

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Підпис

Кваліфікаційна наукова
праця на правах рукопису

Власенко Олександр Миколайович

УДК 004.032.26

ДИСЕРТАЦІЯ

**Методи та моделі інтелектуалізації процесів оперативного аналізу
ризиків на основі м'яких обчислень**

05.13.23 – системи та засоби штучного інтелекту

Подається на здобуття наукового ступеня кандидата технічних наук

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

Науковий керівник

Бодянський Євгеній Володимирович,
доктор технічних наук, професор

Цей примірник дисертаційної роботи
ідентичний за змістом з іншими,
поданими до спеціалізованої вченої
ради

Учений секретар спецради Д 64.052.01

Підпис

Є.І. Литвинова

Харків – 2021

АНОТАЦІЯ

Власенко О.М. Методи та моделі інтелектуалізації процесів оперативного аналізу ризиків на основі м'яких обчислень. – Кваліфікаційна наукова робота на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук (доктора філософії) за спеціальністю **05.13.23 - системи та засоби штучного інтелекту**. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2021.

Мета дослідження – розробка методів та моделей інтелектуалізації процесів оперативного аналізу ризиків в умовах неповної інформації на основі гібридних м'яких обчислень, які вирізняються високою точністю та обчислювальною ефективністю.

Задачі дослідження: 1) Провести аналіз існуючих методів та підходів в задачах прогнозування, ідентифікації та аналізу, які використовуються у оперативному аналізі ризиків, виявити їх недоліки та можливості для поліпшення за рахунок використання м'яких обчислень. 2) Розробити нейро-фаззі модель зі швидким навчанням для використання у задачах прогнозування, ідентифікації та діагностики, яка вирізняється здатністю представляти складні залежності у вхідних даних та компактним розміром. 3) Розробити ансамбль нейро-фаззі моделей для зменшення похибки та спрощення процедури селекції гіперпараметрів. 4) Розробити метод навчання, який використовує метод емпіричної модової декомпозиції для зниження шумів. 5) Створити ієрархічну модель підтримки прийняття рішень, яка би інкорпоровувала запропоновані моделі. 6) Провести експериментальні дослідження розроблених методів, вирішити за їх допомогою практичні задачі.

Об'єкт дослідження – процес інтелектуального аналізу даних для виявлення та оцінки ризиків.

Предмет дослідження – методи та моделі інтелектуального аналізу даних на базі м'яких обчислень.

Методи дослідження – базуються на теорії м'яких обчислень, а саме на теорії штучних нейронних мереж та теорії нечіткої логіки для побудови нейро-фаззі моделей та їхніх ансамблів; теорії оптимізації для синтезу методів навчання. Імітаційне моделювання використовується для перевірки ефективності розроблених моделей.

Вперше запропоновано п'ятишарову гібридну нейро-фаззі модель з багатовимірними гаусіанами у консеквентному шарі та швидкий метод її навчання з варіантами на базі квадратичної похибки та спеціалізованого критерію, що характеризується високою точністю, швидкістю обробки та обчислювальною швидкістю у процесі навчання за рахунок зменшення налаштованих параметрів моделі.

Вперше запропоновано ансамбль гібридних нейро-фаззі моделей з багатовимірними гаусіанами у консеквентному шарі та метод його синтезу, що характеризується зменшеною похибкою навчання, покращеними узагальнюючими можливостями та спрощеною процедурою селекції гіперпараметрів.

Удосконалено метод навчання нейро-фаззі моделей та їх ансамблів застосуванням емпіричної модової декомпозиції, що відрізняється від аналогів можливістю зниження рівня шуму у випадку високодинамічних даних.

Набув подальшого розвитку метод побудови ієрархічних знання-орієнтованих нечітких систем шляхом включення оцінок ймовірностей у процес логічного висновування та використання нейро-фаззі моделей як компонентів ієрархічної структури, що на відміну від існуючих методів дозволяє опрацьовувати малоймовірні критичні значення у задачах оперативного аналізу ризиків.

Проведено низку імітаційних експериментів на основі тестових та реальних даних, результати яких підтверджують доцільність застосування

запропонованого підходу для вирішення задач інтелектуалізації процесів оперативного аналізу ризиків. Розв'язано практичну задачу на базі розроблених методів і моделей для вирішення задачі оперативного керування фінансовим портфоліо.

Практична значимість отриманих результатів. Використання запропонованих методів та моделей дозволяє підвищити ефективність інтелектуальних систем аналізу ризиків в умовах неповних даних. Розроблені модулі з запропонованими методами підтвердили свою ефективність у задачах динамічного прогнозування складних процесів та ідентифікації потенційних ризиків.

Розроблене програмне забезпечення, що реалізує запропоновані і досліджені в дисертації методи, впроваджено

– у ТОВ «САЙТОСС» при розробці програмного забезпечення для управління ризиками інвестиційного портфоліо (акт впровадження від 14.01.2021);

– у Харківському національному університеті радіоелектроніки на кафедрі штучного інтелекту в освітній процес з курсу «Нейромеревеві методи обчислювального інтелекту» та «Штучні нейронні мережі: архітектура, навчання, застосування».

Публікації. За тематикою дослідження опубліковано 15 наукових праць, з них 1 розділ у колективній монографії, що входить до науково-метричної бази SCOPUS, 2 статті за кордоном, що входить до науково-метричної бази SCOPUS, 2 статті у виданнях, які зазначені в переліках фахових видань України з технічних наук, 10 публікацій у матеріалах конференцій (3 включено до науково-метричної бази даних SCOPUS).

Ключові слова: нейро-фаззі моделі, градієнтні методи навчання, метод емпіричної модової декомпозиції, ансамблі нейро-фаззі моделей.

