередньо навченій ШНМ.

3. Апроксимація значення відносної похибки роботи першої ШНМ $t_u^{pred} = f_2(u)$ для поточного вектора $u = (u_1, u_2, \dots, u_m)$, шляхом його застосування на другій попередньо навчений ШНМ.
4. Корекція результати для u -го вхідного вектору за формулою:

$$y_u^{(1)} = \frac{y_u^{pred}}{t_u^{pred} + 1}. \quad (5.45)$$

У випадку, якщо точність корекції не відповідає вимогам користувача, її можна підвищити за рахунок можливості ітеративної компенсації похибки моделі в процесі інтелектуального аналізу коротких вибірок даних. Для цього, $y_i^{(1)}$ з (5.44) підставити замість y_i^{pred} з (5.39) і виконати ітераційний процес $y_i^{(iter)}$, $iter = 2, 3, \dots$ компенсації похибки моделі регресії відповідно до обраного користувачем критерію зупинки.

Структурно-функційну схему розробленого каскадного ансамблю наведено на рис. 5.4.

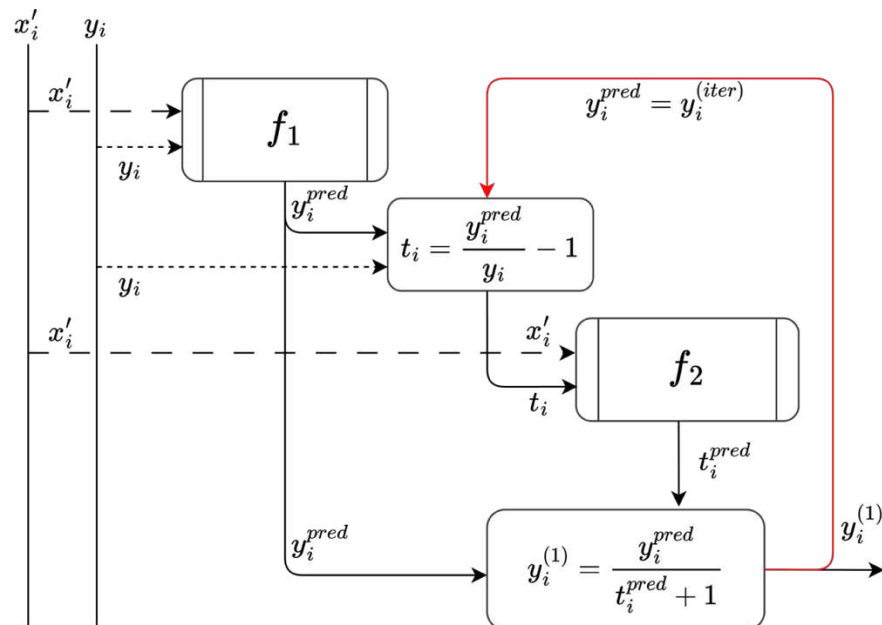


Рисунок 5.4 – Структурно-функційна схема прямого методу компенсації похибки моделі регресії з можливістю ітеративного підвищення точності розв’язання задач регресії в умовах обмежених обсягів даних

Такий підхід забезпечує можливість підвищення точності апроксимації за рахунок зменшення залишкової похибки в кожному наступному повторенні.

Ряд основних критеріїв зупинки для реалізації ітераційної процедури прямого методу компенсації похибки моделі регресії, які орієнтовано на зменшення глобальної похибки та стабілізації якості апроксимації, наведено у таблиці В.1 додатку В.

Прямий метод компенсації похибки моделі регресії з використанням раціональних дробів поєднує в собі кілька важливих переваг [5]:

- він не вимагає знання внутрішньої структури або способу та параметрів реалізації початкової моделі $f_1(\cdot)$. Процедура навчання другого рівня опирається лише на вхідні вектори вибірки $\hat{A}$ та відносну помилку попередньої апроксимації t_i . Це дозволяє застосовувати корекцію до будь-якої ШНМ/методу машинного навчання, який може бути представлений у вигляді “чорної” чи “сірої скриньки”.

- запропонована модель корекції у вигляді раціонального дробу
$$y_i \approx \frac{y_i^{pred}}{1+f_2(x_{i,j})}$$
 забезпечує коригування з інваріантністю до зміни масштабу.

В цьому випадку похибка коригується відносно величини самого результату. Якщо початковий результат апроксимації близький до істинного значення ($f_2(x_{i,j}) \approx 0$), тоді корекція мінімальна, а у разі великих систематичних зсувів – відповідно більша. Це дозволяє зберегти стійкість моделі на різних ділянках простору ознак.

- навчання другого рівня каскадної конструкції на відносній похибці виду $t_i = \frac{y_i^{pred}}{y_i} - 1$ фокусується саме на тих особливостях чи характеристиках даних, які у першій ШНМ мають гірші властивості. Завдяки цьому коригувальний механізм у вигляді другої ШНМ навчається компенсувати

саме ті похибки, які залишилися поза увагою першої ШНМ, і не витрачає ресурсів на відтворення уже добре апроксимованих залежностей.

- розроблений каскадний ансамбль легко масштабується. Зокрема, при необхідності можна додати кілька ітераційних повторень (для випадку аналізу малих даних) чи кілька нових рівнів каскаду (для випадку аналізу даних середніх та великих обсягів), де кожен наступний каскадний ансамбль навчається на залишкових похибках попереднього. Такий підхід дає змогу послідовно знижувати залишкову похибку й адаптувати модель до найдрібніших особливостей даних.

- розроблений метод зберігає початкову вимірність простору вхідних даних на кожному новому рівні каскадного ансамблю. Такий підхід забезпечує можливість аналізу високовимірних даних обмежених обсягів, особливо у випадках коли $m \sim n$, що доволі часто зустрічається в процесі розв'язання прикладних задач біомедичної інженерії.

Враховуючи все вищесказане, прямий метод компенсації похибки моделі регресії є особливо корисним у задачах з малою вибіркою даних, яка містить шум, що характерно для різних біомедичних задач. Він дозволяє максимізувати корисну інформацію з даних обмежених обсягів, підвищуючи робастність і точність апроксимації без необхідності суттєвого збільшення обчислювальних ресурсів.

5.4.2. Зворотний метод компенсації похибки моделі регресії

У випадках, коли прямий метод (п. 5.4.1) демонструє обмежену здатність до корекції, зокрема, залишкова похибка має мультиплікативний характер або апроксимовані значення сильно відрізняються за масштабом від реальних доцільно застосувати зворотній метод [37].

Цей метод розроблено щоб:

- подолати вади прямої корекції, що полягають у неможливості адекватно підсилити або послабити вплив апроксимації залежно від масштабу вихідної змінної;
- забезпечити лінійне подання коригувальних ознак, яке спростить навчання другої моделі ШНМ;
- надати гнучкий механізм каскадного уточнення з використанням ітераційної процедури, що дозволить зменшувати залишкові відхилення навіть на дуже малих вибірках.

В основі зворотного методу компенсації похибки моделі регресії покладено можливість зворотної корекції початкового результату апроксимації з використанням раціональних дробів [37]. Розглянемо його детальніше.

Нехай для кожного нормалізованого вектора матриці $\acute{A}$:

$$\acute{x}_i = \mathcal{N}(x_i) \in \mathbb{R}^m, \quad i = \overline{1, n} \quad (5.46)$$

перша ШНМ дає:

$$y_i^{pred} = f_1(\acute{x}_i) \quad (5.47)$$

і необхідно, щоб:

$$y_i^{pred} \approx y_i. \quad (5.48)$$

Згідно зворотного підходу компенсація похибки апроксимації виконується за формулою:

$$y_i \approx y_i^{pred} (1 + g(\acute{x}_i)), \quad (5.49)$$

де: $g: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ – друга ШНМ.

Етап навчання для зворотного методу компенсації похибки моделі регресії полягає у послідовному виконанні таких кроків [37]:

1. Обчислення цільових значень залишкової функції:

$$r_i = \frac{y_i}{y_i^{pred}} - 1, i = \overline{1, n}. \quad (5.50)$$

2. Формування вибірки даних A^+ для реалізації процедури навчання другої ШНМ (коректувальної функції):

$$A^+ = (\acute{x}_i, r_i) \in \mathbb{R}^{n \times (m+1)}. \quad (5.51)$$

3. Навчання другої ШНМ $\mathcal{g}(A^+)$ і отримання апроксимованого значення r_i^{pred} для кожного вектора матриці A^+ :

$$r_i^{pred} = \mathcal{g}(\acute{x}_i), i = \overline{1, n}. \quad (5.52)$$

4. Формування уточненого результату:

$$y_i^{(1)} = y_i^{pred} (1 + r_i^{pred}). \quad (5.53)$$

Слід зазначити, що крок 4 виконується лише для оцінки похибки на навчальній вибірці даних.

Структурно-функційну схему зворотного методу компенсації похибки моделі регресії на основі використання раціональних дробів подамо у вигляді схеми дворівневого каскадного ансамблю (рис. 5.5).

Етап застосування зворотного методу компенсації похибки моделі регресії полягає у послідовному виконанні таких кроків [37]:

1. Нормалізація $\acute{u} = \mathcal{N}(u)$, поточного вектора $u = (u_1, u_2, \dots, u_m)$ з невідомим значенням вихідного сигналу функцією нормалізації $\mathcal{N}(\cdot)$, яка використовувалася для навчальної вибірки даних.
2. Апроксимація значення $y_u^{pred} = f_1(\acute{u})$ для нормалізованого вектора $\acute{u} = (\acute{u}_1, \acute{u}_2, \dots, \acute{u}_m)$ шляхом його застосування на першій попередньо-навченій ШНМ.

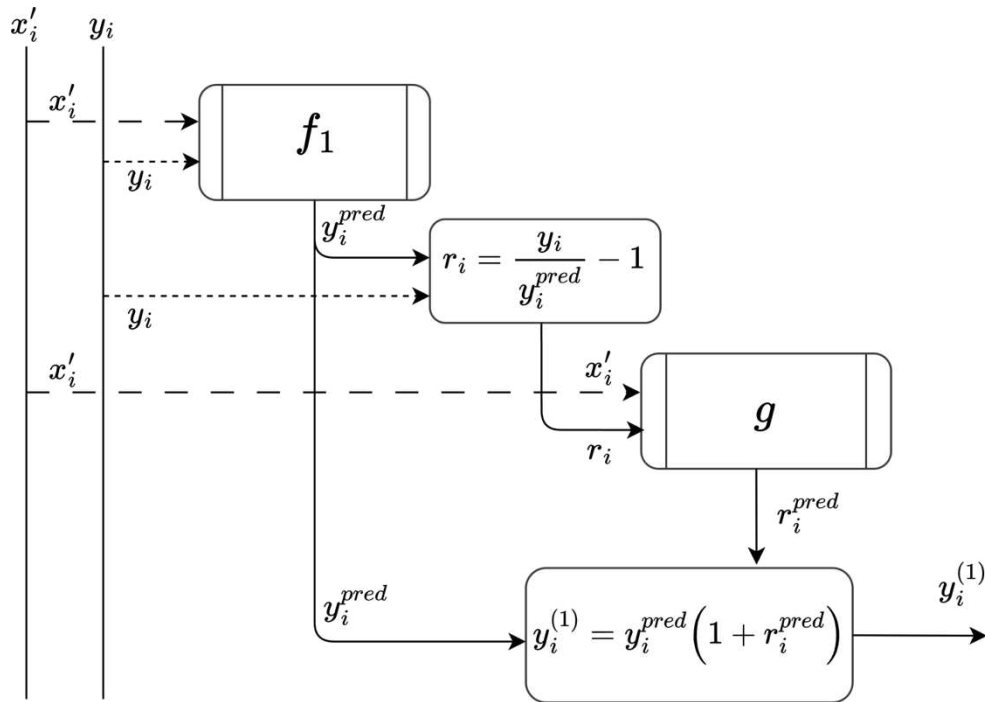


Рисунок 5.5 – Структурно-функційна схема зворотного методу компенсації похибки моделі регресії на основі використання раціональних дробів

3. Оцінка відносної корекції $r_u^{pred} = g(u)$ для u -го вектора шляхом його застосування на другій попередньо навченій ШНМ:

$$r_u^{pred} = g(u). \quad (5.54)$$

4. Корекція результатів для u -го вхідного вектора за формулою:

$$y_u^{(1)} = y_u^{pred} (1 + r_u^{pred}). \quad (5.55)$$

У випадку, коли точність корекції не відповідає вимогам користувача, її можна підвищити за рахунок можливості ітеративного процесу компенсації похибки моделі в процесі інтелектуального аналізу коротких наборів даних. Для цього $y_i^{(1)}$ з (5.55) підставляємо замість y_i^{pred} з (5.49) і виконуємо ітераційний процес $y_i^{(iter)}$, $iter = 2, 3, \dots$ компенсації похибки моделі регресії відповідно до обраного користувачем критерію зупинки.

Структурно-функційну схему такого каскадного ансамблю подано на рис. 5.6.

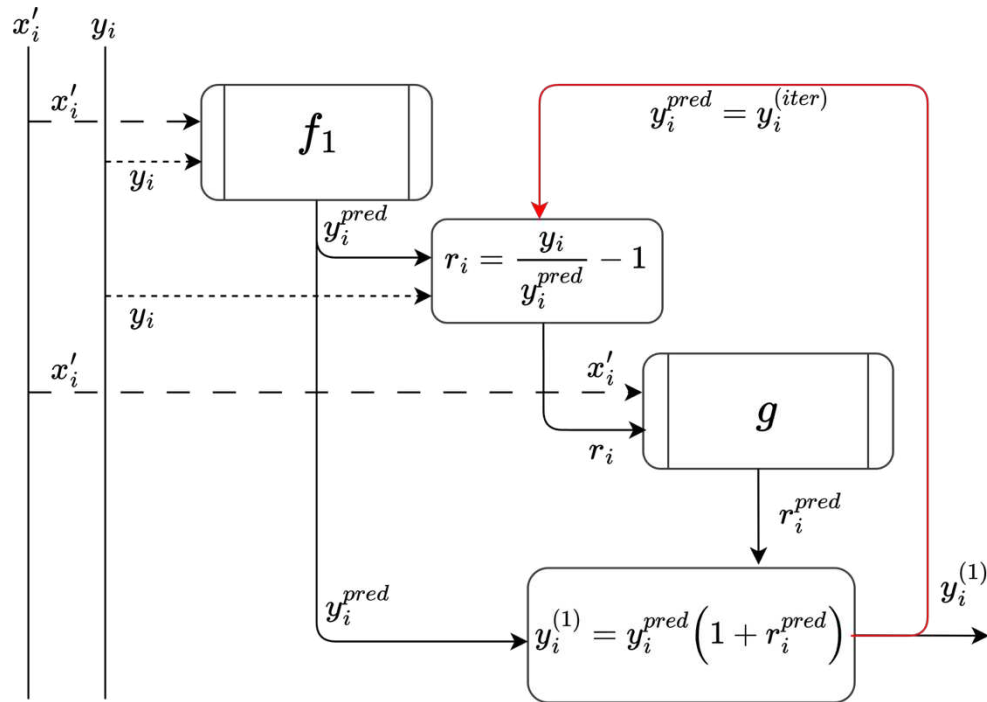


Рисунок 5.6 – Структурно-функційна схема зворотного методу компенсації похибки моделі регресії з можливістю ітеративного підвищення точності розв’язання задач регресії в умовах обмежених обсягів даних

Такий підхід забезпечує можливість підвищення робастності та точності апроксимації за рахунок зменшення залишкової похибки в кожній наступній ітерації.

Ряд основних критеріїв зупинки для ітераційної процедури зворотного методу компенсації похибки моделі регресії, які орієнтовано на зменшення залишкової похибки та стабілізацію коригувальної моделі, подано у таблиці В.2 додатку В.

Зворотний метод компенсації похибки моделі регресії з використанням раціональних дробів поєднує кілька важливих переваг [37]:

- він забезпечує більш ефективну компенсацію мультиплікативної похибки. Якщо в прямому методі (підрозділ 5.4.1) уточнення реалізується через додавання коригувального члена в знаменнику (5.44), то зворотний метод моделює початкову похибку як мультиплікативну поправку до

результату апроксимації (5.50)–(5.53). Такий підхід дозволяє природно масштабувати уточнення пропорційно результату апроксимації, і краще враховувати систематичні відхилення, особливо коли похибка зростає разом з величиною y_i^{pred} ;

- він забезпечує прозорість і лінійність моделі уточнення, а також кращу збіжність у порівнянні з дробовими виразами прямого методу. Навчальна вибірка для другої моделі формується як набір векторів виду $A^+ = (x_i, r_i) \in \mathbb{R}^{n \times (m+1)}$ з (5.51). Це дозволяє використовувати стандартні регресійні лінійні або нелінійні методи. На відміну від прямого методу, де функція f_2 оперувала в знаменнику дробу, тут усі ознаки входять до моделі в чисельнику, що спрощує процес навчання та покращує збіжність методу;
- завдяки масштабуванню ознак первинним результатом апроксимації зворотний метод демонструє підвищену робастність на малих і зашумлених вибірках. Оскільки шум та випадкові флуктуації менше впливають на множники в чисельнику, коригувальна модель $\phi(A^+)$ навчається на більш однорідній цільовій змінній. Це знижує ризик перенавчання і покращує властивості узагальнення методу;
- метод зберігає повну автономність першої ШНМ $f_1(\cdot)$. Усі операції другого рівня опираються лише на вхідні вектори без потреби розкривати його внутрішню архітектуру або параметри. Це робить його універсальним доповненням до будь-якої існуючої моделі регресії;
- на відміну від ряду існуючих методів розширення ознак [285] зворотний підхід не потребує збільшення вимірності простору ознак, працюючи у $\mathbb{R}^m$. Це забезпечує його ефективність у задачах, де кількість ознак m порівняна з кількістю спостережень n , що є типовою ситуацією в біомедичних задачах;

- каскадна структура зворотного методу компенсації похибки моделі регресії легко масштабується та може повторно уточнювати результати. Це забезпечує можливість послідовно зменшувати залишкову похибку, що дозволяє добитися потрібного рівня точності навіть за обмежених обсягів даних. Така гнучкість методу забезпечує можливість його використання в процесі розв'язання різноманітних задач аналізу зашумлених біомедичних даних малих обсягів.

5.5 Висновки до розділу

1. У розділі проаналізовано ключові причини зниження точності роботи штучних нейронних мереж, що обумовлені обмеженістю вибірки, наявністю шуму чи складністю моделі. Обґрунтовано необхідність розроблення та вдосконалення методів часткової компенсації похибок ШНМ, що забезпечують підвищення точності моделювання. Запропоновано низку методів компенсації похибок вихідного сигналу ШНМ на основі каскадних ансамблів, які дозволять підвищити точність апроксимації, стабільність результатів та робастність моделі в процесі аналізу зашумлених біомедичних даних малих обсягів.

2. Використання каскаду двох нейронних мереж узагальненої регресії дозволяє забезпечити ефективну компенсацію систематичної похибки первинної моделі. Це обґрунтовано на основі аналізу залишкових похибок та можливості їх моделювання. Серед основних переваг розробленого методу слід відзначити можливість використання різних підходів до корекції результатів апроксимації, зокрема, сумування виходів двох нейронних мереж, зваженого сумування з застосуванням коректуючої нейроподібної структури МПГП та нелінійного розширення вихідних сигналів обох НМУР за допомогою полінома Колмогорова-Габора. Це дозволить забезпечити додаткову гнучкість моделі та дозволить урахувати особливості у даних і характер похибок. Крім того, розроблений метод

підтримує можливість ітеративного уточнення результатів апроксимації, що забезпечує підвищення точності в умовах нестабільних даних.

3. Удосконалений метод глобально локальної апроксимації, в основі якого покладено можливість розділення сигналу на глобальну та локальну компоненти, надає можливість підвищення точності за рахунок незалежного моделювання кожної компоненти двома окремими НМУР ансамблю. Такий підхід дозволить більш ефективно моделювати як загальні закономірності у даних обмежених обсягів, так і локальні особливості, що є критично важливим для аналізу біомедичних вибірок. Удосконалений метод дозволить знизити загальну дисперсію результату без значного збільшення систематичного зсуву, що досягається завдяки узгодженій роботі двох нейронних мереж узагальненої регресії ансамблю.

4. Прямий та зворотний методи компенсації похибки моделі регресії на основі раціональних дробів можуть забезпечити підвищення точності апроксимації за рахунок використання корекції, яка інваріантна до масштабних змін, а також можливості ітеративного уточнення результатів. Прямий підхід дозволяє адаптивно зменшувати похибку шляхом корекції прогнозованого значення залежно від його величини. Це забезпечить підвищену робастність моделі навіть у присутності значних систематичних відхилень. Зворотний підхід, завдяки використанню зворотного множника, забезпечить ефективну компенсацію мультиплікативної похибки та покращену збіжність алгоритму. Це дозволить моделі доволі гнучко враховувати особливості вхідних даних і забезпечить послідовне зменшення залишкової похибки.

РОЗДІЛ 6

ВЕРИФІКАЦІЯ РОЗРОБЛЕНИХ АНСАМБЛЕВИХ МЕТОДІВ НА МОДЕЛЬНИХ ДАНИХ

У розділі подано результати верифікації розроблених нами та протестованих на модельних даних ряду методів. Зокрема, продемонстровано ефективність поліноміального розширення входів за допомогою членів полінома Колмогорова-Габора, що суттєво підвищує точність результатів і запобігає перенавчанню. Показано, що використання удосконаленого методу класифікації на основі формування комбінованої поверхні відгуку з використанням кластерного аналізу забезпечує високу точність результатів при збереженні низької обчислювальної складності. Продemonстровано, що використання методів розв'язання задач регресії/класифікації на основі декартового квадрата множини вхідних компонент векторів демонструє високу ефективність в процесі аналізу екстремально коротких вибірок даних. Встановлено, що метод формування аугментованих вибірок зменшеного обсягу з використанням кластеризації дозволяє зменшити обчислювальну складність усіх методів цього класу без значної втрати точності. Це відкриває широкі перспективи щодо подальшого застосування цих методів для розв'язання задач у різних напрямках та спеціалізаціях біомедичної інженерії.

Розроблені програмні засоби, які використовуються в цьому розділі для проведення експериментальних досліджень, захищено авторським правом на комп'ютерні програми [54], [55], [56]. Отримані результати досліджень опубліковано у розділі колективної монографії [6], високореєтингових журналах інших держав, які включено до наукометричних баз Scopus та Web of Science: [14], [34], а також апробовано і опубліковано у матеріалах українських та закордонних конференцій: [18], [21], [27], [36], [49], [57].

6.1 Експериментальне дослідження методів нелінійного перетворення вхідних компонент векторів для формування простору ознак вищої вимірності

У цьому підрозділі подано результати моделювання розвинутих та удосконалених методів нелінійного перетворення вхідних компонент векторів та порівняння їх ефективності з ефективністю існуючих. Дослідження виконувалися з використанням авторського програмного забезпечення, написаного мовою Python. Такий підхід забезпечив можливість реалізації авторських методів та легкої інтеграції існуючих методів машинного навчання чи ШНМ з доступних бібліотек, розроблених на цій мові. На розроблені програмні засоби отримано свідоцтва про авторське право на комп'ютерну програму (додаток Д).

6.1.1 Оцінювання результатів методу нелінійного доповнення входів нейроподібних структур МПГП членами полінома Колмогорова-Габора

Метод нелінійного розширення входів лінійної неітеративної нейроподібної структури МПГП (Метод 2.1) протестовано на задачі у сферах медичної економіки та управління охороною здоров'я.

Задача. Метод 2.1 використано в процесі розв'язання задачі оцінки вартості особистого медичного страхування пацієнта. Ефективний розв'язок цієї задачі має значну практичну цінність, оскільки дозволяє оптимізувати витрати на медичні послуги, забезпечуючи більш ефективне управління ресурсами медичних установ і зниження фінансового навантаження на пацієнтів. Це забезпечує також можливість розроблення персоналізованих підходів до лікування, що у свою чергу зменшує витрати на надання медичних послуг і підвищує їх доступність. Крім того, високоточний результат розв'язання цієї задачі зменшує фінансові ризики для пацієнтів та підвищує загальну ефективність медичної допомоги, що є важливим для розвитку та удосконалення технологій управління охороною здоров'я.

Опис та попереднє опрацювання вибірки даних. Початкова вибірка даних з [286] для розв’язання поставленої задачі містить 1338 спостережень, кожне з яких включає шість вхідних атрибутів та одну вихідну змінну – індивідуальні витрати на медичне страхування. Вхідні атрибути включають вік страхувальника, стать, індекс маси тіла, кількість утриманців, статус куріння та місце проживання бенефіціара в США. Вік страхувальників варіюється від 18 до 64 років, середній вік складає 39,2 роки. Індекс маси тіла в середньому становить 30,66, що вказує на надмірну вагу. Кількість утриманців коливається від 1 до 5. У вибірці є 1064 курці та 274 некурці. Місце проживання бенефіціара поділяється на чотири регіони США: північний захід, північний схід, південний захід та південний схід.

Для попереднього опрацювання даних виконано перетворення текстових змінних у числову форму за допомогою бінарного кодування. Зокрема, стать страхувальника та статус курець/не курець були подані двома окремими стовпцями (чоловік/жінка та курець/некурець). Стовпець місця проживання бенефіціара був перетворений на чотири окремі стовпці для кожного з регіонів США: південний захід, південний схід, північний захід та північний схід. Після виконання усіх процедур попереднього опрацювання фінальна вибірка даних містить 11 числових вхідних атрибутів.

Моделювання та налаштування параметрів методу. Моделювання реалізовано з використанням полінома Колмогорова-Габора другого степеня (розділ 2) та лінійної неітеративної нейроподібної структури МПГП [19, розділ 1]. Використання полінома Колмогорова-Габора другого степеня для нелінійного розширення входів МПГП зумовлює ріст кількості незалежних атрибутів, яка обчислюється за формулою (2.7).

Необхідною оцінкою роботи запропонованого методу є показник співвідношення складності моделі до точності її роботи. Складність моделі в даному випадку визначається двома параметрами: ступенем полінома Колмогорова-Габора (саме тому було обрано поліном другого ступеня) і кількістю

нейронів прихованого шару лінійної нейроподібної структури МПГП. Саме останній є єдиним параметром цього методу, який слід підібрати.

На рис.6.1 наведено динаміку значень показників ефективності (похибки МАРЕ, час навчання) при зміні кількості нейронів прихованого шару лінійної нейроподібної структури МПГП на проміжну [25:50] з кроком п'ять.

Проведені експериментальні дослідження показали, що оптимальний результат методу досягається при 35 нейронах у прихованому шарі (рис. 6.1а). Подальше збільшення кількості нейронів прихованого шару лінійної нейроподібної структури МПГП не виглядає доречним, оскільки дещо зменшує точність роботи методу, його властивості узагальнення, а також суттєвим чином збільшує тривалість його роботи (рис. 6.2б, таблиця 6.1).

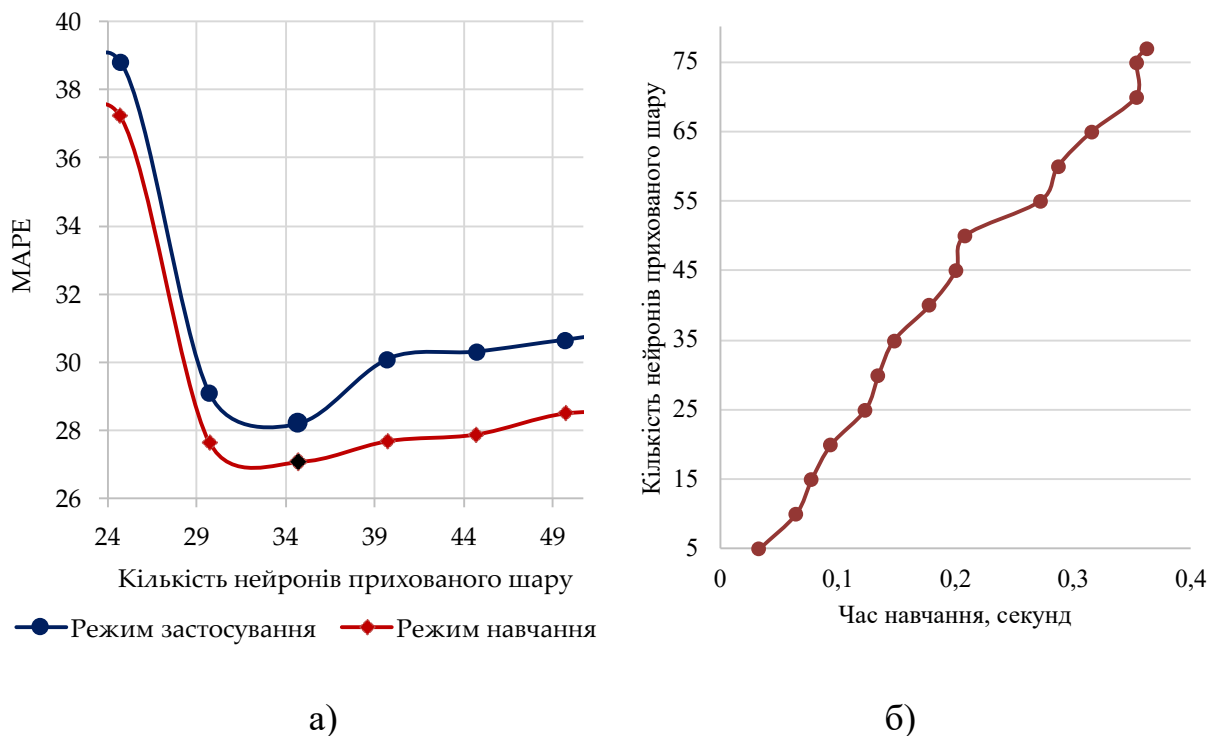


Рисунок 6.1 – Результати підбору оптимальної кількості нейронів прихованого шару лінійної нейроподібної структури МПГП на основі: а) похибки роботи методу (МАРЕ); б) тривалості процедури навчання методу

Результати. Результати Методу 2.1 до та після оптимізації на основі

показників з Додатку А та значення оптимальних параметрів його роботи подано у таблиці 6.1.

Таблиця 6.1 – Результати роботи методу нелінійного розширення входів лінійної неітеративної нейроподібної структури МПГП (Метод 2.1.) в процесі розв’язання поставленої задачі

Режим роботи методу / Показник ефективності	MAPE	RMSE	MAE	Тривалість навчання, секунд
<i>до оптимізації</i>				
Режим навчання	28.4187	4639.78	2767.61	0.362
Режим застосування	30,823	5050.3	3099.7	-
<i>після оптимізації</i>				
Режим навчання	27.0794	4696.1	2783.14	0.148
Режим застосування	28,203	5029.5	3077.6	-
Значення оптимальних параметрів роботи методу	<i>Степінь полінома Колмогорова-Габора: 2</i> <i>Базова ШНМ для реалізації процедури навчання:</i> Лінійна неітеративна нейроподібна структура МПГП. <i>Кількість нейронів прихованого шару: 35.</i>			

Як видно з результатів, поданих у таблиці 6.1, оптимізована версія Методу 2.1 демонструє вищу точність роботи та узагальнюючі властивості в порівнянні з неоптимізованою, і, відповідно, суттєво меншу (майже у 2.5 разів) тривалість методу навчання.

Порівняння. Оцінювання ефективності Методу 2.1 виконувалося шляхом порівняння його точності з точністю за існуючими лінійними методами.

На рис. 6.2 подано порівняння похибок на основі MAE (додаток А) усіх досліджуваних методів. Як видно з рис. 6.2, найменшу точність демонструє регресор на основі Мащини опорних векторів з лінійним ядром. Суттєво кращу точність отримано з використанням лінійної нейроподібної структури МПГП. Проте найменшу похибку, майже у 1,4 рази демонструє розроблений метод нелінійного розширення входів лінійної неітеративної нейроподібної структури з використанням членів полінома Колмогорова-Габора другого степеня. Окрім цього, результат розв’язання поставленої задачі можна представити у поліноміальній формі [20], синтезувавши коефіцієнти при членах полінома методом діагональної матриці [43].

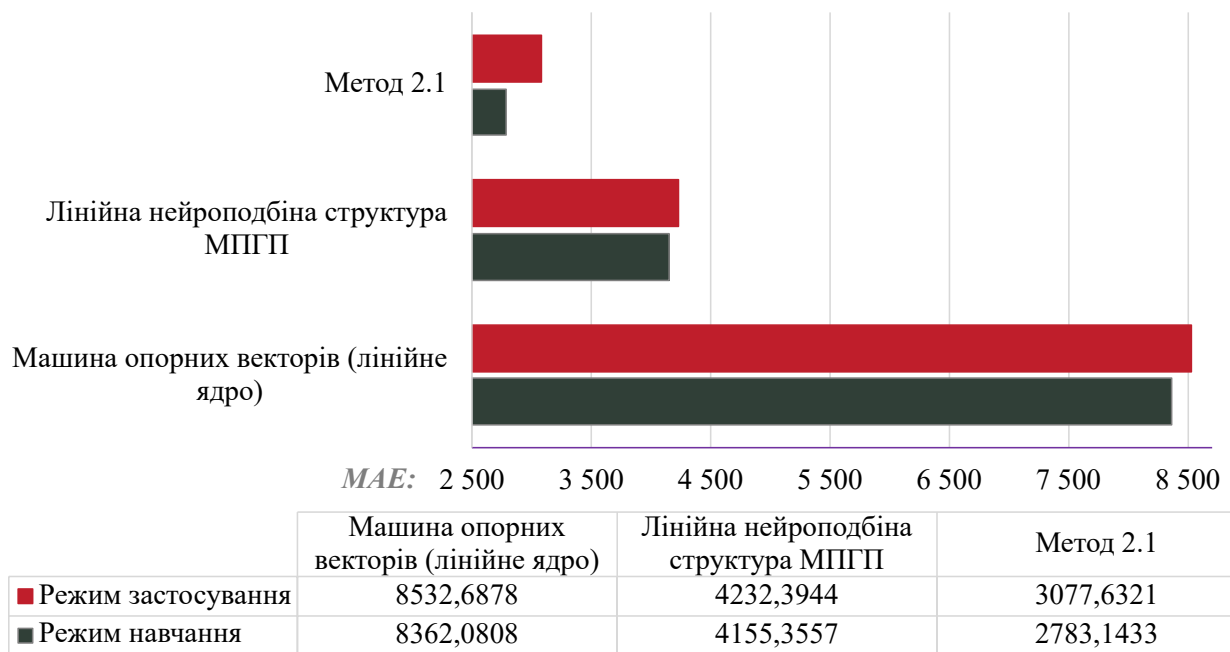


Рисунок 6.2 – Порівняння похибок (MAE) Методу 2.1 з похибками за існуючими методами

Такий підхід забезпечує ряд переваг в процесі використання розробленого методу для розв’язання різноманітних прикладних задач в галузі біомедичної інженерії.

6.1.2 Оцінювання результатів методу класифікації даних на основі формування комбінованої поверхні відгуку в області кардіології

Метод класифікації даних на основі формування комбінованої поверхні відгуку з використанням кластерного аналізу було протестовано на прикладі задачі в галузі кардіології. Розроблено спеціалізоване програмне забезпечення мовою Python, що забезпечує реалізацію авторських підходів та інтеграцію з наявними методами машинного навчання і штучних нейронних мереж, доступними у відкритих бібліотеках. На створене програмне забезпечення отримано свідоцтва про реєстрацію авторського права на комп'ютерну програму (додаток Д).

Задача та її практична цінність. Удосконалений метод з розділу 2 (метод 2.4.1) використано для розв'язання задачі класифікації потенційного ризику серцевого нападу на основі різноманітних медичних даних пацієнта. Ефективне розв'язання цієї задачі, зокрема, методами машинного навчання дозволить підвищити точність та швидкість виявлення потенційних пацієнтів з високим ризиком серцевого нападу. Практичне значення результатів використання такого підходу полягає у тому, що вони дозволять: стати додатковим джерелом інформації для лікарів. Це зменшить ймовірність лікарських помилок в процесі діагностування; здійснити раннє втручання для підбору вірного підходу до профілактики та лікування пацієнта. Це значно знизить ймовірність важких ускладнень чи летальних випадків; персоналізувати лікування, адаптуючи стратегії до індивідуальних ризиків пацієнтів. Це зробить процес профілактики чи лікування більш ефективним.

Опис та попереднє опрацювання вибірки даних. Для моделювання використано вибірку даних малого обсягу з [287]. Вибірka складається з 302 спостережень, які характеризуються 14 незалежними атрибутами, включаючи вік, стать, тип болю в грудях, артеріальний тиск в процесі відпочинку, рівень холестерину в сироватці крові, рівень глюкози в крові натще, результати ЕКГ у спокої, максимальну частоту серцевих скорочень, ангіну, викликану фізичним

навантаженням, депресію ST-сегмента, індуковану фізичним навантаженням відносно стану спокою, нахил пікового ST-сегмента, кількість основних судин, пофарбованих за допомогою флуороскопії, тип дефекту (нормальний, фіксований, оборотний). Залежний атрибут – діагноз серцево-судинного захворювання (бінарна класифікація). Досліджувана вибірка даних містить чималу кількість пропусків, категорійні змінні тощо. Після виконання процедур попередньої обробки, які включали видалення стовпців, що містять багато пропусків, перетворення категоріальних ознак у числові, а також нормалізацію з використанням `MaxAbsScaler` [50], кінцевий набір даних містив 18 незалежних атрибутів і 294 спостереження.

Моделювання та налаштування параметрів методу. Моделювання методу виконувалося з використанням п'ятикратної кросвалідації. Класифікаторами першого та другого рівнів удосконаленого методу були дві лінійні нейроподібні структури МППП з неітеративним методом навчання [19].

Для кластеризації даних використовувався метод `k-means`. Такий вибір обумовлено як простотою, так і швидкістю його роботи. Пошук оптимальної кількості кластерів методу `k-means` виконувався з використанням двох різних методів. Вони продемонстрували різні результати (від трьох до семи кластерів) [4], [41]. Саме тому в роботі проведено експерименти щодо впливу різної кількості кластерів на точність роботи удосконаленого методу на проміжку [3:7] з кроком 1. Результати цього експерименту наведено на рис. 6.3.

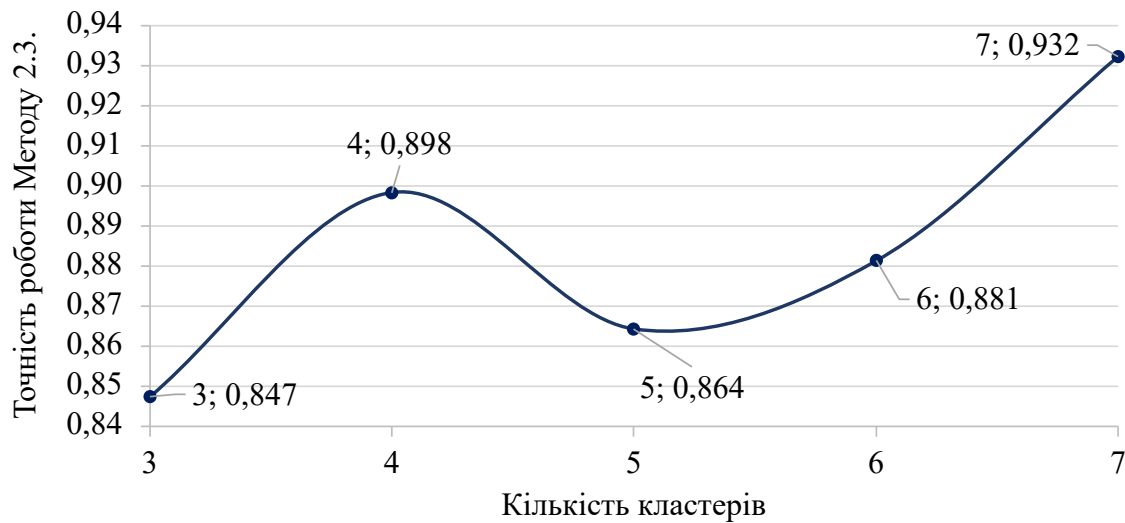


Рисунок 6.3 – Зміна точності роботи удосконаленого методу при зміні кількості кластерів методу k-means на проміжку [3:7] з кроком 1

Як видно з рис. 6.3, найвищу точність роботи удосконаленого методу отримано при використанні семи кластерів.