СПИСОК ОПУБЛІКОВАНИХ РОБІТ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "A novel neuro-fuzzy model for multivariate time-series prediction", *Data*, 3(4), 62, 2018. (Входить до міжнародних науково-метричних баз **Scopus** та **Web of Science**)
2. A. Vlasenko, N. Vlasenko, O. Vynokurova, Y. Bodyanskiy and D. Peleshko. "A Novel ensemble neuro-fuzzy model for financial time series forecasting", *Data*, 4(3), 126, 2019. (Індексується міжнародними науково-метричними базами **Scopus** та **Web of Science**)
3. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "An Empirical Mode Decomposition Based Method to Synthesize Ensemble Multidimensional Gaussian Neuro-Fuzzy Models in Financial Forecasting", *Communications in Computer and Information Science*, volume 1158, pp. 140-149, 2020. Springer, Cham. (Індексується міжнародною науково-метричною базою **Scopus**)
4. А. Власенко, Е. Кучеренко, «Иерархическая нечетко-вероятностная модель в задачах управления рисками», *Системи обробки інформації*, Вип. 1, С. 145-149, 2013. (Входить до міжнародної науково-метричної бази **Index Copernicus International**)
5. А. Власенко, Е. Кучеренко. «Метод логического вывода в иерархических нечетко-вероятностных моделях сложных систем», *Бионика интеллекта*, Вып. 2 (81), С. 47–52, 2013.
6. A. Vlasenko, O. Vynokurova, N. Vlasenko and M. Peleshko, "A Hybrid Neuro-Fuzzy Model for Stock Market Time-Series Prediction", in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, 2018. (Індексується міжнародними науково-метричними базами **Scopus** та **Web of Science**)
7. A. Vlasenko, Y. Rashkevych, N. Vlasenko, D. Peleshko and O. Vynokurova, "A Hybrid EMD - Neuro-Fuzzy Model for Financial Time Series

Analysis", in *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, Ukraine, 2020. (Індексується міжнародною науково-метричною базою **Scopus**)

8. A. Vlasenko, N. Vlasenko, O. Vynokurova and Y. Bodyanskiy, "An Enhancement of a Learning Procedure in Neuro-Fuzzy Model", in *2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC)*, Kyiv, Ukraine, 2018. (Індексується міжнародними науково-метричними базами **Scopus** та **Web of Science**)

9. A. Vlasenko, Y. Bodyanskiy, «An approach to overcome prediction delay problem in hybrid neuro-fuzzy networks for financial time-series forecasting», в *Матеріали VI Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології»*, Одеса, 2017.

10. Y. Bodyanskiy, A. Vlasenko, O. Vynokurova and I. Pliss, "A hybrid neuro-fuzzy network for financial time-series prediction", в *Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту»*, Залізний порт, Україна, 2017.

11. А. Власенко, Е. Кучеренко, «Применение нечеткой кластеризации в задачах структурной идентификации нечетких моделей Такаги-Сугено-Канга» в *Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту»*, Євпаторія, Україна, 2012.

12. А. Власенко, Е. Кучеренко, «Архитектура иерархической системы логического вывода с нечетко-вероятностной базой знаний», в *Материалы XI международной научно-практической конференции "Математическое обеспечение интеллектуальных систем"* Днепропетровск, 2013.

13. А. Власенко, Е. Кучеренко, «Иерархические системы нечеткого вывода Такаги-Сугено в задачах управления рисками», в *Материалы XVI*

международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2012.

14. А. Власенко, Е. Кучеренко, «Реализация иерархической нечетко-вероятностной модели логического вывода на базе библиотеки Aforge.Net», в *Материалы XVII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2013.*

15. А. Власенко, Е. Кучеренко, «Реляционная база нечетко-вероятностных знаний для логического вывода в задачах анализа рисков», в *Материалы XVIII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2014.*

ABSTRACT

Oleksandr Vlasenko. Soft Computing methods and models of risk analysis processes intellectualization.

Dissertation for obtaining the scientific degree of the candidate of technical sciences in the specialty **05.13.23 – Systems and tools of artificial intelligence.** – Kharkiv National University of Radio Electronics, Ministry of Education and Science of Ukraine, Kharkiv, 2021.

The purpose of the study is to develop hybrid soft computing methods and models for intellectualizing the processes of operational risk analysis in conditions of incomplete information, which are characterized by high accuracy and computational efficiency.

Research objectives: 1) Analyze the existing methods and approaches in the tasks of forecasting, identification and analysis applied to operational risk analysis in order to identify their weaknesses and points for improvement through the use of soft computing. 2) Develop a fast-learning neuro-fuzzy model for use in prediction, identification and diagnostic tasks, which is characterized by the ability to represent

complex dependencies in the input data and compact size. 3) Develop an ensemble of neuro-fuzzy models to reduce error and simplify the procedure of selection of hyperparameters. 4) Develop a neuro-fuzzy models learning method that uses the method of empirical mode decomposition to reduce noise. 5) Create a hierarchical decision support model, which would incorporate the proposed models. 6) Perform experiments with the developed methods, solve practical problems with their software implementation.

The object of research is the process of intellectual data analysis to identify and assess risks.

The subject of research is soft computing methods and models for intellectualizing the processes of operational risk analysis.

Research methods are based on the theory of soft computing, namely on the theory of artificial neural networks and the theory of fuzzy logic for the construction of neuro-fuzzy models and their ensembles, optimization theory for the synthesis of learning methods. Software simulation is used to test the effectiveness of developed models.

for the first time:

- a novel five-layer hybrid neuro-fuzzy model with multidimensional Gaussians in the consequent layer and a fast learning method with variants based on standard quadratic error and a specialized criterion are proposed. It is distinguished by high accuracy, processing speed and computational effectiveness.

- an ensemble of hybrid neuro-fuzzy models with multidimensional Gaussians in a consequent layer and a method of its synthesis are proposed, characterized by reduced learning error, improved generalization capabilities and a simplified procedure for selecting hyperparameters.

were improved:

- the neuro-fuzzy models learning methods are improved by using empirical mode decomposition as a tool for preliminary data decomposition and noise

reduction, which differs from competitors by the advanced noise reduction in the case of highly dynamic data.

– the method of constructing hierarchical knowledge-oriented fuzzy systems by incorporating probability estimates into the process of logical inference and using neuro-fuzzy models as components of a hierarchical structure have been further developed, which, unlike existing methods, allows to handle critical values in operational risk analysis.

A number of simulation experiments based on synthetic and real data were performed, the results of which confirm the feasibility of applying the proposed techniques to solve the operational risk analysis processes intellectualization problem. The practical task of financial portfolio operative risk management is solved on the basis of the developed methods and models.

The significant practical results are obtained. The use of the proposed methods and models allows to increase the efficiency of intelligent risk analysis systems in the conditions of incomplete data. The developed modules with the proposed methods have confirmed their effectiveness in the tasks of dynamic forecasting of complex processes and identification of potential risks.

Developed software implementations of the proposed methods and models were used:

– in LLC "SITEOSS" in the development of software for investment portfolio risk management (implementation act from 14.01.2021);

– at the Kharkiv National University of Radio Electronics at the Department of Artificial Intelligence in the educational process, in the courses "Neural Network Methods of Computational Intelligence" and "Artificial Neural Networks: Architecture, Learning, Applications".

15 scientific works have been published on the subject of the research, including 1 book chapter, indexed in Scopus, 2 articles in foreign journals, indexed in Scopus and Web of Science, 2 articles in journals recommended by the Ministry

of Education and Science of Ukraine in Technical Sciences, 10 publications in conference proceedings (3 indexed in SCOPUS and Web of Science).

Keywords: neuro-fuzzy models, gradient optimization methods, empirical mode decomposition, ensembles of neuro-fuzzy models

LIST OF PUBLICATIONS

The list of publications, which reflect the main scientific results of the thesis:

1. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "A novel neuro-fuzzy model for multivariate time-series prediction", *Data*, 3(4), 62, 2018. (indexed in **Scopus** та **Web of Science**)

2. A. Vlasenko, N. Vlasenko, O. Vynokurova, Y. Bodyanskiy and D. Peleshko. "A Novel ensemble neuro-fuzzy model for financial time series forecasting", *Data*, 4(3), 126, 2019. (indexed in **Scopus** та **Web of Science**)

3. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "An Empirical Mode Decomposition Based Method to Synthesize Ensemble Multidimensional Gaussian Neuro-Fuzzy Models in Financial Forecasting", *Communications in Computer and Information Science*, volume 1158, pp. 140-149, 2020. Springer, Cham. (indexed in **Scopus**)

4. А. Власенко, Е. Кучеренко, «Иерархическая нечетко-вероятностная модель в задачах управления рисками», *Системы обработки інформації*, Вип. 1, С. 145-149, 2013. (indexed in **Index Copernicus International**)

5. А. Власенко, Е. Кучеренко. «Метод логического вывода в иерархических нечетко-вероятностных моделях сложных систем», *Бионика интеллекта*, Вып. 2 (81), С. 47–52, 2013.

Results that confirm the approbation of the thesis:

6. A. Vlasenko, O. Vynokurova, N. Vlasenko and M. Peleshko, "A Hybrid Neuro-Fuzzy Model for Stock Market Time-Series Prediction", in 2018 IEEE

Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2018. (indexed in **Scopus** та **Web of Science**)

7. A. Vlasenko, Y. Rashkevych, N. Vlasenko, D. Peleshko and O. Vynokurova, "A Hybrid EMD - Neuro-Fuzzy Model for Financial Time Series Analysis", in 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 2020. (indexed in **Scopus**)

8. A. Vlasenko, N. Vlasenko, O. Vynokurova and Y. Bodyanskiy, "An Enhancement of a Learning Procedure in Neuro-Fuzzy Model", in 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 2018. (indexed in **Scopus** та **Web of Science**)

9. A. Vlasenko, Y. Bodyanskiy, «An approach to overcome prediction delay problem in hybrid neuro-fuzzy networks for financial time-series forecasting», в Матеріали VI Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології», Одеса, 2017.

10. Y. Bodyanskiy, A. Vlasenko, O. Vynokurova and I. Pliss, "A hybrid neuro-fuzzy network for financial time-series prediction", в Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту», Залізний порт, Україна, 2017.

11. А. Власенко, Е. Кучеренко, «Применение нечеткой кластеризации в задачах структурной идентификации нечетких моделей Такаги-Сугено-Канга» в Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту», Євпаторія, Україна, 2012.

12. А. Власенко, Е. Кучеренко, «Архитектура иерархической системы логического вывода с нечетко-вероятностной базой знаний», в Матеріали XI международной научно-практической конференции "Математическое обеспечение интеллектуальных систем" Днепропетровск, 2013.

13. А. Власенко, Е. Кучеренко, «Иерархические системы нечеткого вывода Такаги-Сугено в задачах управления рисками», в Матеріали XVI

международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2012.

14. А. Власенко, Е. Кучеренко, «Реализация иерархической нечетко-вероятностной модели логического вывода на базе библиотеки Aforge.Net», в Материалы XVII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2013.

15. А. Власенко, Е. Кучеренко, «Реляционная база нечетко-вероятностных знаний для логического вывода в задачах анализа рисков», в Материалы XVIII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке», Харьков, 2014.

ЗМІСТ

| | |
|--|-----------|
| ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ | 15 |
| ВСТУП..... | 16 |
| 1 ОГЛЯД СТАНУ ПРОБЛЕМИ І ПОСТАНОВКА ЗАВДАННЯ ДОСЛІДЖЕННЯ | 22 |
| 1.1 Статистичні методи у аналізі ризиків | 23 |
| 1.2 М'які обчислення | 26 |
| 1.2.1 Нечітка логіка та системи нечіткого логічного висновування..... | 27 |
| 1.2.2 Штучні нейронні мережі | 32 |
| 1.2.3 Нейро-фаззі моделі..... | 35 |
| 1.3 Методи математичної оптимізації у машинному навчанні. | 38 |
| 1.4 Перетворення сигналів в задачах обробки даних для аналізу ризиків..... | 41 |
| 1.5 Висновки до першого розділу..... | 43 |
| 2 ГІБРИДНА НЕЙРО-ФАЗЗІ МОДЕЛЬ У ЗАДАЧАХ ПРОГНОЗУВАННЯ ТА АНАЛІЗУ | 45 |
| 2.1 Гібридна п'ятишарова нейро-фаззі модель с багатовимірними гаусіанами у шарі консеквентів та одним виходом..... | 45 |
| 2.1.1 Модифікація з навчанням рецепторів фазифікації | 54 |
| 2.1.2 Спеціалізований критерій навчання для подолання проблеми затримки у прогнозуванні | 57 |
| 2.2 Узагальнена модель для декількох вихідних значень | 59 |
| 2.3 Висновки до розділу | 65 |
| 3 АНСАМБЛЬ НЕЙРО-ФАЗЗІ МОДЕЛЕЙ У ЗАДАЧАХ ОПЕРАТИВНОГО АНІЛІЗУ РИЗИКІВ | 66 |
| 3.1 Архітектура ансамблю нейро-фаззі моделей | 66 |
| 3.3 Висновки до розділу | 74 |
| 4 ПОЄДНАННЯ ЕМПІРИЧНОЇ МОДОВОЇ ДЕКОМПОЗИЦІЇ ТА НЕЙРО-ФАЗЗІ МОДЕЛЕЙ, ІЄРАРХІЧНІ МОДЕЛІ..... | 75 |