Результати. Результати удосконаленого методу подано у таблиці 6.2. Слід зазначити, що у таблиці 6.2 зведено точність роботи удосконаленого методу у режимах як навчання, так і застосування для двох типів кластеризації: з використанням значення залежного атрибуту (2.33) та без його використання (2.35).

Як видно з результатів, поданих у таблиці 6.2, удосконалений метод демонструє високі властивості узагальнення. Окрім цього, використання (2.35) забезпечило суттєве підвищення точності роботи удосконаленого методу в порівнянні з використанням (2.33) при мінімальній зміні тривалості процедури його навчання.

Таблиця 6.2 – Результати методу класифікації даних на основі формування комбінованої поверхні відгуку в процесі розв’язання поставленої задачі

Режим роботи методу / Показник ефективності	Точність	Тривалість процедури навчання (секунд)
<i>Класифікація виконувалася без використання вихідного атрибуту</i>		
Режим навчання	0.872	0.10
Режим застосування	0.847	-
<i>Класифікація виконувалася з використанням вихідного атрибуту</i>		
Режим навчання	0.948	0.11
Режим застосування	0.932	-
Значення оптимальних параметрів роботи методу	<i>Метод класифікації: k-means.</i> <i>Тип класифікації: з використанням вихідного атрибуту.</i> <i>Оптимальна кількість кластерів: 7.</i> <i>Базова ШНМ для реалізації процедури навчання на першому рівні (ШНМ-1): Лінійна неітеративна нейроподібна структура МПГП.</i> <i>Кількість нейронів прихованого шару ШНМ-1: 18.</i> <i>Базова ШНМ для реалізації процедури навчання на другому рівні (ШНМ-2): Лінійна неітеративна нейроподібна структура МПГП.</i> <i>Кількість нейронів прихованого шару ШНМ-2: 8.</i>	

Порівняння. Оцінювання точності удосконаленого методу виконувалося шляхом порівняння з точністю за існуючим методом з [229], а також з лінійним

класифікатором на основі базової лінійної неітеративної нейроподібної структури, яку покладено в основу реалізації обох методів.

На рис. 6.4 подано результати порівняння значень точності роботи усіх досліджуваних методів.

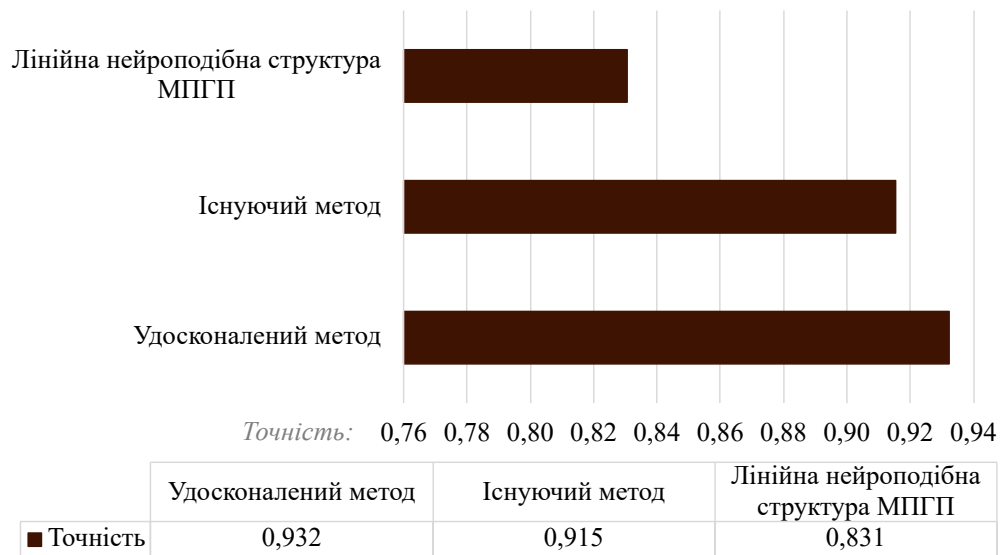


Рис. 6.4 – Порівняння точності роботи досліджуваних методів

Як видно з рис. 6.4, найменшу точність продемонстрував лінійний класифікатор на основі неітеративної нейроподібної структури МПГП. Значно кращу точність отримано з використанням існуючого методу з [229]. Проте найвищу точність отримано з використанням удосконаленого методу класифікації даних на основі формування комбінованої поверхні відгуку з використанням кластерного аналізу (Метод 2.4.1 з розділу 2). Більше того, він забезпечив суттєве скорочення кількості незалежних атрибутів задачі і зменшив тривалість процедури навчання фінального класифікатора. Зменшення кількості незалежних атрибутів задачі потенційно збільшує властивості узагальнення удосконаленого методу в порівнянні з існуючим. Окрім цього він забезпечує можливість аналізу коротких високовимірних вибірок даних, що у випадку використання існуючого методу з

[229] не завжди є можливим через суттєве зростання вимірності вхідних даних задачі.

6.2 Експериментальне дослідження методів розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів

У цьому підрозділі подано результати обчислювальних експериментів та порівняння ефективності розроблених методів з ефективністю за існуючими підходами. Експериментальне дослідження ефективності усіх методів цього класу виконувалося за допомогою розробленої на мові Python бібліотеки методів аналізу екстремально коротких вибірок даних. Розроблені методи передбачають можливість використання як нейроподібних структуру МПГП з РБФ-розширенням входів так і існуючих методів машинного навчання та ШНМ, які вільно доступні в інших бібліотеках. ЦЕ розширює потенційні можливості використання програмного засобу для реалізації усіх методів з третього розділу. Розроблене програмне забезпечення, захищене свідоцтвом про авторське право на комп’ютерну програму [54] (додаток Д).

6.2.1 Оцінювання результатів базового методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів

Базовий метод розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів з пункту 3.2 (Метод 3.1) протестований на штучній вибірці даних.

Генерація штучної вибірки даних. Для моделювання цього методу згенеровано штучну вибірку даних. Вона містить два незалежні атрибути та один залежний. Для її генерації використано нелінійну функцію виду:

$$F(a, b) = (1 - a^2) + 2(1 - b)^2, \quad (6.1)$$

і синтезовано 30 спостережень, кожне з яких генерувалося при випадкових значеннях a, b на проміжку $[-1, 1]$, та для яких обчислено значення цільової функції (6.1).

Моделювання та налаштування параметрів методу. Моделювання методу реалізовано з використанням п'ятикратної кросвалідації. Такий підхід забезпечує отримання більш надійної та об'єктивної оцінки ефективності моделі, особливо у випадку аналізу малих даних. Окрім цього він дозволяє визначити наскільки добре модель працює на різних підмножинах даних, а також допомагає зрозуміти її здатність узагальнювати на нових даних.

В якості базової нейронної мережі для реалізації навчання методу використано неітеративну нейроподібну структуру Моделі Послідовних Геометричних Перетворень з РБФ-розширенням входів [288], описаної у першому розділі. Підбір оптимальних параметрів роботи методу (кількість РБФ- центрів та значення розмаху гаусівської функції), які по суті є параметрами нейроподібної структури МПГП з РБФ-розширенням входів, виконувався з використанням методу диференційної еволюції.

Результати. Результати розробленого методу з використанням штучної вибірки даних на основі метрик з Додатку А та значення оптимальних параметрів його роботи подано у таблиці 6.3.

Як видно з результатів, поданих у таблиці 6.3, розроблений метод демонструє високу точність роботи, високі властивості узагальнення та повністю уникає проблеми перенавчання, яка властива існуючим методам в процесі аналізу екстремально коротких вибірок даних.

Порівняння. Оцінювання ефективності розробленого методу проведено шляхом порівняння його результатів з результатами за кількома існуючими методами. Для цього було використано:

- нейроподібну структуру МПГП з РБФ-ядром [288] (значення оптимальних параметрів: 10 центрів рбф, значення розмаху гаусівської функції = 1.2) ;

- нейронну мережу узагальненої регресії [157] (значення оптимального параметра: значення розмаху гаусівської функції = 0.4);
- лінійну неітеративну нейроподібну структуру МПГП [19] (значення оптимальних параметрів: кількість нейронів вхідного і прихованого шарів становить два).

Таблиця 6.3 – Результати роботи методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів (Метод 3.1)

Режим роботи методу / Показник ефективності	MAPE	RMSE	MAE	R ²
Режим навчання	0,041	0,0021	0,001	0,99
Режим застосування	0,043	0,0023	0,001	0,99
Значення оптимальних параметрів роботи методу	<i>Базова ШНМ для реалізації процедури навчання: Неітеративна нейроподібна структура МПГП з РБФ-ядром. Кількість РБФ-центрів: 29. Значення розмаху гаусівської функції: 9.46. Метод оптимізації – диференційна еволюція.</i>			

На рис. 6.5 подано порівняння похибок роботи (на основі RMSE) різних досліджуваних методів.

Як видно з рис. 6.5, лінійний метод з [19] демонструє найбільшу похибку в процесі аналізу заданої екстремально-короткої вибірки даних. Дещо кращі результати показують нейромережеві методи з [288] та [157]. Проте найкращі результати отримано саме в процесі використання Методу 3.1, який демонструє в 150 разів менше значення похибки RMSE за найбільш точний метод з [288] серед досліджуваних.

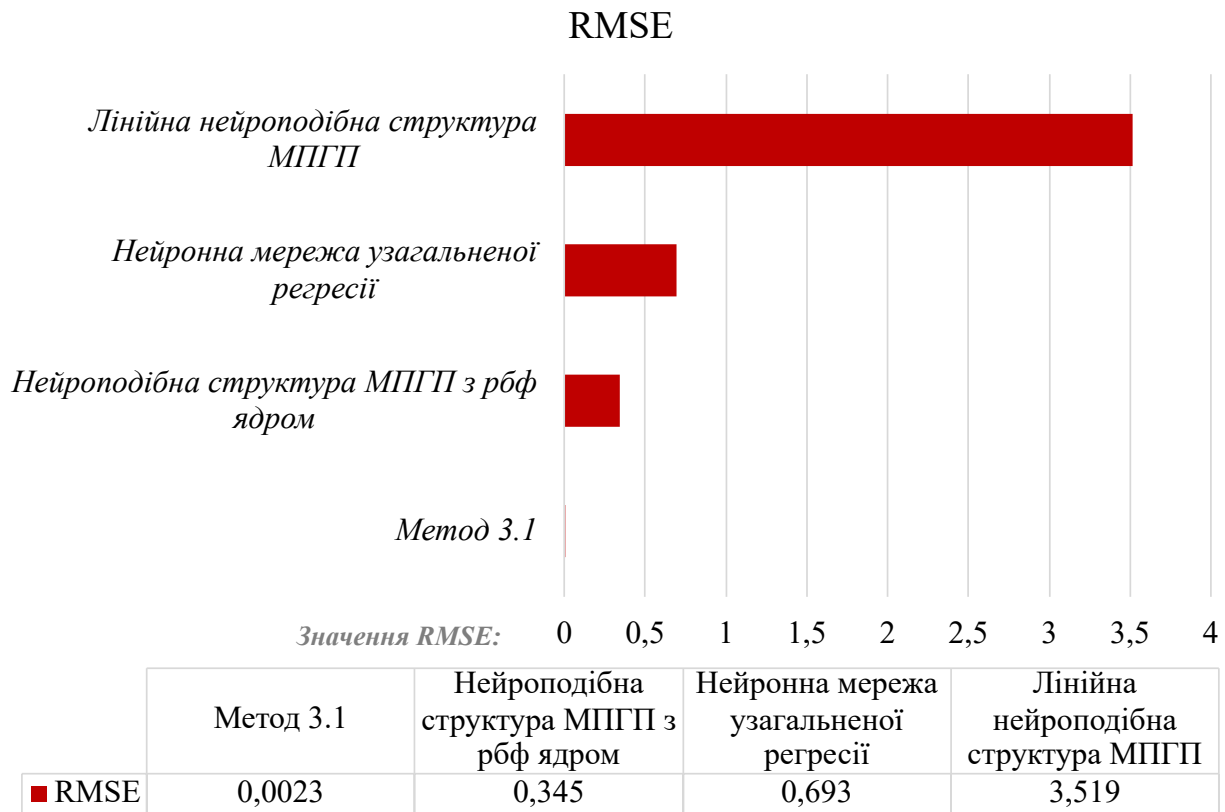


Рисунок 6.5 – Порівняння похибки (RMSE) методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів (Метод 3.1) з похибками за існуючими методами

Зазначені висновки підтверджуються з використанням усіх інших показників з Додатку А.

6.2.2 Оцінювання результатів адитивного методу розв’язання задачі регресії в галузі ревматології

Адитивний метод розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів з пункту 3.3 третього розділу роботи (Метод 3.2) протестовано в процесі розв’язання задачі в галузі ревматології.

Задача та її практична цінність. Метод 3.2 використано в процесі розв’язання задачі оцінювання міцності трабекулярної кістки при стиску (в МПа) у

пацієнтів, уражених остеоартрозом. Це дозволяє визначити максимальну силу, яку кісткова тканина може витримати до початку її деформації або руйнування в процесі стискання. Рішення цієї задачі надає важливу інформацію для оцінки механічних властивостей кісткової тканини, що є критично важливим для медичних досліджень і клінічної практики. Отримані результати можуть бути використані для більш точної оцінки ризику переломів, розробки персоналізованих методів лікування, оцінюванні ефективності терапії, вдосконалення методів діагностики та лікування остеоартрозу, а також для створення нових матеріалів для імплантів, що підвищать їх ефективність та довговічність.

Опис вибірки даних. Вибірка даних з відкритого джерела [86] містить 35 зразків параметрів для 17 чоловіків та 18 жінок різного віку (вік і стать – два незалежні атрибути). Вона отримана від пацієнтів, які перенесли тотальне ендопротезування кульшового суглоба через важкий остеоартроз.

Дані включають характеристики трабекулярної тканини з головки стегнової кістки, відібрані з основної компресійної ділянки. Для кожного зразка авторами вибірки даних було проведено серію експериментів, включаючи мікрокомп'ютерну томографію (мікроКТ) з роздільною здатністю пікселів 19-5 мкм та деформаційні випробування за допомогою екстензометра. Для оцінки структурних характеристик були розраховані пористість тканини, фактор трабекулярної товщини та індекс структурної моделі, який використовується для характеристики морфології кісткової тканини (ще три незалежні атрибути). Усі вимірювання забезпечували високу точність, з похибкою не більше 0-7%. Таким чином, вибірка даних, яку подано у [86], містила п'ять незалежних атрибутів і один залежний.

Моделювання та налаштування параметрів методу. Моделювання методу реалізовано з використанням п'ятикратної кросвалідації. Серед процедур попереднього опрацювання відмітимо використання лише нормалізацію даних. Вона виконувалася шляхом поділу кожного значення певної ознаки на максимальне

за абсолютною величиною значення цієї ознаки, зберігаючи при цьому знак значень і переводячи їх у діапазон від -1 до 1 [44].

Зважаючи на квадратичне збільшення обсягу вибірки для навчання, відповідно до Методу 3.2, для реалізації методу навчання використано неітеративну нейроподібну структуру МПГП з РБФ-розширенням входів [288]. Підбір оптимальних параметрів методу, а саме кількість РБФ-центрів та значення розмаху гаусівської функції виконувався з використанням методу диференційної еволюції.

Результати. Результати розробленого методу (Метод 3.2) на основі показників з Додатку А та значення його оптимальних параметрів подано у таблиці 6.4.

Таблиця 6.4 – Результати адитивного методу в процесі розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів (Метод 3.2)

Режим роботи методу / Показник ефективності	MAPE	RMSE	MAE	R ²
Режим навчання	14,56	1,89	1,533	0,93
Режим застосування	17,87	1,95	1,574	0,91
Значення оптимальних параметрів роботи методу	<i>Базова ШНМ для реалізації процедури навчання: Неітеративна нейроподібна структура МПГП з РБФ ядром. Кількість РБФ-центрів: 11. Значення розмаху гаусівської функції: 1.1. Метод оптимізації – диференційна еволюція.</i>			

Як видно з результатів, поданих у таблиці 6.4, Метод 3.2 демонструє високу точність роботи та високі властивості узагальнення в процесі розв’язання задачі

оцінки міцності трабекулярної кістки при стиску (в МПа) у випадку аналізу екстремально-короткої вибірки даних.

Порівняння. Оцінку точності Методу 3.2 здійснено шляхом порівняння його результатів з точністю реалізації Методів 3.1 та 3.2, які ґрунтуються на використанні ітеративних методів машинного навчання або нейронних мереж і не вимагають процесу навчання:

- метод 3.2 з використанням нейронної мережі на основі радіально-базисних функцій для реалізації процедури навчання [2] (*Метод 3.2 на основі РБФ ШНМ*);
- метод 3.1 з використанням нейронної мережі на основі радіально-базисних функцій для реалізації процедури навчання [6] (*Метод 3.1 на основі РБФ ШНМ*);
- метод 3.1 з використанням нейроподібної структури МППП з РБФ-розширенням входів (*Метод 3.1*);
- метод 3.1 з використанням нейронної мережі узагальненої регресії (НМУР) [8] (*Метод 3.1 на основі НМУР*);
- метод 3.1 з використанням машини опорних векторів з РБФ-розширенням входів [7] (*Метод 3.1 на основі МОП (РБФ ядро)*);
- метод 3.1 з використанням машини опорних векторів з поліноміальним ядром [11] (*Метод 3.1 на основі МОП (поліноміальне ядро)*).

Окрім цього процедура порівняння включала використання існуючих штучних нейронних мереж та методів машинного навчання:

- нейронна мережа узагальненої регресії [14] (*НМУР*);
- нейронна мережа на основі радіально-базисних функцій (*РБФ ШНМ*);
- машина опорних векторів з РБФ-ядром (*МОВ (РБФ-ядро)*);
- машина опорних векторів з поліноміальним ядром (*МОВ (поліноміальне ядро)*).

Усі досліджувані методи і моделі, а також значення їхніх оптимальних параметрів зведено у таблиці Б.1 додатку Б.

На рис. 6.6 подано порівняння похибок на основі RMSE різних досліджуваних методів.

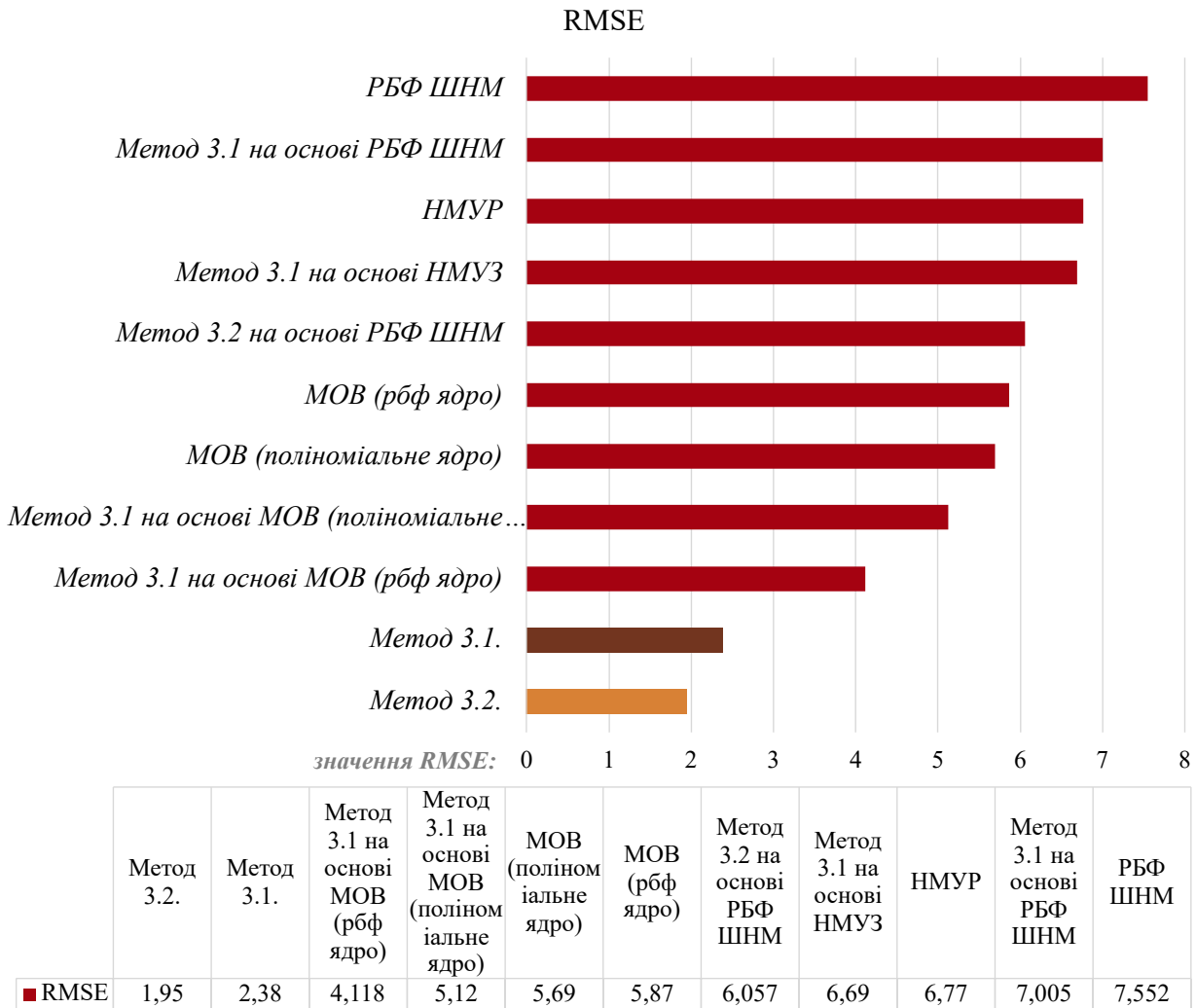


Рисунок 6.6 – Порівняння похибок (RMSE) адитивного методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів (Метод 3.2) з похибками за іншими методами

З результатів, представлених на рис. 6.6, можна зробити такі висновки. Найвищу похибку демонструє нейронна мережа на основі радіально-базисних

функцій (РБФ ШНМ), що обумовлено екстремально короткою вибіркою даних, яка не дозволяє ефективно використовувати ШНМ на практиці. Однак значно кращі результати отримані при застосуванні Методу 3.2 на основі РБФ ШНМ та Методу 3.1 на основі РБФ ШНМ. Зокрема, Метод 3.1 на основі РБФ ШНМ показав майже 8% покращення, а Метод 3.2 на основі РБФ ШНМ – майже 14% покращення в порівнянні з РБФ ШНМ. Це свідчить про потенційну можливість використання різних ШНМ ітеративного типу для навчання розробленого Методу 3.2. Незважаючи на це, точність роботи усіх трьох методів є незадовільною, що унеможливорює їх практичне використання в процесі розв’язання поставленої задачі.

Дещо кращі результати в процесі розв’язання поставленої задачі демонструє нейронна мережа узагальненої регресії (НМУР), зокрема, в порівнянні з РБФ ШНМ. Проте реалізація Методу 3.1 на основі НМУР не показала значного покращення точності (зменшення похибки RMSE лише 1.2%). Це свідчить про недоцільність використання інтерполяційних штучних нейронних мереж для реалізації Методу 3.1 і як наслідок – Методу 3.2.

Незважаючи на обмежений обсяг вибірки даних для реалізації навчання, використання машини опорних векторів з нелінійними ядрами (МОП з РБФ-ядром та МОВ з поліноміальним ядром) показало кращі результати за точність двох досліджуваних типів штучних нейронних мереж (РБФ ШНМ, НМУР). Метод 3.1 на основі МОП (РБФ-ядро) та Метод 3.1 на основі МОП (поліноміальне ядро) демонструють суттєве зменшення похибки RMSE (майже на 30% та майже на 13% відповідно). Це свідчить про потенційну можливість використання нелінійних ітеративних методів машинного навчання для реалізації навчання в Методах 3.1. та 3.2. Незважаючи на це, точність роботи усіх цих методів не є достатньою для розв’язання задачі оцінювання міцності трабекулярної кістки при стисканні на основі обмеженої вибірки даних.

Найкращі результати розв’язання поставленої задачі отримано з

використанням Методу 3.1 та Методу 3.2, які використовують неітеративну нейроподібну структуру з РБФ-розширенням входів для реалізації навчання. Зокрема, застосування цих методів дозволило зменшити значення похибки RMSE на 52% та на 42% відповідно в порівнянні з найближчим методом на рис. 3.11 (Метод 3.1 на основі МОВ з РБФ-ядром). Ці висновки підтверджуються також через аналіз інших похибок роботи досліджуваних методів.

Зважаючи на найнижче значення похибки RMSE Методу 3.2 серед усіх досліджуваних методів і моделей, високий коефіцієнт детермінації та відмінні властивості узагальнення цього методу (таблиця 6.4), він цілком підходить для розв'язання задачі оцінки міцності трабекулярної кістки при стиску у пацієнтів, які страждають на остеоартроз, за умов аналізу екстремально коротких вибірок даних.

6.2.3 Оцінювання результатів методу класифікації даних з використанням принципів каскадування в області трансплантології

Метод класифікації на основі декартового квадрата множини вхідних компонент векторів з застосуванням принципів каскадування методів машинного навчання, викладений у пункті 3.4 (Метод 3.3), протестовано в процесі розв'язання задачі в області трансплантології.

Задача та її практична цінність. Задача полягає у визначенні успіху або невдачі трансплантації стовбурових клітин у дітей на основі різних клінічних параметрів, таких як тип захворювання, сумісність груп крові, доза клітин CD34+ та інші фактори [289]. Аналіз цих даних дозволяє зрозуміти, як конкретні умови впливають на ефективність процедури трансплантації, а також на тривалість та якість життя пацієнтів після її проведення.

Високоточне розв'язання цієї задачі допоможе оптимізувати лікування, покращити результати трансплантацій і зменшити ризики ускладнень після операції. Окрім цього це дозволить лікарям точніше обирати стратегії лікування та

знизити ймовірність відторгнення трансплантата. Це, в свою чергу, підвищить ефективність лікування дітей, зменшить кількість ускладнень і покращить якість їх життя після процедури трансплантації.

Опис вибірки даних. Вибірка даних складається з 187 записів і 36 ознак і містить інформацію про дітей, які проходили процедуру трансплантації стовбурових клітин кісткового мозку. Усі пацієнти отримали трансплантацію від донорів, з якими не було генетичної спорідненості, тобто трансплантати були взяті від непов'язаних осіб. Процедура трансплантації проводилась без попереднього втручання або маніпуляцій з клітинами донора.

У пацієнтів були різні захворювання як злоякісні, наприклад, лейкози та мієлодиспластичний синдром, так і незлоякісні, такі як апластична анемія та анемія Фанконі. Злоякісні захворювання характеризуються наявністю ракових клітин, які активно ростуть і можуть поширюватися в організмі, тоді як незлоякісні захворювання не супроводжуються онкологічними змінами.

Основні характеристики даних включають інформацію про донорів і реципієнтів, сумісність груп крові, тип стовбурових клітин, що були використані для трансплантації, а також стан пацієнтів після процедури [290].

Моделювання та налаштування параметрів методу. Моделювання Методу 3.3 реалізовано з використанням п'ятикратної кросвалідації. Серед процедур попереднього опрацювання даних слід відмітити заміну пропусків у даних середнім значенням та використання нормалізації даних згідно з методом `MaxAbsScaler`, що масштабує ознаки до інтервалу $[-1; 1]$ без зміни знаку.

Процедура аугментації виконувалася з використанням ЙНМ як ШНМ першого рівня каскаду. Зважаючи на квадратичне збільшення обсягу вибірки для навчання, згідно Методу 3.3 для реалізації методу навчання використали неітеративну нейроподібну структуру МППП з РБФ-ядром [288]. Підбір оптимальних параметрів роботи цієї неітеративної ШНМ, а саме: кількість РБФ-

центрів та значення розмаху гаусівської функції, значення розмаху гаусівської функції для ЙМН виконувався з використанням методу диференційної еволюції.

Результати. Результати, отримані за допомогою розробленого методу (Метод 3.3), на основі різних показників з Додатку А зведені в таблиці 6.5.

Таблиця 6.5 – Результати методу класифікації даних на основі декартового квадрата множини вхідних компонент векторів з використанням принципів каскадування (Метод 3.3.)

Показник	Режим навчання	Режим застосування	Значення оптимальних параметрів роботи методу
Точність	0.98	0.97	<i>ШНМ для аугментації даних:</i> Ймовірнісна нейронна мережа. <i>Базова ШНМ для реалізації процедури навчання:</i> Неітеративна нейроподібна структура МППП з РБФ розширенням входів. <i>Кількість РБФ центрів:</i> 93. <i>Значення розмаху гаусівської функції для МППП:</i> 0,85678349. <i>Відстань:</i> Евклідова. <i>Значення розмаху гаусівської функції для ЙМН:</i> 0.69442567. <i>Метод оптимізації</i> – метод імітації відпалу.
Влучність	0.99	0.99	
Повнота	0.98	0.97	
F1-міра	0.99	0.98	
Коефіцієнт кореляції Метьюса	0.97	0.95	
Каппа Коена	0.96	0.95	
Час навчання (секунд)	440.2	-	

Оскільки вибірка даних є дещо незбалансована (коефіцієнт дисбалансу 1.2), особливу увагу приділяли значенням F1-міри, коефіцієнта кореляції Метьюса та каппі Коена. Це дозволило здійснити всебічну оцінку отриманих результатів,

оскільки в науковій літературі [291], [292], [293], [294], [295] досі триває дискусія щодо того, який з цих показників найбільш повно характеризує ефективність моделі у випадку незбалансованої вибірки даних. Крім того в таблиці 6.5 наведено значення оптимальних параметрів, що забезпечують ефективну роботу Методу 3.3. Як видно з результатів, поданих у таблиці 6.5, Метод 3.3 демонструє високі властивості узагальнення та високу точність роботи на основі усіх трьох найважливіших показників: F1-міри (0.98), коефіцієнта кореляції Метьюса (0.95) та Каппи Коена (0.95) в процесі розв'язання задачі класифікації успіху або невдачі трансплантації стовбурових клітин у дітей на основі різних клінічних параметрів у випадку аналізу короткої вибірки даних.

Порівняння. Оцінювання ефективності Методу 3.3 здійснено шляхом порівняння його результатів з результатами кількох різних його реалізацій, зокрема, з варіантами, що включають як ітеративні методи машинного навчання, так і ймовірнісну нейронну мережу, а також з базовим Методом 3.1:

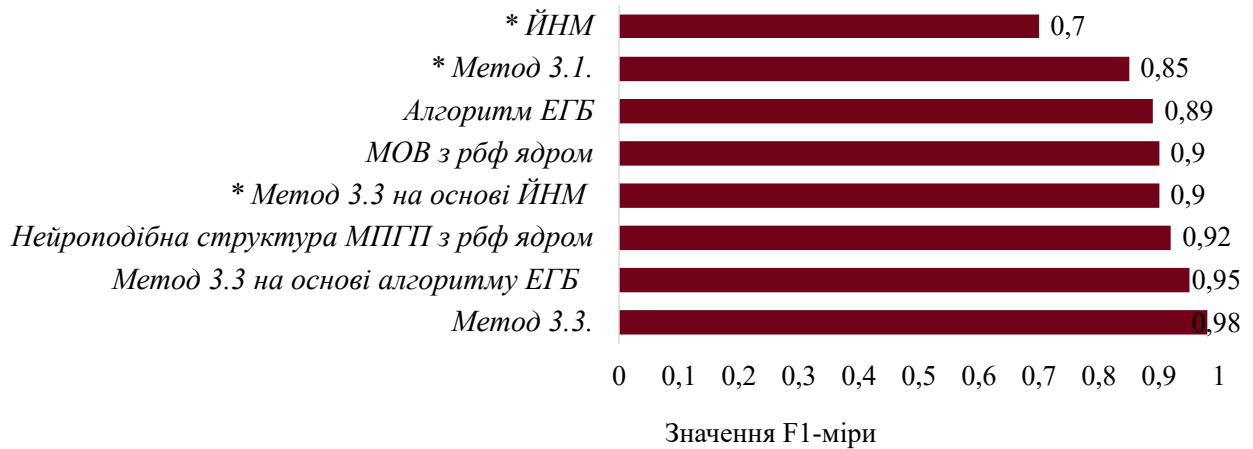
- метод 3.1 для класифікації на основі ймовірнісної нейронної мережі (*Метод 3.1*);
- метод 3.3 з використанням двох ймовірнісних нейронних мереж (*Метод 3.3 на основі ЙНМ*);
- метод 3.3 з використанням алгоритму екстремального градієнтного бустінгу -ЕГБ (*Метод 3.3 на основі алгоритму ЕГБ*).

Крім того процедура порівняння передбачала використання вже існуючих нейронних мереж та методів машинного навчання:

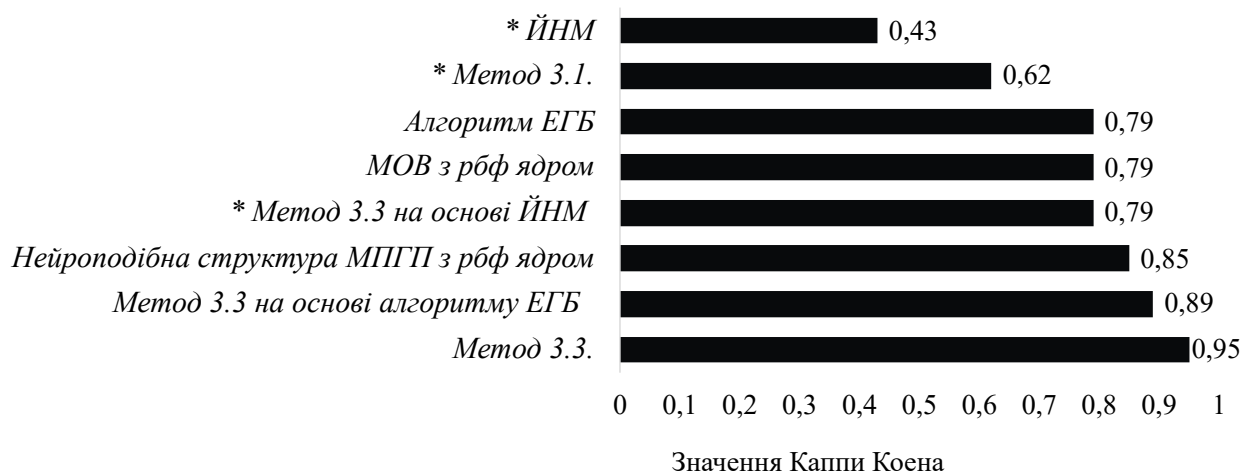
- ймовірнісна нейронна мережа (*ЙНМ*);
- неітеративна нейроподібна структура МППП з РБФ-розширенням входів ядром (*Нейроподібна структура МППП з РБФ-ядром*);
- машина опорних векторів (МОВ) з РБФ ядром (*МОВ (РБФ ядро)*);
- алгоритм екстремального градієнтного бустінгу (*Алгоритм ЕГБ*).

Усі досліджувані методи, а також значення їх оптимальних параметрів подано у таблиці Б.2 додатку Б.

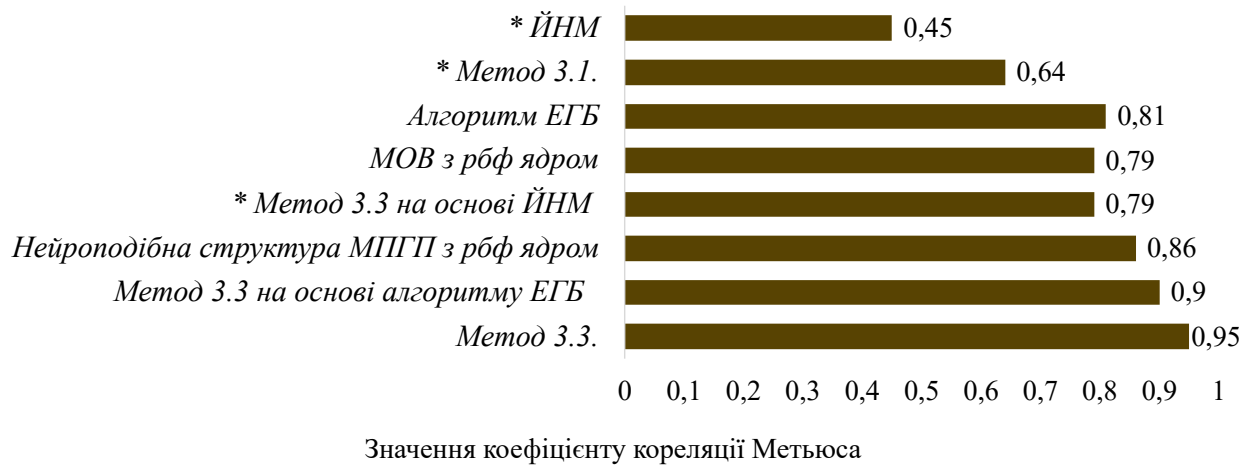
На рис. 6.7 подано порівняння основних показників ефективності для усіх досліджуваних методів.



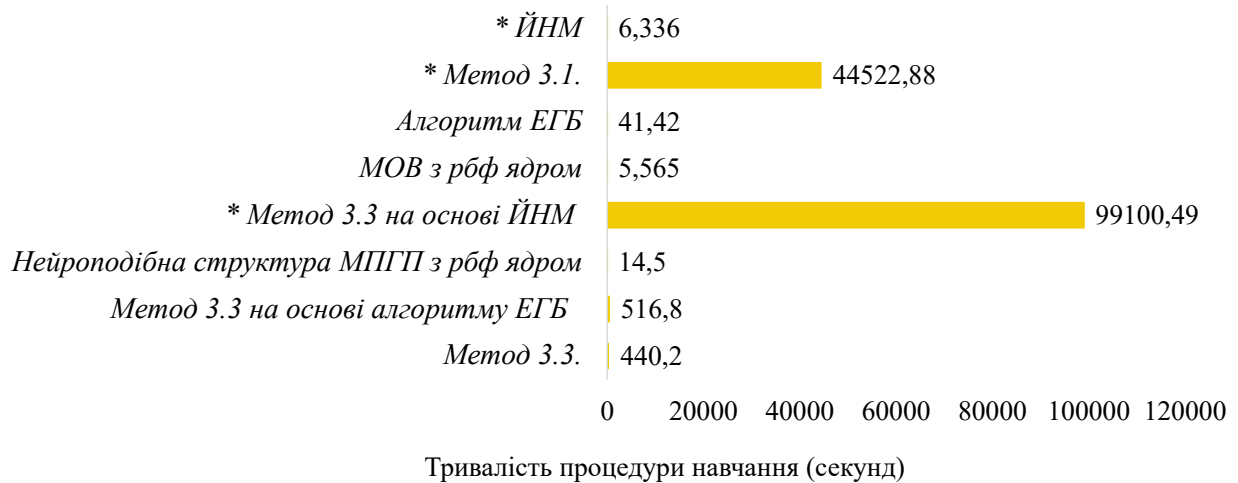
а)



б)



в)



г)

* в цих випадках брався до уваги увесь час реалізації відповідного методу.

Рис. 6.7 – Порівняння методу класифікації даних на основі каскадування (Метод 3.3) з іншими методами з використанням: а) F1-міри; б) Каппи Коена; в) коефіцієнта кореляції Метьюса; г) тривалості процедури навчання (реалізації) методу в секундах

З рис. 6.7 можна зробити наступні висновки. Метод 3.3 (з неітеративним навчанням) демонструє найкращі результати за всіма основними показниками. Ці результати свідчать про високу ефективність цього методу, а загальна висока оцінка

за усіма показниками сукупно вказує на стабільність та надійність роботи цього методу для розв'язання поставленої задачі.