| | |
|--|-----|
| 4.2 Синтез ансамблю нейро-фаззі моделей за допомогою методу емпіричної модової декомпозиції..... | 80 |
| 4.4 Висновки до розділу | 91 |
| 5 ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ ТА РОЗВ'ЯЗАННЯ ПРАКТИЧНИХ ЗАДАЧ | 92 |
| 5.1 Моделювання нейро-фаззі моделі з багатовимірними гаусіанами у консеквентному шарі..... | 93 |
| 5.2 Експериментальне моделювання навчання функцій належності першого шару | 96 |
| 5.3 Експериментальне моделювання нейро-фаззі моделі з декількома вихідними змінними. | 101 |
| 5.4 Експериментальне моделювання ансамблів моделей | 107 |
| 5.4 Експериментальні результати застосування емпіричної модової декомпозиції у навчанні нейро-фаззі моделей..... | 111 |
| 5.5 Ієрархчна інтелектуальна система підтримки прийняття рішень у оперативному аналізі ризиків. | 115 |
| ВИСНОВКИ..... | 119 |
| ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ..... | 121 |
| ДОДАТОК А..... | 134 |
| СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ | 134 |
| ДОДАТОК В..... | 137 |
| ДОКУМЕНТИ, ЩО ПІДТВЕРДЖУЮТЬ ВПРОВАДЖЕННЯ | 137 |

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

ANN – Artificial Neural Network(s)

ANFIS – Adaptive Neuro-Fuzzy Inference System (Adaptive Network-based Fuzzy Inference System)

ARMA – Autoregressive Moving Average

ARIMA – Autoregressive Integrated Moving Average

GARCH – Generalized Autoregressive Conditional Heteroskedasticity

EMD – Empirical Mode Decomposition

LSTM – Long Short Term Memory

ВСТУП

Актуальність дослідження. У багатьох сферах людської діяльності виникає проблема оперативного реагування на різноманітні ризики, і сучасні інтелектуальні технології підтримки прийняття рішень здатні поліпшити цей процес у автоматичному та напівавтоматичному режимах. Задача побудови таких систем включає в себе побудову моделей для прогнозування, ідентифікації та аналізу в умовах неповних та зашумлених даних зі складною динамікою.

Протягом десятиліть класичні статистичні моделі домінували у моделюванні складних процесів. Однак вони накладають суттєві обмеження на стаціонарність даних, у багатьох випадках потребують складної попередньої обробки вхідних даних та апарату верифікації гіпотез.

Методи штучного інтелекту успішно конкурують із усталеними статистичними методами у задачах обробки нелінійних та нестационарних даних. Серед них популярність отримали моделі м'яких обчислень, особливо штучні нейронні мережі та нечітка логіка, які у свою чергу часто поєднуються в гібридні моделі, здатні поєднувати сильні сторони та зменшувати прояви недоліків один одного.

До недоліків нейро-фаззи моделей можна віднести проблему великої розмірності у багатьох прикладних проблемах.

Таким чином, на сьогоднішній день є актуальною задача розробки нових методів та моделей аналізу даних в умовах інформації мінливого хаотичного характеру та високочастотної динаміки, та які відрізняються невеликим розміром та обчислювальною ефективністю.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана в рамках держбюджетних тем «Динамічний інтелектуальний аналіз послідовностей нечіткої інформації за умов суттєвої невизначеності на основі гібридних систем обчислювального інтелекту»

(№ДР0116U002539) та “Глибинні гібридні системи обчислювального інтелекту для аналізу потоків даних та їх швидке навчання” (ДР №0119U001403), які виконувалися у Харківському національному університеті радіоелектроніки, згідно наказів Міністерства освіти і науки України за результатами конкурсного відбору проектів наукових досліджень, у яких автор брав участь як виконавець.

Мета та задачі дослідження. Розробка методів та моделей інтелектуалізації процесів оперативного аналізу ризиків в умовах неповної інформації на основі гібридних м’яких обчислень, які вирізняються високою точністю та обчислювальною ефективністю.

Відповідно до поставленої мети у дисертаційній роботі необхідно вирішити такі завдання:

- провести аналіз існуючих методів та підходів в задачах прогнозування, ідентифікації та аналізу, які використовуються у оперативному аналізі ризиків, виявити їх недоліки та можливості для поліпшення за рахунок використання м’яких обчислень;

- розробити нейро-фаззі модель зі швидким навчанням для використання у задачах прогнозування, ідентифікації та діагностики, яка вирізняється здатністю представляти складні залежності у вхідних даних та компактним розміром;

- розробити ансамбль нейро-фаззі моделей для зменшення похибки та спрощення процедури селекції гіперпараметрів;

- розробити метод навчання, який використовує метод емпіричної модової декомпозиції для зниження шумів;

- створити ієрархічну модель підтримки прийняття рішень, яка би інкорпоровала запропоновані моделі;

- провести експериментальні дослідження розроблених методів, вирішити за їх допомогою практичні задачі.

Об'єкт дослідження – процес інтелектуального аналізу даних для виявлення та оцінки ризиків.

Предмет дослідження – методи та моделі інтелектуального аналізу даних на базі м'яких обчислень.

Методи дослідження – базуються на теорії м'яких обчислень, а саме на теорії штучних нейронних мереж та теорії нечіткої логіки для побудови нейро-фаззі моделей та їхніх ансамблів; теорії оптимізації для синтезу методів навчання. Імітаційне моделювання використовується для перевірки ефективності розроблених моделей.

Наукова новизна отриманих результатів. До нових, одержаних особисто автором, належать такі результати:

1. *Вперше запропоновано* п'ятишарову гібридну нейро-фаззі модель з багатовимірними гаусіанами у консеквентному шарі та швидкий метод її навчання з варіантами на базі квадратичної похибки та спеціалізованого критерію, що характеризується високою точністю, швидкістю обробки та обчислювальною швидкодією у процесі навчання за рахунок зменшення налаштованих параметрів моделі.

2. *Вперше запропоновано* ансамбль гібридних нейро-фаззі моделей з багатовимірними гаусіанами у консеквентному шарі та метод його синтезу, що характеризується зменшеною похибкою навчання, покращеними узагальнюючими можливостями та спрощеною процедурою селекції гіперпараметрів.

3. *Удосконалено* метод навчання нейро-фаззі моделей та їх ансамблів застосуванням емпіричної модової декомпозиції, що відрізняється від аналогів можливістю зниження рівня шуму у випадку високодинамічних даних.

4. *Набув подальшого розвитку* метод побудови ієрархічних знання-орієнтованих нечітких систем шляхом включення оцінок ймовірностей у процес логічного висновування та використання нейро-фаззі моделей як компонентів ієрархічної структури, що на відміну від існуючих методів

дозволяє опрацьовувати малоймовірні критичні значення у задачах оперативного аналізу ризиків.