Метод 3.3 на основі алгоритму ЕГБ (з ітеративним методом навчання) також показав хороші результати. Хоча цей метод досягнув ідеальної влучності, дещо нижча повнота вказує на те, що він може не виявляти всі позитивні випадки. Окрім цього він поступається в точності Методу 3.3 (з неітеративним навчанням), зокрема, на більш, ніж 5% на основі Каппи Коена, а також через ітеративність вимагає тривалішої процедури навчання.

Неітеративна нейроподібна структура МПГП з РБФ-розширенням входів показала хороші результати. Однак ряд її показників були суттєво гіршими в порівнянні з Методом 3.3 та його варіантом на основі ЕГБ. Це свідчить про те, що хоча ця модель має гарну здатність до точної класифікації, вона може бути менш чутлива до деяких класів, що негативно позначається на її повноті.

Машина опорних векторів (МОВ) з РБФ-ядром та Алгоритм ЕГБ продемонстрували подібні результати. Проте використання МОВ з РБФ-ядром демонструє хорошу здатність моделі до точної класифікації, але з певною втратою чутливості до деяких класів. В той же час застосування Алгоритму ЕГБ досягає ідеальної влучності, але може пропускати значну кількість позитивних випадків.

Метод 3.3 на основі ЙНМ показав значно слабші результати порівняно з іншими методами. Аналогічно Метод 3.1 на основі ЙНМ та сама ЙНМ продемонстрували найгірші показники серед усіх досліджуваних варіантів. Це свідчить про обмежену ефективність цих методів через низьку гнучкість ЙНМ, яку покладено в основу їх роботи. Крім того Методи 3.3 і 3.1 на основі ЙНМ потребують значних часових витрат на реалізацію, особливо при роботі з великими обсягами даних (квадратично збільшена вибірка даних), що обумовлено особливостями функціонування ЙНМ (див. розділ 2). Враховуючи такі недоліки, ці методи є непридатними для практичного використання у розв'язанні поставленої задачі.

6.2.4 Оцінювання результатів методу формування аугментованої вибірки даних обмеженого обсягу шляхом кластеризації

Метод формування аугментованої вибірки даних зменшеного обсягу за допомогою кластеризації, описаний у пункті 3.6 (Метод 3.5), який можна застосовувати до всіх раніше згаданих методів цього класу, був протестований на штучній вибірці даних.

Задача та її практична цінність. Процедура аугментації, застосована до всіх методів, описаних у пунктах 3.2–3.5, передбачає квадратичне збільшення обсягу вибірки даних для реалізації процесу навчання (рис. 6.8).

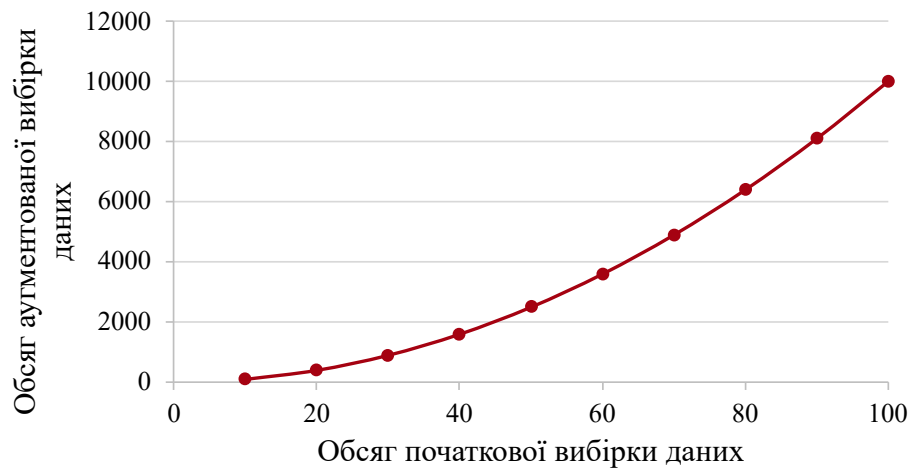


Рисунок 6.8 – Квадратичне збільшення обсягу вибірки даних при реалізації процедури аугментації для методів, описаних у підрозділах 3.2–3.5 третього розділу

У випадку аналізу малих вибірок даних, де кількість спостережень у початковій вибірці перевищує 100, розроблені методи вимагатимуть суттєво більших часових витрат на їх реалізацію. У ситуаціях, коли час на прийняття рішення є критичним, наприклад, у задачах експрес-діагностики, це може знизити ефективність застосування таких методів. Зважаючи на це, виникає необхідність зменшення обсягу аугментованої вибірки для методів розв’язання задач регресії

класифікації, що ґрунтуються на декартовому квадраті множини вхідних компонент векторів, при цьому зберігаючи високу точність результатів їх роботи.

Генерація штучної вибірки даних. Для моделювання синтезованої штучної вибірки даних, що містить два незалежних атрибути та один залежний, генерація даних здійснювалася за допомогою нелінійної функції Франке:

$$f(a, b) = 0.75 \exp\left(-\frac{(9a-2)^2}{4} - \frac{(9b-2)^2}{4}\right) + 0.75 \exp\left(-\frac{(9a+1)^2}{49} - \frac{(9b+1)^2}{10}\right) + 0.5 \exp\left(-\frac{(9a-7)^2}{4} - \frac{(9b-3)^2}{4}\right) - 0.2 \exp(-(9a-4) - (9b-7)^2). \quad (6.2)$$

Загалом було синтезовано 100 спостережень, кожне з яких генерувалося при випадкових значеннях параметрів a, b , що належать проміжку $[-1, 1]$. Для кожного з цих спостережень обчислювалося значення цільової функції згідно з формулою (6.2).

Моделювання та налаштування параметрів методу. Моделювання виконувалося на прикладі зменшення обсягу аугментованої вибірки для Методу 3.1. Кластеризація вибірки даних здійснювалася за допомогою k-means. Оптимальну кількість кластерів можна визначити за допомогою одного з обраних методів, описаних у другому розділі, або ж з використанням заданого критерію зупинки для методу розв'язання задач регресії чи класифікації при зміні кількості кластерів. Критерій зупинки ґрунтується на декартовому квадраті множини вхідних компонент векторів.

Результати. Результати реалізації Методу 3.5 на прикладі зменшення обсягу аугментованої вибірки даних Методу 3.1 на основі різних показників з Додатку А та значення оптимальних параметрів його роботи подано у таблиці 6.6.

Згідно з результатами, наведеними в таблиці 6.6, розроблений метод показує відмінну точність, володіє чудовими властивостями узагальнення та успішно уникає перенавчання, яке часто виникає в процесі використання існуючих методів при роботі з екстремально малими вибірками даних.

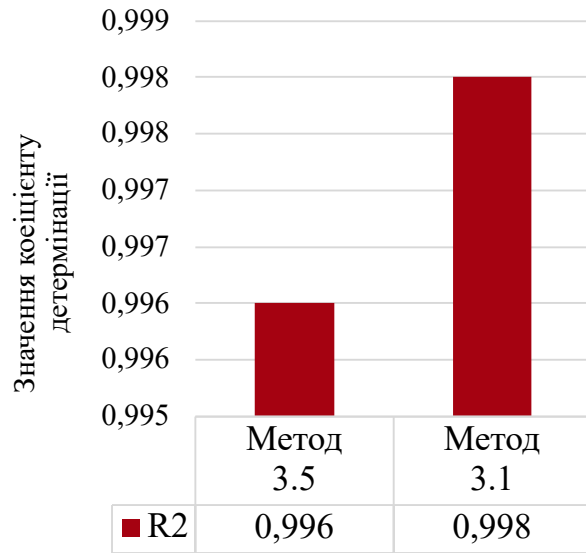
Таблиця 6.6 – Результати методу зменшення вимірності аугментованої вибірки даних за рахунок кластеризації (Метод 3.5)

Показник ефективності	Режим навчання	Режим застосування	Значення оптимальних параметрів роботи методу
MAPE	1,970	2,011	Кількість кластерів: 9 Базова ШНМ для реалізації процедури навчання: Неітеративна нейроподібна структура МППП з РБФ ядром. Кількість РБФ центрів: 11. Значення розмаху гаусівської функції: 3,63.
RMSE	0,021	0,023	
MAE	0,013	0,015	
R2	0,997	0,996	
MaxError	0,060	0,071	
Кількість векторів в аугментованій вибірці даних	720	-	
Тривалість процедури навчання (в секундах)	0,086	-	

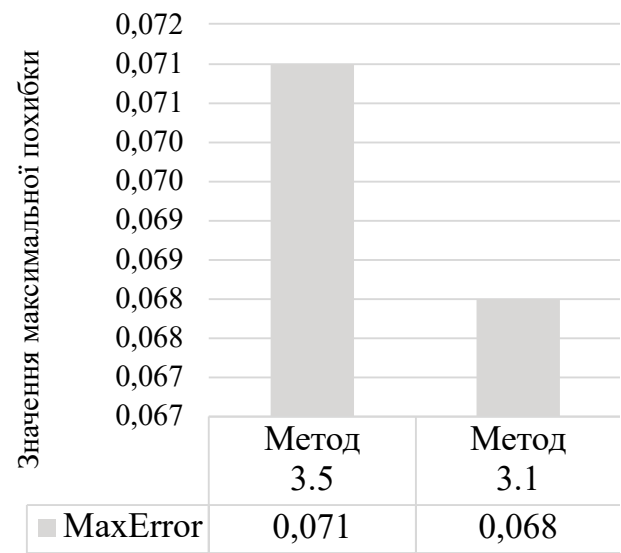
Порівняння. Для оцінки ефективності Методу 3.5 проведено порівняння результатів його роботи з результатами за існуючим Методом 3.1 за такими критеріями:

- точність роботи,
- обсяг аугментованої вибірки даних для реалізації процедури навчання,
- тривалість процедури навчання (в секундах).

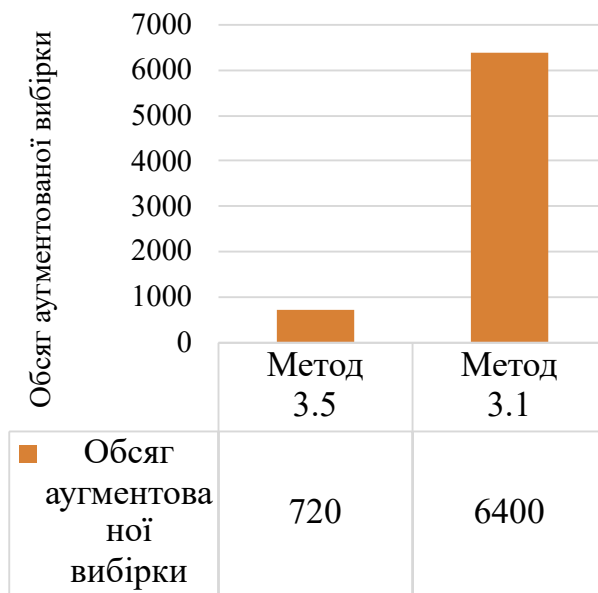
На рис. 6.9 подано значення коефіцієнта детермінації, максимальної похибки, обсягу аугментованої вибірки та тривалості навчання для обох методів. З результатів, поданих на рис. 6.9, видно, що використання Методу 3.5 супроводжується суттєвим скорочення обсягу аугментованої вибірки даних майже у дев'ять разів. Це очевидно відображається на тривалості процедури навчання. Однак такий процес супроводжується невеликим зменшенням точності роботи Методу 3.5. в порівнянні з точністю за Методом 3.1.



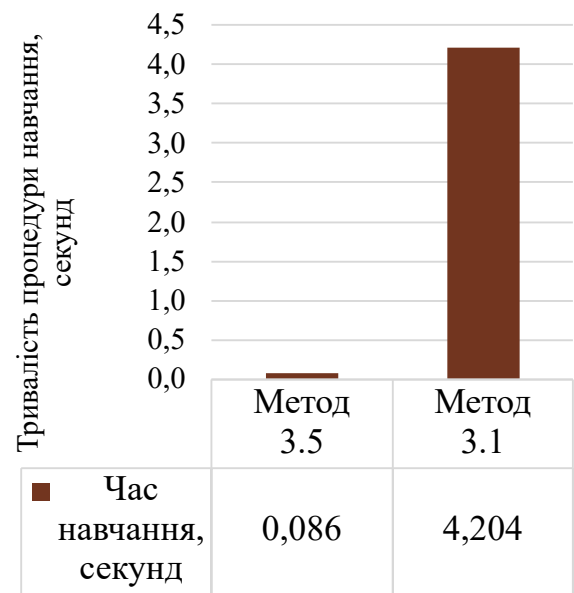
а)



б)



в)



г)

Рисунок 6.9 – Порівняння точності Методу 3.5 для зменшення вимірності аугментованої вибірки даних шляхом кластеризації з базовим Методом 3.1 на основі: а) коефіцієнта детермінації; б) максимальної похибки; в) обсягу аугментованого набору даних; г) часу навчання методу (в секундах).

Подібні результати можна отримати, застосувавши Метод 3.5. для зменшення обсягу вибірки даних для навчання Методів 3.2-3.4.

Вибір щодо можливості використання Методу 3.5 в процесі практичної реалізації Методів 3.1-3.4 залежить від конкретних вимог до тривалості навчання та точності, які висуваються в процесі розв’язання певної прикладної задачі, а також від обсягу початкової вибірки даних. Якщо пріоритет надається швидкодії, а початкова вибірка даних відносно великого обсягу (100-1000 спостережень), то застосування Методу 3.5 є доцільним. Якщо ж точність є критично важливим показником, а вибірка даних є екстремально короткою, то слід віддати перевагу використанню базових варіантів Методів 3.1-3.4, незважаючи на вищі вимоги до часу їх навчання та обчислювальних ресурсів (пам’яті).

6.3 Експериментальне дослідження методів часткової компенсації похибок роботи моделей машинного навчання на основі каскадних ансамблів в задачі регресії

У цьому підрозділі подано результати обчислювальних експериментів та порівняння ефективності розроблених та удосконалених методів з ефективністю за існуючими підходами. Експериментальне дослідження ефективності усіх методів виконувалося за допомогою розробленого програмного забезпечення на мові Python. Зокрема, реалізовані каскадні ансамблі дозволяють використовувати довільні методи машинного навчання чи ШНМ, доступні у бібліотеках Python.

6.3.1 Оцінювання результатів методу часткової компенсації похибки роботи нейронної мережі узагальненої регресії

Розроблений метод часткової компенсації похибки роботи нейронної мережі узагальненої регресії був протестований на вибірці даних з галузі екологічного моніторингу.

Задача та її практична цінність. Розроблений метод використано в процесі розв'язання задачі вірогідного відновлення частково пропущених або втрачених даних у частині вибірки даних моніторингу стану повітряного середовища. Ефективний розв'язок цієї задачі забезпечить можливість підвищити якість і надійність подальшого інтелектуального аналізу даних у системах моніторингу стану повітряного середовища.

Опис вибірки даних. Частину вибірки даних взято з відкритого джерела [296]. Вона містить 11 незалежних змінних, зібраних хімічними сенсорами пристрою IoT, що вимірюють різні параметри якості повітря. Цільовим атрибутом обрано концентрацію чадного газу (Carbon monoxide, CO), оскільки він містить найбільшу кількість пропущених значень в досліджуваній вибірці даних.

Моделювання та налаштування параметрів методу. Моделювання розробленого методу проводилося з використанням трьох різних способів агрегування результатів: 1) простим підсумовуванням відповідно до формули (5.23), 2) використанням коригувальної лінійної нейроподібної структури МППП згідно з (5.24), 3) розширенням простору вхідних даних (виходи першої та другої НМУР ансамблю) за допомогою полінома Колмогорова-Габора другого степеня з подальшим застосуванням лінійної нейроподібної структури МППП відповідно до (5.25). Як процедуру попередньої обробки було застосовано нормалізацію, яка полягала у діленні кожного значення ознаки на її максимальне за модулем значення.

Результати і порівняння. Результати реалізації розробленого методу з трьома різними способами агрегування результатів, а також порівняння його роботи з існуючими методами (випадковий ліс, адаптивний бустинг, нейронна мережа узагальненої регресії) наведено на рис. 6.10.

Аналіз отриманих результатів підтверджує високу ефективність розробленого методу при застосуванні всіх трьох способів агрегування результатів. Зменшення RMSE становить понад 4%, 6% та 8% відповідно для першого, другого та третього способів агрегування. Найкращі показники досягнуто при використанні

третього способу розширення простору вхідних даних (виходи першої та другої НМУР ансамблю) за допомогою полінома Колмогорова-Габора другого степеня з подальшим застосуванням лінійної нейроподібної структури МППП згідно з формулою (5.25).

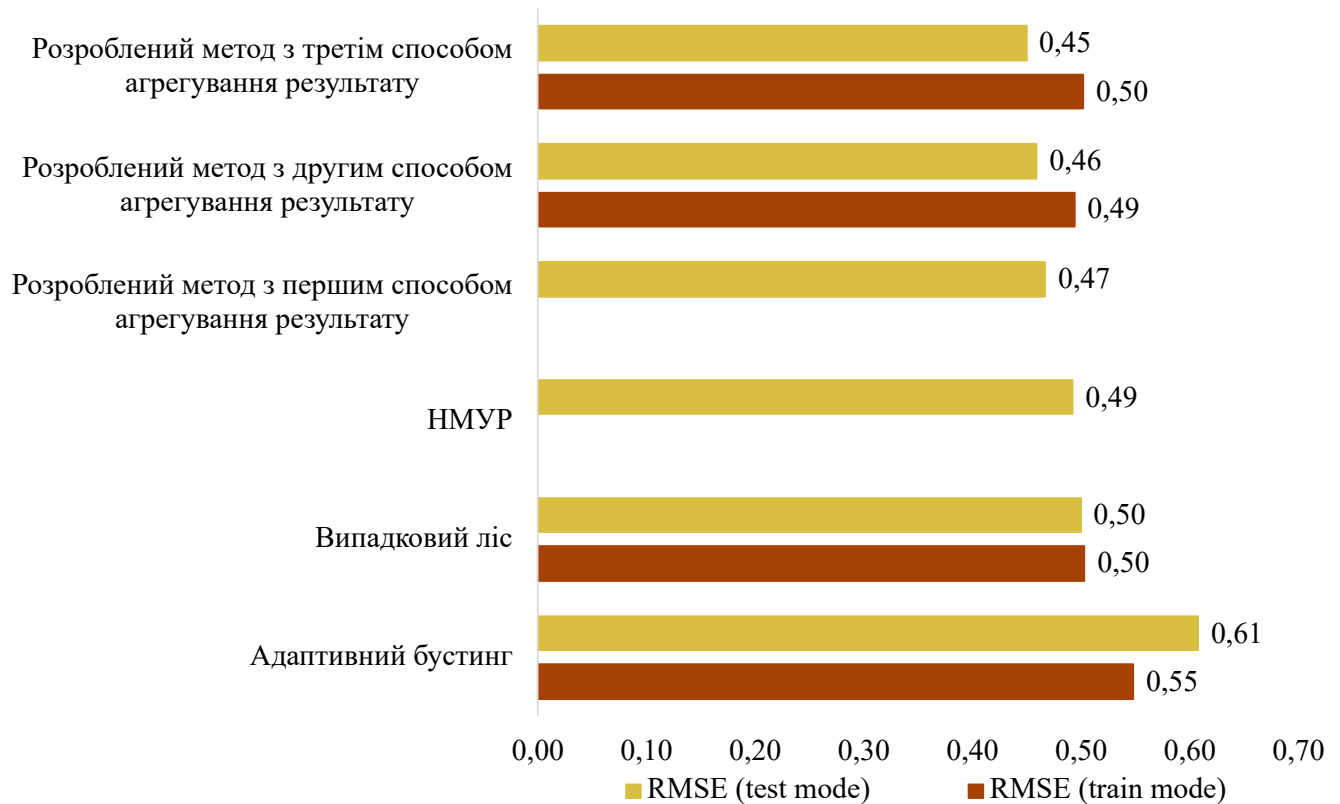


Рисунок 6.10 – Порівняння похибок RMSE для усіх досліджуваних методів.

Це свідчить про практичну доцільність використання розробленого підходу як окремого методу, так і як складової частини ансамблевих алгоритмів, що базуються на НМУР.

6.3.2 Оцінювання результатів методу глобально локальної апроксимації даних малих обсягів в галузі біомедичної інформатики

Удосконалений метод глобально локальної апроксимації даних малих обсягів був протестований на вибірці даних з відкритого джерела в процесі розв'язання задачі в галузі біомедичної інформатики.

Задача та її практична цінність. Удосконалений метод використано в процесі розв'язання задачі оцінювання (задача регресії) відсотка жирової тканини тіла на основі антропометричних та лабораторних вимірювань. Отримані результати можуть бути використані для ранньої діагностики ожиріння, визначенні ризику розвитку метаболічних ускладнень, індивідуалізації планів харчування, а також у підвищенні ефективності підтримки клінічних рішень у сфері профілактичної та персоналізованої медицини.

Опис вибірки даних. Вибірка даних з відкритого джерела [297] містить 252 вектори, кожен з яких представлений тринадцятьма антропометричними показниками, що характеризують морфометричні параметри тіла людини. До незалежних атрибутів належать вік, маса тіла, зріст, окружність шиї, грудей, талії, стегон, а також обхвати стегна, коліна, щиколотки, біцепса, передпліччя та зап'ястка [298]. Усі вимірювання виконані відповідно до стандартних антропометричних протоколів. Як цільову (залежну) змінну використано відсоток жирової тканини організму, отриманий за допомогою методу підводного зважування [298]. Така структура даних дає змогу формалізувати задачу оцінки жирового компонента тіла як задачу регресії, релевантну для досліджень у галузі персоналізованої превентивної медицини та оцінювання ризику метаболічних порушень.

Моделювання та налаштування параметрів методу. Моделювання удосконаленого методу здійснювалося з застосуванням п'ятикратної кросвалідації з метою оцінки стабільності та узагальнювальної здатності побудованої моделі.

Серед процедур попереднього опрацювання даних було використано нормалізацію, що проводилась шляхом поділу кожного значення окремої ознаки на її максимальне за модулем значення [44].

Результати. Результати удосконаленого методу на основі показників з Додатку А та значення його оптимальних параметрів подано у таблиці 6.7.

Таблиця 6.7 – Результати удосконаленого методу глобально локальної апроксимації

Метрика	Режим навчання	Режим застосування	Значення оптимальних параметрів роботи методу
MAE	2,57	3,05	<i>ШНМ для виділення глобальної компоненти: НМУР-1.</i>
MAPE	0,29	0,32	
MSE	11,43	15,8	
MaxError	8,41	8,8	<i>Значення розмаху гаусівської функції: 0.1.</i>
MedianAE	2,26	2,82	
RMSE	3,24	3,81	
R ²	0,87	0,84	<i>ШНМ для виділення глобальної компоненти: НМУР-2.</i> <i>Значення розмаху гаусівської функції: 0.04.</i> <i>Метод оптимізації – диференційна еволюція.</i>

Як видно з результатів, поданих у таблиці 6.7, удосконалений метод демонструє високу точність роботи та високі властивості узагальнення в процесі розв’язання задачі оцінювання відсотка жирової тканини тіла на основі антропометричних та лабораторних вимірювань.

Порівняння. Оцінювання ефективності удосконаленого методу здійснено шляхом порівняння його результатів з точністю існуючих методів.

На рис. 6.11 подано порівняння похибок на основі RMSE різних досліджуваних методів.



Рисунок 6.11 – Порівняння точності удосконаленого методу з точністю за існуючими методами на основі RMSE

З результатів, представлених на рис. 6.11, можна зробити такі висновки. Результати моделювання свідчать про перевагу удосконаленого методу у розв’язанні поставленої задачі, що підтверджується найменшим значенням середньоквадратичної похибки (RMSE) серед усіх розглянутих методів. Базовий варіант глобально локальної апроксимації, орієнтований на аналіз великих обсягів даних, продемонстрував значно гірші результати: зростання похибки на 13 % за показником RMSE. Модифікована версія методу, що виконує послідовне визначення глобальної та локальної складових сигналу, показала на 7 % нижчу точність за RMSE у порівнянні з удосконаленим методом. Існуючі методи, такі як випадковий ліс, машина опорних векторів з РБФ-ядром, нейронна мережа узагальненої регресії та лінійна нейроподібна структура, продемонстрували вищі значення похибок. Це свідчить про їх нижчу ефективність порівняно з удосконаленим методом при роботі з біомедичними вибірками малих обсягів.

6.3.3 Оцінювання результатів методів компенсації похибок моделі регресії з використанням раціональних дробів

Розроблені прямий та зворотний методи компенсації похибки довільної моделі регресії були протестовані на вибірці даних з галузі біоматеріалознавства.

Задача та її практична цінність. Розроблені методи використано в процесі розв’язання задачі оцінювання твердості алюмінієвої кераміки, виготовленої з порошків оксиду алюмінію Al_2O_3 , що застосовується як компонента третьових пар у кульшових ендопротезах. Ефективний розв’язок цієї задачі забезпечить високоточне та оперативне оцінювання механічних властивостей матеріалу без необхідності проведення тривалих і затратних експериментальних випробувань. Це дозволить оптимізувати технологічні процеси виготовлення порошкових композицій, а також підвищити якість і надійність ендопротезів.

Опис вибірки даних. Вибірка даних зібрана на кафедрі матеріалознавства та інженерії матеріалів Національного університету “Львівська політехніка”. Вона містить 115 зразків твердості алюмінієвої кераміки, які характеризуються за 11 фізико-хімічними та морфологічними параметрами.

Моделювання та налаштування параметрів методу. Для моделювання обох методів компенсації похибок довільної моделі регресії обрано машину опорних векторів, яка показала високу точність роботи і використовується фахівцями цієї кафедри. Попередньо дані було нормалізовані шляхом поділу кожного значення окремої ознаки на її максимальне за модулем значення [44]. В якості моделі корекції обрано лінійну неітеративну нейроподібну структуру МПГП.

Результати і порівняння. Результати базової моделі регресії (машини опорних векторів з РБФ-ядром), та прямого і зворотного методів компенсації похибок її роботи на основі показників з Додатку А подано у таблиці 6.8.

Таблиця 6.8 – Результати для досліджуваних методів

Метрика	Машина опорних векторів з РБФ-ядром (MOB)	Прямий метод компенсації похибок MOB	Зворотний метод компенсації похибок MOB
MAE	1,046	0,963	0,952
MSE	2,46	2,24	2,19
MaxError	4,02	3,85	3,81
MedianAE	0,62	0,6	0,59
RMSE	1,53	1,49	1,48
R ²	0,87	0,89	0,89

Аналіз отриманих результатів підтверджує ефективність застосування розроблених методів компенсації похибок у задачі регресії на основі машини опорних векторів з радіально-базисною функцією (РБФ-розширенням входів). Використання як прямого, так і зворотного методів дозволило досягти систематичного покращення розглянутих метрик точності. Зокрема, середньоквадратична похибка MSE зменшилася майже на 9 % при використанні прямого методу та більш, ніж на 11 % при застосуванні зворотного. Зворотний метод демонструє дещо кращу ефективність, що зумовлено його здатністю враховувати масштаб початкового прогнозу, забезпечувати більш рівномірне навчання коригувальної моделі та реалізовувати масштабну компенсацію результатів. Загалом отримані результати підтверджують доцільність використання обох методів компенсації похибок моделі регресії в умовах роботи з даними обмеженого обсягу. Ці методи можуть бути масштабовані для аналізу даних більшого обсягу.

6.4 Висновки до розділу

1. Верифікацію методу нелінійного доповнення входів лінійної неітеративної нейроподібної структури МПГП з використанням членів полінома Колмогорова–Габора було проведено на задачі про оцінювання індивідуальних витрат на медичне страхування. Результати моделювання засвідчили, що застосування полінома другого степеня дозволяє суттєво підвищити точність у порівнянні з базовою лінійною моделлю, а оптимізація кількості нейронів прихованого шару додатково покращила властивість узагальнення моделі та скоротила час навчання майже в 2,5 рази. Порівняльний аналіз продемонстрував перевагу розробленого методу над класичними лінійними регресорами та базовою структурою МПГП за всіма показниками точності. Важливою перевагою методу є можливість аналітичного представлення результату у формі поліноміального виразу. Це істотно підвищує рівень інтерпретованості моделі та довіру до її висновків з боку клінічних експертів.

2. Верифікація удосконаленого методу класифікації даних на основі формування комбінованої поверхні відгуку з застосуванням кластерного аналізу проводилося на задачі про оцінювання ризику серцевого нападу. Результати показали суттєве покращення точності класифікації порівняно з існуючими підходами. Метод забезпечив високу властивість узагальнення, зокрема, при використанні кластеризації з врахуванням вихідного атрибуту. Оптимізація кількості кластерів дозволила досягнути максимальних показників точності при збереженні низької обчислювальної складності. Порівняльний аналіз підтвердив перевагу методу над базовою лінійною нейроподібною структурою МПГП і відомими методами як за критерієм точності, так і за здатністю до ефективного опрацювання біомедичних вибірок даних малих обсягів. Додатковою перевагою методу є зменшення вимірності вхідного простору даних для лінійного

класифікатора другого рівня. Це забезпечило підвищення швидкодії та стійкості до перенавчання в задачах медичного діагностування.

3. Верифікація базового методу розв'язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів показала його високу ефективність при аналізі екстремально коротких вибірок даних. Результати моделювання на штучно згенерованій вибірці даних засвідчили значне підвищення точності та відсутність перенавчання. Це було досягнуто завдяки використанню нелінійного розширення входів і оптимізації параметрів за допомогою диференційної еволюції. У порівнянні з лінійним та нелінійними моделями ШНМ розроблений підхід продемонстрував, зокрема, в 150 разів менше значення похибки RMSE на штучних даних. Це підтверджує ефективність його використання в умовах аналізу вибірок даних малих обсягів.

4. Верифікація адитивного методу розв'язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів виконувалося для задачі про оцінювання міцності трабекулярної кістки при стисканні в ревматології. Результати засвідчили його високу точність, властивості узагальнення та робастність навіть за умов екстремально-короткої вибірки. Завдяки сумісному використанню аугментації даних та неітеративної нейроподібної структури з РБФ-розширенням входів метод значно перевершив класичні ШНМ та методи машинного навчання. Зокрема, він зменшив похибку RMSE на 42–52% у порівнянні з найближчими аналогами. Такі результати підтверджують доцільність і ефективність застосування розробленого методу для розв'язання реальних прикладних задач у біомедичних дослідженнях за умов аналізу вибірок даних обмежених обсягів.

5. Метод класифікації даних на основі декартового квадрата множини вхідних компонент векторів з використанням принципів каскадування продемонстрував високі точність та властивість узагальнення при розв'язанні задачі про оцінювання результату трансплантації стовбурових клітин у дітей на

основі клінічних параметрів. Найкращі результати досягнуто завдяки поєднанню аугментації даних, використанню ймовірнісної нейронної мережі як ШНМ першого рівня і навчанням неітеративною нейроподібною структурою з РБФ-розширенням входів як ШНМ другого рівня. Це забезпечило високу ефективність за всіма ключовими метриками: F1-мірою, коефіцієнтом Метьюса та каппою Коена. Порівняльний аналіз з альтернативними реалізаціями цього методу та класичними ШНМ і методами машинного навчання підтвердив суттєву його перевагу. Отримані результати підтверджують доцільність використання каскадної структури для розробленого методу при аналізі медичних даних обмежених обсягів у критично важливих клінічних завданнях.

6. Розроблений метод розв'язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів з формуванням комбінованої поверхні відгуку продемонстрував високу ефективність розв'язання задачі оцінювання в'язкості руйнування цирконієвої кераміки на етапі спікання. Використання кластеризації для формування комбінованої поверхні відгуку та неітеративного навчання нейроподібною структурою РБФ-розширенням входів дозволило досягнути точності, що перевищує базовий рівень на понад 10%. Отримані результати підтвердили доцільність застосування цього методу для задач аналізу матеріалів у медичній стоматології, особливо у випадку екстремально коротких вибірок даних. Він забезпечує високі властивості узагальнення, зменшення потреби в експериментальних ресурсах та стабільність моделювання за умов аналізу даних малих обсягів.

7. Розроблений метод формування аугментованої вибірки даних зменшеного обсягу на основі кластеризації продемонстрував здатність істотно скоротити обсяг даних для навчання майже в дев'ять разів без значної втрати точності. Проведене моделювання підтвердило, що навіть за умов скорочення аугментованої вибірки даних метод зберігає високу якість апроксимації, високі властивості узагальнення та уникає перенавчання. Такий підхід є доцільним для

задач, у яких критично важливо зменшити обчислювальну складність або прискорити метод навчання моделей, особливо у випадках, коли початкові вибірки даних перевищують 100 спостережень. Водночас для екстремально коротких вибірок даних доцільно зберігати повний обсяг аугментованих даних, надаючи перевагу базовим методам цього класу.

8. Розроблений метод часткової компенсації похибки нейронної мережі узагальненої регресії ефективно відновлює пропущені дані у вибірці моніторингу якості повітря, що підтверджується зменшенням RMSE більш, ніж на 4%, 6% та 8% відповідно при застосуванні трьох різних способів агрегування результатів в порівнянні з базовою НМУР. Найкращі результати досягнуто шляхом розширення простору вхідних даних поліномом Колмогорова-Габора другого степеня з подальшим використанням лінійної нейроподібної структури. Це свідчить про доцільність його застосування як окремого методу, так і в ансамблевих підходах для підвищення точності відновлення даних у системах екологічного моніторингу.

9. Верифікація удосконаленого методу глобально локальної апроксимації даних малих обсягів показала високу ефективність при розв'язанні задачі оцінювання відсотка жирової тканини тіла за антропометричними та лабораторними показниками. Удосконалений метод забезпечує стабільність та високу точність результату, що підтверджується найнижчими значеннями середньоквадратичної похибки RMSE у порівнянні з базовими та класичними методами машинного навчання. Отримані результати свідчать про доцільність застосування удосконаленого методу для задач біомедичної інформатики, особливо в умовах обмежених обсягів даних. Це відкриває перспективи його використання у персоналізованій медицині та превентивних клінічних дослідженнях.

10. Аналіз отриманих результатів свідчить про високий потенціал використання розроблених прямого і зворотного методів компенсації похибок для підвищення точності моделей регресії. Зокрема, обидва підходи показали суттєве зниження різних похибок роботи машини опорних векторів з РБФ-ядром. MSE

зменшується майже на 9 % при застосуванні прямого методу та понад 11 % – у разі використання зворотного методу. Перевага останнього полягає у здатності враховувати масштаб початкової оцінки та забезпечувати більш збалансоване навчання коригувальної моделі, що сприяє ефективнішій компенсації помилок. Це робить розроблені методи особливо корисними для роботи з обмеженими вибірками біомедичних даних у різних практичних задачах.

РОЗДІЛ 7

ТЕСТУВАННЯ І РЕЗУЛЬТАТИ ВПРОВАДЖЕННЯ РОЗРОБЛЕНИХ МЕТОДІВ НА ДАНИХ БІОМЕДИЧНИХ ДОСЛІДЖЕНЬ

У цьому розділі приведені результати впровадження розроблених та удосконалених у дисертаційній роботі методів, орієнтованих на розв'язання задач регресії чи класифікації з використанням біомедичних даних малого обсягу. Описано реальні прикладні задачі у різних галузях біомедичної інженерії, включаючи біоматеріалознавство, ревматологію, стоматологію, травматологію та трансплантологію. Подано результати моделювання розроблених методів. Проведено порівняння ефективності їх роботи. Таке порівняння продемонструвало перевагу над існуючими методами показало доцільність практичного використання розроблених методів в процесі розв'язання практичних задач. Результати впроваджень підтверджено відповідними актами від провідних медичних установ України. Апробація ряду методів виконувалася також в межах виконання міжнародного гранту, що фінансувався Європейською Комісією за програмою Horizon Europe.

Розроблені програмні засоби, які використовувалися в цьому розділі для проведення експериментальних досліджень, захищено авторським правом на комп'ютерні програми [54], [55], [56]. Отримані результати досліджень опубліковано у високореєтингових журналах інших держав, які включено до наукометричних баз Scopus та Web of Science: [14], [15], [23], [34], а також апробовано і опубліковано у матеріалах українських та закордонних конференцій: [18], [21], [27], [29], [30], [36], [49], [57].

7.1 Апробація методу розширення входів лінійної нейроподібної структури з використанням вихідних сигналів ймовірнісної нейронної мережі в галузі біоматеріалознавства

Метод класифікації даних на основі розширення входів лінійної нейроподібної структури вихідними сигналами шару підсумовування ймовірнісної нейронної мережі апробовано в процесі розв'язання задачі в галузі біоматеріалознавства. В процесі практичної реалізації розробленого методу створено веб-інтерфейс. Він надає можливість проводити експерименти у режимі реального часу. Функціонал програмного засобу для класифікації матеріалів з порошкових титанових сплавів сферичної та несферичної форми передбачає вибір варіанту формування ймовірностей у шарі підсумовування ймовірнісної нейронної мережі згідно з формулами (2.15), (2.17) або (2.18), а також вибір навчальних вибірок та ініціацію навчання лінійних нейроподібних структур МППП. Програму реалізовано мовою Python. Результатом роботи веб-системи є мітка, що вказує на належність конкретного спостереження до відповідного класу. Графічний інтерфейс розробленого веб-засобу наведено на рис. 7.1.

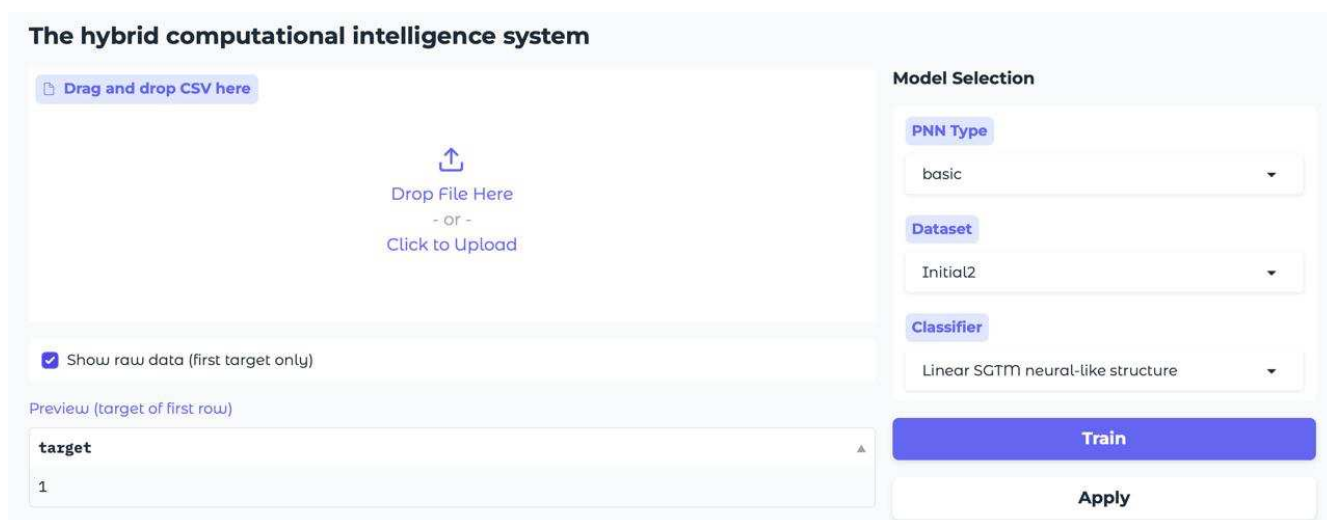


Рисунок 7.1 – Графічний інтерфейс веб-засобу для класифікації матеріалів з порошкових титанових сплавів сферичної та несферичної форми

Результати впроваджено на кафедрі стоматології дитячого віку та імплантології Харківського національного медичного університету впродовж 2022–2024 р. р. Це підтверджується відповідним Актом впровадження від 04.03.2024 р., виданим Харківським національним медичним університетом (додаток Г).

Задача та її практична цінність. Розроблений метод використано для класифікації матеріалів з порошкових титанових сплавів сферичної (рис. 7.2а) та несферичної (рис. 7.2б) форми [34]. Біосумісні вироби з таких сплавів виготовляються за допомогою адитивних технологій.

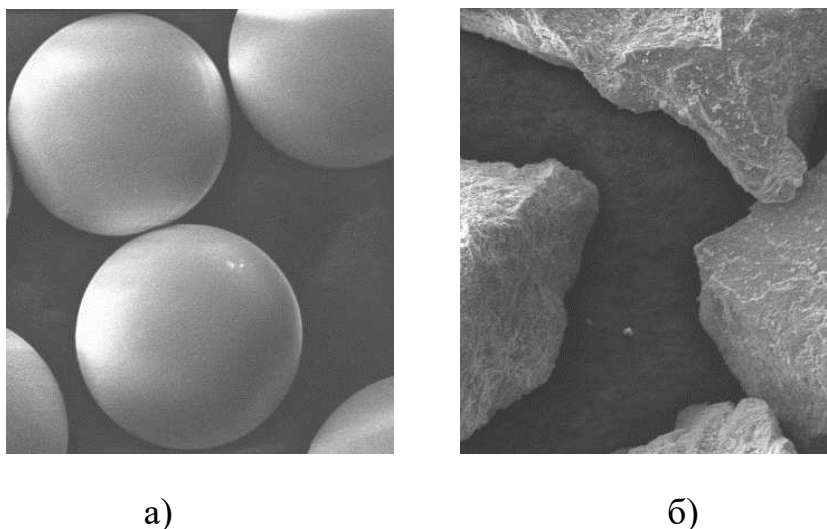


Рисунок 7.2 – Фрагменти зображень, що демонструють морфологію структури порошків титанових сплавів: а) сферичної форми; б) несферичної форми [34]

Завдання полягає в ідентифікації властивостей (чотири класи) попередньо підготованих сумішей порошкових матеріалів різного фракційного складу методами машинного навчання ще до виготовлення виробів шляхом пресування та спікання (рис. 7.3) [35].



Рисунок 7.3 – Класи біосумісності для матеріалів з титанових сплавів [35]

При цьому були враховані і характеристики порошків титанових сплавів, такі як хімічний склад, насипна густина, структура та фізичні властивості (рис. 7.4) [94].



Рисунок 7.4 – Набір незалежних атрибутів [94]

Ця задача є критично важливою, зокрема, для виробництва високоякісних медичних імплантатів, оскільки точна класифікація матеріалів ще до формування виробу та наступного його спікання гарантує високу якість, мінімізує поруватість, підвищує біосумісність, ергономічність та довговічність таких виробів. Використання не якісно підготованого та погано відтестованого по вхідних

параметрах порошкового матеріалу може призвести до відторгнення надрукованого імпланту, інфекцій чи інших серйозних ускладнень в організмі людини [223].

Ефективний розв'язок цієї задачі дозволяє зменшити ризики виникнення дефектів на всіх стадіях технологічного процесу виробництва, таких як зміни габаритних розмірів, форми імплантатів та невідповідність їх експлуатаційних властивостей через непередбачувану поведінку порошкового матеріалу в процесі спікання. Це важливо для забезпечення стабільності і надійності кінцевого продукту – виробів медичного призначення. Застосування Методу 2.2 дозволяє також оптимізувати виробничі процеси, зменшити витрати часу, матеріальних та енергетичних ресурсів, що значно підвищить ефективність виробництва [34], [42], [225].

Опис та попереднє опрацювання вибірки даних. Зібрана вибірка даних для моделювання отримана на кафедрі матеріалознавства та інженерії матеріалів Національного університету “Львівська політехніка” [299] на основі попередніх досліджень морфології, елементного та гранулометричного складу сферичних і несферичних порошків титанових сплавів різних систем легування [90]. Вибірка складається з 480 спостережень. Вони характеризуються 20 незалежними атрибутами, серед яких основними є хімічний і фазовий склад, розмір та форма частинок, полідисперсність і супутні компоненти [94], [223], [300]. Вихідний атрибут – мітка належності до одного з чотирьох класів: матеріал з відмінними властивостями, матеріал з оптимальними властивостями, матеріал з можливими дефектами, дефективний матеріал (див. рис. 7.1). Така мітка сформована на основі таких основних характеристик: середній діаметр частинок порошку та їх полідисперсність [301]. Ці характеристики мають вирішальне значення для визначення якості матеріалу, оскільки вони впливають на його структуру, механічні та біологічні властивості. Зокрема, середній діаметр частинок порошку визначає ефективність з'єднання часток в процесі спікання, що безпосередньо позначається на міцності і стабільності кінцевого матеріалу. Водночас полідисперсність

порошку відображає рівень варіативності розмірів частинок. Це впливає на щільність і однорідність матеріалу, а також на його здатність до адаптації в процесі виготовлення імплантатів за допомогою адитивних технологій [223].

Моделювання та налаштування параметрів методу. Моделювання виконувалося з використанням п'ятикратної кросвалідації. Розширення простору вхідних даних лінійних неітеративних нейроподібних структур МППП виконувалося з використанням вихідних сигналів шару сумування ймовірнісної нейронної мережі, де ймовірності у шарі сумування ЙМН формуються згідно (2.15). Для реалізації навчання і застосування розробленого методу обрано лінійну нейроподібну структуру МППП з неітеративним методом навчання [19].

Метод передбачав необхідність підбору значення єдиного оптимального параметра – розмаху гаусівської функції для ймовірнісної нейронної мережі, який у цій роботі виконувався з використанням методу диференційної еволюції.

Результати. Результати застосування методу на основі різних показників та значення оптимальних параметрів його роботи подано у таблиці 7.1.

Таблиця 7.1 – Результати застосування методу розширення входів з використанням вихідних сигналів шару підсумовування ймовірнісної нейронної мережі

Режим роботи методу / Показник ефективності	Точність	Влучність	Повнота	F1-міра
Режим навчання	0.93	0.94	0.96	0.95
Режим застосування	0,91	0,92	0,94	0,93
Значення оптимальних	<i>Формування набору ймовірностей ймовірнісної нейронної мережі: формула (2.15).</i>			

параметрів роботи методу	<i>Значення розмаху гаусівської функції для ЙМН: ($\sigma = 0.125922178$).</i> <i>Базова ШНМ для реалізації процедури навчання: Лінійна неітеративна нейроподібна структура МППП.</i> <i>Оптимізована кількість нейронів прихованого шару: 17.</i>

Як видно з результатів, поданих у таблиці 7.1, розроблений метод демонструє високі властивості узагальнення, а також високу точність класифікації по усіх досліджуваних показниках.

Порівняння. Оцінювання ефективності розробленого методу виконувалося порівнянням його роботи з ефективністю за рядом існуючих лінійних методів машинного навчання, а також його реалізації, в основі якої покладено використання ітеративного методу машинного навчання лінійного типу. Оптимальні параметри їх роботи подано у [223].

На рис. 7.5 подано результати порівняння значень F1-міри усіх досліджуваних методів.

Як видно з рис. 7.5, найменшу точність демонструють існуючі лінійні методи машинного навчання ітеративного типу та лінійна нейроподібна структура МППП з неітеративним навчанням (без розширення входів). Суттєво кращу точність, 85% за F1-мірою, показала класична ймовірнісна нейронна мережа, вихідні сигнали шару сумування якої формуються на основі використання (2.15). Значне підвищення точності також продемонструвала реалізація розробленого методу, побудована на основі машини опорних векторів з лінійним ядром (розроблений метод на основі МОВ з лінійним ядром). Тут значення F1-міри суттєво зросло в порівнянні з базовим класифікатором.

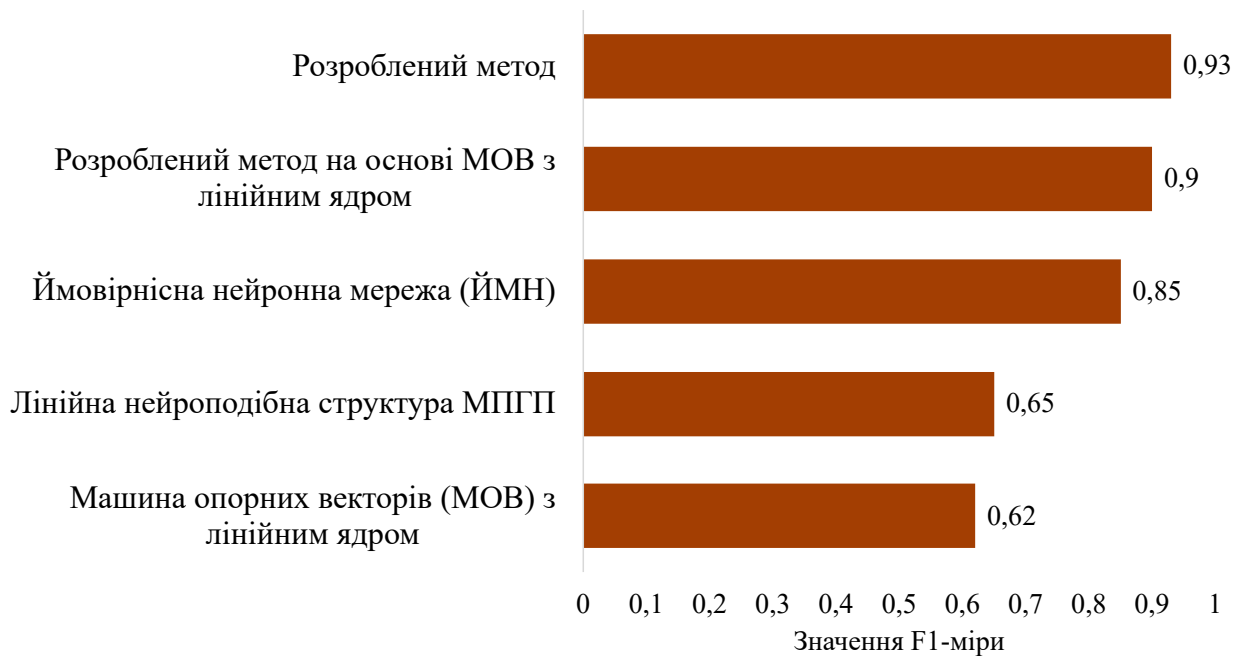


Рисунок 7.5 – Порівняння значень F1-міри для розробленого методу та інших досліджуваних методів

Однак найвищу точність за F1-мірою отримано з використанням лінійної неітеративної нейроподібної структури МПГП. Вона використовувалася для навчання і застосування розробленого методу. Це забезпечило можливість ефективного використання розробленого методу для розв’язання задачі класифікації якості матеріалів з титанових сплавів, призначених для виробництва медичних імплантатів.

7.2 Апробація методу кусково-лінійної апроксимації багатовимірних поверхонь відгуку в галузі ревматології

Метод кусково-лінійної апроксимації багатовимірних поверхонь відгуку з пункту 2.5 апробовано в процесі розв’язання задачі в галузі ревматології. Для реалізації цього методу розроблено спеціалізоване програмне забезпечення мовою програмування Python. Воно забезпечує побудову локальних лінійних моделей, автоматичне розбиття простору ознак на області апроксимації, вивід різних

показників ефективності як для режиму навчання, так і для режиму застосування для одного вектора даних з невідомим значенням вихідного сигналу. Застосування розробленого програмного засобу дозволило дослідити ефективність розвиненого методу в процесі аналізу вибірок даних малих обсягів.

Результати впроваджено в практичну діяльність Комунального некомерційного підприємства Харківської обласної ради “Обласна клінічна лікарня” у період з жовтня 2022 року по листопад 2024 року. Це підтверджено відповідним Актом про впровадження від 02.12.2024 р., виданим зазначеним підприємством (додаток Г).

Задача та її практична цінність. Розвинений метод застосовано для розв’язання задачі регресії для оцінювання біологічного віку зразка трабекулярної кістки людини (у роках) на основі структурно-функціональних характеристик тканини. Така оцінка є важливою складовою в рамках сучасного підходу до ранньої діагностики дегенеративних змін кісткової тканини, зокрема, остеопорозу та остеоартрозу. Ефективне розв’язання цієї задачі забезпечить можливість не тільки діагностувати наявність дегенеративних змін, але й прогнозувати їх прогресію. Це важливо для своєчасного втручання та корекції лікування. Метод відкриває також можливості для персоналізованого планування ортопедичного втручання, зокрема, індивідуального підбору імплантатів на основі оціненого стану кісткової тканини пацієнта.

Опис вибірки даних. Вибірка даних для моделювання була сформована на основі матеріалу, зібраного у відділенні травматології Комунального некомерційного підприємства Харківської обласної ради “Обласна клінічна лікарня”. Вона включає 72 зразки трабекулярної кістки, отримані від пацієнтів віком від 37 до 85 років (35 чоловіків, 37 жінок), яким було проведено тотальну ендопротезну заміну кульшового суглоба у зв’язку з клінічно діагностованим тяжким остеоартрозом. Кожен зразок описується п’ятьма вхідними ознаками: стисною міцністю, пористістю, статтю пацієнта, товщиною трабекул та індексом

моделі структури (SMI). Цільовим атрибутом для задачі регресії виступає біологічний вік кісткової тканини, визначений в роках..

Моделювання. Для моделювання розвинутого методу вхідні дані було попередньо нормалізовано шляхом ділення кожного значення на максимальне абсолютне значення відповідної ознаки в межах кожного стовпця. У якості базової моделі для реалізації методу навчання застосовано лінійну неітеративну нейроподібну структуру МППП, детальний опис якої наведено в першому розділі. Для оцінювання якості моделі використовувалась п'ятикратна кросвалідація, що дозволило забезпечити достовірність отриманих результатів та мінімізувати вплив випадкового розподілу вибірки..

Результати. Результати розвинутого методу на основі різних показників з Додатку А, а також значення оптимальних параметрів його роботи подано у таблиці 7.2.

Таблиця 7.2 – Результати роботи методу кусково-лінійної апроксимації багатовимірної поверхні відгуку для розв'язання задачі оцінювання біологічного віку зразка трабекулярної кістки людини (у роках) на основі структурно-функціональних характеристик тканини

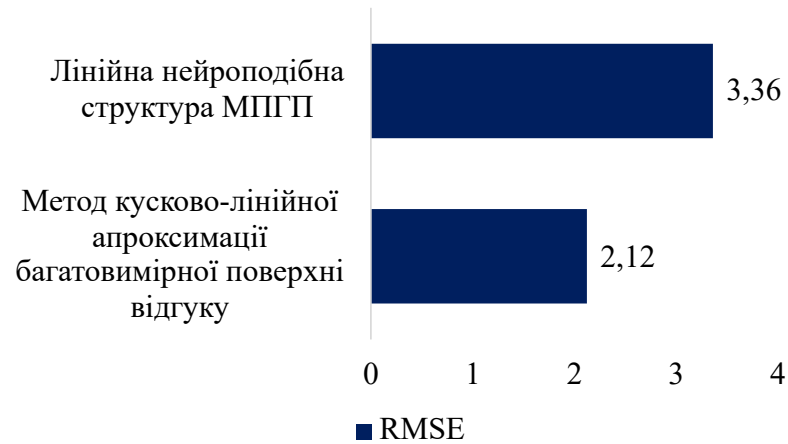
Показник	Режим навчання	Режим застосування	Значення оптимальних параметрів роботи методу
Коефіцієнт детермінації (R^2)	0.93	0.92	<i>Базова ШНМ для реалізації процедури навчання: Лінійна неітеративна нейроподібна структура МППП. Метод нормалізації за максимальним абсолютним значенням.</i>
Квадратний корінь з середньоквадратичної помилки (RMSE)	1.9	2.12	
Середня абсолютна відносна помилка (MAPE)	0.07	0.09	

Показник	Режим навчання	Режим застосування	Значення оптимальних параметрів роботи методу
Максимальна залишкова помилка (MaxError)	5.5	5.8	<i>Кількість частин на які розбивається кожна ознака: 2.</i>

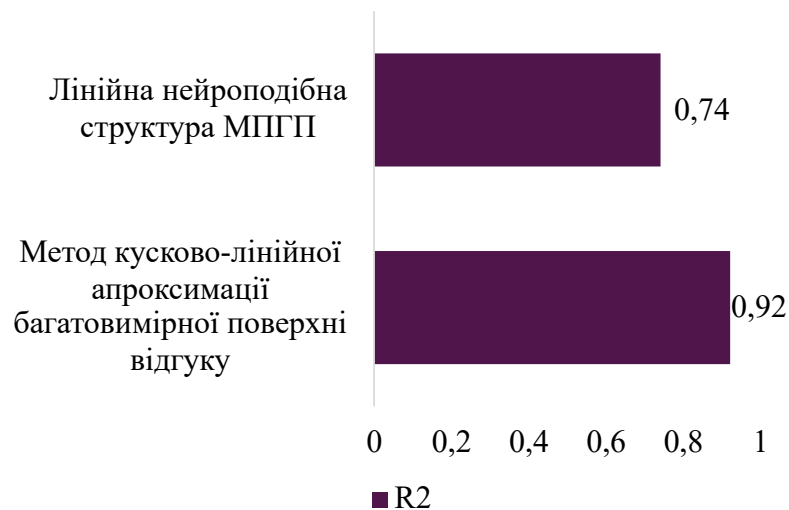
Як видно з результатів, поданих у таблиці 7.2, розвинутий метод демонструє високу точність роботи при збереженні належних властивостей узагальнення та повністю уникає проблеми перенавчання, яка властива існуючим методам нелінійного розширення входів в процесі аналізу коротких вибірок даних.

Порівняння. Оцінювання ефективності роботи розробленого методу проведено шляхом порівняння його результатів з результатами лінійної нейроподібної структури МПГП. Результати порівняння наведено на рис. 7.6.

Як видно з отриманих результатів, розвинутий метод забезпечує майже на 58% підвищення точності відповідно до показника RMSE порівняно з лінійною неітеративною нейроподібною структурою МПГП без розширення простору вхідних ознак. Це свідчить про те, що ізоморфне розбиття кожної ознаки початкової вибірки на дві частини дало змогу реалізувати локальне коригування поверхні відгуку. Це, у свою чергу, забезпечило суттєве підвищення точності розв'язання задачі при збереженні високих властивостей узагальнення, притаманної лінійним неітеративним нейроподібними структурам МПГП. Результати практичних експериментів також показали, що подальше збільшення кількості фрагментів ознак (до трьох і більше) не приводить до істотного покращення точності, натомість демонструє помітне зниження властивостей моделі до узагальнення.



а)



б)

Рисунок 7.6 – Порівняння точності роботи методу кусково-лінійної апроксимації багатовимірної поверхні відгуку з точністю за лінійною нейроподібною структурою МПГП: а) RMSE; б) R^2

Серед обмежень досліджуваного методу відмітимо незначне збільшення обчислювальних витрат, зумовлене розширенням вхідного простору. Проте з урахуванням аналізу екстремально-короткої вибірки даних, неітеративної природи лінійних нейроподібних структур МПГП та двократного зростання вимірності

простору ознак цей недолік є допустимим і не має істотного впливу на практичну ефективність методу.

7.3 Апробація методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів з використанням сформованої комбінованої поверхні відгуку в галузі медичної стоматології

Метод розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент векторів з використанням сформованої комбінованої поверхні відгуку, викладений у пункті 3.5 третього розділу (Метод 3.4). Він апробований в процесі розв’язання прикладної задачі в галузі медичної стоматології. Для реалізації цього та інших методів, наведених у третьому розділі, розроблено спеціалізовану бібліотеку засобів мовою програмування Python. Її захищено авторським правом на комп’ютерну програму (додаток Д). Бібліотека підтримує реалізацію всіх п’яти методів, розроблених у третьому розділі, і забезпечує можливість використання різних базових моделей навчання, зокрема, машини опорних векторів з РБФ-ядром, РБФ-нейронної мережі, нейронної мережі узагальненої регресії, ймовірнісної нейронної мережі, неітеративних нейроподібних структур з РБФ-розширенням вхідних ознак.

Цей метод, як і всі інші розроблені в роботі методи розв’язання задач регресії та класифікації на основі декартового квадрата множини вхідних компонент векторів, розроблено та використано в межах виконання наукового гранту EURIZON “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” № 871072, 01.04.2024 – 31.03.2025р., який фінансувався Європейською Комісією за програмою *Horizon Europe*. Практичне використання результатів досліджень підтверджено відповідним актом використання наукових результатів, №17244/3 від 17.03.2025 р., виданим Українським науково-технологічним центром (додаток Г).

Задача та її практична цінність. Розроблений метод використаний для оцінювання значення механічної властивості цирконієвої кераміки – в'язкості руйнування (тріщиностійкості). Такий підхід дозволяє оцінити цю характеристику матеріалу ще на етапі використання суміші порошків на основі оксиду цирконію, які потім спікаються для отримання кінцевого продукту. Ефективний розв'язок цієї задачі не лише покращує ще на етапі спікання якість виробів і продовжує їх термін служби, але й значно скорочує час, фінансові та матеріальні ресурси, зокрема, на високотехнологічне обладнання для експериментального визначення цієї механічної характеристики традиційними методами [91]. Це забезпечує виготовлення більш надійних функціональних матеріалів, що є критично важливим для медичних досліджень і клінічної практики [92].

Точні розв'язки для цієї задачі важливі для виготовлення стійких зубних реставрацій, імплантатів та коронок з цирконієвих матеріалів, стабілізованих ітрієм (Y-TZP). Висока тріщиностійкість цирконієвих матеріалів робить їх ідеальними для заміни металевих компонент у стоматології, забезпечуючи довговічність і естетичні переваги. Такий підхід відкриває нові можливості для покращення якості лікування та створення біомедичних імплантатів з покращеними механічними властивостями. Це гарантує їх зносостійкість, довговічність та здатність витримувати екстремальні умови експлуатації.

Опис вибірки даних. Екстремально-коротка вибірка даних, що використовувалась для моделювання, зібрана на кафедрі матеріалознавства та інженерії матеріалів Національного університету “Львівська політехніка” [302]. Вона містить 33 спостереження, кожне з яких включає кілька ключових параметрів. Вони характеризують оксидні керамічні матеріали на основі цирконію. Зокрема, досліджувана вибірка даних складається з 15 незалежних атрибутів, серед яких: фракції оксидів ZrO_2 , Y_2O_3 , MgO , CeO_2 , Al_2O_3 , TiO_2 , HfO_2 , SiO_2 та інші, середній розмір зерна, густина та температура спікання. Ці параметри подано у вигляді числових значень, де мол.% використовується для оксидних фракцій, мкм для

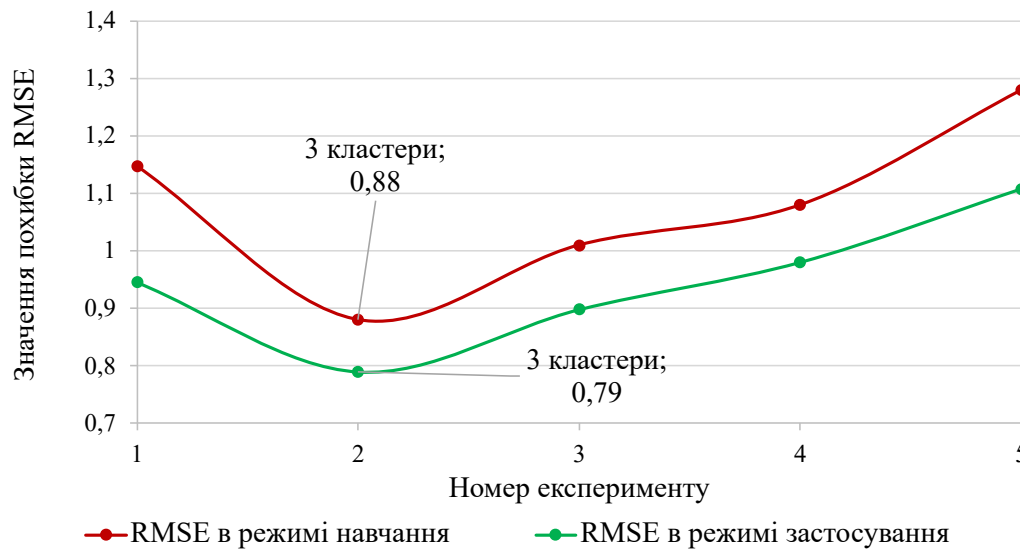
середнього розміру зерна, $\text{г}\cdot\text{см}^{-3}$ для густини та $^{\circ}\text{C}$ для температури спікання. Для фазового складу (кубічна, тетрагональна і/або моноклінна фази цирконію) застосовано бінарну систему, де одиниця позначає домінуючу фазу, а нуль – інші фази [92]. Сукупність цих параметрів дозволяє оцінювати механічні та функціональні властивості матеріалів, їх стійкість до зовнішніх навантажень і температурних змін [91]. Залежний атрибут – в'язкість руйнування (тріщиностійкість) цирконієвої кераміки. Високоточна оцінка його значення дає можливість оцінити властивості цирконієвої кераміки ще на етапі використання суміші порошків на основі оксиду цирконію, які спікаються для утворення кінцевого матеріалу.

Моделювання та налаштування параметрів методу. Моделювання розробленого методу реалізовано з використанням п'ятикратної кросвалідації. Обсяг вибірки даних становить 33 спостереження. Дані також попередньо нормалізувалися з використанням MaxAbsScaler.

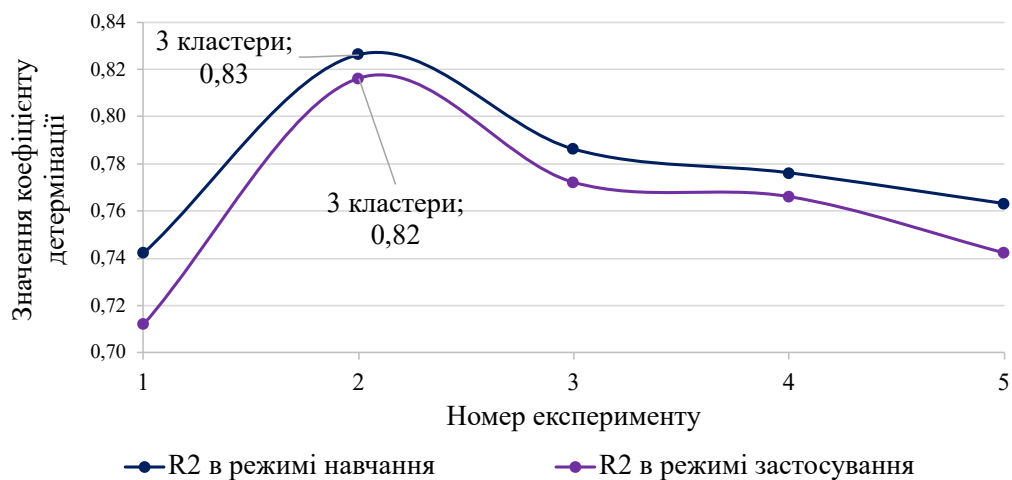
Для реалізації навчання використано неітеративну нейроподібну структуру МППП з РБФ-розширенням входів [288]. Для виокремлення кластерів у заданій вибірці даних обрано метод k-means через його простоту та швидкодію, легкість у налаштуванні та ефективність в процесі роботи з різними типами даних.

Оскільки розроблений метод базується на додатковому використанні кластеризації, підбір його оптимальних параметрів включав підбір оптимальної кількості РБФ-центрів, значення розмаху гаусівської функції та кількості кластерів. Оптимізація виконувалася з використанням методу диференційної еволюції.

Результати підбору оптимальної кількості кластерів на проміжку [2:6] з використанням RMSE і коефіцієнта детермінації R^2 для максимізації точності роботи методу приведено на рис. 7.7. Слід зазначити, що на осі абсцис відображено номер експерименту, де перший експеримент приведений для двох кластерів, другий – для трьох і т. д.



а)



б)

Рисунок 7.7 – Вибір оптимальної кількості кластерів для Методу 3.4 на основі: а) RMSE; б) коефіцієнту детермінації

Мінімальну похибку RMSE розробленого методу при розв’язанні задачі отримано для рьох кластерів (рис. 7.7а). Подальше збільшення їх кількості зменшувало точність роботи методу (рис. 7.7б).

Результати. Результати методу на основі показників параметрів його роботи подано у таблиці 7.3.

Таблиця 7.3 – Результати методу розв’язання задачі регресії на основі декартового квадрата множини вхідних компонент з використанням комбінованої поверхні відгуку

Режим роботи методу / Показник ефективності	MAPE	RMSE	MAE	R ²
Режим навчання	0.088	0.883	0.554	0.828
Режим застосування	0.097	0,791	0.468	0.816
Значення оптимальних параметрів роботи методу	<i>Кількість кластерів: 3.</i> <i>Базова ШИМ: Неітеративна нейроподібна структура МПГП з РБФ-розширенням входів.</i> <i>Кількість РБФ центрів: 46.</i> <i>Значення розмаху гаусівської функції: 4.5133646.</i> <i>Метод оптимізації – диференційна еволюція.</i>			

Як видно з результатів, поданих у таблиці 7.3, Метод 3.4 демонструє високу точність та високі властивості узагальнення в процесі розв’язання задачі оцінювання механічної властивості цирконієвої кераміки – в'язкості руйнування (тріщиностійкості) у випадку аналізу екстремально-короткої вибірки даних.

Порівняння. Оцінювання ефективності методу проведено на основі порівняння його результатів з результатами його реалізацій на основі ітеративних методів та базового методу 3.1 з розділу 3:

- розроблений метод з використанням машини опорних векторів з РБФ-ядром (*Розроблений метод на основі МОВ з РБФ-ядром*);
- метод 3.1 з використанням машини опорних векторів (МОВ) з РБФ-ядром (*Базовий метод на основі МОВ з РБФ-ядром*).

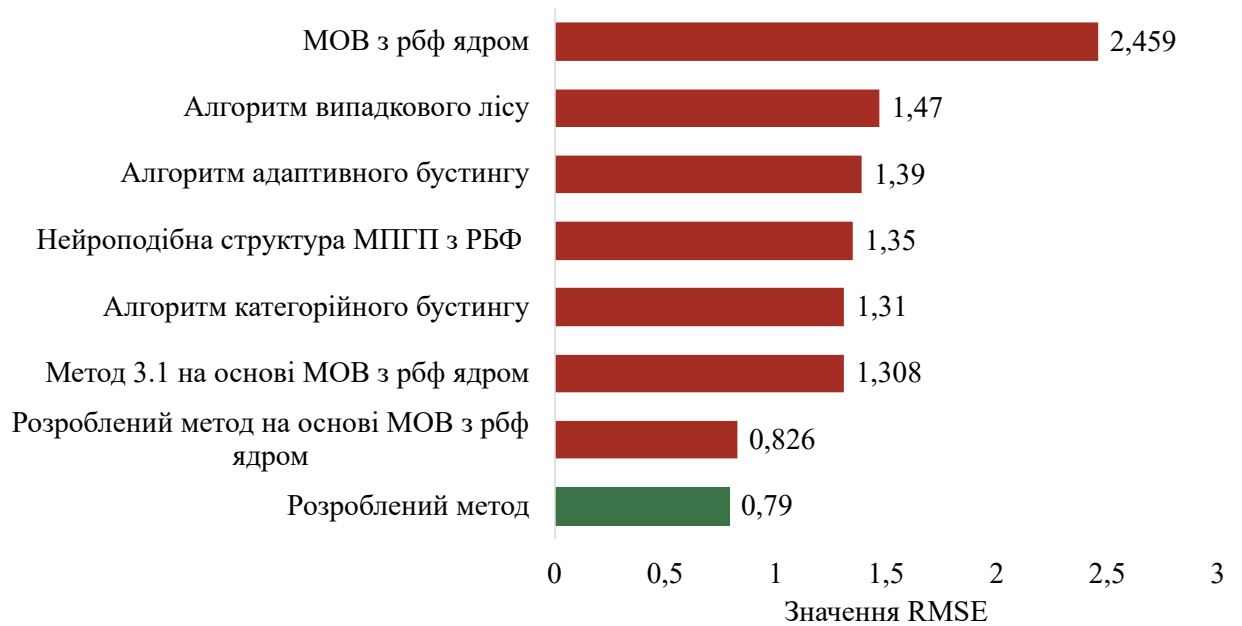
У процедурі порівняльного аналізу використовувалися також існуючі моделі нейронних мереж та методи машинного навчання:

- алгоритм категорійного бустінгу;
- алгоритм випадкового лісу;
- неітеративна нейроподібна структура Моделі Послідовних Геометричних Перетворень (МППГ) з РБФ-розширенням входів (*Нейроподібна структура МППГ з РБФ-ядром*);
- машина опорних векторів (МОВ) з РБФ-ядром (*МОВ (РБФ-ядро)*);
- алгоритм екстремального градієнтного бустінгу (*Алгоритм ЕГБ*).

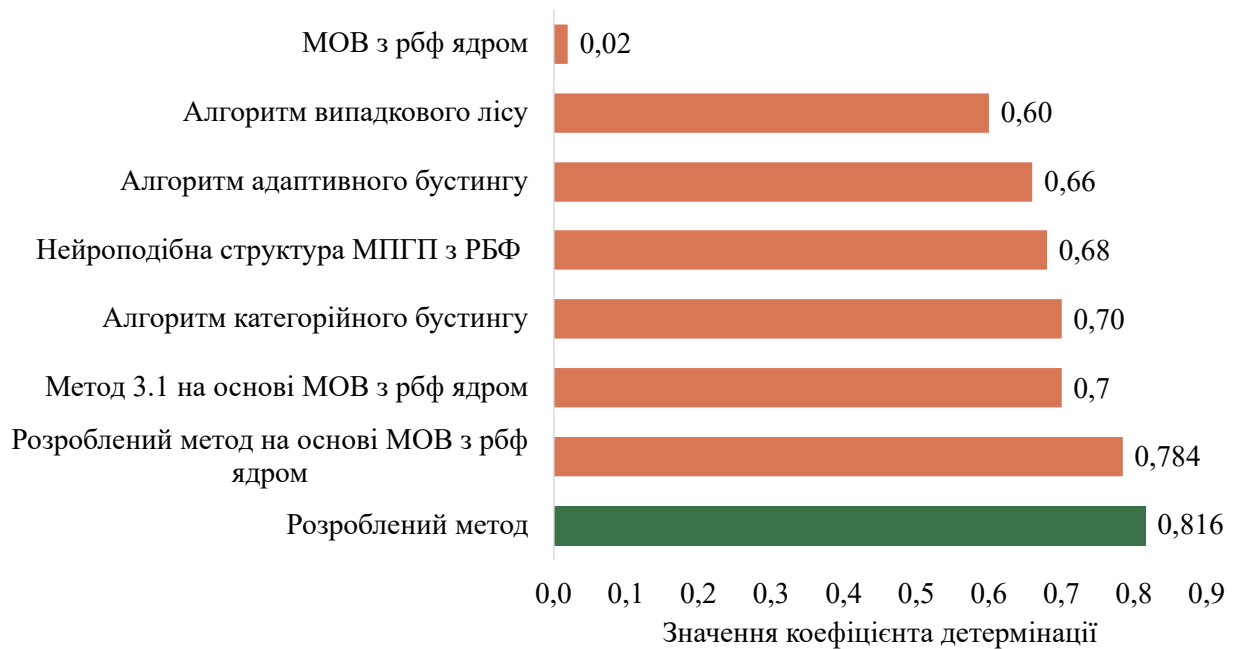
На рис. 7.8 подано порівняння двох показників ефективності для усіх досліджуваних методів. Їх оптимальні параметри наведено в додатку Б.

З результатів, поданих на рис. 7.8, можна зробити наступні висновки. Метод на основі МОВ з РБФ-ядром демонструє найнижчу точність і, зважаючи на значення коефіцієнта детермінації, не є придатним для розв'язання п задачі. Однак, Методи 3.1 та розроблений метод на основі МОВ з РБФ-ядром показали суттєве покращення точності, понад 70%. Це свідчить про потенційну ефективність використання ітеративних методів машинного навчання для реалізації навчання в розробленому методі.

Застосування існуючих методів ансамблевого навчання з різних класів, включаючи алгоритми випадкового лісу, адаптивного бустінгу та категорійного бустінгу, а також нейроподібної структури МППГ з РБФ-розширенням входів демонструє точність роботи згідно з коефіцієнтом детермінації на рівні 70% і менше. Однак у роботі [91], яка спеціалізується на дослідженні порошків на основі оксиду цирконію, зазначено, що така точність є достатньою для практичного використання методів машинного навчання.



а)



б)

Рисунок 7.8 – Порівняння точності розробленого методу з точністю за іншими методами на основі значень: а) RMSE; б) коефіцієнту детермінації

Розроблений метод з неітеративним навчанням забезпечив найвищу точність роботи як за показником RMSE, так і за коефіцієнтом детермінації. Застосування цього методу дозволило також підвищити точність розв'язання задачі на понад 16%, порівняно з існуючим методом який показує 70% точності [91]. Ураховуючи ці результати та інші показники ефективності, наведені в таблиці 7.3, використання розробленого методу є доцільним для розв'язання задач передбачення механічної властивості цирконієвої кераміки – в'язкості руйнування (тріщиностійкості), особливо у випадку аналізу екстремально коротких вибірок даних.

7.4 Апробація методу класифікації даних на основі ансамблю ймовірнісних нейронних мереж з використанням симетричних випадкових зміщень у галузі травматології

Метод розв'язання задачі класифікації даних на основі ансамблю ймовірнісних нейронних мереж з використанням симетричних випадкових зміщень апробовано в галузі травматології. Для його реалізації розроблено програмний засіб мовою Python. Створене програмне забезпечення підтримує можливість синтезу гомогенних ансамблів на основі нейронної мережі узагальненої регресії для задач регресії, ймовірнісної нейронної мережі для задач класифікації, неітеративних нейроподібних структур з РБФ-розширенням вхідних ознак для задач регресії і класифікації. Програмний засіб також реалізує механізми формування декорельованих вибірок з використанням методів випадкових і симетричних випадкових зміщень. Агрегування результатів реалізоване трьома способами: середнім арифметичним для регресійних задач, голосуванням більшості для задач класифікації, з використанням мета-моделі у вигляді лінійної неітеративної нейроподібної структури МПГП для обох типів задач. Програмний засіб підтримує також можливість включення початкових вхідних ознак у мета-вибірку для побудови мета-моделі. Це потенційно підвищує апроксимаційні властивості синтезованого ансамблю [57].

Результати впроваджено в діяльність Військово-медичного клінічного центру Західного регіону в період з січня 2024 року по січень 2025 року. Впровадження підтверджено відповідним Актом впровадження від 22.04.2025 р., виданим зазначеним медичним закладом (додаток Г).

Задача та її практична цінність. Розроблений метод використовувався для прогнозування летальності пацієнтів з політравмами на основі 56 клінічних та лабораторних аналізів для оцінювання ризику смерті, оптимізації тріажу в умовах інтенсивної терапії, моніторингу ефективності лікування та підтримки прийняття рішень лікарями.

Опис вибірки даних. Екстремально-коротка вибірка даних, що використовувалась для моделювання, зібрана фахівцями кафедри анестезіології та інтенсивної терапії Харківської міської клінічної лікарні швидкої та невідкладної медичної допомоги ім. проф. О. І. Мещанінова. Вибірка містить 73 зразки пацієнтів-чоловіків, госпіталізованих внаслідок політравм. У вибірці подано 56 незалежних атрибутів, що охоплюють різноманітні клінічні та лабораторні показники. Залежною змінною є летальність, позначена значеннями один (летальний випадок) та нуль (виживання) із загальної кількості 73 випадки: 31 виявився летальним, а 42 – вижили.

Екстремально-коротка вибірка даних, використана для моделювання, сформована фахівцями кафедри анестезіології та інтенсивної терапії Харківської міської клінічної лікарні швидкої та невідкладної медичної допомоги імені професора О. І. Мещанінова. Вибірка містить 73 записи, що стосуються пацієнтів чоловічої статі, які були госпіталізовані з політравмами різного ступеня тяжкості. Для кожного пацієнта зібрано 56 незалежних змінних, які охоплюють широкий спектр клініко-фізіологічних та лабораторно-діагностичних показників, що зазвичай використовуються у практиці інтенсивної терапії для оцінки стану пацієнтів у критичних станах.