Практичне значення отриманих результатів. Використання запропонованих методів та моделей дозволяє підвищити ефективність інтелектуальних систем аналізу ризиків в умовах неповних даних. Розроблені модулі з запропонованими методами підтвердили свою ефективність у задачах динамічного прогнозування складних процесів та ідентифікації потенційних ризиків.

Розроблене програмне забезпечення, що реалізує запропоновані і досліджені в дисертації методи, впроваджено

– у ТОВ «САЙТОСС» при розробці програмного забезпечення для управління ризиками інвестиційного портфоліо (акт впровадження від 14.01.2021);

– у Харківському національному університеті радіоелектроніки на кафедрі штучного інтелекту в освітній процес з курсу «Нейромережеві методи обчислювального інтелекту» та «Штучні нейронні мережі: архітектура, навчання, застосування».

Особистий внесок здобувача. Усі положення, що виносяться на захист, основні результати теоретичних та експериментальних досліджень отримані здобувачем особисто. У публікаціях, написаних у співавторстві, автору належать такі результати: [1] – адаптивна нейро-фаззі модель для декількох вихідних змінних та метод її навчання, [2] – ансамбль нейро-фаззі моделей та метод його синтезу, [3] – метод синтезу нейро-фаззі моделей на базі методу емпіричної модової декомпозиції, [4] – ієрархічна нечітко ймовірнісна модель, [5] – модифікація методу логічного виведення, [6] - адаптивна нейро-фаззі модель, [7] - адаптивна нейро-фаззі модель з навчанням на базі методу емпіричної модової декомпозиції, [8] – розширення методу онлайн навчання адаптивної нейро-фаззі моделі, [9] – модифікація методу навчання адаптивної нейро-фаззі моделі на базі спеціалізованого критерію, [10] - адаптивна нейро-

фаззі модель та метод її навчання, [11] – аналіз та практичне застосування методів нечіткої кластеризації, [12] – архітектура нечіткої ієрархічної моделі логічного виведення, [13] - ієрархічна нечітко модель, [14] – розробка імплементації та імітаційне моделювання нечіткої ієрархічної моделі логічного виведення, [15] – створення бази даних для представлення правил у ієрархічних нечітко ймовірнісних моделях.

Апробація результатів дисертації. Основні результати дисертаційної роботи доповідалися й обговорювалися на конференціях: Second International Conference on Data Stream Mining & Processing (21-25 August 2018, Lviv); Third International Conference on Data Stream Mining & Processing (21-25 August 2020, Lviv); First International Conference on System Analysis & Intelligent Computing (8-12 October 2018, Kyiv); VI Міжнародна науково-практична конференція «Інформаційні Управляючі Системи та Технології» (20 – 22 вересня 2017 р, Одеса); Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту-ISDMCI-2016 (24-28 травня 2016 р., Залізний порт, м. Херсон); Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту-ISDMCI-2012 (21-25 травня 2012 р., Євпаторія); XVI Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (17-19 квітня 2012 р., м. Харків); XVII Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (22-24 квітня 2013 р., м. Харків); XVIII Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (14-16 квітня 2014 р., м. Харків); XI Міжнародна науково-практична конференція "Математичне забезпечення інтелектуальних систем" 2013, 20-22 листопада 2013 г, Дніпро.

Публікації. За тематикою дослідження опубліковано 15 наукових праць, з них 1 розділ у колективній монографії, що входить до науково-метричної бази Scopus, 2 статті за кордоном, що входить до науково-метричної бази SCOPUS, 2 статті у виданнях, які зазначені в переліках фахових видань

України з технічних наук, 10 публікацій у матеріалах конференцій (3 включено до науково-метричної бази даних SCOPUS).

Структура та обсяг дисертації. Дисертація складається із вступу, п'яти розділів, висновків, що містять основні результати, списку використаних джерел і додатку. Загальний обсяг дисертації складає 139 сторінок (з них 120 - основного тексту), містить 70 рисунків, 15 таблиць, список використаних джерел, що включає 120 найменувань та займає 13 сторінок, 2 додатки на 6 сторінках.

1 ОГЛЯД СТАНУ ПРОБЛЕМИ І ПОСТАНОВКА ЗАВДАННЯ ДОСЛІДЖЕННЯ

У багатьох сферах людської діяльності виникає проблема оперативного аналізу ризиків. Вирішення цієї проблеми часто потребує обробки нелінійних та нестационарних даних з мінливим та хаотичним характером та високочастотною динамікою. Тому розробка методів та моделей інтелектуалізації процесів оперативного аналізу ризиків є важливою практичною задачею. Прогнозування майбутніх значень, трендів, аналіз взаємозалежностей на основі зібраних історичних даних є вирішальним елементом у багатьох моделях та методах прийняття рішень. Зазвичай існує багато додаткових обмежень, які накладаються на моделі аналізу ризиків, крім точності - обчислювальна продуктивність, складність моделі, мінімальний розмір набору даних тощо.

Ризик являє собою певну подію або фактор, здатний принести кому-небудь шкоду або збиток. Ризик можна математично представити як структуру вигляду [1]:

$$R = \{(s_i, p_i, x_i)\}, i = 1, 2, \dots, N \quad (1.1)$$

де s_i – опис ризику,

p_i – ймовірність ризикової події,

x_i – наслідки або оцінки збитків.

Управління ризиками – це формальні процеси, пов'язані з ідентифікацією, аналізом ризиків та прийняттям рішень, які включають максимізацію позитивних і мінімізацію негативних наслідків настання ризикових подій [2]. Побудова формальних моделей, орієнтованих на знання в задачах управління ризиками стикається з проблемами неточності, неповноти і суперечності даних предметної області.

Мета розділу – проаналізувати стан проблеми побудови методів обробки даних та моделей складних систем та процесів, які використовуються у оперативному аналізі ризиків, визначити мету та задачі дослідження.

Завдання:

- 1) провести огляд існуючих методів та моделей у оперативному аналізі ризиків, проаналізувати їх практичні застосування, обмеження та недоліки;
- 2) розглянути основні принципи м'яких обчислень - нечіткої логіки та систем нечіткого логічного висновування, штучних нейронних мереж;
- 3) проаналізувати найбільш поширені методи оптимізації у машинному навчанні, з'ясувати основні проблеми та обмеження їх чисельних реалізацій;
- 4) сформулювати задачу дослідження.