Цільовою (залежною) змінною обрано показник летальності, який подано у

бінарному вигляді. Таким чином, вибірка характеризується не тільки обмеженим обсягом, але й помірним дисбалансом класів. Це накладає додаткові обмеження на використання традиційних статистичних підходів та обґрунтовує доцільність застосування спеціалізованих методів машинного навчання, зокрема, ансамблевих, адаптованих до роботи з даними малих обсягів.

Моделювання та налаштування параметрів методу. Моделювання розробленого методу реалізовано двома варіантами: шляхом усереднення та з використанням стекового підходу до агрегування вихідних сигналів моделей ЙНМ ансамблю. В якості мета-моделі використано лінійну неітеративну нейроподібну структуру МППП.

Базовими моделями ансамблю використано ймовірнісні нейронні мережі, параметри яких узгоджувалися попередньо. Зокрема, для визначення оптимального значення параметра розмаху гаусівської функції, що використовується у ймовірнісній нейронній мережі, проведено п'ятикратну крос-валідацію на початковій навчальній вибірці. Отримане оптимальне значення цього параметра використовувалося для базових моделей ансамблю.

Перед навчанням моделі вибірку даних було нормалізовано шляхом поділу значень кожного параметра на його максимальне абсолютне значення у відповідному стовпці таблиці даних. Для створення унікальних, декорельованих вибірок даних використовувався метод випадкових симетричних зміщень (4.43)-(4.47).

Результати. Результати реалізації методу на основі показників (додаток А) та значення оптимальних параметрів його роботи подано у таблиці 7.4.

Як видно з результатів, поданих у таблиці 7.3, обидва ансамблі (усередненням та стекінговий ансамбль) досягли максимальної повноти (1.00). Це вказує на їх здатність виявляти всі істинно позитивні класи. Проте стекінговий ансамбль ймовірнісних нейронних мереж забезпечив вищу точність за усіма показниками.

Таблиця 7.4 – Результати реалізації методу розв’язання задачі класифікації даних на основі ансамблів ймовірнісних нейронних мереж з використанням симетричних випадкових зміщень у галузі травматології

Метод	Точність	Влучність	Повнота	F1-міра	Коефіцієнт кореляції Метьюса	Каппа Коена
Розроблений ансамбль (усереднення)	0.9	0.8	1	0.89	0.88	0.87
Розроблений ансамбль (стекинг)	0.92	0.84	1	0.91	0.9	0.89
Оптимальні параметри	<i>Базова ШНМ: Ймовірнісна нейронна мережа.</i> <i>Значення розмаху гаусівської функції для усереднення: 0.283894427.</i> <i>Значення розмаху гаусівської функції для стекингового ансамблю: 0.345562311.</i> <i>Метод декореляції: метод симетричних випадкових відхилень.</i> <i>Величина відхилень: 0.1.</i> <i>Кількість елементів ансамблю: 11 (для ансамблю з усередненням).</i> <i>Кількість елементів ансамблю: 5 (для стекингового ансамблю).</i> <i>Мета-модель: лінійна неітеративна нейроподібна структура МПГП (для стекингового ансамблю).</i> <i>Метод оптимізації: диференційна еволюція.</i>					

Вищі показники стекінгової моделі пояснюються здатністю мета-моделі ефективно комбінувати вихідні сигнали базових моделей з урахуванням їх сильних і слабких сторін. Водночас стекінг вимагає більш обережного застосування в умовах екстремально коротких вибірок через ризик перенавчання мета-моделі, тоді як усереднення забезпечує вищу стійкість, але дещо нижчу точність.

Порівняння. Оцінювання ефективності методу проведено на основі порівняння його результатів з результатами ряду відомих методів.

На рис. 7.9 подано порівняння двох показників ефективності для усіх досліджуваних методів.

З аналізу результатів, наведених на рис. 7.9, можна зробити такі висновки. Метод опорних векторів з РБФ-ядром продемонстрував найнижчу точність і, з огляду на значення коефіцієнта кореляції Метьюса, не є придатним для розв'язання поставленої задачі. Ймовірнісна нейронна мережа забезпечила доволі високі показники точності класифікації. Ансамбль, побудований на основі усереднення результатів декількох ймовірнісних нейронних мереж, продемонстрував вищу ефективність за коефіцієнтом кореляції Метьюса при збереженні подібного рівня влучності, повноти та F1-міри. Найкращі результати показав ансамбль, реалізований з використанням стекінгового механізму агрегації результатів, у якому вибірки для базових моделей формувалися з використанням симетричних випадкових відхилень. Цей підхід забезпечив найвищу точність за всіма досліджуваними метриками. Зокрема, він забезпечив підвищення точності за коефіцієнтом кореляції Метьюса на майже 6% порівняно з базовою ймовірнісною нейронною мережею.

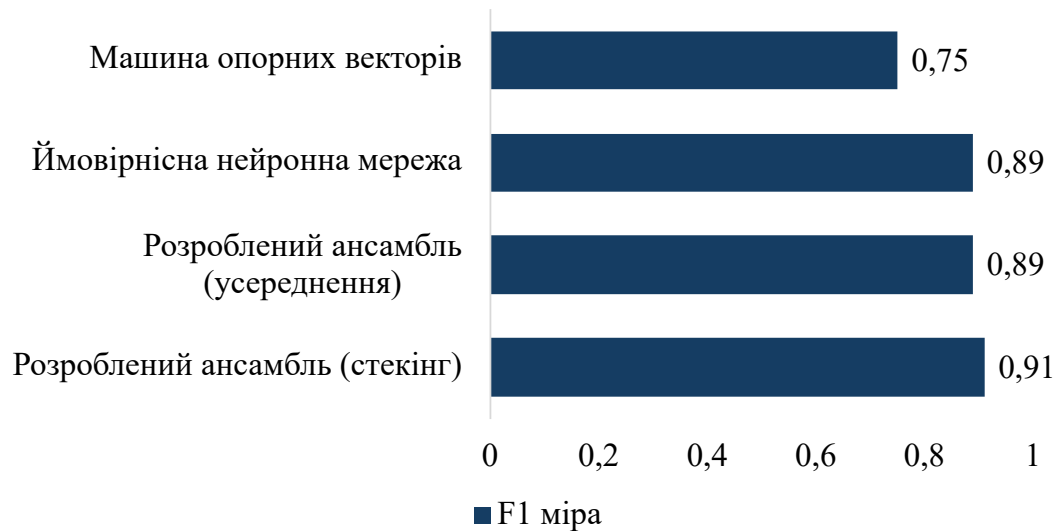
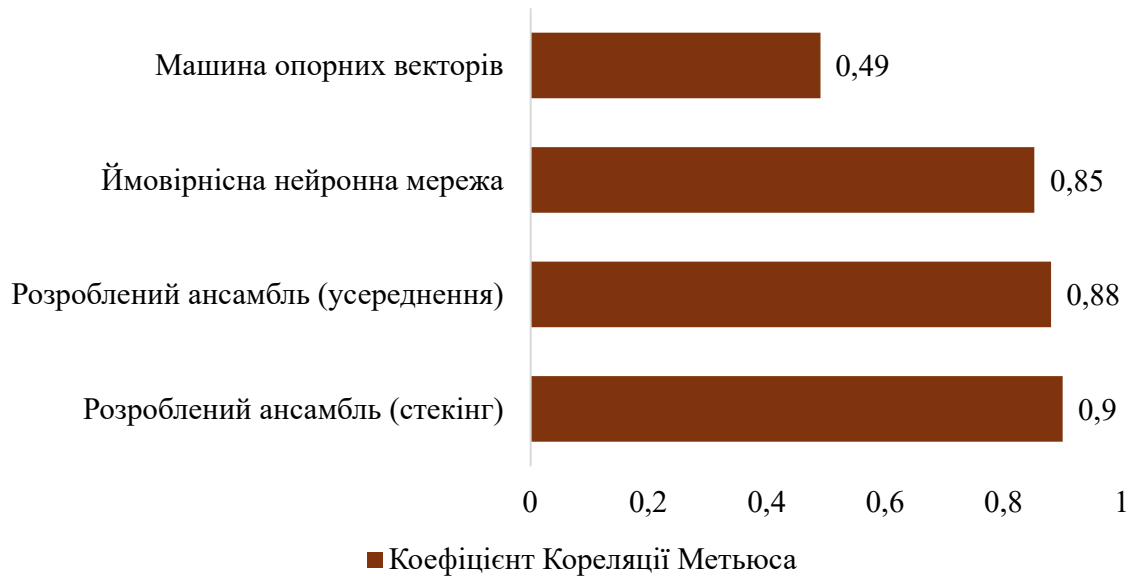


Рисунок 7.9 – Порівняння ефективності двох варіантів розробленого ансамблю з ефективністю за іншими методами на основі значень: а) коефіцієнта кореляції Метьюса; б) F1–міри.

Аналіз отриманих результатів підтвердив, що саме цей метод може бути обраний для практичного впровадження при розв’язанні поставленої задачі.

7.5 Апробація методу класифікації даних на основі ансамблю неітеративних нейроподібних структур з РБФ-розширенням входів з використанням симетричних випадкових зміщень у галузі трансплантології

Метод розв’язання задачі класифікації на основі ансамблю неітеративних нейроподібних структур з РБФ-розширенням входів, що використовує симетричні випадкові зміщення для декореляції вибірок даних, апробовано в галузі трансплантології. Для його реалізації розроблено програмне забезпечення мовою Python. Воно підтримує два підходи до формування незалежних підвибірок даних для кожного члена ансамблю відповідно до (4.27) – (4.31) або (4.43) – (4.47). В якості базових моделей можуть використовуватись нейронна мережа узагальненої регресії, ймовірнісна нейронна мережа та неітеративна нейроподібна структура з РБФ-розширенням входів. У ролі мета-моделі передбачено використання лінійної неітеративної нейроподібної структури МПГП. Програмне рішення також дозволяє використовувати початкові вхідні ознаки задачі в процесі формування вибірки для мета-моделі. Це може забезпечити підвищення точності апроксимації, особливо в умовах обмеженої кількості незалежних атрибутів у вхідних даних. Розроблене програмне забезпечення захищено свідоцтвом про авторське право на комп’ютерну програму (додаток Д).

Результати впроваджено в лікувальну роботу Комунального некомерційного підприємства Харківської обласної ради “Обласний медичний клінічний центр урології і нефрології ім. В. І. Шаповала” у період з жовтня 2022 року по вересень 2024 року. Впровадження підтверджено відповідним Актом про впровадження, виданим зазначеним закладом (додаток Г).

Задача та її практична цінність. Розроблений метод використовувався для оцінювання виживаності трансплантату нирки впродовж одного місяця після трансплантації. Ця задача надзвичайно актуальна у сучасній трансплантології,

оскільки саме цей період є критичним для виникнення можливості відторгнення органу, загострення хронічних захворювань або ускладнень, пов'язаних з імуносупресивною терапією. Точне та своєчасне оцінювання ризиків відторгнення трансплантату дозволяє медичним фахівцям приймати обґрунтовані рішення щодо модифікації терапевтичних протоколів, індивідуалізації лікування та покращення післяопераційного моніторингу.

Застосування методів машинного навчання у цій задачі забезпечує можливість виявлення прихованих закономірностей у високовимірних клінічних даних, які складно або неможливо виявити за допомогою традиційного статистичного аналізу.

Опис вибірки даних. Вибірка даних містить 146 пацієнтів, які отримали HLA-сумісні трансплантати нирки в період з 1992 року по 2020 рік [303]. Усі 146 пацієнтів проходили первинну трансплантацію: 55 жінок (37,6%) і 91 чоловік (62,3%) з середнім віком $32,8 \pm 8,4$ року в діапазоні від 18 до 60 років на момент проведення операції. Усі пацієнти отримували амбулаторне лікування у відділенні нефрології та діалізу Львівської обласної клінічної лікарні (ЛОКЛ) через один місяць після трансплантації нирки [304], [305].

Вибірка містить групу пацієнтів, які пережили перший місяць після трансплантації нирки, отримували підтримувальну імуносупресивну терапію та проходили першу трансплантацію. У 79,8% випадків операції з пересадки нирки проводилися в Україні (Київ, Запоріжжя, Харків, Львів, Одеса), ще у 20,2% випадків – в інших країнах (Італія, Білорусь, Пакистан, Польща, Туреччина, Китай). Усі пацієнти, які проходили трансплантацію в Україні, отримали нирку від родинного донора, тоді як 14,5% пацієнтів (усі з-за кордону) – від посмертного донора. Вибірка даних містить 40 ознак та один залежний атрибут [304], [305].

Клінічне обстеження пацієнтів (збір скарг, медичного та життєвого анамнезу, об'єктивне обстеження) проводилося у відділенні нефрології та діалізу ЛОКЛ через місяць після трансплантації. Лабораторні дослідження загальноклінічних

показників (загальні аналізи крові й сечі, біохімічні аналізи крові) виконувалися у центральній лабораторії ЛОКЛ відповідно до інструкцій діагностичних наборів. Дані щодо попередніх порушень функції нирок та інші медичні відомості отримані з особистих історій хвороби пацієнтів [304], [305].

Наявність вірусних інфекцій визначалася за допомогою імуноферментного аналізу та стандартних методів полімеразної ланцюгової реакції у приватних імунологічних лабораторіях [304], [305].

Моделювання та налаштування параметрів методу. Моделювання розробленого методу проведено з використанням стекінгового підходу до агрегування вихідних сигналів членів ансамблю. В якості мета-моделі використано лінійну неітеративну нейроподібну структуру МПГП.

Базовими моделями ансамблю використано нейроподібні структури МПГП з РБФ-розширенням входів, параметри яких визначалися попередньо. Зокрема, для визначення кількості РБФ-центрів та оптимального значення параметра розмаху гаусівської функції виконано п'ятикратну крос-валідацію на початковій навчальній вибірці даних. Отримане оптимальне значення цього параметра використовувалося для всіх базових моделей ансамблю.

Перед навчанням моделі вибірку даних нормалізовано шляхом поділу значень кожного параметра на його максимальне абсолютне значення у відповідному стовпці таблиці даних. Для створення унікальних, декорельованих вибірок даних використовувався метод випадкових симетричних зміщень відповідно до (4.43)-(4.47).

Результати. Результати методу на основі показників (додаток А) та значення оптимальних параметрів його роботи подано у таблиці 7.5.

Як видно з результатів, поданих у таблиці 7.3, розроблений стекінговий ансамбль неітеративних нейроподібних структур МПГП з РБФ-розширенням входів забезпечує високу точність та властивості узагальнення за усіма показниками.

Таблиця 7.5 – Результати методу розв’язання задачі класифікації даних на основі ансамблю неітеративних нейроподібних структур з РБФ-розширенням входів з використанням симетричних випадкових зміщень у галузі трансплантології

Метод	Точність	Влучність	Повнота	F1-міра	Коефіцієнт кореляції Метьюса	Каппа Коена
Розроблений ансамбль (режим навчання)	0.94	0.92	0.94	0.93	0.88	0.87
Розроблений ансамбль (режим застосування)	0.92	0.90	0.92	0.91	0.85	0.84
Оптимальні параметри	<i>Базова ШНМ:</i> Неітеративна нейроподібна структура МПГП з РБФ-розширенням входів. <i>Кількість РБФ-центрів:</i> 49. <i>Значення розмаху гаусівської функції:</i> 0.194587956. <i>Метод декореляції:</i> метод симетричних випадкових відхилень. <i>Величина відхилень:</i> 0.1. <i>Кількість елементів ансамблю:</i> 9. <i>Мета-модель:</i> лінійна неітеративна нейроподібна структура МПГП. <i>Метод оптимізації:</i> диференційна еволюція.					

Порівняння. Оцінювання ефективності методу проведено на основі порівняння його результатів з результатами за іншими методами. На рис. 7.10 подано порівняння двох показників ефективності для досліджуваних методів.

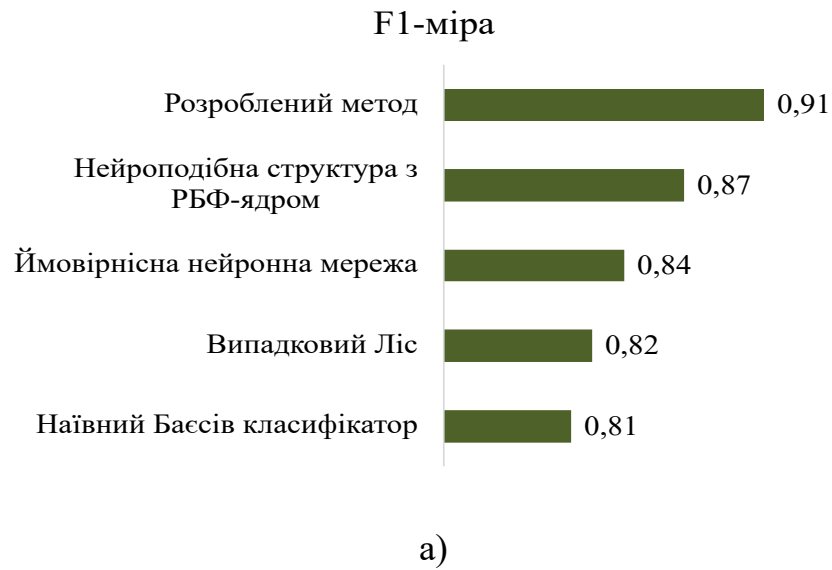


Рисунок 7.10 – Порівняння ефективності розробленого ансамблю з ефективністю за іншими методами на основі значень: а) F1-міри; б) коефіцієнта кореляції Метьюса

З аналізу результатів, наведених на рис. 7.10, можна зробити такі висновки. Наївний Баєсів класифікатор продемонстрував найнижчу точність і, з огляду на значення F1-міри, непридатний для розв'язання задачі. Дещо вищу точність демонструє метод Випадкового Лісу та ймовірнісна нейронна мережа. Ітеративна нейроподібна структура МПГП з РБФ-розширенням входів забезпечила високу

точність розв'язання задачі класифікації, а ансамбль з таких структур дозволив отримати точність класифікації більше 91% на основі F1-міри. Розроблений метод забезпечив найвищу точність за всіма іншими досліджуваними метриками. Окрім цього, використання ансамблю суттєвим чином зменшує дисперсію в порівнянні з використанням однієї моделі. Також він є більш робастним до аномалей та викидів у даних за рахунок синтезу вибірок для кожного члена ансамблю методом симетричних випадкових зміщень. З огляду на отримані результати саме цей метод був обраний для практичного впровадження при розв'язанні поставленої задачі.

7.6 Етичні аспекти застосування розроблених та удосконалених методів інтелектуального аналізу біомедичних даних малих обсягів

У процесі застосування методів інтелектуального аналізу біомедичних даних, особливо в умовах малих вибірок даних, виникають важливі етичні виклики, пов'язані з конфіденційністю персональної медичної інформації, надійністю отриманих висновків та їх клінічною інтерпретацією. Зокрема, обмеженість обсягу даних підвищує ризик помилок другого роду (повнота). Це може призводити, наприклад, до хибної класифікації критичних станів пацієнтів. Окрім цього, навіть точні моделі можуть мати недостатню властивість узагальнення. Це створює ряд потенційних загроз при використанні розроблених методів на нових невідомих даних. Тому обов'язковим є дотримання принципів етичної відповідальності в обробці біомедичних даних, включаючи верифікацію моделей незалежними фахівцями, зокрема, з залученням лікарів-експертів, а також оцінку меж застосування методів та проведення додаткових клінічних перевірок.

Розроблені та удосконалені в дисертаційній роботі методи базуються на використанні аугментації даних, декореляції даних, принципів каскадування та компенсації похибок. Все це загалом підвищує не тільки точність, але й стійкість отриманих результатів. Такі методи зменшують ризики, пов'язані з некоректною інтерпретацією результатів, і забезпечують більшу прозорість та інтерпретованість

моделей. Для етичного впровадження в клінічну практику розроблені програмні засоби пройшли апробацію в медичних установах та супроводжувалися інструкціями щодо інтерпретації результатів. Окрім цього, розроблені програмні засоби не передбачають автоматичного ухвалення клінічних рішень, а натомість функціонують як допоміжні інструменти підтримки лікарських рішень, формування первинних гіпотез та висновків. Такий підхід відповідає принципу дотримання автономії пацієнта та професійної відповідальності медичних фахівців. В процесі апробації розроблених методів, зокрема, у співпраці з медичними установами забезпечено дотримання етичних норм, включно з процедурою анонімізації персональних даних. Таким чином, етичні аспекти були враховані як на етапі моделювання, так і в процесі практичного впровадження розроблених та удосконалених методів. Це забезпечує відповідальне їх використання в біомедичних дослідженнях.

7.7 Висновки до розділу

1. Результати проведеного моделювання підтверджують ефективність розробленого методу класифікації матеріалів з порошкових титанових сплавів, що базується на розширенні входів лінійної неітеративної нейроподібної структури МПГП вихідними сигналами шару підсумовування ймовірнісної нейронної мережі. Розроблений метод забезпечив високі показники точності, повноти, влучності та F1-міри у режимах як навчання, так і застосування, суттєво перевищивши результати базових та ітеративних лінійних моделей. Застосування методу для розв'язання поставленої задачі дозволяє підвищити точність класифікації матеріалів ще на етапі підготовки сировини. Це є критично важливим для виробництва біосумісних медичних імплантатів з необхідними експлуатаційними властивостями. Результати впроваджено на кафедрі стоматології дитячого віку та

імплантології Харківського національного медичного університету впродовж 2022–2024 років.

2. Результати апробації методу кусково-лінійної апроксимації багатовимірної поверхні відгуку показали його високу ефективність для задачі оцінювання (задача регресії) біологічного віку трабекулярної кістки людини на основі структурно-функціональних характеристик тканини в умовах аналізу вибірки даних малого обсягу. Проведене моделювання на основі реальної вибірки даних з 72 зразків кісткової тканини, зібраних у клінічних умовах, продемонструвало високу точність прогнозування ($R^2 = 0.92$) та майже на 58% зниження середньоквадратичної помилки порівняно з лінійною нейроподібною структурою МПГП без розширення простору ознак. Отримані результати підтверджують, що просте ізоморфне розбиття простору ознак (кожна ознака розбивалася на дві частини) дозволяє забезпечити адаптивну локальну корекцію поверхні відгуку при збереженні високих властивостей узагальнення моделі. Результати впроваджено в практичну діяльність Комунального некомерційного підприємства Харківської обласної ради “Обласна клінічна лікарня” у період з жовтня 2022 року по листопад 2024 року.

3. Результати проведеного моделювання підтверджують ефективність використання розробленого методу розв’язання задачі оцінювання механічних властивостей цирконієвої кераміки на основі декартового квадрата множини вхідних компонент векторів з формуванням комбінованої поверхні відгуку в умовах екстремально коротких вибірок даних. Метод продемонстрував високу точність ($R^2 = 0.816$) в порівнянні з існуючими методами машинного навчання та ітеративною його реалізацією (на основі машини опорних векторів з РБФ ядром). Окрім цього, розроблений метод демонструє на 16% вищу точність на основі коефіцієнта детермінації порівняно з існуючим рішенням ($R^2 = 0.7$). Це підтверджує доцільність

використання методу для підвищення точності оцінювання тріщиностійкості цирконієвих матеріалів в умовах аналізу даних обмеженого обсягу. Метод розроблено та використано в межах виконання наукового гранту EURIZON: “Artificial Intelligence in Medicine: Small Data Approach (SmallAIM)” (№ 871072, 01.04.2024 – 31.03.2025), який фінансувався Європейською Комісією за програмою *Horizon Europe*.

4. Результати апробації розробленого методу класифікації на основі ансамблю ймовірнісних нейронних мереж з використанням симетричних випадкових зміщень підтверджують його ефективність для задачі оцінювання летальності пацієнтів з політравмами в умовах екстремально коротких вибірок даних (задача класифікації). Розроблений метод досяг максимальної повноти (1.0) для обох типів ансамблів і забезпечив найвищі значення точності, F1-міри та коефіцієнта кореляції Метьюса у стекінговій конфігурації серед усіх досліджуваних методів. Результати впроваджено в практичну діяльність Військово-медичного клінічного центру Західного регіону в період з січня 2024 року по січень 2025 року.

5. Результати апробації методу класифікації на основі ансамблю неітеративних нейроподібних структур з РБФ-розширенням входів, що використовує симетричні випадкові зміщення для декореляції вибірок, показали його високу ефективність при розв’язанні задачі оцінювання виживаності трансплантату нирки впродовж одного місяця після операції. Проведене моделювання на основі реальної клінічної вибірки даних обмеженого обсягу засвідчило високу точність класифікації у режимах як навчання, так і застосування з одночасним підвищенням значень коефіцієнта кореляції Метьюса (до 0.85) та показника Каппи Коена (до 0.84) порівняно з існуючими методами. Важливою перевагою розробленого методу є його робастність до аномалій і викидів у даних. Це досягається завдяки формуванню незалежних підвибірок з використанням

методу симетричних випадкових зміщень. Результати впроваджено в лікувальну роботу Комунального некомерційного підприємства Харківської обласної ради “Обласний медичний клінічний центр урології і нефрології ім. В. І. Шаповала” у період з жовтня 2022 року по вересень 2024 року.

6. Розроблені програмні засоби, які реалізують досліджувані методи в цьому розділі, захищено авторським правом на комп’ютерні програми. Для етичного впровадження в клінічну практику програмні засоби супроводжувалися інструкціями щодо інтерпретації результатів. Розроблені програмні засоби не призначені для автоматичного ухвалення клінічних рішень, а натомість функціонують як допоміжні інструменти підтримки лікарських рішень, формування первинних гіпотез і попередніх висновків. Етичні аспекти були враховані як на етапі моделювання, так і в процесі практичного впровадження розроблених та удосконалених методів, що забезпечує безпечне і відповідальне їх використання в біомедичних дослідженнях.

ВИСНОВКИ

У дисертаційній роботі на основі виконання теоретичних та експериментальних досліджень вирішено важливу теоретичну проблему розвитку основ теорії синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів. При цьому отримано такі основні результати та висновки:

1. Проведено аналіз методів та засобів інтелектуального опрацювання структурованих біомедичних даних малого обсягу та обґрунтовано необхідність синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням для підвищення точності і повторюваності результатів їх обробки.

2. Набув подальшого розвитку метод кусково-лінійної апроксимації для випадку аналізу багатовимірних поверхонь відгуку, який за рахунок ізоморфного подання кожного значення ознаки у вигляді її локальних компонент дав можливість збільшити вимірність простору вхідних даних та забезпечити локальну корекцію параметрів поверхні відгуку лінійною неітеративною нейроподібною структурою МППП, що дозволяє зменшити похибку лінійної моделі на 58% згідно RMSE.

3. Удосконалено метод формування комбінованої поверхні відгуку з застосуванням кластерного аналізу, який на відміну від існуючого доповнює простір початкових ознак тільки значенням залежного атрибута центра кластера, що забезпечує мінімальне зростання вимірності простору ознак та підвищує точність моделей регресії і класифікації на малих вибірках без втрати здатності до узагальнення.

4. Розроблено метод нелінійного розширення простору незалежних ознак біомедичної вибірки даних, який внаслідок їх доповнення вихідними сигналами шару підсумовування ймовірнісної нейронної мережі дозволяє підвищити точність класифікації за рахунок використання лінійних неітеративних нейроподібних структур МПГП на 43% та 9% згідно F1-міри у порівнянні з обома формуючими ШНМ при збереженні властивостей узагальнення.

5. Розроблена методологія розв'язання задач регресії та бінарної класифікації на екстремально-коротких вибірках біомедичних даних шляхом аугментації даних на основі декартового квадрата множини вхідних компонент векторів дозволяє застосовувати різнотипні нелінійні моделі машинного навчання, зокрема, неітеративні нейроподібні структури МПГП з RBF-розширенням входів для реалізації навчання та забезпечує:

- покращення властивостей узагальнення та зниження фактору перенавчання базового методу завдяки непрямій постановці задачі регресії, яка дозволяє моделі змістити акценти на відносні відмінності між спостереженнями і підвищити точність розв'язання задач регресії нелінійними ШНМ;
- підвищення робастності до аномалій та викидів у даних з використанням адитивного методу за рахунок можливості часткової взаємної компенсації похибок з різними знаками при формуванні значення шуканого сигналу методом усереднення або голосування більшістю, а також зменшення похибки RMSE методу до 59 % в процесі розв'язання задач бінарної класифікації порівняно з методом опорних векторів;
- підвищену здатність до апроксимації суттєво нелінійних поверхонь відгуку з використанням каскадного методу за рахунок додаткового розширення вхідного простору початкової вибірки даних

результатами роботи ймовірнісної нейронної мережі на першому етапі методу, а також підвищення точності базової ШНМ на 10 % відповідно до значень коефіцієнта кореляції Метьюса;

- можливість апроксимації суттєво нелінійної поверхні відгуку, заданої екстремально короткою вибіркою даних, шляхом формування комбінованої поверхні відгуку, що дозволяє точніше моделювати залежності з урахуванням локальної структури даних та підвищити адаптивність моделі до складних випадків, зокрема, коли класи частково перекриваються або спостерігається значна варіативність значення залежних атрибутів;
- зменшення обсягу аугментованої вибірки даних базового методу майже у 9 разів при зниженні точності менше 1% за коефіцієнтом детермінації завдяки використанню методу формування аугментованої вибірки необхідного обсягу на основі кластеризації.

6. Розроблено метод синтезу ансамблю лінійних нейроподібних структур з неітеративним машинним навчанням, який завдяки агрегуванню кускових лінійно апроксимованих поверхонь відгуку кожного його елемента забезпечує підвищення точності розв'язання задачі регресії на 5% відповідно до значень коефіцієнта детермінації при збереженні високої швидкодії його функціонування та простоти реалізації.

7. Розроблено метод синтезу ансамблів неітеративних нейроподібних структур з РБФ-розширенням входів, який шляхом агрегування мета-моделлю апроксимованих поверхонь відгуку, сформованих на незалежних вибірках з використанням симетричних випадкових зміщень, забезпечує ефективну аугментацію, декореляцію та часткову компенсацію похибок з протилежними знаками, що підвищує точність розв'язання задачі регресії майже на 6 % відповідно до значень коефіцієнта кореляції Метьюса.

8. Розроблено метод часткового усунення систематичної похибки нейронної мережі узагальненої регресії, який за рахунок використання ансамблю з двох таких нейронних мереж та трьох способів агрегування їх результатів дозволяє зменшити похибки функціонування ШНМ на 4 – 8% відповідно до RMSE в процесі розв'язання задачі регресії у випадку аналізу біомедичних даних малих обсягів.

9. Удосконалено метод глобально локальної апроксимації біомедичних даних малих обсягів, який на відміну від існуючого методу виконує незалежне виділення глобальної та локальної компонент шуканого сигналу, використовуючи ансамбль з двох нейронних мереж узагальненої регресії з можливістю ітеративного уточнення результату. Це забезпечує зменшення похибки RMSE на 13% при розв'язанні задачі регресії в умовах аналізу біомедичних даних обмеженого обсягу.

10. Розроблено прямий та зворотний методи компенсації похибки моделі регресії, які завдяки використанню перетворень залежностей у вигляді раціональних дробів дозволяють застосовувати ітеративне уточнення результатів, підвищуючи точність довільної моделі регресії без потреби у знанні її внутрішньої структури або параметрів, забезпечуючи зменшення похибки MSE, наприклад, машини опорних векторів на 9% та 11% відповідно.

11. Розвинуто теорію синтезу нейроподібних ансамблевих структур з неітеративним машинним навчанням, яка ґрунтується на розроблених методах нелінійного розширення простору ознак, аугментації даних, декореляції вибірок та компенсації похибок, що дає можливість підвищити точність розв'язання задач регресії та класифікації на вибірках біомедичних даних малих обсягів.

12. Розроблені та захищені авторським правом на комп'ютерні програми програмні засоби забезпечують можливість експериментального оцінювання ефективності розроблених та вдосконалених методів як на модельних, так і на реальних даних біомедичних досліджень. Отримані результати апробовані при

розв'язанні низки прикладних задач у різних напрямках біомедичних досліджень, зокрема, в біоматеріалознавстві, ревматології, травматології, стоматології та трансплантології. Їх впровадження виконано у співпраці з медичними закладами, науково-освітніми установами та в межах виконання наукового-дослідних робіт в Україні та міжнародних грантових проєктів за кордоном. Це засвідчено відповідними актами впровадження та використання результатів досліджень. Етичні аспекти враховані як на етапі моделювання, так і в процесі практичного впровадження розроблених та удосконалених методів, що забезпечує безпечне і відповідальне їх використання в біомедичних дослідженнях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] I. Izonin, A. K. Kazantzi, R. Tkachenko, and S.-A. Mitoulis, ‘GRNN-based cascade ensemble model for non-destructive damage state identification: small data approach’, *Engineering with Computers*, vol. 41, no. 1, pp. 723–738, Feb. 2025, doi: 10.1007/s00366-024-02048-1.
- [2] I. Izonin, I. Nesterenko, A. K. Kazantzi, R. Tkachenko, R. Muzyka, and S. A. Mitoulis, ‘Enhanced ANN-based ensemble method for bridge damage characterization using limited dataset’, *Sci Rep*, vol. 14, no. 1, p. 24395, Oct. 2024, doi: 10.1038/s41598-024-73738-5.
- [3] I. Izonin, R. Tkachenko, S. A. Mitoulis, A. Faramarzi, I. Tsmots, and D. Mashtalir, ‘Machine learning for predicting energy efficiency of buildings: a small data approach’, *Procedia Computer Science*, vol. 231, pp. 72–77, 2024, doi: 10.1016/j.procs.2023.12.173.
- [4] I. Izonin, ‘An unsupervised-supervised ensemble technology with non-iterative training algorithm for small biomedical data analysis’, *International scientific journal ‘Computer systems and information technologies’*, no. 4, pp. 67–74, Dec. 2023, doi: 10.31891/csit-2023-4-9.
- [5] I. Izonin, ‘An Ensemble Method for the Regression Model Parameter Adjustments: Direct Approach’, *Mathematical and computer modelling. Series: Technical sciences*, vol. 24, pp. 35–44, Dec. 2023, doi: 10.32626/2308-5916.2023-24.35-44.
- [6] I. Izonin and R. Tkachenko, ‘Universal intraensemble method using nonlinear AI techniques for regression modeling of small medical data sets’, in *Cognitive and Soft Computing Techniques for the Analysis of Healthcare Data*, Elsevier, 2022, pp. 123–150. doi: 10.1016/B978-0-323-85751-2.00002-5.
- [7] І. В. Ізонін, Я. Р. Бейда, М. А. Федорчук, Д. Б. Машталір, та Р. Р. Підкостельний, ‘Інтелектуальний аналіз коротких вибірок даних: методи аугментації даних’, в *Економічний і соціальний розвиток України в ХХІ столітті: національна візія та виклики глобалізації: збірник тез доповідей ХХ Міжнародної науково-практичної конференції молодих вчених*, Тернопіль, May 2023, pp. 705–707.

- [8] I. Izonin, R. Tkachenko, S. Fedushko, D. Koziy, K. Zub, and O. Vovk, ‘PБФ-Based Input Doubling Method for Small Medical Data Processing’, in *Advances in Artificial Systems for Logistics Engineering*, vol. 82, Z. Hu, Q. Zhang, S. Petoukhov, and M. He, Eds, in Lecture Notes on Data Engineering and Communications Technologies, vol. 82. , Cham: Springer International Publishing, 2021, pp. 23–31. doi: 10.1007/978-3-030-80475-6_3.
- [9] I. Izonin, R. Tkachenko, M. Gregus, K. Zub, and N. Lotoshynska, ‘Input Doubling Method based on SVR with PБФ kernel in Clinical Practice: Focus on Small Data’, in *proceedings of the 12th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops*, in Procedia Computer Science, vol. 184. Warsaw, Poland, Mar. 2021, pp. 606–613. doi: 10.1016/j.procs.2021.03.075.
- [10] I. Izonin, R. Tkachenko, M. Gregus ml., K. Zub, and P. Tkachenko, ‘A GRNN-based Approach towards Prediction from Small Datasets in Medical Application’, in *proceedings of the 12th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops*, in Procedia Computer Science, vol. 184. Warsaw, Poland: Elsevier, Mar. 2021, pp. 242–249. doi: 10.1016/j.procs.2021.03.033.
- [11] I. Izonin, R. Tkachenko, N. Horbal, M. Greguš, V. Verhun, and Y. Tolstyak, ‘An Approach Toward Numerical Data Augmentation and Regression Modeling Using Polynomial-Kernel-Based SVR’, in *Proceedings of International Conference on Data Science and Applications*, vol. 288, M. Saraswat, S. Roy, C. Chowdhury, and A. H. Gandomi, Eds, in Lecture Notes in Networks and Systems, vol. 288. , Singapore: Springer Singapore, 2022, pp. 771–781. doi: 10.1007/978-981-16-5120-5_58.
- [12] I. Izonin, R. Tkachenko, I. Dronyuk, P. Tkachenko, M. Gregus, and M. Rashkevych, ‘Predictive modeling based on small data in clinical medicine: PБФ-based additive input-doubling method’, *Mathematical Biosciences and Engineering*, vol. 18, no. 3, pp. 2599–2613, 2021, doi: 10.3934/mbe.2021132.

- [13] I. Izonin, R. Tkachenko, N. Shakhovska, and N. Lotoshynska, ‘The Additive Input-Doubling Method Based on the SVR with Nonlinear Kernels: Small Data Approach’, *Symmetry*, vol. 13, no. 4, p. 612, Apr. 2021, doi: 10.3390/sym13040612.
- [14] I. Izonin, R. Tkachenko, O. Berezhsky, I. Krak, M. Kováč, and M. Fedorchuk, ‘Improvement of the ANN-Based Prediction Technology for Extremely Small Biomedical Data Analysis’, *Technologies*, vol. 12, no. 7, p. 112, July 2024, doi: 10.3390/technologies12070112.
- [15] I. Izonin, R. Tkachenko, N. Hovdysh, O. Berezhsky, K. Yemets, and I. Tsmots, ‘Cascade-Based Input-Doubling Classifier for Predicting Survival in Allogeneic Bone Marrow Transplants: Small Data Case’, *Computation*, vol. 13, no. 4, p. 80, Mar. 2025, doi: 10.3390/computation13040080.
- [16] I. Izonin, R. Tkachenko, P. Yendyk, I. Pliss, Y. Bodyanskiy, and M. Gregus, ‘Enhanced Input-Doubling Method Leveraging Response Surface Linearization to Improve Classification Accuracy in Small Medical Data Processing’, *Computation*, vol. 12, no. 10, p. 203, Oct. 2024, doi: 10.3390/computation12100203.
- [17] I. Izonin, R. Tkachenko, K. Yemets, M. Gregus, Y. Tomashy, and I. Pliss, ‘An Approach Towards Reducing Training Time of the Input Doubling Method via Clustering for Middle-Sized Data Analysis’, in *proceedings of the 19th International Conference on Future Networks and Communications/ 21th International Conference on Mobile Systems and Pervasive Computing/14th International Conference on Sustainable Energy Information Technology*, in *Procedia Computer Science*, vol. 241. Marshall University, Huntington, WV, USA: Elsevier, 2024, pp. 32–39. doi: 10.1016/j.procs.2024.08.007.
- [18] R. Tkachenko, I. Izonin, N. Kryvinska, V. Chopyak, N. Lotoshynska, and D. Danylyuk, ‘Piecewise-linear Approach for Medical Insurance Costs Prediction using SGTm Neural-Like Structure’, in *proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018)*, in *CEUR Workshop Proceedings*, vol. vol–2255: Lviv, Ukraine, Nov. 2018, pp. 170–179.
- [19] R. Tkachenko and I. Izonin, ‘Model and Principles for the Implementation of Neural-Like Structures Based on Geometric Data Transformations’, in *Advances in*

- Computer Science for Engineering and Education*, vol. 754, Z. Hu, S. Petoukhov, I. Dychka, and M. He, Eds, in *Advances in Intelligent Systems and Computing*, vol. 754., Cham: Springer International Publishing, 2019, pp. 578–587. doi: 10.1007/978-3-319-91008-6_58.
- [20] I. Izonin, R. Tkachenko, N. Kryvinska, P. Tkachenko, and M. Greguš ml., ‘Multiple Linear Regression Based on Coefficients Identification Using Non-iterative SGTМ Neural-like Structure’, in *Advances in Computational Intelligence*, vol. 11506, I. Rojas, G. Joya, and A. Catala, Eds, in *Lecture Notes in Computer Science*, vol. 11506., Cham: Springer International Publishing, 2019, pp. 467–479. doi: 10.1007/978-3-030-20521-8_39.
- [21] І. В. Ізонін, Б. Ільчишин, М. Гаврилюк, та О. Бойчук, ‘Ансамбль однорідних методів машинного навчання з усередненням результату’, в *Інформаційні технології і автоматизація - 2023 : матеріали XVI Міжнародної науково-практичної конференції*, Одеса, Oct. 2023, pp. 339–340.
- [22] I. Izonin, R. Tkachenko, P. Vitynskyi, K. Zub, P. Tkachenko, and I. Dronyuk, ‘Stacking-based GRNN-SGTМ Ensemble Model for Prediction Tasks’, in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain: IEEE, Nov. 2020, pp. 326–330. doi: 10.1109/DASA51403.2020.9317124.
- [23] R. O. Tkachenko, I. V. Izonin, V. M. Danylyk, and V. Yu. Mykhalevych, ‘Stacking of the SGTМ neural-like structure with РБФ layer based on generation of a random curtain of its hyperparameters for prediction tasks’, *Ukrainian Journal of Information Technology*, vol. 3, no. 1, pp. 49–55, 2021, doi: 10.23939/ujit2021.03.049.
- [24] П. В. Вітинський, Р. О. Ткаченко, І. В. Ізонін, та Н. О. Кустра, ‘Ансамблі нейроподібних структур МПГП з РБФ розширенням входів для задач регресії та класифікації’, *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 4, no. 275, pp. 72–79, 2019.
- [25] I. V. Izonin, R. O. Tkachenko, and O. L. Semchyshyn, ‘An Ensemble Method for the Analysis of Small Biomedical Data based on a Neural Network Without Training’,

Elektronnoe modelirovanie, vol. 45, no. 6, pp. 65–76, Dec. 2023, doi: 10.15407/emodel.45.06.065.

- [26] П. В. Вітинський, Р. О. Ткаченко, та І. В. Ізонін, ‘Ансамбль мереж GRNN для розв’язання задач регресії з підвищеною точністю’, *Науковий вісник НЛТУ України*, vol. 29, no. 8, Art. no. 8, Oct. 2019, doi: 10.36930/40290822.
- [27] I. Izonin, R. Tkachenko, B. Trush, and S. Nahirnyi, ‘Homogeneous ensemble models based on ANNs with noniterative training: classification and regression tasks’, in *Reliability engineering and computational intelligence: book of abstracts of workshops RECI 2022 and ACeSYRI 2022*, Delft, the Netherlands: <https://ki.fri.uniza.sk/RECI2022/Abstracts.pdf>, 15.11 2022, p. 27. [Online]. Available: <https://ki.fri.uniza.sk/RECI2022/Abstracts.pdf>
- [28] R. Tkachenko, I. Izonin, M. Greguš ml., P. Tkachenko, and I. Dronyuk, ‘Committee of the SGTM Neural-Like Structures with Extended Inputs for Predictive Analytics in Insurance’, in *Big Data Innovations and Applications*, vol. 1054, M. Younas, I. Awan, and S. Benbernou, Eds, in Communications in Computer and Information Science, vol. 1054. , Cham: Springer International Publishing, 2019, pp. 121–132. doi: 10.1007/978-3-030-27355-2_9.
- [29] R. Tkachenko, P. Tkachenko, I. Izonin, P. Vitynskyi, N. Kryvinska, and Y. Tsymbal, ‘Committee of the Combined ПФ-SGTM Neural-Like Structures for Prediction Tasks’, in *Mobile Web and Intelligent Information Systems*, vol. 11673, I. Awan, M. Younas, P. Ünal, and M. Aleksy, Eds, in Lecture Notes in Computer Science, vol. 11673. , Cham: Springer International Publishing, 2019, pp. 267–277. doi: 10.1007/978-3-030-27192-3_21.
- [30] I. Izonin, R. Tkachenko, N. Kryvinska, M. Gregus, P. Tkachenko, and P. Vitynskyi, ‘Committee of SGTM Neural-Like Structures with ПФ kernel for Insurance Cost Prediction Task’, in *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Lviv, Ukraine: IEEE, July 2019, pp. 1037–1040. doi: 10.1109/UKRCON.2019.8879905.
- [31] I. Izonin, R. Tkachenko, V. Verhun, and K. Zub, ‘An approach towards missing data management using improved GRNN-SGTM ensemble method’, *Engineering Science*

- and Technology, an International Journal*, vol. 24, no. 3, pp. 749–759, June 2021, doi: 10.1016/j.jestch.2020.10.005.
- [32] I. V. Izonin, P. O. Tkachenko, та X. B. Зуб, ‘Застосування бустингу в ансамблях нейронних мереж узагальненої регресії для підвищення точності розв’язання задач апроксимації’, в *На шляху до індустрії 4.0: інформаційні технології, моделювання, штучний інтелект, автоматизація: колективна монографія*, Одеса: Астропринт, 2021, pp. 294–409.
- [33] R. Tkachenko, I. Izonin, N. Kryvinska, I. Dronyuk, and K. Zub, ‘An Approach towards Increasing Prediction Accuracy for the Recovery of Missing IoT Data Based on the GRNN-SGTM Ensemble’, *Sensors*, vol. 20, no. 9, Art. no. 9, Jan. 2020, doi: 10.3390/s20092625.
- [34] I. Izonin, R. Tkachenko, M. Gregus, Z. Duriagina, and N. Shakhovska, ‘PNN-SVM Approach of Ti-Based Powder’s Properties Evaluation for Biomedical Implants Production’, *Computers, Materials & Continua*, vol. 71, no. 3, pp. 5933–5947, 2022, doi: 10.32604/cmc.2022.022582.
- [35] I. Izonin, R. Tkachenko, Z. Duriagina, N. Shakhovska, V. Kovtun, and N. Lotoshynska, ‘Smart Web Service of Ti-Based Alloy’s Quality Evaluation for Medical Implants Manufacturing’, *Applied Sciences*, vol. 12, no. 10, p. 5238, May 2022, doi: 10.3390/app12105238.
- [36] I. Izonin, R. Tkachenko, and M. Greguš, ‘I-PNN: An Improved Probabilistic Neural Network for Binary Classification of Imbalanced Medical Data’, in *Database and Expert Systems Applications*, vol. 13427, C. Strauss, A. Cuzzocrea, G. Kotsis, A. M. Tjoa, and I. Khalil, Eds, in *Lecture Notes in Computer Science*, vol. 13427, Cham: Springer International Publishing, 2022, pp. 147–157. doi: 10.1007/978-3-031-12426-6_12.
- [37] I. Izonin, R. Tkachenko, O. Shcherbii, O. Berezhsky, I. Krak, and M. Oliynyk, ‘An Approximation Cascade Scheme via Rational Fractions for Biomedical Data Analysis’, in *2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT)*, Lviv, Ukraine: IEEE, Oct. 2023, pp. 1–4. doi: 10.1109/CSIT61576.2023.10324122.

- [38] I. Izonin, R. Tkachenko, O. Shcherbii, R. Muzyka, A. Faramarzi, and S. A. Mitoulis, ‘Enhanced Cascade Schemes for Advancing Machine Learning-Based Prediction of Heating and Cooling Loads in Residential Buildings’, in *2024 Mediterranean Smart Cities Conference (MSCC)*, Martil - Tetuan, Morocco: IEEE, May 2024, pp. 1–5. doi: 10.1109/MSCC62288.2024.10697024.
- [39] I. Izonin, R. Tkachenko, R. Bliakhar, M. Kovac, Y. Bodyanskiy, and O. Chala, ‘An improved ANN-based global-local approximation for small medical data analysis’, *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, Sept. 2023, doi: 10.4108/eetpht.9.3320.
- [40] I. В. Ізонін, Р. С. Бляхар, та Р. О. Ткаченко, ‘Нейромережевий метод послідовної глобально локальної апроксимації коротких наборів медичних даних’, в *Нейромережні технології та їх застосування : збірник наукових праць XXII Міжнародної наукової конференції НМТЗ-2023*, Краматорськ-Вінниця-Тернопіль, 2023, pp. 87–89.
- [41] I. Izonin, R. Tkachenko, M. Sydor, S. Chretien, I. Pliss, and T. Vovk, ‘Modification of Combined Unsupervised-Supervised Cascade Scheme for Small Biomedical Data Classification’, in *2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT)*, Lviv, Ukraine: IEEE, Sept. 2023, pp. 107–112. doi: 10.1109/ELIT61488.2023.10310830.
- [42] I. Izonin, T. Tepla, D. Danylyuk, R. Tkachenko, Z. Duriagina, and I. Lemishka, ‘Towards an Intelligent Decision Making of Ti-based Powders Selection for Medical Manufacturing’, in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain: IEEE, Nov. 2020, pp. 1–5. doi: 10.1109/DASA51403.2020.9316866.
- [43] R. Tkachenko, I. Izonin, P. Vitynskyi, N. Lotoshynska, and O. Pavlyuk, ‘Development of the Non-Iterative Supervised Learning Predictor Based on the Ito Decomposition and SGTM Neural-Like Structure for Managing Medical Insurance Costs’, *Data*, vol. 3, no. 4, p. 46, Oct. 2018, doi: 10.3390/data3040046.
- [44] I. Izonin, R. Tkachenko, N. Shakhovska, B. Ilchyshyn, and K. K. Singh, ‘A Two-Step Data Normalization Approach for Improving Classification Accuracy in the Medical

- Diagnosis Domain’, *Mathematics*, vol. 10, no. 11, p. 1942, June 2022, doi: 10.3390/math10111942.
- [45] I. Izonin, R. Tkachenko, K. Yemets, and M. Havryliuk, ‘An interpretable ensemble structure with a non-iterative training algorithm to improve the predictive accuracy of healthcare data analysis’, *Scientific Reports*, vol. 14, no. 1, p. 12947, June 2024, doi: 10.1038/s41598-024-61776-y.
- [46] I. Izonin, R. Muzyka, R. Tkachenko, I. Dronyuk, K. Yemets, and S.-A. Mitoulis, ‘A Method for Reducing Training Time of ML-Based Cascade Scheme for Large-Volume Data Analysis’, *Sensors*, vol. 24, no. 15, p. 4762, July 2024, doi: 10.3390/s24154762.
- [47] I. Izonin, R. Tkachenko, I. Krak, O. Berezhsky, I. Shevchuk, and S. K. Shandilya, ‘A cascade ensemble-learning model for the deployment at the edge: case on missing IoT data recovery in environmental monitoring systems’, *Frontiers in Environmental Science*, vol. 11, p. 1295526, Oct. 2023, doi: 10.3389/fenvs.2023.1295526.
- [48] I. Izonin, R. Tkachenko, O. Gurbych, M. Kovac, L. Rutkowski, and R. Holoven, ‘A non-linear SVR-based cascade model for improving prediction accuracy of biomedical data analysis’, *Mathematical Biosciences and Engineering*, vol. 20, no. 7, pp. 13398–13414, 2023, doi: 10.3934/mbe.2023597.
- [49] I. Izonin, R. Tkachenko, L. Ryvak, K. Zub, M. Rashkevych, and O. Pavliuk, ‘Addressing Medical Diagnostics Issues: Essential Aspects of the PNN-based Approach’, in *CEUR-WS.org*, vol. 2753: *Proceedings of the 3rd International Conference on Informatics & Data-Driven Medicine Växjö, Sweden, November 19 - 21, 2020*, Växjö, Sweden, Nov. 2020, pp. 209–218.
- [50] I. Izonin, R. Tkachenko, N. Shakhovska, B. Ilchyshyn, M. Gregus, and C. Strauss, ‘Towards Data Normalization Task for the Efficient Mining of Medical Data’, in *Proceedings of the 2022 12th International Conference on Advanced Computer Information Technologies*, Ruzomberok, Slovakia: IEEE, 2022, pp. 480–484. doi: 10.1109/ACIT54803.2022.9913112.
- [51] I. Izonin, N. Kryvinska, R. Tkachenko, K. Zub, and P. Vitynskyi, ‘An Extended-Input GRNN and its Application’, in *proceedings of the 10th International Conference on*

- Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2019) / The 9th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2019) / Affiliated Workshops*, in *Procedia Computer Science*, vol. 160. Coimbra, Portugal: Elsevier, Nov. 2019, pp. 578–583. doi: 10.1016/j.procs.2019.11.044.
- [52] I. Izonin, R. Tkachenko, R. Pidkostelnyy, O. Pavliuk, V. Khavalko, and A. Batyuk, ‘Experimental Evaluation of the Effectiveness of ANN-based Numerical Data Augmentation Methods for Diagnostics Tasks’, in *proceedings of the 4th International Conference on Informatics & Data-Driven Medicine*, in *CEUR Workshop Proceedings*, vol. 3038. Valencia, Spain, Nov. 2021, pp. 223–232.
- [53] I. Izonin and R. Tkachenko, ‘An approach towards the response surface linearization via ANN-based cascade scheme for regression modeling in Healthcare’, in *in proceedings of the 12th International Conference on Emerging Ubiquitous Systems and Pervasive Networks / 11th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare*, in *Procedia Computer Science*, vol. 198. Leuven, Belgium: Elsevier, 2022, pp. 724–729. doi: 10.1016/j.procs.2021.12.313.
- [54] І. В. Ізонін, Р. О. Ткаченко, Томаши Є. О., та Гентош Л. І., ‘Бібліотека методів інтелектуального аналізу малих даних на основі декартового добутку впорядкованих множин входів: комп’ютерна програма’, Свідоцтво про реєстрацію авторського права № 127079, June 04, 2024
- [55] Вітинський П. Б., Р. О. Ткаченко, та І. В. Ізонін, ‘Комбінована нейроподібна система для задач аналітики в умовах коротких вибірок даних: комп’ютерна програма’, Свідоцтво про реєстрацію авторського права № 93094, Oct. 16, 2019
- [56] І. В. Ізонін, Нагірний С. Б., та Ткаченко Р. О., ‘Гомогенні стекінгові ансамблі штучних нейронних мереж без навчання (Multi-GRNN): комп’ютерна програма’, Свідоцтво про реєстрацію авторського права № 114033., Aug. 09, 2022
- [57] I. Izonin, R. Tkachenko, P. Vitynskyi, K. Zub, P. Tkachenko, and I. Dronyuk, ‘Stacking-based GRNN-SGTM Ensemble Model for Prediction Tasks’, in *2020*

- International Conference on Decision Aid Sciences and Application (DASA)*, Nov. 2020, pp. 326–330. doi: 10.1109/DASA51403.2020.9317124.
- [58] M. Kuzmynych, O. Matviyiv, and M. Lobur, ‘Analysis of Electromyography Methods for Healthcare Monitoring’, in *2017 XIIIth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH)*, Lviv, Ukraine: IEEE, Apr. 2017, pp. 101–104. doi: 10.1109/memstech.2017.7937543.
- [59] O. Berezsky, O. Pitsun, P. Liashchynskyi, B. Derysh, and N. Batryn, ‘Computational Intelligence in Medicine’, in *Lecture Notes in Data Engineering, Computational Intelligence, and Decision Making*, vol. 149, S. Babichev and V. Lytvynenko, Eds, in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 149, Cham: Springer International Publishing, 2023, pp. 488–510. doi: 10.1007/978-3-031-16203-9_28.
- [60] W.-H. Jian *et al.*, ‘Glycosaminoglycan-based hybrid hydrogel encapsulated with polyelectrolyte complex nanoparticles for endogenous stem cell regulation in central nervous system regeneration’, *Biomaterials*, vol. 174, pp. 17–30, Aug. 2018, doi: 10.1016/j.biomaterials.2018.05.009.
- [61] Y. David and T. Judd, ‘Evidence-based impact by clinical engineers on global patients outcomes’, *Health Technol.*, vol. 10, no. 2, pp. 517–535, Mar. 2020, doi: 10.1007/s12553-019-00345-0.
- [62] M. Koohi-Moghadam *et al.*, ‘MetaMarker: a pipeline for *de novo* discovery of novel metagenomic biomarkers’, *Bioinformatics*, vol. 35, no. 19, pp. 3812–3814, Oct. 2019, doi: 10.1093/bioinformatics/btz123.
- [63] Y. Qiao, Q. Zhang, Q. Wang, Y. Li, and L. Wang, ‘Filament-anchored hydrogel layer on polypropylene hernia mesh with robust anti-inflammatory effects’, *Acta Biomaterialia*, vol. 128, pp. 277–290, July 2021, doi: 10.1016/j.actbio.2021.04.013.
- [64] H. A. Al-Swadi, A. R. A. Usman, A. S. Al-Farraj, M. I. Al-Wabel, M. Ahmad, and A. Al-Faraj, ‘Sources, toxicity potential, and human health risk assessment of heavy metals-laden soil and dust of urban and suburban areas as affected by industrial and mining activities’, *Sci Rep*, vol. 12, no. 1, p. 8972, May 2022, doi: 10.1038/s41598-022-12345-8.

- [65] G. Das and I. Ghosh, ‘Biomedical Data Types, Sources, Content, and Retrieval’, in *Text Mining Approaches for Biomedical Data*, A. Sharan, N. Malik, H. Imran, and I. Ghosh, Eds, in Transactions on Computer Systems and Networks. , Singapore: Springer Nature Singapore, 2024, pp. 19–54. doi: 10.1007/978-981-97-3962-2_2.
- [66] J. Luo, M. Wu, D. Gopukumar, and Y. Zhao, ‘Big Data Application in Biomedical Research and Health Care: A Literature Review’, *Biomed Inform Insights*, vol. 8, p. BII.S31559, Jan. 2016, doi: 10.4137/BII.S31559.
- [67] O. Duda, N. Kunanets, S. Martsenko, V. Nykytyuk, and V. Pasichnyk, ‘Information technology platform for the selection and analytical processing of information on COVID-19’, in *2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)*, LVIV, Ukraine: IEEE, Sept. 2021, pp. 231–238. doi: 10.1109/csit52700.2021.9648839.
- [68] O. Duda, V. Pasichnyk, N. Kunanets, R. Antonii, and O. Matsiuk, ‘Multidimensional Representation of COVID-19 Data Using OLAP Information Technology’, in *2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT)*, Zbarazh, Ukraine: IEEE, Sept. 2020, pp. 277–280. doi: 10.1109/csit49958.2020.9321889.
- [69] C. A. Federico and A. A. Trotsyuk, ‘Biomedical Data Science, Artificial Intelligence, and Ethics: Navigating Challenges in the Face of Explosive Growth’, *Annual Review of Biomedical Data Science*, vol. 7, no. 1, pp. 1–14, Aug. 2024, doi: 10.1146/annurev-biodatasci-102623-104553.
- [70] R. Bellazzi, ‘Big Data and Biomedical Informatics: A Challenging Opportunity’, *Yearb Med Inform*, vol. 23, no. 01, pp. 08–13, Aug. 2014, doi: 10.15265/IY-2014-0024.
- [71] A. R. Sonawane, S. T. Weiss, K. Glass, and A. Sharma, ‘Network Medicine in the Age of Biomedical Big Data’, *Front. Genet.*, vol. 10, Apr. 2019, doi: 10.3389/fgene.2019.00294.
- [72] V. Berisha *et al.*, ‘Digital medicine and the curse of dimensionality’, *npj Digit. Med.*, vol. 4, no. 1, p. 153, Oct. 2021, doi: 10.1038/s41746-021-00521-5.

- [73] Q. Chang *et al.*, ‘Mining multi-center heterogeneous medical data with distributed synthetic learning’, *Nat Commun*, vol. 14, no. 1, p. 5510, Sept. 2023, doi: 10.1038/s41467-023-40687-y.
- [74] P. C. Austin, I. R. White, D. S. Lee, and S. Van Buuren, ‘Missing Data in Clinical Research: A Tutorial on Multiple Imputation’, *Canadian Journal of Cardiology*, vol. 37, no. 9, pp. 1322–1331, Sept. 2021, doi: 10.1016/j.cjca.2020.11.010.
- [75] O. Faust, M. Salvi, P. D. Barua, S. Chakraborty, F. Molinari, and U. R. Acharya, ‘Issues and Limitations on the Road to Fair and Inclusive AI Solutions for Biomedical Challenges’, *Sensors*, vol. 25, no. 1, p. 205, Jan. 2025, doi: 10.3390/s25010205.
- [76] S. Subbotin, G. Tabunshchyk, P. Arras, D. Tabunshchyk, and E. Trotsenko, ‘Intelligent Data Analysis for Individual Hypertensia Patient’s State Monitoring and Prediction’, in *2021 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, Nur-Sultan, Kazakhstan: IEEE, Apr. 2021, pp. 1–4. doi: 10.1109/SIST50301.2021.9465989.
- [77] P. Radiuk, O. Kovalchuk, V. Slobodzian, E. Manziuk, and I. Krak, ‘Human-in-the-Loop Approach Based on MRI and ECG for Healthcare Diagnosis’, *CEUR-WS.org*, vol. 3302, pp. 9–20, 2022.
- [78] I. Krak, O. Barmak, and E. Manziuk, ‘Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology’, *Computational Intelligence*, vol. 38, no. 3, pp. 921–946, June 2022, doi: 10.1111/coin.12289.
- [79] V. N. Vapnik and A. Ya. Chervonenkis, ‘On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities’, *Theory Probab. Appl.*, vol. 16, no. 2, pp. 264–280, Jan. 1971, doi: 10.1137/1116025.
- [80] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 1995. doi: 10.1007/978-1-4757-2440-0.
- [81] T. O. Nazirova and O. B. Kostenko, ‘Нейромережева інформаційна технологія опрацювання медичних даних’, *SBUNFU*, vol. 28, no. 8, pp. 141–145, Oct. 2018, doi: 10.15421/40280828.

- [82] Ye. Bodyankiy, K. Petrov, and A. A. Deineko, ‘Evolving radial basis neural network and extreme learning of its parameters’, *Scientific journal “Bionics of Intelligence”*, vol. 2, no. 91, pp. 9–15, Dec. 2018, doi: 10.30837/bi.2018.2(91).02.
- [83] I. Krak, O. Barmak, E. Manziuk, and A. Kulias, ‘Data Classification Based on the Features Reduction and Piecewise Linear Separation’, in *Intelligent Computing and Optimization*, vol. 1072, P. Vasant, I. Zelinka, and G.-W. Weber, Eds, in Advances in Intelligent Systems and Computing, vol. 1072. , Cham: Springer International Publishing, 2020, pp. 282–289. doi: 10.1007/978-3-030-33585-4_28.
- [84] S. Subbotin, ‘Radial-Basis Function Neural Network Synthesis on the Basis of Decision Tree’, *Opt. Mem. Neural Networks*, vol. 29, no. 1, pp. 7–18, Jan. 2020, doi: 10.3103/S1060992X20010051.
- [85] I. Izonin, N. Kryvinska, P. Vitynskyi, R. Tkachenko, and K. Zub, ‘GRNN Approach Towards Missing Data Recovery Between IoT Systems’, in *Advances in Intelligent Networking and Collaborative Systems*, vol. 1035, L. Barolli, H. Nishino, and H. Miwa, Eds, in Advances in Intelligent Systems and Computing, vol. 1035. , Cham: Springer International Publishing, 2020, pp. 445–453. doi: 10.1007/978-3-030-29035-1_43.
- [86] T. Shaikhina and N. A. Khovanova, ‘Handling limited datasets with neural networks in medical applications: A small-data approach’, *Artificial Intelligence in Medicine*, vol. 75, pp. 51–63, Jan. 2017, doi: 10.1016/j.artmed.2016.12.003.
- [87] J. Yang, A. A. S. Soltan, and D. A. Clifton, ‘Machine learning generalizability across healthcare settings: insights from multi-site COVID-19 screening’, *npj Digit. Med.*, vol. 5, no. 1, p. 69, June 2022, doi: 10.1038/s41746-022-00614-9.
- [88] O. Bisikalo, V. Kharchenko, V. Kovtun, I. Krak, and S. Pavlov, ‘Parameterization of the Stochastic Model for Evaluating Variable Small Data in the Shannon Entropy Basis’, *Entropy*, vol. 25, no. 2, p. 184, Jan. 2023, doi: 10.3390/e25020184.
- [89] A. Iasonos and J. O’Quigley, ‘Randomised Phase 1 clinical trials in oncology’, *Br J Cancer*, vol. 125, no. 7, pp. 920–926, Sept. 2021, doi: 10.1038/s41416-021-01412-y.
- [90] І. А. Лемішка, ‘Оптимізація параметрів мікроструктури і властивостей композиційних матеріалів на основі титанових сплавів для адитивних

- технологій’, дисертація на здобуття наукового степеня кандидата технічних наук, Луцький національний технічний університет, Луцьк, Україна, 2018. [Online]. Available: <https://drive.google.com/file/d/1o37g5nWtN0U5kICq6-Bm548yT36u-hl4/view>
- [91] В. І. Ваврух, ‘Вплив легування на стабілізацію фазового стану та властивостей оксидної кераміки на основі ZrO_2 ’, дисертація на здобуття наукового степеня доктора філософії, Національний університет ‘Львівська політехніка’, м. Львів, 2025. [Online]. Available: <https://lpnu.ua/sites/default/files/2025/radaphd/29948/disertaciya-vavrukh.pdf>
- [92] V. Kulyk *et al.*, ‘Prediction of hardness, flexural strength, and fracture toughness of ZrO_2 based ceramics using ensemble learning algorithms’, *Acta Metall Slovaca*, vol. 29, no. 2, pp. 93–103, June 2023, doi: 10.36547/ams.29.2.1819.
- [93] Z. A. Duriagina *et al.*, ‘Determination of the best microstructure and titanium alloy powders properties using neural network’, *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 1, no. 87, pp. 25–31, Mar. 2018, doi: 10.5604/01.3001.0012.0736.
- [94] R. Tkachenko, Z. Duriagina, I. Lemishka, I. Izonin, and A. Trostianchyn, ‘Development of machine learning method of titanium alloy properties identification in additive technologies’, *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 12 (93), pp. 23–31, June 2018, doi: 10.15587/1729-4061.2018.134319.
- [95] Y. Wu, X. Ding, Y. Wang, and D. Ouyang, ‘Harnessing the power of machine learning into tissue engineering: current progress and future prospects’, *Burns & Trauma*, vol. 12, p. tkae053, Jan. 2024, doi: 10.1093/burnst/tkae053.
- [96] S. N. Flesher *et al.*, ‘Intracortical microstimulation of human somatosensory cortex’, *Sci. Transl. Med.*, vol. 8, no. 361, Oct. 2016, doi: 10.1126/scitranslmed.aaf8083.
- [97] E. B. Hekler, P. Klasnja, G. Chevance, N. M. Golaszewski, D. Lewis, and I. Sim, ‘Why we need a small data paradigm’, *BMC Med*, vol. 17, no. 1, p. 133, July 2019, doi: 10.1186/s12916-019-1366-x.
- [98] R. Kitchin and T. P. Lauriault, ‘Small data in the era of big data’, *GeoJournal*, vol. 80, no. 4, pp. 463–475, Aug. 2015, doi: 10.1007/s10708-014-9601-7.

- [99] D. Estrin, ‘Small data, where $n = me$ ’, *Commun. ACM*, vol. 57, no. 4, pp. 32–34, Apr. 2014, doi: 10.1145/2580944.
- [100] J. Zhou, Y. Ye, and J. Jiang, ‘Kernel principal components based cascade forest towards disease identification with human microbiota’, *BMC Med Inform Decis Mak*, vol. 21, no. 1, p. 360, Dec. 2021, doi: 10.1186/s12911-021-01705-5.
- [101] P. Kokol, M. Kokol, and S. Zagoranski, ‘Machine learning on small size samples: A synthetic knowledge synthesis’, *Science Progress*, vol. 105, no. 1, Jan. 2022, doi: 10.1177/00368504211029777.
- [102] C. Li, S. Li, Y. Feng, K. Gryllias, F. Gu, and M. Pecht, ‘Small data challenges for intelligent prognostics and health management: a review’, *Artif Intell Rev*, vol. 57, no. 8, July 2024, doi: 10.1007/s10462-024-10820-4.
- [103] P. Xu, X. Ji, M. Li, and W. Lu, ‘Small data machine learning in materials science’, *npj Comput Mater*, vol. 9, no. 1, Mar. 2023, doi: 10.1038/s41524-023-01000-z.
- [104] B. Dou *et al.*, ‘Machine Learning Methods for Small Data Challenges in Molecular Science’, *Chem. Rev.*, vol. 123, no. 13, pp. 8736–8780, July 2023, doi: 10.1021/acs.chemrev.3c00189.
- [105] D. McNeish, ‘Two-Level Dynamic Structural Equation Models with Small Samples’, *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 26, no. 6, pp. 948–966, Nov. 2019, doi: 10.1080/10705511.2019.1578657.
- [106] C. J. Morgan, ‘Use of proper statistical techniques for research studies with small samples’, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, vol. 313, no. 5, pp. L873–L877, Nov. 2017, doi: 10.1152/ajplung.00238.2017.
- [107] ‘Chebyshev’s Inequality & Rule in Statistical Analysis’. Accessed: May 18, 2025. [Online]. Available: <https://www.turing.com/kb/deep-dive-into-chebyshevs-inequality-in-statistics>
- [108] S. Turney, ‘Central Limit Theorem | Formula, Definition & Examples’, Scribbr. Accessed: May 18, 2025. [Online]. Available: <https://www.scribbr.com/statistics/central-limit-theorem/>

- [109] J. Kim, ‘Central Limit Theorem: The Good, the Bad, and the Ugly’, Medium. Accessed: May 18, 2025. [Online]. Available: <https://medium.com/data-and-beyond/central-limit-theorem-the-good-the-bad-and-the-ugly-5ce7e82668fc>
- [110] Х. М. Федорович, ‘Клініко-лабораторні особливості кардіоваскулярного статусу пацієнтів з ревматоїдним артритом та оптимізація принципів його корекції’, дисертація на здобуття наукового степеня доктора філософії, ІваноФранківський національний медичний університет МОЗ України, Івано-Франківськ, Україна, 2024. [Online]. Available: https://www.ifnmu.edu.ua/wp-content/uploads/2024/08/dis_%D0%A4%D0%B5%D0%B4%D0%BE%D1%80%D0%BE%D0%B2%D0%B8%D1%87-26.02.24-1.pdf
- [111] О. В. Каніщева, ‘Оптимізація ведення пацієнтів з артеріальною гіпертензією на підставі вивчення короткострокової варіабельності артеріального тиску’, дисертація на здобуття наукового степеня доктора філософії, Харківський національний університет імені В.Н. Каразіна, Харків, Україна, 2025. [Online]. Available: https://karazin.ua/storage/static-content/source/documents/aspirantura/zakhysty/kanishcheva/Kanishcheva_dis.pdf
- [112] О. В. Соміло, ‘Прогностичні критерії перебігу рецидивуючоремітуючого розсіяного склерозу’, дисертація на здобуття наукового степеня доктора філософії, Дніпровський державний медичний університет, Дніпро, Україна, 2023. [Online]. Available: <https://repo.dma.dp.ua/8749/1/%D0%A1%D0%BE%D0%BC%D1%96%D0%BB%D0%BE%20%D0%9E.%D0%92.%20%D0%94%D0%B8%D1%81%D0%B5%D1%80%D1%82%D0%B0%D1%86%D1%96%D1%8F.pdf>
- [113] A. Ultsch and J. Lötsch, ‘Augmenting small biomedical datasets using generative AI methods based on self-organizing neural networks’, *Briefings in Bioinformatics*, vol. 26, no. 1, p. bbae640, Nov. 2024, doi: 10.1093/bib/bbae640.
- [114] G. Varoquaux, ‘Cross-validation failure: Small sample sizes lead to large error bars’, *NeuroImage*, vol. 180, pp. 68–77, Oct. 2018, doi: 10.1016/j.neuroimage.2017.06.061.

- [115] I. Krak, V. Kuznetsov, S. Kondratiuk, L. Azarova, O. Barmak, and P. Padiuk, ‘Analysis of Deep Learning Methods in Adaptation to the Small Data Problem Solving’, in *Lecture Notes in Data Engineering, Computational Intelligence, and Decision Making*, vol. 149, S. Babichev and V. Lytvynenko, Eds, in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 149. , Cham: Springer International Publishing, 2023, pp. 333–352. doi: 10.1007/978-3-031-16203-9_20.
- [116] S. Huang and H. Deng, *Data analytics: a small data approach*, First edition. Boca Raton: CRC Press, 2021.
- [117] G. Cybenko, ‘Approximation by superpositions of a sigmoidal function’, *Math. Control Signal Systems*, vol. 2, no. 4, pp. 303–314, Dec. 1989, doi: 10.1007/BF02551274.
- [118] K. Hornik, ‘Approximation capabilities of multilayer feedforward networks’, *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991, doi: 10.1016/0893-6080(91)90009-T.
- [119] R. E. Bellman, *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961. doi: 10.1515/9781400874668.
- [120] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, 1st edn. in *Wiley Series in Probability and Statistics*. Wiley, 2002. doi: 10.1002/9781119013563.
- [121] J. M. Jerez *et al.*, ‘Missing data imputation using statistical and machine learning methods in a real breast cancer problem’, *Artificial Intelligence in Medicine*, vol. 50, no. 2, pp. 105–115, Oct. 2010, doi: 10.1016/j.artmed.2010.05.002.
- [122] О. С. Міщук, ‘Нейроподібні методи та засоби прогнозування параметрів забруднення атмосферного повітря’, Національний університет ‘Львівська політехніка’, Львів, Україна, 2021. [Online]. Available: <https://lpnu.ua/sites/default/files/2021/dissertation/9219/mischukdisertaciya0904211.pdf>
- [123] V. Hodge and J. Austin, ‘A Survey of Outlier Detection Methodologies’, *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, Oct. 2004, doi: 10.1023/B:AIRE.0000045502.10941.a9.

- [124] L. Breiman, ‘Bagging predictors’, *Mach Learn*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/BF00058655.
- [125] T. G. Dietterich, ‘Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms’, *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998, doi: 10.1162/089976698300017197.
- [126] Y. Xu, X. Lu, Y. Fei, and Y. Huang, ‘Iterative self-transfer learning: A general methodology for response time-history prediction based on small dataset’, *Journal of Computational Design and Engineering*, vol. 9, no. 5, pp. 2089–2102, Oct. 2022, doi: 10.1093/jcde/qwac098.
- [127] C. M. Bishop, *Pattern recognition and machine learning*. in Information science and statistics. New York: Springer, 2006.
- [128] J. O. Ogutu, T. Schulz-Streeck, and H.-P. Piepho, ‘Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions’, *BMC Proc*, vol. 6, no. S2, p. S10, Dec. 2012, doi: 10.1186/1753-6561-6-S2-S10.
- [129] M. A. Arshad, S. Shahriar, and K. Anjum, ‘The Power Of Simplicity: Why Simple Linear Models Outperform Complex Machine Learning Techniques -- Case Of Breast Cancer Diagnosis’, 2023, *arXiv*. doi: 10.48550/ARXIV.2306.02449.
- [130] R. F. Woolson and W. R. Clarke, *Statistical Methods for the Analysis of Biomedical Data*, 2nd ed. in Wiley Series in Probability and Statistics, no. v. 371. Hoboken: John Wiley & Sons, 2011.
- [131] J. Mandel, ‘Use of the Singular Value Decomposition in Regression Analysis’, *The American Statistician*, vol. 36, no. 1, pp. 15–24, Feb. 1982, doi: 10.1080/00031305.1982.10482771.
- [132] Р. О. Ткаченко, ‘Нейронні мережі прямого поширення з неітераційним навчанням’, дисертація на здобуття наукового степеня доктора технічних наук, Державний науково-дослідний інститут інформаційної інфраструктури, Львів, Україна, 2000.

- [133] Р. О. Ткаченко, П. Р. Ткаченко, and І. В. Ізонін, *Нейромережеві засоби штучного інтелекту*. Львів: Видавництво Львівської політехніки, 2017.
[Online]. Available: <https://vlp.com.ua/node/17380>
- [134] R. Tkachenko, ‘An Integral Software Solution of the SGTm Neural-Like Structures Implementation for Solving Different Data Mining Tasks’, in *Lecture Notes in Computational Intelligence and Decision Making*, vol. 77, S. Babichev and V. Lytvynenko, Eds, in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 77., Cham: Springer International Publishing, 2022, pp. 696–713. doi: 10.1007/978-3-030-82014-5_48.
- [135] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 6th edn. Wiley, 2021. Accessed: May 22, 2025. [Online]. Available: https://www.perlego.com/book/2237423/introduction-to-linear-regression-analysis-pdf?utm_source=google&utm_medium=cpc&campaignid=20933451054&adgroupid=162926082692&gad_source=1&gad_campaignid=20933451054&gbraid=0AAAAADN2nNKk258a7a904CCwhYejucBHv&gclid=Cj0KCQjwlrBBhDnARIsAHEQgOQF7E98UReEfDKRbRwxsEqgVN5ggDaMhg0s-cgnOv4-u4_I8nbpqdMaAvYsEALw_wcB
- [136] I. T. Jolliffe and J. Cadima, ‘Principal component analysis: a review and recent developments’, *Phil. Trans. R. Soc. A.*, vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.
- [137] F. Gabbiani, *Mathematics for Neuroscientists*. Saint Louis: Elsevier Science & Technology, 2010.
- [138] R. Tkachenko, H. Kutucu, I. Izonin, A. Doroshenko, and Y. Tsymbal, ‘Non-Iterative Neural-Like Predictor for Solar Energy in Libya’, p. 11.
- [139] N. A. Schuster, J. J. M. Rijnhart, J. W. R. Twisk, and M. W. Heymans, ‘Modeling non-linear relationships in epidemiological data: The application and interpretation of spline models’, *Front. Epidemiol.*, vol. 2, p. 975380, Aug. 2022, doi: 10.3389/fepid.2022.975380.