1.1 Статистичні методи у аналізі ризиків

Теорія ймовірностей та статистика - це математичний фундамент для представлення невизначеної інформації. Вони забезпечують засоби кількісної оцінки невизначеності, а також аксіоми для висновування нових невизначених тверджень.

Теорія ймовірностей є фундаментальним інструментом багатьох дисциплін науки та інженерії та протягом десятиліть ймовірнісні та статистичні моделі домінували у галузі прогнозування та аналізу в управлінні ризиками.

Модель авторегресії — ковзного середнього (ARMA) також відома як модель Бокса-Дженкінса, запропонована у [2] для аналізу та прогнозування стаціонарних часових рядів та довела свою ефективність у багатьох прикладних задачах [3] [4].

Вона складається з авторегресії та поліноміальної функція ковзного середнього:

$$AR(p): y_t = \mu + \sum_{i=1}^p (\phi_i y_{t-i}) + \omega_t \quad (1.2)$$

$$MA(q): y_t = \mu + \sum_{i=1}^q (\theta_i \omega_{t-i}) + \omega_t,$$

де y_t – прогнозоване значення;

ϕ_i – коефіцієнти авторегресії;

θ_j – коефіцієнти ковзного середнього;

μ – середнє значення;

ω_t – білий шум.

Модель ARMA (p, q) складається з двох основних многочленів, які є AR (p) та MA (q).

$$y_t = \mu + \sum_{i=1}^p (\phi_i y_{t-i}) + \sum_{j=1}^q (\theta_j \omega_{t-j}) + \omega_t \quad (1.3)$$

Узагальненням цієї моделі є ARIMA, дуже розповсюджена її модифікація що враховує сезонність – SARIMA.

Прикладами успішного застосування є прогнозування ризиків - нової хвилі COVID-19 при пом'якшенні обмежень [5], інвестування у фондові індекси [6], соціального страхування [7] та ін.

Однак, такі моделі вимагають, щоб дані були стаціонарними, і це досягається за допомогою різних методів диференціації, ефективність яких викликає сумніви через високий ступінь нестационарності реальних даних [8] [9].

VaR (Value at Risk) є вартісною мірою ризику та визначається таким чином, що ймовірність втрат, більших за VaR, дорівнює (максимум) p , тоді як ймовірність втрат, менших ніж VaR, становить (принаймні) $1 - p$. Прикладом застосування такої моделі до аналізу ризиків [10] де багатовимірний VaR for VaR модель була використана для моделювання нафтового та фондового ринків. Розповсюдженими варіаціями є Conditional Value-at-Risk (CVaR) та entropic value at risk (EVaR).

Узагальнені авторегресивні умовно гетероскедастичні моделі (GARSH) враховують дисперсію похибки часових рядів і дуже популярні у аналізі ризиків у фінансових даних - робота [11] містить приклад застосування 9 поширених різновидів цієї моделі для аналізу ризиків волатильності фондових ринків. Модель має наступну форму:

$$a_t = \sigma_t \varepsilon_t, \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (1.4)$$

де $a_t = r_t - \mu_t$;

ε_t – похибка;

α_i – ARCH параметр;

β_j – GARCH параметр.

Альтернативними до інтерпретаційних статистичних методів є методи, побудовані на принципах доказової ймовірності, серед яких найбільш розповсюдженими є Байєсові. Вони побудовані на висновуванні апостеріорної ймовірності за допомогою теореми Байєса та похідних від неї формалізмів, як логічного наслідку двох передумов, апіорної ймовірності та функції правдоподібності.

Теорема Байєса для параметру θ , даних D та гіпотези H має наступний вигляд (1.5):

$$\overset{\text{posterior}}{\hat{P}}(\theta|D, H) = \frac{\overset{\text{likelihood}}{P(D|\theta, H)} \overset{\text{prior}}{P(\theta|H)}}{\underset{\text{marginal likelihood}}{P(D|H)}} \quad (1.5)$$

Найбільш розповсюдженими практичними застосуваннями є ієрархічні моделі, в тому числі у аналізі ризиків [12].

Теорія Демстера Шафнера є крайньою формою напрямку суб'єктивізму у теорії імовірностей, де функція довіри Bel та функція можливості Pl виражені як

$$\text{Bel}(B) = \sum_{A_k | A_k \subseteq B} m(A_k) \quad (1.6)$$

$$\text{Pl}(B) = \sum_{A_k | A_k \cap B \neq \emptyset} m(A_k),$$

де $B \in \mathfrak{R}(\Omega)$.

Bel(B) виражає мінімальну (або необхідну) підтримку гіпотези B, коли Pl(B) виражає максимальну (або потенційну) підтримку гіпотези B у випадку отримання підкріплюючих свідчень.

Прикладами застосування методів та моделей на базі цієї теорії у аналізі ризиків є роботи [13], де було суттєво покращено модель режиму відмов і аналізу наслідків та [14], у якій вдосконалено оцінки невизначеності у аналізі ризиків хімічної токсичності.

1.2 М'які обчислення

Методи штучного інтелекту десятиліттями успішно конкурують із ustalеними статистичними методами у різноманітних складних прикладних

задачах, включаючи прогнозування часових рядів. Серед них популярність отримали моделі М'яких Обчислень (Soft Computing), особливо штучні нейронні мережі (ANN), Нечітка Логіка (FL) та Еволюційні Обчислення (EC). Вони часто поєднуються в гібридні моделі, що дозволяє у багатьох випадках подолати або зменшити недоліки кожного конкретного підходу.

1.2.1 Нечітка логіка та системи нечіткого логічного висновування

Теорія нечіткої логіки та нечітких множин, які починають свою історію у роботі [15] розширюють та узагальнюють класичну математичну логіку та теорію множин, були поширені на різні математичні об'єкти шляхом введення неперервного ступеня належності та вплинули на розвиток методів та моделей аналізу невизначеності [16] [17].

Нечітка множина визначається своїми елементами $x \in X$, де x – елемент, і може бути представлена упорядкованими парами наступного вигляду:

$$A = \{(x, \mu_A(x)) \mid x \in U\} \quad (1.7)$$

де U – універсальна множина всіх елементів що можуть належати нечіткій множині A ;

$\mu_A(x)$ – функція належності, яка є розширенням характеристичної функції у класичній теорії множин, може набувати значення у закритому інтервалі $[0,1]$ для кожного елементу $u \in U$.

Наприклад, одним з найбільш поширених, є нечітке число трикутного вигляду, у якого функція належності $\mu(x)$ має наступний вигляд:

$$\mu(x) = \begin{cases} (x - a)/(b - a), & a \leq x \leq b \\ (x - c)/(b - c), & b \leq x \leq c \\ 0, & \text{otherwise} \end{cases} \quad (1.8)$$

де a – мінімум;

b – найбільше значення;

c – максимум.

Нечіткі числа добре підходять для відображення неточного характеру суджень, таких як лінгвістичні терміни, що використовуються експертами-людьми.

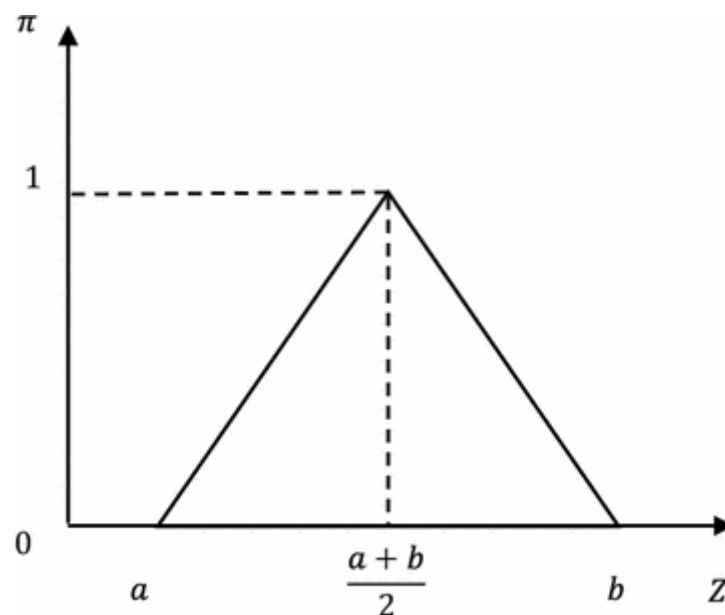


Рисунок 1.1 – Трикутна функція належності

Найбільш поширеними нечіткими моделями є моделі нечіткого логічного виведення, побудовані на правилах “IF ... THEN ...”.

При побудові правил висновування використовуються оператори, такі як "AND", "OR", у деяких випадках "NOT". Об'єднання операторів називають t-нормами. Визначення нечіткого оператора "AND" може бути подано у наступних формах:

$$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)]$$

(1.9)

$$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)],$$

де x – міра належності.

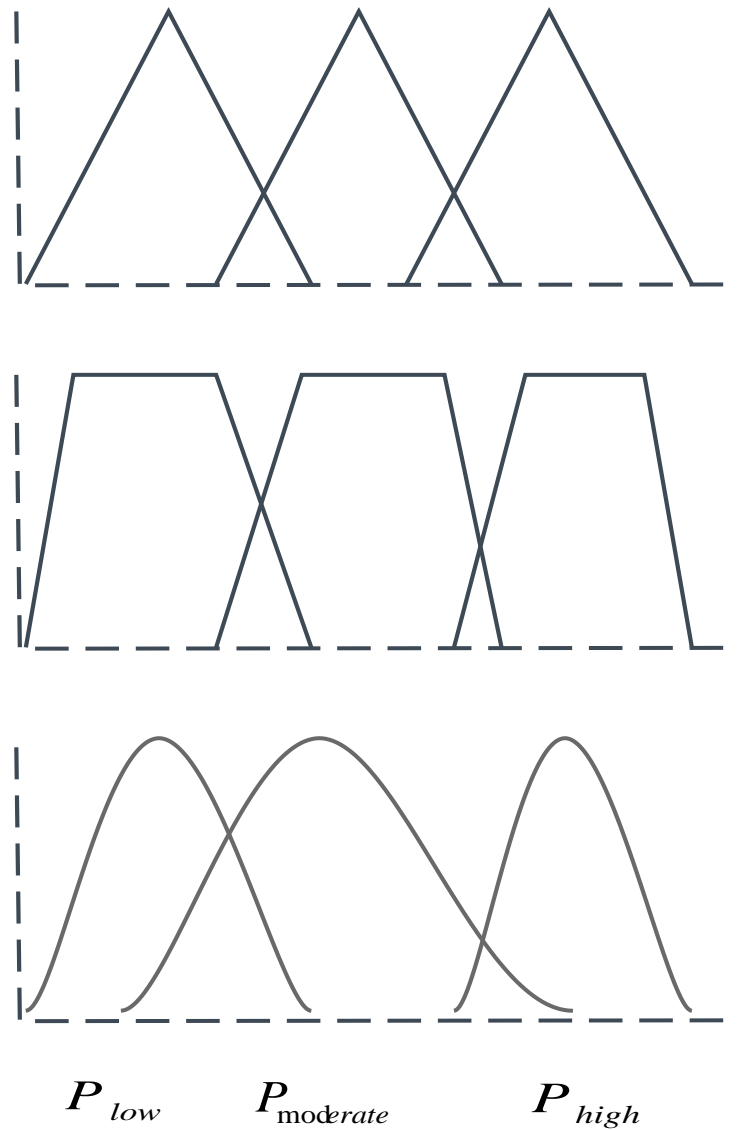


Рисунок 1.2 – Приклад моделювання різними видами функцій належності

Важливим питанням у побудові таких систем є вибір типу нечіткого логічного висновування. Нечітке логічне висновування Мамдані [18] є історично першим запропонованим і довго був найбільш поширеним типом логічного висновування на базі нечіткої логіки.

Правила мають такий вигляд:

$$\text{Rule}_q: \text{If } x \text{ is } A_q \text{ then } y \text{ is } B_q \quad (1.10)$$

де x – вхідна змінна;

A_q – нечітка множина антецеденту;

B_q – нечітка множина консеквенту.

Дефаззіфікований чіткий результат отримується за допомогою методу центроїдів, який визначено нижче:

$$f(x) = \frac{\sum_{j=1}^N \theta_j \prod_{i=1}^n \mu_{F_i^j}(x_i)}{\sum_{j=1}^N [\prod_{i=1}^n \mu_{F_i^j}(x_i)]} \quad (1.11)$$

x – вхідна змінна;

$f(x)$ – вихідне значення;

$\mu_{F_i^j}(x_i)$ – міра належності j -го правила;

θ_j – центроїд j -го консеквенту.

Причиною популярності даного типу нечітких систем є поєднання відносної простоти, прийнятних результатів і лінгвістичної інтерпретується правил.