- [140] D. Chumachenko, I. Meniailov, K. Bazilevych, T. Chumachenko, and S. Yakovlev, ‘Investigation of Statistical Machine Learning Models for COVID-19 Epidemic Process Simulation: Random Forest, K-Nearest Neighbors, Gradient Boosting’, *Computation*, vol. 10, no. 6, p. 86, May 2022, doi: 10.3390/computation10060086.
- [141] D. Chumachenko, P. Piletskiy, M. Sukhorukova, and T. Chumachenko, ‘Predictive Model of Lyme Disease Epidemic Process Using Machine Learning Approach’, *Applied Sciences*, vol. 12, no. 9, p. 4282, Apr. 2022, doi: 10.3390/app12094282.
- [142] Y.-H. Pao, *Adaptive pattern recognition and neural networks*. Reading, Mass.: Addison-Wesley, 1989.
- [143] T. M. Cover, ‘Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition’, *IEEE Trans. Electron. Comput.*, vol. EC-14, no. 3, pp. 326–334, June 1965, doi: 10.1109/PGEC.1965.264137.
- [144] E. Torre, S. Marelli, P. Embrechts, and B. Sudret, ‘Data-driven polynomial chaos expansion for machine learning regression’, *Journal of Computational Physics*, vol. 388, pp. 601–623, July 2019, doi: 10.1016/j.jcp.2019.03.039.
- [145] I. Izonin, R. Tkachenko, N. Kryvinska, M. Gregus, P. Tkachenko, and P. Vitynskyi, ‘Committee of SGTM Neural-Like Structures with $PB\Phi$ kernel for Insurance Cost Prediction Task’, in *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Lviv, Ukraine: IEEE, July 2019, pp. 1037–1040. doi: 10.1109/UKRCON.2019.8879905.
- [146] R. Tkachenko, P. Tkachenko, I. Izonin, P. Vitynskyi, N. Kryvinska, and Y. Tsymbal, ‘Committee of the Combined $PB\Phi$ -SGTM Neural-Like Structures for Prediction Tasks’, in *Mobile Web and Intelligent Information Systems*, vol. 11673, I. Awan, M. Younas, P. Ünal, and M. Aleksy, Eds, in *Lecture Notes in Computer Science*, vol. 11673. , Cham: Springer International Publishing, 2019, pp. 267–277. doi: 10.1007/978-3-030-27192-3_21.
- [147] V. E. Papadimitriou, G. N. Aretoulis, and J. Papathanasiou, ‘Radial Basis Function ($PB\Phi$) and Multilayer Perceptron (MLP) Comparative Analysis on Building

- Renovation Cost Estimation: The Case of Greece’, *Algorithms*, vol. 17, no. 9, p. 390, Sept. 2024, doi: 10.3390/a17090390.
- [148] M. Iwayama, S. Wu, C. Liu, and R. Yoshida, ‘Functional Output Regression for Machine Learning in Materials Science’, *J. Chem. Inf. Model.*, vol. 62, no. 20, pp. 4837–4851, Oct. 2022, doi: 10.1021/acs.jcim.2c00626.
- [149] A. G. Ivakhnenko, ‘Polynomial Theory of Complex Systems’, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-1, no. 4, pp. 364–378, Oct. 1971, doi: 10.1109/TSMC.1971.4308320.
- [150] J. Rahnenführer *et al.*, ‘Statistical analysis of high-dimensional biomedical data: a gentle introduction to analytical goals, common approaches and challenges’, *BMC Med*, vol. 21, no. 1, p. 182, May 2023, doi: 10.1186/s12916-023-02858-y.
- [151] I. Izonin, M. Greguš ml., R. Tkachenko, M. Logoyda, O. Mishchuk, and Y. Kynash, ‘SGD-Based Wiener Polynomial Approximation for Missing Data Recovery in Air Pollution Monitoring Dataset’, in *Advances in Computational Intelligence*, vol. 11506, I. Rojas, G. Joya, and A. Catala, Eds, in Lecture Notes in Computer Science, vol. 11506. , Cham: Springer International Publishing, 2019, pp. 781–793. doi: 10.1007/978-3-030-20521-8_64.
- [152] V. A. Popov, ‘Local approximation of a function’, *Mathematical Notes of the Academy of Sciences of the USSR*, vol. 17, no. 3, pp. 213–221, Mar. 1975, doi: 10.1007/BF01149009.
- [153] J. Schmidt-Hieber, ‘The Kolmogorov–Arnold representation theorem revisited’, *Neural Networks*, vol. 137, pp. 119–126, May 2021, doi: 10.1016/j.neunet.2021.01.020.
- [154] M. Havryliuk, N. Hovdysh, Y. Tolstyak, V. Chopyak, and N. Kustra, ‘Investigation of PNN Optimization Methods to Improve Classification Performance in Transplantation Medicine’, *CEUR-WS.org*, vol. 3609, pp. 338–345.
- [155] S. Somvanshi, S. A. Javed, M. M. Islam, D. Pandit, and S. Das, ‘A Survey on Kolmogorov-Arnold Network’, 2024, *arXiv*. doi: 10.48550/ARXIV.2411.06078.

- [156] R. Schaback and H. Wendland, ‘Kernel techniques: From machine learning to meshless methods’, *Acta Numerica*, vol. 15, pp. 543–639, May 2006, doi: 10.1017/S0962492906270016.
- [157] D. F. Specht, ‘A general regression neural network’, *IEEE Trans. Neural Netw.*, vol. 2, no. 6, pp. 568–576, Nov. 1991, doi: 10.1109/72.97934.
- [158] D. F. Specht, ‘Probabilistic neural networks’, *Neural Networks*, vol. 3, no. 1, pp. 109–118, Jan. 1990, doi: 10.1016/0893-6080(90)90049-Q.
- [159] I. Izonin, N. Kryvinska, R. Tkachenko, K. Zub, and P. Vitynskyi, ‘An Extended-Input GRNN and its Application’, *Procedia Computer Science*, vol. 160, pp. 578–583, Jan. 2019, doi: 10.1016/j.procs.2019.11.044.
- [160] I. Izonin, R. Tkachenko, V. Verhun, and K. Zub, ‘An approach towards missing data management using improved GRNN-SGTM ensemble method’, *Engineering Science and Technology, an International Journal*, p. S2215098620342348, Nov. 2020, doi: 10.1016/j.jestch.2020.10.005.
- [161] R. Tkachenko, I. Izonin, I. Dronyuk, M. Logoyda, and P. Tkachenko, ‘Recover Missing Sensor Data with GRNN-based Cascade Scheme’, *International Journal of Sensors, Wireless Communications and Control*, pp. 1–10, 2020, doi: (in press).
- [162] A. Sebastian, A. Pannone, S. Subbulakshmi Radhakrishnan, and S. Das, ‘Gaussian synapses for probabilistic neural networks’, *Nat Commun*, vol. 10, no. 1, p. 4199, Sept. 2019, doi: 10.1038/s41467-019-12035-6.
- [163] H. Pavard, K. Zormpas-Petridis, H. Clayton, S. Burge, and M. Crispin-Ortuzar, ‘Radiology and multi-scale data integration for precision oncology’, *npj Precis. Onc.*, vol. 8, no. 1, p. 158, July 2024, doi: 10.1038/s41698-024-00656-0.
- [164] S. Chen, E. Dobriban, and J. H. Lee, ‘A Group-Theoretic Framework for Data Augmentation’, *Journal of Machine Learning Research*, vol. 21, no. 245, pp. 1–71, 2020.
- [165] L. Cui, H. Li, K. Chen, L. Shou, and G. Chen, ‘Tabular Data Augmentation for Machine Learning: Progress and Prospects of Embracing Generative AI’, 2024, *arXiv*. doi: 10.48550/ARXIV.2407.21523.

- [166] Y. Gordienko, N. Gordienko, V. Taran, A. Rojbi, S. Telenyk, and S. Stirenko, ‘Effect of natural and synthetic noise data augmentation on physical action classification by brain–computer interface and deep learning’, *Front. Neuroinform.*, vol. 19, Feb. 2025, doi: 10.3389/fninf.2025.1521805.
- [167] Д. І. Чумаченко, ‘Словник термінів у сфері Штучного інтелекту’. Міністерство цифрової трансформації України, 2024. [Online]. Available: <https://thedigital.gov.ua/storage/uploads/files/page/community/docs/%D0%A1%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%20%D1%82%D0%B5%D1%80%D0%BC%D1%96%D0%BD%D1%96%D0%B2%20%D1%83%20%D1%81%D1%84%D0%B5%D1%80%D1%96%20%D0%A8%D0%86%20%281%29.pdf>
- [168] S. Liaskovska, S. Tyskyi, Y. Martyn, A. T. Augousti, and V. Kulyk, ‘Systematic Generation and Evaluation of Synthetic Production Data for Industry 5.0 Optimization’, *Technologies*, vol. 13, no. 2, Art. no. 2, Feb. 2025, doi: 10.3390/technologies13020084.
- [169] N. Shakhovska, B. Sydor, S. Liaskovska, O. Duran, Y. Martyn, and V. Vira, ‘High-Fidelity Synthetic Data Generation Framework for Unique Objects Detection’, *Computation*, vol. 13, no. 5, Art. no. 5, May 2025, doi: 10.3390/computation13050120.
- [170] H. Chen, Y. Ji, and D. Evans, ‘Finding Friends and Flipping Frenemies: Automatic Paraphrase Dataset Augmentation Using Graph Theory’, in *Findings of the Association for Computational Linguistics: EMNLP 2020*, Online: Association for Computational Linguistics, 2020, pp. 4741–4751. doi: 10.18653/v1/2020.findings-emnlp.426.
- [171] D. Snow, ‘DeltaPy: A Framework for Tabular Data Augmentation in Python’, Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3582219, Apr. 2020. doi: 10.2139/ssrn.3582219.
- [172] W. Wang and T.-W. Pai, ‘Enhancing Small Tabular Clinical Trial Dataset through Hybrid Data Augmentation: Combining SMOTE and WCGAN-GP’, *Data*, vol. 8, no. 9, p. 135, Aug. 2023, doi: 10.3390/data8090135.

- [173] P. Teutsch, M. S. Ghazijahani, F. Heyder, C. Cierpka, J. Schumacher, and P. Mäder, ‘Large-scale-aware data augmentation for reduced-order models of high-dimensional flows’, *APL Machine Learning*, vol. 3, no. 1, p. 016104, Jan. 2025, doi: 10.1063/5.0213700.
- [174] H.-X. Dou, X.-S. Lu, C. Wang, H.-Z. Shen, Y.-W. Zhuo, and L.-J. Deng, ‘PatchMask: A Data Augmentation Strategy with Gaussian Noise in Hyperspectral Images’, *Remote Sensing*, vol. 14, no. 24, p. 6308, Dec. 2022, doi: 10.3390/rs14246308.
- [175] O. Habibi, M. Chemmakha, and M. Lazaar, ‘Imbalanced tabular data modelization using CTGAN and machine learning to improve IoT Botnet attacks detection’, *Engineering Applications of Artificial Intelligence*, vol. 118, p. 105669, Feb. 2023, doi: 10.1016/j.engappai.2022.105669.
- [176] Z. Zhao, A. Kunar, R. Birke, H. Van Der Scheer, and L. Y. Chen, ‘CTAB-GAN+: enhancing tabular data synthesis’, *Front. Big Data*, vol. 6, p. 1296508, Jan. 2024, doi: 10.3389/fdata.2023.1296508.
- [177] M. Miletic and M. Sariyar, ‘Challenges of Using Synthetic Data Generation Methods for Tabular Microdata’, *Applied Sciences*, vol. 14, no. 14, p. 5975, July 2024, doi: 10.3390/app14145975.
- [178] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, ‘Modeling Tabular data using Conditional GAN’, July 2019, Accessed: Dec. 26, 2020. [Online]. Available: <https://arxiv.org/abs/1907.00503v2>
- [179] Y. Tan, H. Zhu, J. Wu, and H. Chai, ‘DPTVAE: Data-driven prior-based tabular variational autoencoder for credit data synthesizing’, *Expert Systems with Applications*, vol. 241, p. 122071, May 2024, doi: 10.1016/j.eswa.2023.122071.
- [180] A. X. Wang, M. Q. Le, H.-T. Duong, B. N. Van, and B. P. Nguyen, ‘CTVAE: Contrastive Tabular Variational Autoencoder for imbalance data’, *Knowl Inf Syst*, vol. 67, no. 6, pp. 5335–5354, June 2025, doi: 10.1007/s10115-025-02377-7.
- [181] M. Hajihosseini, A. Maghsoudi, and R. Ghezelbash, ‘Stacking: A novel data-driven ensemble machine learning strategy for prediction and mapping of Pb-Zn

- prospectivity in Varcheh district, west Iran’, *Expert Systems with Applications*, vol. 237, p. 121668, Mar. 2024, doi: 10.1016/j.eswa.2023.121668.
- [182] W. Hao, S. Pang, B. Yang, and J. Xue, ‘Tensor-based multi-view clustering with consistency exploration and diversity regularization’, *Knowledge-Based Systems*, vol. 252, p. 109342, Sept. 2022, doi: 10.1016/j.knosys.2022.109342.
- [183] M. S. Meor Yahaya and J. Teo, ‘Data augmentation using generative adversarial networks for images and biomarkers in medicine and neuroscience’, *Front. Appl. Math. Stat.*, vol. 9, p. 1162760, May 2023, doi: 10.3389/fams.2023.1162760.
- [184] ‘What is ensemble learning? | IBM’. Accessed: May 21, 2025. [Online]. Available: <https://www.ibm.com/think/topics/ensemble-learning>
- [185] H. Shteingart *et al.*, ‘Majority Voting and the Condorcet’s Jury Theorem’, 2020, *arXiv*. doi: 10.48550/ARXIV.2002.03153.
- [186] F. Galton, ‘Vox Populi’, *Nature*, vol. 75, no. 1949, pp. 450–451, Mar. 1907, doi: 10.1038/075450a0.
- [187] J. Surowiecki, *The Wisdom of Crowds*. Westminster: Knopf Doubleday Publishing Group, 2005.
- [188] R. S. Smith and T. Windeatt, ‘A Bias-Variance Analysis of Bootstrapped Class-Separability Weighting for Error-Correcting Output Code Ensembles’, in *2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey: IEEE, Aug. 2010, pp. 61–64. doi: 10.1109/ICPR.2010.24.
- [189] D. Wood, T. Mu, A. Webb, H. Reeve, M. Luján, and G. Brown, ‘A Unified Theory of Diversity in Ensemble Learning’, 2023, doi: 10.48550/ARXIV.2301.03962.
- [190] P. M. Granitto, P. F. Verdes, and H. A. Ceccatto, ‘Neural network ensembles: Evaluation of aggregation algorithms’, 2005, *arXiv*. doi: 10.48550/ARXIV.CS/0502006.
- [191] M. a. H. Akhand and K. Murase, ‘Neural Network Ensemble Training by Sequential Interaction’, in *Artificial Neural Networks – ICANN 2007*, Springer, Berlin, Heidelberg, Sept. 2007, pp. 98–108. doi: 10.1007/978-3-540-74690-4_11.
- [192] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, ‘A Neural Network Ensemble With Feature Engineering for Improved Credit Card Fraud

- Detection’, *IEEE Access*, vol. 10, pp. 16400–16407, 2022, doi: 10.1109/ACCESS.2022.3148298.
- [193] N. García-Pedrajas, D. Ortiz-Boyer, R. del Castillo-Gomariz, and C. Hervás-Martínez, ‘Cascade Ensembles’, in *Computational Intelligence and Bioinspired Systems*, Springer, Berlin, Heidelberg, 2005, pp. 598–603. doi: 10.1007/11494669_73.
- [194] Z. Chen, A. N. Joseph Raj, V. Rajangam, W. Li, V. G. V. Mahesh, and Z. Zhuang, ‘Twofold dynamic attention guided deep network and noise-aware mechanism for image denoising’, *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 3, pp. 87–102, Mar. 2023, doi: 10.1016/j.jksuci.2023.02.003.
- [195] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, ‘Ensemble deep learning: A review’, *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105151, Oct. 2022, doi: 10.1016/j.engappai.2022.105151.
- [196] A. Alotaibi, ‘Ensemble Deep Learning Approaches in Health Care: A Review’, *CMC*, vol. 82, no. 3, pp. 3741–3771, 2025, doi: 10.32604/cmc.2025.061998.
- [197] P. B. Vitynskyi, R. O. Tkachenko, and I. V. Izonin, ‘Ансамбль нейромереж GRNN на підставі зміщених поверхонь відгуку для задач електронної комерції’, *SBUNFU*, vol. 29, no. 9, pp. 142–146, Dec. 2019, doi: 10.36930/40290925.
- [198] I. Izonin, R. Tkachenko, M. Havryliuk, M. Gregus, P. Yendyk, and Y. Tolstyak, ‘An Adaptation of the Input Doubling Method for Solving Classification Tasks in Case of Small Data Processing’, *Procedia Computer Science*, vol. 241, pp. 171–178, 2024, doi: 10.1016/j.procs.2024.08.024.
- [199] X. Li *et al.*, ‘BEDS: Bagging ensemble deep segmentation for nucleus segmentation with testing stage stain augmentation’, 2021, *arXiv*. doi: 10.48550/ARXIV.2102.08990.
- [200] I. Kandel, M. Castelli, and A. Popovič, ‘Comparing Stacking Ensemble Techniques to Improve Musculoskeletal Fracture Image Classification’, *Journal of Imaging*, vol. 7, no. 6, Art. no. 6, June 2021, doi: 10.3390/jimaging7060100.
- [201] O. Galchonkov, O. Baranov, S. Antoshchuk, O. Maslov, and M. Babych, ‘Development of a neural network with a layer of trainable activation functions for

- the second stage of the ensemble classifier with stacking’, *EEJET*, vol. 5, no. 9 (131), Art. no. 9 (131), Oct. 2024, doi: 10.15587/1729-4061.2024.311778.
- [202] N. Gordienko, Y. Gordienko, and S. Stirenko, ‘Synchronized Multi-Augmentation with Multi-Backbone Ensembling for Enhancing Deep Learning Performance’, *ASI*, vol. 8, no. 1, p. 18, Jan. 2025, doi: 10.3390/asi8010018.
- [203] T. R. Mahesh *et al.*, ‘AdaBoost Ensemble Methods Using K-Fold Cross Validation for Survivability with the Early Detection of Heart Disease’, *Comput Intell Neurosci*, vol. 2022, p. 9005278, Apr. 2022, doi: 10.1155/2022/9005278.
- [204] M. T. Ramakrishna, V. K. Venkatesan, I. Izonin, M. Havryliuk, and C. R. Bhat, ‘Homogeneous Adaboost Ensemble Machine Learning Algorithms with Reduced Entropy on Balanced Data’, *Entropy*, vol. 25, no. 2, p. 245, Jan. 2023, doi: 10.3390/e25020245.
- [205] Y. V. Bodyanskiy and O. K. Tyshchenko, ‘A Hybrid Cascade Neural Network with Ensembles of Extended Neo-Fuzzy Neurons and Its Deep Learning’, in *Information Technology, Systems Research, and Computational Physics*, Springer, Cham, July 2018, pp. 164–174. doi: 10.1007/978-3-030-18058-4_13.
- [206] F. Geche, O. Mitsa, O. Mulesa, and P. Horvat, ‘Synthesis of a Two Cascade Neural Network for Time Series Forecasting’, in *2022 IEEE 3rd International Conference on System Analysis & Intelligent Computing (SAIC)*, Kyiv, Ukraine: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/SAIC57818.2022.9922991.
- [207] A. Trostianchyn, Z. Duriagina, I. Izonin, R. Tkachenko, V. Kulyk, and O. Pavliuk, ‘Sm-Co alloys coercivity prediction using stacking heterogeneous ensemble model’, *Acta Metall Slovaca*, vol. 27, no. 4, pp. 195–202, Dec. 2021, doi: 10.36547/ams.27.4.1173.
- [208] Z. P. Brodeur, J. D. Herman, and S. Steinschneider, ‘Bootstrap Aggregation and Cross-Validation Methods to Reduce Overfitting in Reservoir Control Policy Search’, *Water Resources Research*, vol. 56, no. 8, p. e2020WR027184, Aug. 2020, doi: 10.1029/2020WR027184.

- [209] U. Michelucci and F. Venturini, ‘Estimating Neural Network’s Performance with Bootstrap: A Tutorial’, *MAKE*, vol. 3, no. 2, pp. 357–373, Mar. 2021, doi: 10.3390/make3020018.
- [210] L. Mochurad, K. Shakhovska, and S. Montenegro, ‘Parallel Solving of Fredholm Integral Equations of the First Kind by Tikhonov Regularization Method Using OpenMP Technology’, in *Advances in Intelligent Systems and Computing IV*, vol. 1080, N. Shakhovska and M. O. Medykovsky, Eds, in *Advances in Intelligent Systems and Computing*, vol. 1080. , Cham: Springer International Publishing, 2020, pp. 25–35. doi: 10.1007/978-3-030-33695-0_3.
- [211] L. Breiman, ‘Random Forests’, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [212] L. Rokach, ‘Ensemble-based classifiers’, *Artif Intell Rev*, vol. 33, no. 1–2, pp. 1–39, Feb. 2010, doi: 10.1007/s10462-009-9124-7.
- [213] R. O. Tkachenko and S. M. Demyanchuk, ‘Побудова емпіричних формул за допомогою багатосферних нейроподібних структур геометричних перетворень’, *Scientific Bulletin of UNFU*, vol. 25, no. 3, Art. no. 3, Mar. 2015.
- [214] О. І. Чумаченко, ‘Структурно-параметричний синтез гібридних нейронних мереж’, дисертація на здобуття наукового степеня доктора технічних наук, Національний технічний університет України “Київський політехнічний інститут ім. І. Сікорського”, Київ, Україна, 2018. [Online]. Available: https://ela.kpi.ua/bitstream/123456789/29639/1/Chumachenko_diss.pdf
- [215] N. Y. Nikolaev and H. Iba, ‘Polynomial harmonic GMDH learning networks for time series modeling’, *Neural Networks*, vol. 16, no. 10, pp. 1527–1540, Dec. 2003, doi: 10.1016/S0893-6080(03)00188-6.
- [216] P. Vitynskyi, R. Tkachenko, I. Izonin, and H. Kutucu, ‘Hybridization of the SGTM Neural-Like Structure Through Inputs Polynomial Extension’, in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, Ukraine: IEEE, Aug. 2018, pp. 386–391. doi: 10.1109/DSMP.2018.8478456.
- [217] H. R. Madala, H. R. Madala, and A. G. Ivachnenko, *Inductive learning algorithms for complex systems modeling*. Boca Raton, Fla.: CRC Press, 1994.

- [218] J. Yeh and J. Yeh, *Real analysis: theory of measure and integration*, 2nd ed. Hackensack, NJ: World Scientific, 2006.
- [219] Ю. П. Зайченко, *Основи проектування інтелектуальних систем. Навчальний посібник*. Київ: Видавничий дім “Слово”, 2004.
- [220] В. Є. Снитюк, *Прогнозування. Моделі. Методи. Алгоритми: Навчальний посібник*. Київ: “Маклаут”, 2008.
- [221] B. O. M. Berezsky O. M., P. O. Yo. Pitsun O. Yo., M. G. M. Melnyk G. M., and T. V. Datsko, ‘Застосування методу лінійної регресії для аналізу кількісних характеристик цитологічних зображень’, *UJIT*, vol. 3, no. 1, pp. 73–77, 2021, doi: 10.23939/ujit2021.03.073.
- [222] I. Izonin, R. Tkachenko, N. Kryvinska, P. Tkachenko, and M. Greguš ml., ‘Multiple Linear Regression Based on Coefficients Identification Using Non-iterative SGTm Neural-like Structure’, in *Advances in Computational Intelligence*, vol. 11506, I. Rojas, G. Joya, and A. Catala, Eds, in Lecture Notes in Computer Science, vol. 11506, Cham: Springer International Publishing, 2019, pp. 467–479. doi: 10.1007/978-3-030-20521-8_39.
- [223] T. Tepla, *Biocompatible materials selection via new supervised learning methods*. LAP LAMBERT Academic Publishing, 2019. Accessed: Mar. 31, 2022. [Online]. Available: <https://nbn-resolving.org/urn:nbn:de:101:1-2019021206023800111337>
- [224] K. Zub, P. Zhezhnych, and C. Strauss, ‘Two-Stage PNN–SVM Ensemble for Higher Education Admission Prediction’, *BDCC*, vol. 7, no. 2, p. 83, Apr. 2023, doi: 10.3390/bdcc7020083.
- [225] I. Izonin, R. Tkachenko, M. Gregus, L. Ryvak, V. Kulyk, and V. Chopyak, ‘Hybrid Classifier via PNN-based Dimensionality Reduction Approach for Biomedical Engineering Task’, in *proceedings of the 18th International Conference on Mobile Systems and Pervasive Computing (MobiSPC), The 16th International Conference on Future Networks and Communications (FNC), The 11th International Conference on Sustainable Energy Information Technology*, in Procedia Computer Science, vol. 191. Leuven, Belgium: Elsevier, Aug. 2021, pp. 230–237. doi: 10.1016/j.procs.2021.07.029.

- [226] А. С. Довбиш, А. В. Васильєв, та В. О. Любчак, *Інтелектуальні інформаційні технології в електронному навчанні: монографія*. Суми: Сумський державний університет, 2013.
- [227] Н. І. Мельникова, ‘Моделі та методи підтримки персоналізованих рішень у медичних системах’, дисертація на здобуття наукового степеня доктора технічних наук, Національний університет ‘Львівська політехніка’, Львів, Україна, 2023. [Online]. Available: <https://lpnu.ua/sites/default/files/2023/dissertation/22379/melnykova-ni-dyser.pdf>
- [228] I. Izonin, R. Tkachenko, S. A. Mitoulis, A. Faramarzi, I. Tsmots, and D. Mashtalir, ‘Machine learning for predicting energy efficiency of buildings: a small data approach’, *Procedia Computer Science*, vol. 231, pp. 72–77, 2024, doi: 10.1016/j.procs.2023.12.173.
- [229] О. С. Міщук, ‘Нейроподібні методи та засоби прогнозування параметрів забруднення атмосферного повітря’, дисертація на здобуття наукового степеня кандидата технічних наук, Національний університет ‘Львівська політехніка’, Львів, Україна, 2021. [Online]. Available: <https://lpnu.ua/sites/default/files/2021/dissertation/9219/mischukdisertaciya0904211.pdf>
- [230] P. J. Rousseeuw, ‘Silhouettes: A graphical aid to the interpretation and validation of cluster analysis’, *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, Nov. 1987, doi: 10.1016/0377-0427(87)90125-7.
- [231] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, 1st edn. in Wiley Series in Probability and Statistics. Wiley, 1990. doi: 10.1002/9780470316801.
- [232] R. Tibshirani, G. Walther, and T. Hastie, ‘Estimating the Number of Clusters in a Data Set Via the Gap Statistic’, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 63, no. 2, pp. 411–423, July 2001, doi: 10.1111/1467-9868.00293.

- [233] T. Calinski and J. Harabasz, ‘A dendrite method for cluster analysis’, *Comm. in Stats. - Theory & Methods*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.
- [234] D. L. Davies and D. W. Bouldin, ‘A Cluster Separation Measure’, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979, doi: 10.1109/TPAMI.1979.4766909.
- [235] I. Izonin, R. Tkachenko, V. Kachmar, and M. Havryliuk, ‘Hybrid ensemble learning model to improve the performance and interpretability of medical diagnosis: Small data tasks’, in *Responsible and Explainable Artificial Intelligence in Healthcare*, Elsevier, 2025, pp. 221–258. doi: 10.1016/B978-0-443-24788-0.00009-1.
- [236] M. H. Stone, ‘The Generalized Weierstrass Approximation Theorem’, *Mathematics Magazine*, vol. 21, no. 4, pp. 167–184, 1948, doi: 10.2307/3029750.
- [237] H. Anton, *Elementary linear algebra*, 10. ed. Hoboken, NJ: Wiley, 2010.
- [238] S. J. Wetzel, R. G. Melko, and I. Tamblyn, ‘Twin neural network regression is a semi-supervised regression algorithm’, *Mach. Learn.: Sci. Technol.*, vol. 3, no. 4, p. 045007, Dec. 2022, doi: 10.1088/2632-2153/ac9885.
- [239] M. Tynes *et al.*, ‘Pairwise Difference Regression: A Machine Learning Meta-algorithm for Improved Prediction and Uncertainty Quantification in Chemical Search’, *J. Chem. Inf. Model.*, vol. 61, no. 8, pp. 3846–3857, Aug. 2021, doi: 10.1021/acs.jcim.1c00670.
- [240] M. K. Belaid, M. Rabus, and E. Hüllermeier, ‘Pairwise Difference Learning for Classification’, in *Discovery Science*, vol. 15244, D. Pedreschi, A. Monreale, R. Guidotti, R. Pellungrini, and F. Naretto, Eds, in Lecture Notes in Computer Science, vol. 15244. , Cham: Springer Nature Switzerland, 2025, pp. 284–299. doi: 10.1007/978-3-031-78980-9_18.
- [241] V. Hrytsyk, A. Borkivskyi, and T. Oliynyk, ‘Achieving the Best Symmetry by Finding the Optimal Clustering Filters for Specific Lighting Conditions’, *Symmetry*, vol. 16, no. 9, p. 1247, Sept. 2024, doi: 10.3390/sym16091247.

- [242] I. Ishikawa, T. Teshima, K. Tojo, K. Oono, M. Ikeda, and M. Sugiyama, ‘Universal Approximation Property of Invertible Neural Networks’, *Journal of Machine Learning Research*, vol. 24, pp. 1–68, 2023.
- [243] K.-L. Du, R. Zhang, B. Jiang, J. Zeng, and J. Lu, ‘Understanding Machine Learning Principles: Learning, Inference, Generalization, and Computational Learning Theory’, *Mathematics*, vol. 13, no. 3, p. 451, Jan. 2025, doi: 10.3390/math13030451.
- [244] P. C. Pendharkar, ‘A Radial Basis Function Neural Network for Stochastic Frontier Analyses of General Multivariate Production and Cost Functions’, *Neural Process Lett*, vol. 55, no. 5, pp. 6247–6268, Oct. 2023, doi: 10.1007/s11063-022-11137-5.
- [245] J. Huang, D. Xu, Y. Li, and Y. Ma, ‘Near-Optimal Tracking Control of Partially Unknown Discrete-Time Nonlinear Systems Based on Radial Basis Function Neural Network’, *Mathematics*, vol. 12, no. 8, p. 1146, Apr. 2024, doi: 10.3390/math12081146.
- [246] T. Poggio and F. Girosi, ‘Networks for approximation and learning’, *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, Sept. 1990, doi: 10.1109/5.58326.
- [247] H. Ranglani, ‘Empirical Analysis Of The Bias-Variance Tradeoff Across Machine Learning Models’, *SSRN Journal*, 2025, doi: 10.2139/ssrn.5086450.
- [248] Б. М. Павлишенко, ‘Методи інтелектуального аналізу консолідованих даних для підтримки прийняття рішень’, дисертація на здобуття наукового степеня доктора технічних наук, Харківський національний університет радіоелектроніки, Харків, Україна, 2021. [Online]. Available: https://nure.ua/wp-content/uploads/2021/Disertation/pavlyshenko_dissertation-3-new-08_04_21_compressed.pdf
- [249] Е. А. Мандзюк, ‘Теоретичні та прикладні засади інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом’, дисертація на здобуття наукового степеня доктора технічних наук, Хмельницький національний університет, Хмельницький, Україна, 2022. [Online]. Available: https://nauka.khmnmu.edu.ua/wp-content/uploads/dysertatsiia_manziuk.pdf

- [250] K. Morimoto, ‘Information Use and the Condorcet Jury Theorem’, *Mathematics*, vol. 9, no. 10, p. 1098, May 2021, doi: 10.3390/math9101098.
- [251] D. H. Wolpert, ‘Stacked generalization’, *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992, doi: 10.1016/S0893-6080(05)80023-1.
- [252] B. C. Boehmke and B. Greenwell, *Hands-on machine learning with R*. in Chapman & Hall/CRC the R series. Boca Raton London New York: CRC Press, Taylor & Francis Group, 2020. doi: 10.1201/9780367816377.
- [253] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. in Springer Series in Statistics. New York, NY: Springer New York, 2009. doi: 10.1007/978-0-387-84858-7.
- [254] I. Izonin, R. Tkachenko, P. Vitynskyi, K. Zub, P. Tkachenko, and I. Dronyuk, ‘Stacking-based GRNN-SGTM Ensemble Model for Prediction Tasks’, in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain: IEEE, Nov. 2020, pp. 326–330. doi: 10.1109/DASA51403.2020.9317124.
- [255] I. Izonin, R. Tkachenko, R. Holoven, M. Shavarskyi, S. Bukin, and I. Shevchuk, ‘Multistage SVR-РБФ-Based Model for Heart Rate Prediction of Individuals’, in *Advances in Artificial Systems for Medicine and Education VI*, vol. 159, Z. Hu, Z. Ye, and M. He, Eds, in Lecture Notes on Data Engineering and Communications Technologies, vol. 159. , Cham: Springer Nature Switzerland, 2023, pp. 211–220. doi: 10.1007/978-3-031-24468-1_19.
- [256] Y. Bodyanskiy, Y. Zaychenko, I. Pliss, and O. Chala, ‘Matrix Neural Network with Kernel Activation Function and Its Online Combined Learning’, in *2022 IEEE 3rd International Conference on System Analysis & Intelligent Computing (SAIC)*, Kyiv, Ukraine: IEEE, Oct. 2022, pp. 1–4. doi: 10.1109/SAIC57818.2022.9922993.
- [257] А. Ю. Шафроненко, ‘Адаптивні методи нечіткої кластеризації потоків даних з використанням еволюційного самонавчання’, дисертація на здобуття наукового степеня доктора технічних наук, Харківський національний університет радіоелектроніки, Харків, Україна, 2025. [Online]. Available:

https://drive.google.com/file/d/1Sm3durJsOMdAi6V_7UoAMWzMOL2-CpYt/view

- [258] N. Shakhovska, I. Izonin, and N. Melnykova, ‘The Hierarchical Classifier for COVID-19 Resistance Evaluation’, *Data*, vol. 6, no. 1, p. 6, Jan. 2021, doi: 10.3390/data6010006.
- [259] R. Kaminsky, N. Shakhovska, N. Kryvinska, and M. Younas, ‘Dendrograms-based disclosure method for evaluating cluster analysis in the IoT domain’, *Computers & Industrial Engineering*, vol. 158, p. 107402, Aug. 2021, doi: 10.1016/j.cie.2021.107402.
- [260] N. Shakhovska, R. Kaminskyi, and B. Khudoba, ‘Experimental study and clustering of operating staff of search systems in the sense of stress resistance’, *Front. Big Data*, vol. 6, p. 1239017, Oct. 2023, doi: 10.3389/fdata.2023.1239017.
- [261] О. Г. Івахненко, *Моделі оптимальної складності та принципи самоорганізації*. Київ: Наукова думка, 1979.
- [262] S. Zhou, ‘An Analysis of The Small Sample Datasets Based on Machine Learning’, in *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen China: ACM, Oct. 2022, pp. 1654–1658. doi: 10.1145/3573428.3573720.
- [263] K. D. Sree, ‘Data Analytics: Role of Activation function In Neural Net’, *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 5, pp. 299–302, 2019.
- [264] H. Rehman, A. Ekelhart, and R. Mayer, ‘Backdoor Attacks in Neural Networks – A Systematic Evaluation on Multiple Traffic Sign Datasets’, in *Machine Learning and Knowledge Extraction*, vol. 11713, A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl, Eds, in Lecture Notes in Computer Science, vol. 11713. , Cham: Springer International Publishing, 2019, pp. 285–300. doi: 10.1007/978-3-030-29726-8_18.
- [265] M. Eltrissi and O. Yousef, ‘Predicting System Surface Parameters Using Artificial Neural Network’, in *ADIPEC*, Abu Dhabi, UAE: SPE, Nov. 2024, p. D041S140R007. doi: 10.2118/222605-MS.

- [266] M. Hammad Waseem, M. Sajjad Ahmed Nadeem, I. Rasool Khan, Seong-O-Shim, W. Aziz, and U. Habib, ‘Reinforcing Artificial Neural Networks through Traditional Machine Learning Algorithms for Robust Classification of Cancer’, *Computers, Materials & Continua*, vol. 75, no. 2, pp. 4293–4315, 2023, doi: 10.32604/cmc.2023.036710.
- [267] D. A. Gzar, A. M. Mahmood, and M. K. Abbas, ‘A Comparative Study of Regression Machine Learning Algorithms: Tradeoff Between Accuracy and Computational Complexity’, *MMEP*, vol. 9, no. 5, pp. 1217–1224, Dec. 2022, doi: 10.18280/mmep.090508.
- [268] M. Yuksekgonul, J. Zou, L. Zhang, and C. E. Guestrin, ‘Beyond Confidence: Reliable Models Should Also Consider Atypicality’, in *Advances in Neural Information Processing Systems*, Dec. 2023, pp. 1–34. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/7900318ffaf5e9bc60250f134c6cc3c7-Paper-Conference.pdf
- [269] H. Wang, W. Wu, T. Chen, X. Dong, and G. Wang, ‘An improved neural network for TOC, S1 and S2 estimation based on conventional well logs’, *Journal of Petroleum Science and Engineering*, vol. 176, pp. 664–678, May 2019, doi: 10.1016/j.petrol.2019.01.096.
- [270] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla, and F. Herrera, ‘A unifying view on dataset shift in classification’, *Pattern Recognition*, vol. 45, no. 1, pp. 521–530, Jan. 2012, doi: 10.1016/j.patcog.2011.06.019.
- [271] J. Song, C. E. Romero, Z. Yao, and B. He, ‘A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler’, *Knowledge-Based Systems*, vol. 118, no. C, pp. 4–14, Feb. 2017, doi: 10.1016/j.knosys.2016.11.003.
- [272] D. Bani-Hani and M. Khasawneh, ‘A Recursive General Regression Neural Network (R-GRNN) Oracle for classification problems’, *Expert Systems with Applications*, vol. 135, pp. 273–286, Nov. 2019, doi: 10.1016/j.eswa.2019.06.018.
- [273] I. Izonin, N. Kryvinska, P. Vitynskyi, R. Tkachenko, and K. Zub, ‘GRNN Approach Towards Missing Data Recovery Between IoT Systems’, in *Advances in Intelligent*

- Networking and Collaborative Systems*, L. Barolli, H. Nishino, and H. Miwa, Eds, in *Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing, 2020, pp. 445–453. doi: 10.1007/978-3-030-29035-1_43.
- [274] Z. Nowroozilarki, B. J. Mortazavi, and R. Jafari, ‘Variational Autoencoders for Biomedical Signal Morphology Clustering and Noise Detection’, *IEEE J. Biomed. Health Inform.*, vol. 28, no. 1, pp. 169–180, Jan. 2024, doi: 10.1109/JBHI.2023.3320585.
- [275] S. Geman, E. Bienenstock, and R. Doursat, ‘Neural Networks and the Bias/Variance Dilemma’, *Neural Computation*, vol. 4, no. 1, pp. 1–58, Jan. 1992, doi: 10.1162/neco.1992.4.1.1.
- [276] R. Tkachenko, A. Doroshenko, I. Izonin, Y. Tsymbal, and B. Havrysh, ‘Imbalance Data Classification via Neural-Like Structures of Geometric Transformations Model: Local and Global Approaches’, in *Advances in Computer Science for Engineering and Education*, vol. 754, Z. Hu, S. Petoukhov, I. Dychka, and M. He, Eds, Cham: Springer International Publishing, 2019, pp. 112–122. doi: 10.1007/978-3-319-91008-6_12.
- [277] G. E. Hinton and R. R. Salakhutdinov, ‘Reducing the Dimensionality of Data with Neural Networks’, *Science*, vol. 313, no. 5786, pp. 504–507, July 2006, doi: 10.1126/science.1127647.
- [278] E. W. Cheney and T. H. Southard, ‘A Survey of Methods for Rational Approximation, with Particular Reference to a New Method Based on a Formula of Darboux’, *SIAM Rev.*, vol. 5, no. 3, pp. 219–231, July 1963, doi: 10.1137/1005065.
- [279] V. Peiris, N. Sharon, N. Sukhorukova, and J. Ugon, ‘Generalised rational approximation and its application to improve deep learning classifiers’, *Applied Mathematics and Computation*, vol. 389, p. 125560, Jan. 2021, doi: 10.1016/j.amc.2020.125560.
- [280] A. P. Holub and L. O. Lysenko, ‘Апроксиманти типу Паде деяких класів функцій кількох змінних’, *I*, vol. 69, no. 5, Art. no. 5, May 2017.

- [281] Y. Pelekh, A. Kunynets, H. Beregova, and T. Magerovska, ‘Methods for solving the initial value problem with a two-sided estimate of the local error’, *Fiz.-mat. model. inf. tehnol.*, no. 33, pp. 88–92, Sept. 2021, doi: 10.15407/fmmit2021.33.088.
- [282] Z. Tang, X. Li, and Z. Wang, ‘Stable parameters identification for rational approximation of Single degree of freedom frequency response function of semi-infinite medium’, *Numerical Meth Engineering*, vol. 124, no. 24, pp. 5596–5615, Dec. 2023, doi: 10.1002/nme.7356.
- [283] P. Lietaert, K. Meerbergen, J. Pérez, and B. Vandereycken, ‘Automatic rational approximation and linearization of nonlinear eigenvalue problems’, *IMA Journal of Numerical Analysis*, vol. 42, no. 2, pp. 1087–1115, Apr. 2022, doi: 10.1093/imanum/draa098.
- [284] N. Boullé, A. Herremans, and D. Huybrechs, ‘Multivariate Rational Approximation of Functions with Curves of Singularities’, *SIAM J. Sci. Comput.*, vol. 46, no. 6, pp. A3401–A3426, Dec. 2024, doi: 10.1137/23M1626414.
- [285] I. Izonin, R. Tkachenko, R. Holoven, K. Yemets, M. Havryliuk, and S. K. Shandilya, ‘SGD-Based Cascade Scheme for Higher Degrees Wiener Polynomial Approximation of Large Biomedical Datasets’, *MAKE*, vol. 4, no. 4, pp. 1088–1106, Nov. 2022, doi: 10.3390/make4040055.
- [286] ‘Medical Cost Personal Datasets’. Accessed: Feb. 17, 2019. [Online]. Available: <https://kaggle.com/mirichoi0218/insurance>
- [287] ‘Heart Attack Prediction’. Accessed: Aug. 16, 2023. [Online]. Available: <https://www.kaggle.com/datasets/imnikhilanand/heart-attack-prediction>
- [288] R. Tkachenko, H. Kutucu, I. Izonin, A. Doroshenko, and Y. Tsymbal, ‘Non-Iterative Neural-Like Predictor for Solar Energy in Libya’, *CEUR Workshop Proceedings*, vol. vol–2105: proceedings of the 14th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume I: Main Conference Kyiv, Ukraine, May 14–17, 2018., pp. 35–45, 2018.
- [289] Ł. Wróbel, A. Gudyś, and M. Sikora, ‘Learning rule sets from survival data’, *BMC Bioinformatics*, vol. 18, no. 1, p. 285, Dec. 2017, doi: 10.1186/s12859-017-1693-x.

- [290] Unknown, ‘Bone marrow transplant: children’. UCI Machine Learning Repository, 2020. doi: 10.24432/C5NP6Z.
- [291] O. Rainio, J. Teuvo, and R. Klén, ‘Evaluation metrics and statistical tests for machine learning’, *Sci Rep*, vol. 14, no. 1, p. 6086, Mar. 2024, doi: 10.1038/s41598-024-56706-x.
- [292] P. Christen, D. J. Hand, and N. Kirielle, ‘A Review of the F-Measure: Its History, Properties, Criticism, and Alternatives’, *ACM Comput. Surv.*, vol. 56, no. 3, pp. 1–24, Mar. 2024, doi: 10.1145/3606367.
- [293] D. Chicco and G. Jurman, ‘A statistical comparison between Matthews correlation coefficient (MCC), prevalence threshold, and Fowlkes–Mallows index’, *Journal of Biomedical Informatics*, vol. 144, p. 104426, Aug. 2023, doi: 10.1016/j.jbi.2023.104426.
- [294] D. Chicco and G. Jurman, ‘The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation’, *BMC Genomics*, vol. 21, no. 1, p. 6, Dec. 2020, doi: 10.1186/s12864-019-6413-7.
- [295] D. Chicco and G. Jurman, ‘The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification’, *BioData Mining*, vol. 16, no. 1, p. 4, Feb. 2023, doi: 10.1186/s13040-023-00322-4.
- [296] S. De Vito, S. D. Vito, E. Massera, M. Piga, L. Martinotto, and G. D. Francia, ‘On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario’, *Sensors and Actuators B: Chemical*, vol. 129, no. 2, pp. 750–757.
- [297] ‘lib.stat.cmu.edu/datasets/bodyfat’. Accessed: June 06, 2025. [Online]. Available: <https://lib.stat.cmu.edu/datasets/bodyfat>
- [298] Z. Fan, R. Chiong, Z. Hu, F. Keivanian, and F. Chiong, ‘Body fat prediction through feature extraction based on anthropometric and laboratory measurements’, *PLoS ONE*, vol. 17, no. 2, p. e0263333, Feb. 2022, doi: 10.1371/journal.pone.0263333.
- [299] ‘Dataset of titanium layers properties after laser sintering depending on chemical and phase composition, size and form of parts, polydispersity and satellites composition. | Request PDF’, ResearchGate. Accessed: Apr. 26, 2022. [Online].

Available:

https://www.researchgate.net/publication/355913560_Dataset_of_titanium_layers_properties_after_laser_sinteringDepending_on_chemical_and_phase_composition_size_and_form_of_parts_polydispersity_and_satellites_composition

- [300] I. Izonin, A. Trostianchyn, Z. Duriagina, R. Tkachenko, T. Tepla, and N. Lotoshynska, ‘The Combined Use of the Wiener Polynomial and SVM for Material Classification Task in Medical Implants Production’, *International Journal of Intelligent Systems and Applications*, vol. 10, no. 9, pp. 40–47, Sept. 2018, doi: 10.5815/ijisa.2018.09.05.
- [301] T. L. Tepla *et al.*, ‘Alloys selection based on the supervised learning technique for design of biocompatible medical materials’, *Archives of Materials Science and Engineering*, vol. 1, no. 93, pp. 32–40, Sept. 2018, doi: 10.5604/01.3001.0012.6944.
- [302] V. Kulyk and I. Izonin, ‘Data for the prediction of fracture toughness of ZrO₂ based ceramics using machine learning algorithms’, ResearchGate. Accessed: Feb. 26, 2023. [Online]. Available: [https://www.researchgate.net/publication/368823655_Data_for_the_prediction_of_fracture_toughness_of_ZrO₂_based_ceramics_using_machine_learning_algorithms](https://www.researchgate.net/publication/368823655_Data_for_the_prediction_of_fracture_toughness_of_ZrO2_based_ceramics_using_machine_learning_algorithms)
- [303] Shakhovska N., ‘Organ transplantation’. figshare, July 04, 2021. doi: 10.6084/m9.figshare.14906241.v1.
- [304] Y. Tolstyak *et al.*, ‘The Ensembles of Machine Learning Methods for Survival Predicting after Kidney Transplantation’, *Applied Sciences*, vol. 11, no. 21, p. 10380, Nov. 2021, doi: 10.3390/app112110380.
- [305] I. Izonin, Y. Tolstyak, V. Kachmar, R. Tkachenko, and V. Chopyak, ‘A Hybrid Two-ML-based Classifier to Predict the Survival of Kidney Transplants One Month after Transplantation’, in *Proceedings of the 6th International Conference on Informatics & Data-Driven Medicine*, Bratislava, Slovakia: CEUR-WS.org, pp. 322–331.
- [306] D. Arpit *et al.*, ‘A Closer Look at Memorization in Deep Networks’, 2017, *arXiv*. doi: 10.48550/ARXIV.1706.05394.

- [307] T. Miseta, A. Fodor, and Á. Vathy-Fogarassy, ‘Surpassing early stopping: A novel correlation-based stopping criterion for neural networks’, *Neurocomputing*, vol. 567, p. 127028, Jan. 2024, doi: 10.1016/j.neucom.2023.127028.
- [308] B. M. Hussein and S. M. Shareef, ‘An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning’, *ITM Web Conf.*, vol. 64, p. 01003, 2024, doi: 10.1051/itmconf/20246401003.
- [309] Y. Bai *et al.*, ‘Understanding and Improving Early Stopping for Learning with Noisy Labels’, presented at the 35th Conference on Neural Information Processing Systems (NeurIPS 2021)., Dec. 2021, pp. 1–12. [Online]. Available: <https://proceedings.neurips.cc/paper/2021/file/cc7e2b878868cbae992d1fb743995d8f-Paper.pdf>
- [310] L. N. Smith and N. Topin, ‘Super-convergence: very fast training of neural networks using large learning rates’, in *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, T. Pham, Ed., Baltimore, United States: SPIE, May 2019, p. 36. doi: 10.1117/12.2520589.

ДОДАТКИ

ДОДАТОК А

**Метрики оцінювання точності розв'язання задач регресії та
класифікації методами машинного навчання**

Таблиця А. 1 – Метрики оцінювання точності розв'язання задачі регресії методами машинного навчання.

№	Назва показника	Формула обчислення показника
1	Максимальна залишкова похибка (ME)	$ME(y_i^{actual}, y_i^{pred}) = \max(y_i^{actual} - y_i^{pred})$
2	Медіана абсолютної похибки (MedAE)	$MedAE(y_i^{actual}, y_i^{pred}) = \text{median}(y_i^{actual} - y_i^{pred} , \dots, y_N^{actual} - y_N^{pred})$
3	Середня абсолютна похибка (MAE)	$MAE(y_i^{actual}, y_i^{pred}) = \frac{1}{N} \sum_i y_i^{actual} - y_i^{pred} $
4	Середньоквадратична похибка (MSE)	$MSE(y_i^{actual}, y_i^{pred}) = \frac{1}{N} \sum_i (y_i^{actual} - y_i^{pred})^2$
5	Середня абсолютна відносна похибка (MAPE)	$MAPE(y_i^{actual}, y_i^{pred}) = \frac{1}{N} \sum_i \frac{ y_i^{actual} - y_i^{pred} }{y_i^{actual}}$
6	Корінь середньоквадратичної похибки (RMSE)	$RMSE(y_i^{actual}, y_i^{pred}) = \sqrt{\frac{\sum_i (y_i^{actual} - y_i^{pred})^2}{N}}$
7	Коефіцієнт детермінації (R ²)	$R^2(y_i^{actual}, y_i^{pred}) = 1 - \frac{\sum_i (y_i^{actual} - y_i^{pred})^2}{\sum_i (y_i^{actual} - \bar{y}_i)^2}, \quad \bar{y}_i = \frac{1}{N} \sum_i y_i^{actual}$

де y^{actual} – фактичне (спостережене) значення цільової (залежної) змінної; y^{pred} – значення цільової (залежної) змінної, отримане із використанням досліджуваної моделі машинного навчання.

Таблиця А. 2 – Метрики оцінювання точності класифікації методами машинного навчання [15]

№	Назва показника	Формула обчислення показника
1	Точність	$Accuracy = \frac{TP + TN}{N}$
2	Влучність	$Precision = \frac{TP}{TP + FP}$
3	Повнота	$Recall = \frac{TP}{TP + FN}$
4	F-1 міра	$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$
5	Коефіцієнт кореляції Метьюса	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}$
6	Каппа Коена	$Cohen's\ Kappa = \frac{P_o - P_e}{1 - P_e},$

де:

- TP (True Positive) – кількість об'єктів позитивного класу, правильно віднесені до цього класу класифікатором на основі машинного навчання;
- TN (True Negative) – кількість об'єктів негативного класу, правильно віднесені до негативного класу класифікатором на основі машинного навчання;
- FP (False Positive) – кількість об'єктів негативного класу, які класифікатор на основі машинного навчання помилково відніс до позитивного класу;
- FN (False Negative) – кількість об'єктів позитивного класу, які класифікатор на основі машинного навчання помилково відніс до негативного класу;
- $P_o = \frac{TP + TN}{TP + TN + FP + FN}$ – спостережена згода;
- $P_e = \frac{(TP + FP)(TP + FN) + (FN + TN)(FP + TN)}{(TP + TN + FP + FN)^2}$ - очікувана згода за випадковим розподілом.

ДОДАТОК Б

**Оптимальні параметри реалізації методів машинного навчання для
розв'язання задач регресії та класифікації**

Таблиця Б.1 – Значення оптимальних параметрів роботи існуючих методів та ітеративних реалізацій методів 3.1 і 3.2, які використовуються для порівняння з методом 3.2.

№	Повна назва методу	Скорочена назва методу	Оптимальні параметри роботи методу
1.	Метод 3.2 із використанням нейронної мережі на основі радіально-базисних функцій	Метод 3.2 на основі РБФ ШНМ	25 центрів РБФ, значення розмаху гаусівської функції (сигма): 0.8300 epoch
2.	Метод 3.1 нейронної мережі на основі радіально-базисних функцій [8]	Метод 3.1 на основі РБФ ШНМ	18 центрів РБФ, значення сигма: 0.448
3.	Метод 3.1 нейронної мережі на основі нейроподібної структури МПГП з РБФ-ядром	Метод 3.1	23 центрів РБФ, значення сигма: 8.1
4.	Метод 3.1 із використанням нейронної мережі узагальненої регресії (НМУР) [10]	Метод 3.1 на основі НМУР	значення сигма: 8.6
5.	Метод 3.1 на основі машини опорних векторів з РБФ-ядром [9]	Метод 3.1 на основі МОВ (РБФ-ядро)	Gamma: 'auto'; Coef0: 0.0; Epsilon: 0.01
6.	Метод 3.1 із використанням машини опорних векторів з поліноміальним ядром [11]	Метод 3.1 на основі МОВ (поліноміальне ядро)	ступінь полінома 3
7.	Нейронна мережа узагальненої регресії [157]	НМУР	значення сигма: 0.076
8.	Нейронна мережа на основі радіально-базисних функцій	РБФ ШНМ	92 центрів РБФ; сигма: 0.44
9.	Машина опорних векторів з РБФ-ядром	МОВ (РБФ-ядро)	Gamma: 'auto'; Coef0: 0.0; Epsilon: 0.001

10.	Машина опорних векторів з поліноміальним ядром	МОВ (поліноміальне ядро)	ступінь полінома 2

Таблиця Б.2 – Значення оптимальних параметрів роботи існуючих методів та ітеративних реалізацій Методу 3.3, які використовуються для порівняння з Методом 3.3.

№	Повна назва методу	Скорочена назва методу	Оптимальні параметри роботи методу
1.	Метод 3.1 для класифікації на основі ймовірнісної нейронної мережі (ЙМН)	Метод 3.1	Евклідова відстань, значення розмаху гаусівської функції (сигма): 0.508605399
2.	Метод 3.3 із використанням двох ймовірнісних нейронних мереж	Метод 3.3 на основі ЙНМ	Евклідова відстань, значення сигм на обох етапах: сигма1=2.69521; сигма2=4.11678
3.	Методу 3.3 із використанням алгоритму екстремального градієнтного бустінгу (ЕГБ)	Метод 3.3 на основі алгоритму ЕГБ	Евклідова відстань, значення сигми: 2.69521, 'n_estimators': 179, 'learning_rate': 0.86680094, 'max_depth': 3
6.	Ймовірнісна нейронна мережа	ЙНМ	Евклідова відстань, значення сигми: 0.559809855
7.	Неітеративна нейроподібна структура Моделі Послідовних Геометричних Перетворень (МПГП) з РБФ-ядром	Нейроподібна структура МПГП з РБФ-ядром	114 центрів РБФ, значення сигми: 3.6563412
8.	Машина опорних векторів (МОВ) з РБФ-ядром	МОВ (РБФ ядро)	'C': 8.488580413820364, 'gamma': 0.003594862073604377
9.	Алгоритм екстремального градієнтного бустінгу	Алгоритм ЕГБ	'n_estimators': 179, 'learning_rate': 0.5598098557162532, 'max_depth': 5

Таблиця Б.3 – Значення оптимальних параметрів роботи існуючих методів та ітеративних реалізацій методу 3.3 методу 3.1, які використовуються для порівняння з методом 3.3.

<i>№</i>	<i>Суть методу</i>	<i>Скорочена назва методу</i>	<i>Оптимальні параметри роботи методу</i>
1.	Метод 3.4 із використанням машини опорних векторів (МОВ) з РБФ-ядром	Метод 3.4 на основі МОВ з РБФ-ядром	Кількість кластерів: 4; C: 10; gamma: 0.01; max_iter: -1
2.	Метод 3.1 із використанням машини опорних векторів (МОВ) з РБФ-ядром	Метод 3.1 на основі МОВ з РБФ-ядром	C: 10; gamma: 0.001; max_iter: -1
3.	Алгоритм категорійного бустінгу	Алгоритм категорійного бустінгу	n_estimators: 64; learning_rate: 0.05; max_depth: 4; min_samples_leaf: 2
4.	Алгоритм випадкового лісу	Алгоритм випадкового лісу	n_estimators: 100; max_depth: 7; min_samples_split: 5; min_samples_leaf: 3
5.	Неітеративна нейроподібна структура Моделі Послідовних Геометричних Перетворень (МПГП) з РБФ-ядром	Нейроподібна структура МПГП з РБФ-ядром	57 центрів РБФ, значення сигми: 1.264589
6.	Машина опорних векторів (МОВ) з РБФ-ядром	МОВ (РБФ-ядро)	C: 0.1; gamma: 0.001; max_iter: -1
7.	Алгоритм екстремального градієнтного бустінгу	Алгоритм ЕГБ	n_estimators: 100; learning_rate: 0.1; max_depth: 5; gamma: 0.2

ДОДАТОК В

Критерії зупинки ітераційних процедур методів компенсації похибки моделі регресії

Таблиця В.1 – Ряд основних критеріїв зупинки для реалізації ітераційної процедури прямого методу компенсації похибки моделі регресії.

<i>Критерій зупинки</i>	<i>Опис</i>	<i>Переваги (“+”) та недоліки (“-”)</i>	<i>Джерело</i>
Фіксована кількість ітерацій	Метод виконує наперед визначену кількість кроків	“+” проста програмна реалізація; “+” апріорне знання часу та обсягу обчислень; “-” досягнення оптимальної точності раніше чи пізніше заданого критерію зупинки.	[306]
Поріг глобальної похибки	Зупинка, коли обраний показник ефективності опускається нижче заданого критерію	“+” чіткий контроль якості; “+” можливість легкої інтерпретації; “-” ігнорування локальних похибок; “-” необхідність надійної оцінки похибки.	[307]
Стабільність крос-валідаційної помилки	Зупинка, коли k-fold CV-помилка стає меншою за задане значення	“+” захист від перенавчання; “+” можливість оцінювання властивостей узагальнення; “-” дуже висока вартість обчислень на кожному кроці.	[308]
Відносне зменшення похибки	Зупинка, якщо відносне зменшення уточненої похибки між двома послідовними ітераціями стає меншим за заданий поріг	“+” автоматичне виявлення точки насичення похибки; “+” уникнення надлишкових ітерацій; “-” потреба в інформації про історію помилок; “-” необхідність підбору правильного порогового значення.	[309]

Таблиця В.2 – Ряд основних критеріїв зупинки для ітераційної процедури зворотного методу компенсації похибки моделі регресії.

<i>Критерій зупинки</i>	<i>Опис</i>	<i>Переваги (“+”) та недоліки (“-”)</i>	<i>Джерело</i>
Фіксована кількість ітерацій	Метод виконується наперед визначену кількість кроків	“+” проста програмна реалізація; “+” апріорне знання часу та обсягу обчислень; “-” досягнення оптимальної точності раніше чи пізніше заданого критерію зупинки.	[306]
Порогова середня величина корекції	Зупинка, якщо середнє абсолютне значення апроксимованих корекцій опускається нижче за встановлений поріг	“+” простота інтерпретації та легкість обчислень; “+” прямий зв’язок із моделлю корекції; “-” ігнорування великих індивідуальних похибок та чутливість до шуму; “-” необхідність правильного вибору порогового значення.	[310]
Зміна параметрів коригувальної моделі	Зупинка, коли зміна ваг або гіперпараметрів другого засобу машинного навчання між ітераціями менша заданого порогу	“+” гарантоване отримання готової моделі; “-” необхідність відстеження внутрішніх параметрів ШНМ.	[310]
Відносне зменшення похибки	Зупинка, якщо відносне зменшення уточненої похибки між двома послідовними ітераціями стає меншим за заданий поріг	“+” автоматичне виявлення точки насичення похибки; “+” уникнення надлишкових ітерацій; “-” потреба у збереженні історії помилок; “-” необхідність підбору порогового значення.	[309]

ДОДАТОК Г**Акти впровадження та використання результатів дисертаційної роботи**



ВНХ.П 17ацц/З
 "17" березень 2025 р.

"ЗАТВЕРДЖУЮ"

Перший Заступник виконавчого директора
 Українського науково-технологічного центру
 Микола ЛУБІВ

АКТ

використання наукових результатів
 дисертаційної роботи Ізоніна Івана Вікторовича,
 представленої на здобуття наукового ступеня доктора технічних наук

Комісія в складі: голови комісії – Першого заступника виконавчого директора Українського науково-технологічного центру, Миколи Лубіва, та членів комісії - к.т.н., с.н.с., Ірини Плісс, к.м.н., Ярослава Толстяка, цим актом підтверджують те, що результати дисертаційної роботи Івана Ізоніна, а саме методологія передбачення і класифікації на основі декартового квадрату множини вхідних компонентів векторів використано у міжнародному науковому проєкті *"Штучний інтелект у медицині: підхід малих даних"* в рамках ініціативи *"Віддалені дослідницькі гранти для українських науковців"*, підтримуваної проєктом EURIZON, який фінансується Європейською Комісією (грантова угода № 871072, <https://www.eurizon-project.eu/>).

Отримані Іваном Ізоніним результати, а саме адитивний метод передбачення на основі декартового квадрату множини вхідних компонентів векторів, методи класифікації на основі декартового квадрату множини вхідних компонентів векторів із використанням принципів каскадування методів машинного навчання, та із використанням комбінованої поверхні відгуку були використані для розв'язання ряду прикладних задач у галузях трансплантології, ревматології та медичної стоматології.

Застосування розроблених методів дозволило отримати результати передбачення чи класифікації підвищеної точності в порівнянні з існуючими алгоритмами машинного навчання та штучними нейронними мережами. Це забезпечило можливість використання отриманих результатів як додаткову інформацію для клініцистів під час діагностики, лікування та реабілітації пацієнтів.

Голова комісії:

Члени комісії:

к.т.н., с.н.с.

к.м.н.



Микола ЛУБІВ

Ірина ПЛІСС

Ярослав ТОЛСТЯК

«ЗАТВЕРДЖУЮ»

Начальник ВМКЦ Західного регіону
полковник медичної служби

Книгиницький В.Ф.

«12»

04

2025 р.



АКТ ВПРОВАДЖЕННЯ

1. Метод класифікації малих даних на основі ансамблю однорідних засобів машинного навчання із використанням випадкових симетричних зміщень для прогнозування ризику летальності пацієнтів з політравмою

найменування пропозиції для впровадження²

2. Національний університет "Львівська політехніка", м. Львів, вул С. Бандери 12

установа, що розробила, її поштова адреса

Ізонін Іван Вікторович

прізвище, ім'я, по-батькові авторів²

3. Джерело інформації: Mortality risk prediction formula for multiple trauma patients admitted to the hospital via machine learning algorithms // Advances in Computer Science for Engineering and Education VII : proceedings of the ICCSEEA 2024, 2024 / eds. Z. Hu, F. Yanovsky, I. Dychka, M. He. – Cham : Springer, 2025. – Vol. 242. – (Lecture Notes on Data Engineering and Communications Technologies). – P. 208–219. – DOI: https://doi.org/10.1007/978-3-031-84228-3_18.

назва, рік видання методичних рекомендацій, інформаційного листа, вихідні дані статті, № а.с., і т. п.²

4. Впроваджено у Військово-медичний клінічний центр Західного регіону

найменування лікувально-профілактичного закладу⁴

5. Термін впровадження⁴ з 01.2024 по 01.2025

6. Загальна кількість спостережень⁴ 73

7. Ефективність впровадження у відповідності з критеріями, викладеними у джерелі інформації (п. 3)

Показники ⁴	За даними	
	авторів, які пропонують впровадження	організації, що впровадила
Клінічна ефективність	на 91%	84%

8. Зауваження, пропозиції⁴

М.М.М.М.

Відповідальний за впровадження

Провідний хірург – заступник начальника Центру
Полковник медичної служби

Прохоренко Г.А.

«12» 04 2025 р.

Начальник клініки нейрохірургії та неврології ВМКЦ ЗР.
Полковник медичної служби

Алексєєв О.М.

«12» 04 2025 р.

посада, підпис, прізвище, ім'я, по-батькові

1 Узагальнені акти впровадження затверджує заст. Завідуючого відділом охорони здоров'я.

2 Заповнюється розробником.

3 Заповнюється організацією, яка впровадила розробку.

4 В акт вдруковуються тільки ті показники, яким відповідає дана розробка.

«ЗАТВЕРДЖУЮ»
 Генеральний директор
 КНП ХОР «ОМКЦУН ІМ.
 В.І. ШАПОВАЛА»
 Владислав ДЕМЧЕНКО



АКТ ПРО ВПРОВАДЖЕННЯ

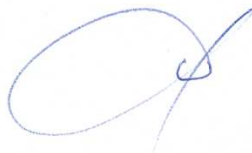
матеріалів дисертаційної роботи Ізоніна Івана Вікторовича
 на здобуття наукового ступеня доктора технічних наук

1. **Найменування пропозиції для впровадження:** - метод класифікації на основі стекінгового об'єднання неітеративних нейроподібних структур з нелінійним розширенням входів на основі радіально-базисних функцій, використовується для прогнозування виживання ниркового трансплантанта через місяць після пересадки.
2. **Найменування установи:** Національний університет «Львівська Політехніка», вул. Степана Бандери, 12, Львів, 79013, Україна.
3. **Автор:** Ізонін Іван Вікторович.
4. **Джерело інформації:** П. В. Вітинський, Р. О. Ткаченко, І. В. Ізонін, and Н. О. Кустра, «Ансамблі нейроподібних структур МППІ з RBF розширенням входів для задач регресії та класифікації», Вісник Хмельницького національного університету. Серія: Технічні науки, vol. 4, no. 275, pp. 72–79, 2019.
5. **Де і коли впроваджено:** впроваджено в лікувальну роботу КНП ХОР «ОМКЦУН ІМ.В.І. ШАПОВАЛА» у жовтні 2022 року – вересні 2024 року.
6. **Результати застосування методу** за період з 01.10.2022 року по 20.09.2024 року – позитивні.
7. **Ефективність впровадження** – метод експрес-діагностики забезпечує високу точність прогнозу і надає додаткову інформацію для лікаря, яка використовується для вибору стратегії лікування, моніторингу стану пацієнтів після трансплантації нирки та своєчасне коригування імуносупресивної терапії.
8. **Зауваження, пропозиції** – немає.

Цей акт складено для подачі до спеціалізованої вченої ради для захисту дисертації і не є підставою для фінансових розрахунків.

Відповідальний за впровадження: завідувач відділення трансплантації нирки КНП ХОР «ОМКЦУН ІМ.В.І. ШАПОВАЛА» В.Л.Лисаченко.

Генеральний директор



Владислав ДЕМЧЕНКО

«ЗАТВЕРДЖУЮ»

Генеральний Директор Комунального
некомерційного підприємства Харківської
обласної ради «Обласна клінічна лікарня»

Врагов Р.С.



02.12.2024

АКТ ПРО ВПРОВАДЖЕННЯ

матеріалів дисертаційної роботи Ізоніна Івана Вікторовича
на здобуття наукового ступеня доктора технічних наук

1. **Найменування пропозиції для впровадження:** - метод кусково-лінійної апроксимації багатовимірних поверхонь відгуку та програмний засіб на його основі для розв'язання задачі прогнозування міцності трабекулярної кістки на основі даних пацієнтів, які страждали від ерозивної форми остеоартриту та перенесли тотальне ендопротезування кульшового суглоба.
2. **Найменування установи:** Національний університет «Львівська Політехніка», вул. Степана Бандери, 12, Львів, 79013, Україна.
3. **Автор:** Ізонін Іван Вікторович.
4. **Джерело інформації:** І. В. Ізонін, Б. Ільчишин, М. Гаврилюк, О. Бойчук, 'Ансамбль однорідних методів машинного навчання з усередненням результату', Інформаційні технології і автоматизація - 2023 : матеріали XVI Міжнародної науково-практичної конференції, Одеса, 2023, ст. 339–340.
5. **Де і коли впроваджено:** впроваджено в роботу Комунального некомерційного підприємства Харківської обласної ради «Обласна клінічна лікарня» у жовтні 2022 року – листопаді 2024 року.
6. **Результати застосування методу** за період з 01.10.2022 року по 15.11.2024 року – позитивні.
7. **Ефективність впровадження** – 89 %.
8. **Зауваження, пропозиції** – немає.

Цей акт складено для подачі до спеціалізованої вченої ради для захисту дисертації і не є підставою для фінансових розрахунків.

Відповідальний за впровадження:

Завідувач відділенням травматології

І.І. Спесивий

«ЗАТВЕРДЖУЮ»

Проректор з наукової роботи
Харківського національного
медичного університету
проф. Валерій М'ЯСОВСЬКИЙ
«04» _____ 2024 р.

АКТ ПРО ВПРОВАДЖЕННЯ

- 1. Найменування пропозиції** (метод профілактики, діагностики, лікування, пристрій, форма організаційної роботи та ін.) Метод класифікації даних на основі розширення входів лінійної нейроподібної структури вихідними сигналами шару сумування ймовірнісної нейронної мережі.
- 2. Ким і коли запропоновано** Ізоніним Іваном Вікторовичем у межах дисертаційної роботи на здобуття наукового ступеня доктора технічних наук
- 3. Джерело інформації** (інформаційний лист, звіт про НДР, дисертація, монографія, з'їзди, конгреси, конференції, семінари тощо)
Стаття Ivan Izonin, Roman Tkachenko, Zoia Duriagina. Smart Web Service of Ti-Based Alloy's Quality Evaluation for Medical Implants Manufacturing. Applied Sciences. 2022; 12(10):5238
- 4. Де і коли впроваджено** На кафедрі стоматології дитячого віку та імплантології Харківського національного медичного університету чинаючи з 2022 року.
- 5. Результати застосування методу за період з 2022 по 2024 рр.**
Метод успішно використовується для визначення властивостей матеріалів з титанових сплавів, призначених для виготовлення біосумісних зубних імплантів.
- 6. Ефективність впровадження за критеріями, висловленими в джерелі інформації (п.3)** Запропонований метод є альтернативою традиційним підходам до оцінки властивостей титанових сплавів, значно знижуючи матеріальні, технологічні та часові витрати. Це забезпечує можливість виготовлення високоякісних титанових зубних імплантів, що відповідають сучасним вимогам медичної практики. Матеріали впроваджуються в навчальний та науковий процес кафедри стоматології дитячого віку та імплантології.
- 7. Зауваження, пропозиції** Подальше вдосконалення методу може включати його адаптацію для аналізу інших біосумісних матеріалів.

Відповідальний(і) за впровадження

Член комісії проф., д. мед.н.,
професорка кафедри стоматології
дитячого віку та імплантології Назарян Р.С.

01.03.2024

(дата)

(підпис)



ЗАТВЕРДЖУЮ

Проректор з наукової роботи
Національного університету
"Львівська політехніка"

Іван ДЕМИДОВ

2025 р.

АКТ

**про використання наукових результатів
дисертаційної роботи Ізоніна Івана Вікторовича
представленої на здобуття наукового ступеня доктора технічних наук**

Комісія в складі: голови комісії – начальника науково-дослідної частини д.т.н., с.н.с. Небесного Р.В. та членів комісії – директора ІКНІ, д.т.н., проф. Шаховської Н.Б., професора кафедри СШ, д.т.н., проф. Яковини В.С., доцента кафедри СШ, к.т.н., доц. Кривенчука Ю.П. та доцента кафедри СШ, к.т.н., доц. Хавалка В.М. цим актом підтверджують, що результати дисертаційного дослідження Ізоніна Івана Вікторовича, використано у науково-дослідних роботах фінансованих Міністерством освіти і науки України, які виконувались на кафедрі систем штучного інтелекту і включено до відповідних звітів за результатами виконання цих робіт.

Зокрема, метод формування комбінованої поверхні відгуку із застосуванням кластерного аналізу використано в межах виконання науково-технічної (експериментальної) розробки "Методи та засоби інформаційної безпеки та гігієни на основі інтерпретованого штучного інтелекту" (№ 0121U107809, 2022–2024) для підвищення точності лінійних моделей регресії та класифікації на вибірках даних малих обсягів без втрати узагальнення.

Розроблені прямий та зворотний методи компенсації похибки моделі регресії на основі використання раціональних дробів використано в межах науково-технічної (експериментальної) розробки "Технології та системи оброблення і зберігання персоналізованих військових медичних даних" (№ 0121U107809, 2022–2023), для підвищення точності та робастності апроксимації в задачах аналізу біомедичних даних малих обсягів.

Розроблений метод синтезу ансамблів гомогенних нелінійних штучних нейронних мереж на основі агрегування апроксимованих поверхонь відгуку, сформованих на незалежних вибірках даних із використанням симетричних випадкових зміщень використано в межах науково-технічної (експериментальної) розробки "Розроблення інформаційної технології оцінювання та прогнозування надійності програмного забезпечення методами машинного навчання" (№ 0121U109527, 2021–2022) для підвищення точності та стійкості моделей регресії та класифікації.

Розроблений метод нелінійного розширення простору незалежних ознак біомедичної

вибірки даних вихідними сигналами шару підсумовування ймовірнісної нейронної мережі використано в межах прикладного дослідження “Інформаційна технологія опрацювання персоналізованої медичної інформації” (№ 0119U002257, 2019–2020) для підвищення точності класифікації даних на лінійних неітеративних нейроподібних структурах.

Голова комісії:
Начальник науково-дослідної частини
д.т.н., с.н.с.



Роман НЕБЕСНИЙ

Члени комісії:

Директор ІКНІ
д.т.н., професор



Наталія ШАХОВСЬКА

Члени комісії:

Професор кафедри СШ
д.т.н., професор



Віталій ЯКОВИНА

Доцент кафедри СШ
к.т.н., доцент

Юрій КРИВЕНЧУК

Доцент кафедри СШ
к.т.н., доцент



Віктор ХАВАЛКО



ЗАТВЕРДЖУЮ

Проректор з наукової роботи
Національного університету
Львівська політехніка

Іван ДЕМИДОВ

2025 р.

АКТ

**про використання наукових результатів
дисертаційної роботи Ізоніна Івана Вікторовича
представленої на здобуття наукового ступеня доктора технічних наук**

Комісія в складі: голови комісії – начальника науково-дослідної частини д.т.н., с.н.с. Небесного Р.В. та членів комісії – директора ІКНІ, д.т.н., проф. Шаховської Н.Б., професора кафедри СШІ, д.т.н., проф. Яковини В.С., доцента кафедри СШІ, к.т.н., Кривенчука Ю.П. та доцента кафедри СШІ, к.т.н., Хавалка В.М. цим актом підтверджують, що результати дисертаційної роботи Ізоніна Івана Вікторовича, використано у науково-дослідних роботах фінансованих Національним Фондом Досліджень України, які виконувались на кафедрі систем пугного інтелекту і включено до відповідних звітів за результатами виконання цих робіт.

Зокрема, розроблений метод часткової компенсації систематичної похибки нейронної мережі узагальненої регресії використано в межах виконання прикладного дослідження *“Система підтримки прийняття рішень моделювання поширення вірусних інфекцій” (№ 211/01.220)* для розроблення ансамблю двох нейронних мереж, який забезпечує підвищення точності розв’язання задач регресії у випадку аналізу біомедичних даних малих обсягів.

Удосконалений метод глобально-локальної апроксимації біомедичних даних малих обсягів, використано в межах виконання прикладного дослідження *“Методи та засоби дослідження маркерів старіння та їх впливу на постковідні ефекти для продовження працездатного періоду” (№ 2021.01/0103)* для підвищення точності моделей регресії та класифікації за рахунок незалежного виділення глобальної та локальної компонент пуганого сигналу із використанням ансамблю з двох нейронних мереж узагальненої регресії, із можливістю додаткового ітеративного уточнення шуканого значення.

Голова комісії:
Начальник науково-дослідної частини
д.т.н., с.н.с.

Роман НЕБЕСНИЙ

Члени комісії:

Директор ІКНІ
д.т.н., професор

Наталія ШАХОВСЬКА

Члени комісії:
Професор кафедри СШІ
д.т.н., професор
Доцент кафедри СШІ
к.т.н., доцент
Доцент кафедри СШІ
к.т.н., доцент

Віталій ЯКОВИНА

Юрій КРИВЕНЧУК

Віктор ХАВАЛКО

“ЗАТВЕРДЖУЮ”



Проректор з науково-педагогічної роботи
Національного університету
"Львівська політехніка"

Олег ДАВИДЧАК

2025 р.

про впровадження в навчальний процес результатів
дисертаційної роботи
Ізоніна Івана Вікторовича

Цей акт складено про те, що результати дисертаційної роботи Ізоніна Івана Вікторовича впроваджено у навчальний процес кафедри "Систем штучного інтелекту" Національного університету "Львівська політехніка".

Впровадження результатів дисертаційної роботи полягає в їхньому використанні при викладанні навчальних дисциплін як окремих розділів лекційних курсів.

Зокрема, при викладанні дисципліни "Основи штучного інтелекту" для студентів освітнього ступеня "бакалавр" за спеціальністю 122 "Комп'ютерні науки" використано такі результати дисертаційного дослідження:

- методи нелінійного розширення простору вхідних даних для підвищення точності розв'язання задач класифікації та регресії;
- методи інтелектуального аналізу структурованих вибірок даних малих обсягів.

Окрім цього, при викладанні дисципліни "Моделювання та аналіз даних у інформаційних системах" для студентів освітнього ступеня "магістр" за спеціальністю 122 "Комп'ютерні науки" використано такі результати дисертаційного дослідження:

- методи синтезу гомогенних ансамблів штучних нейронних мереж лінійного та нелінійного типу з різними підходами до агрегування результатів;
- методи компенсації похибок роботи штучних нейронних мереж.

Голова комісії:

Директор ІКНІ

д.т.н., професор

Наталія ШАХОВСЬКА

Члени комісії:

Доцент кафедри СШІ

к.т.н., доцент

Віктор ХАВАЛІКО

Доцент кафедри СШІ

к.т.н., доцент

Леся ГЕНТОШ

ДОДАТОК Д

Свідоцтва про реєстрацію авторського права на комп'ютерні програми

УКРАЇНА



СВІДОЦТВО

про реєстрацію авторського права на твір

№ 127079

Комп'ютерна програма «Бібліотека методів інтелектуального аналізу малих даних на основі декартового добутку впорядкованих множин входів»

(вид. назва твору)

Автор (співавтори) Ізонін Іван Вікторович, Ткаченко Роман Олексійович, Томаши Євген Олександрович, Гентош Леся Ігорівна

(прізвище, ім'я, по батькові (за наявності), псевдонім (за наявності))

Авторські майнові права належать спільно Ізонін Іван Вікторович, просп. Шевченка, 27, кв. 28, м. Червоноград, Львівська обл., 800100; Ткаченко Роман Олексійович, вул. Йосипа Сліпого, 14, кв. 5, м. Львів, 29017; Томаши Євген Олександрович, вул. Протитанкова, 43, м. Макіївка, Донецька обл., 86111; Гентош Леся Ігорівна, вул. Виговського, 7, кв. 30, м. Львів, 79022

(прізвище, ім'я, по батькові (за наявності) фізичної особи / найменування юридичної особи, адреса)

Дата реєстрації 4 червня 2024 р.

**Директор Державної організації
«Український національний
офіс інтелектуальної власності
та інновацій»**


Олена ОРЛЮК



УКРАЇНА



СВІДОЦТВО

про реєстрацію авторського права на твір

№ 114033

Комп'ютерна програма «Гомогенні стекінгові ансамблі штучних нейронних мереж без навчання (Multi-GRNN)» («Multi-GRNN»)

(вид, назва твору)

Автор(и) Ізонін Іван Вікторович, Нагірний Святослав Богданович, Ткаченко Роман Олексійович

(повне ім'я, псевдонім (за наявності))

Дата реєстрації 9 серпня 2022 р.

В.о. Генерального директора
Державного підприємства
«Український інститут
інтелектуальної власності»

Олександр ОПАНАСЕНКО







СВІДОЦТВО

про реєстрацію авторського права на твір

№ 93094

Комп'ютерна програма "Комбінована нейроподібна система для задач аналітики в умовах коротких вибірок даних"

(вид, назва твору)

Автор(и) Вітинський Павло Богданович, Ткаченко Роман Олексійович, Ізонін Іван Вікторович

(повне ім'я, псевдонім (за наявності))

Дата реєстрації

16.10.2019



**Заступник Міністра розвитку економіки,
торгівлі та сільського господарства
України Д. О. Романович**

Reu