Нечітке логічне висновування Такагі-Сугено-Канга [19] відрізняється від логічного висновування Мамдані використанням «чітких» функцій в консеквента правил, і тому подібні моделі показують хорошу продуктивність

в різних прикладних задачах, але мають меншу лінгвістичну прозорість. Антецедент в даному типі логічного висновування описує нечіткі області в просторі входних значень, в яких функції консеквента правила валідні.

Правило логічного висновування нечітких систем Такагі-Сугено-Канга має наступний вигляд:

$$\begin{aligned}
 &R_j : \text{if } x_1 \text{ is } A_{1j} \text{ and} \\
 &x_2 \text{ is } A_{2j} \text{ and... and } x_n \text{ is } A_{nj} \\
 &\text{Then } y = g_j(x_1, x_2, \dots, x_n), (j = 1, 2, \dots, N),
 \end{aligned}
 \tag{1.12}$$

де $g_j(\cdot)$ є «чіткою» функцією від x_i .

Загальний результат логічного висновування в даній системі може бути представлений як [5]:

$$\frac{\sum_{j=1}^N g_j(\cdot) T_{i=1}^{m_j} \mu_{ij}(x_i)}{\sum_{j=1}^N T_{i=1}^{m_j} \mu_{ij}(x_i)}
 \tag{1.13}$$

де $1 \leq m_j \leq n$ кількість входних змінних;

N – кількість правил;

n – кількість входних змінних;

μ_{ij} – функції належності;

T – T-норма нечіткої кон'юкції.

Нечіткі моделі, засновані на правилах, які здатні ефективно оперувати суб'єктивними і неточними знаннями про предметну область, застосовувались до різних галузей в реальних інженерних, наукових та економічних додатках, включаючи завдання прогнозування та аналізу ризиків, наприклад роботи [20] [21] [22] [23] [24] [25] [26]. Його здібності були застосовані до одного з найскладніших завдань – торгових стратегій [27].

У таблиці 1 наведено приклади лінгвістичних змінних та відповідних нечітких чисел застосованих для моделі ідентифікації та пріорітизації ризиків [28].

Таблиця 1

Лінгвістичні значення для оцінки наслідків у аналізі ризиків

| Лінгвістичне значення | Чисельне значення | Трикутні нечіткі числа |
|------------------------------|--------------------------|-------------------------------|
| Extremely high impact (VH) | 4 | (0.75, 1, 1) |
| High impact (H) | 3 | (0.5, 0.75, 1) |
| Low impact (l) | 2 | (0.25, 0.5, 0.75) |
| Very low impact (VL) | 1 | (0, 0.25, 0.5) |
| No impact (N) | 1 | (0, 0, 0.25) |

Однак вони мають деякі істотні недоліки, в першу чергу, спричинені їхньою залежністю від знань експертів або складних методів вилучення правил.

Також однією з ключових проблем створення систем нечіткого логічного висновування є «прокляття розмірності» – експонентний характер залежності між кількістю вхідних параметрів і кількістю правил. Для вирішення даної проблеми розроблено клас систем нечіткого логічного висновування з ієрархічною структурою [29].

1.2.2 Штучні нейронні мережі

Штучні нейронні мережі натхнені дивовижними здібностями біологічних нейронів вирішувати складні проблеми. Вони складаються з

обчислювальних вузлів – штучних нейронів та їхніх з'єднань. Найпростіша мережа зображена на рис. 1.3.

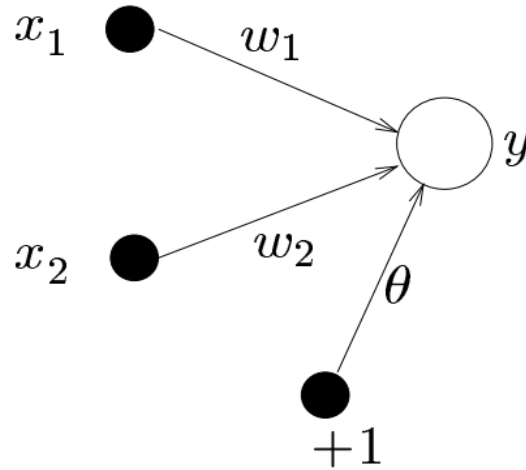


Рисунок 1.3 – Приклад одношарової нейронної мережі з двома вхідними параметрами

Більшість мереж прямого розповсюдження мають нелінійні функції активації як показано на рис. 1.4. та приховані шари як на рис. 1.5 [30].

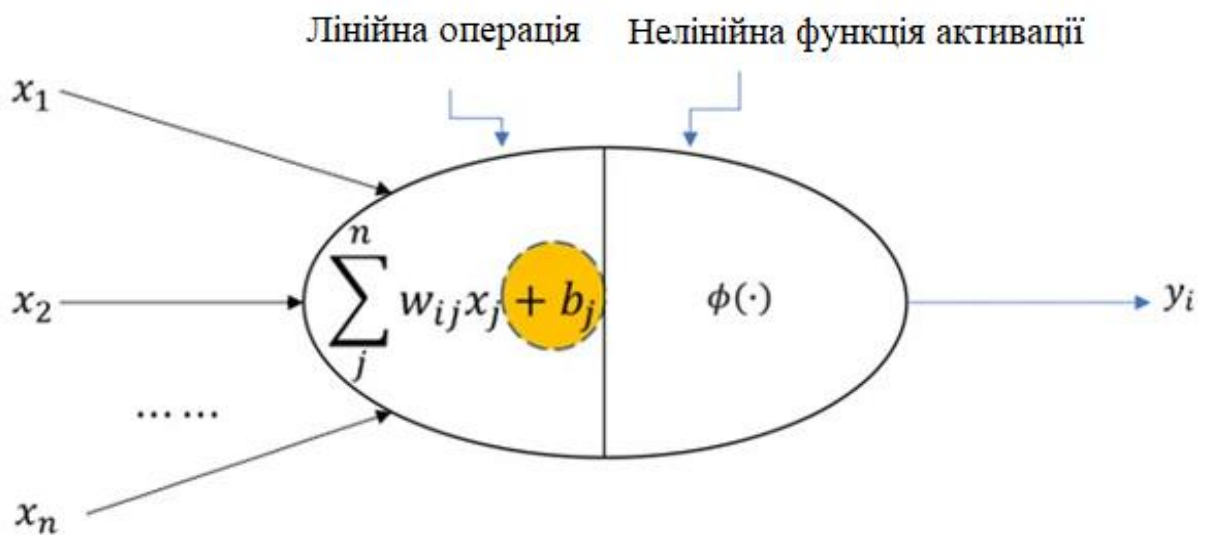


Рисунок 1.4 – Нейронна мережа з нелінійною функцією активації

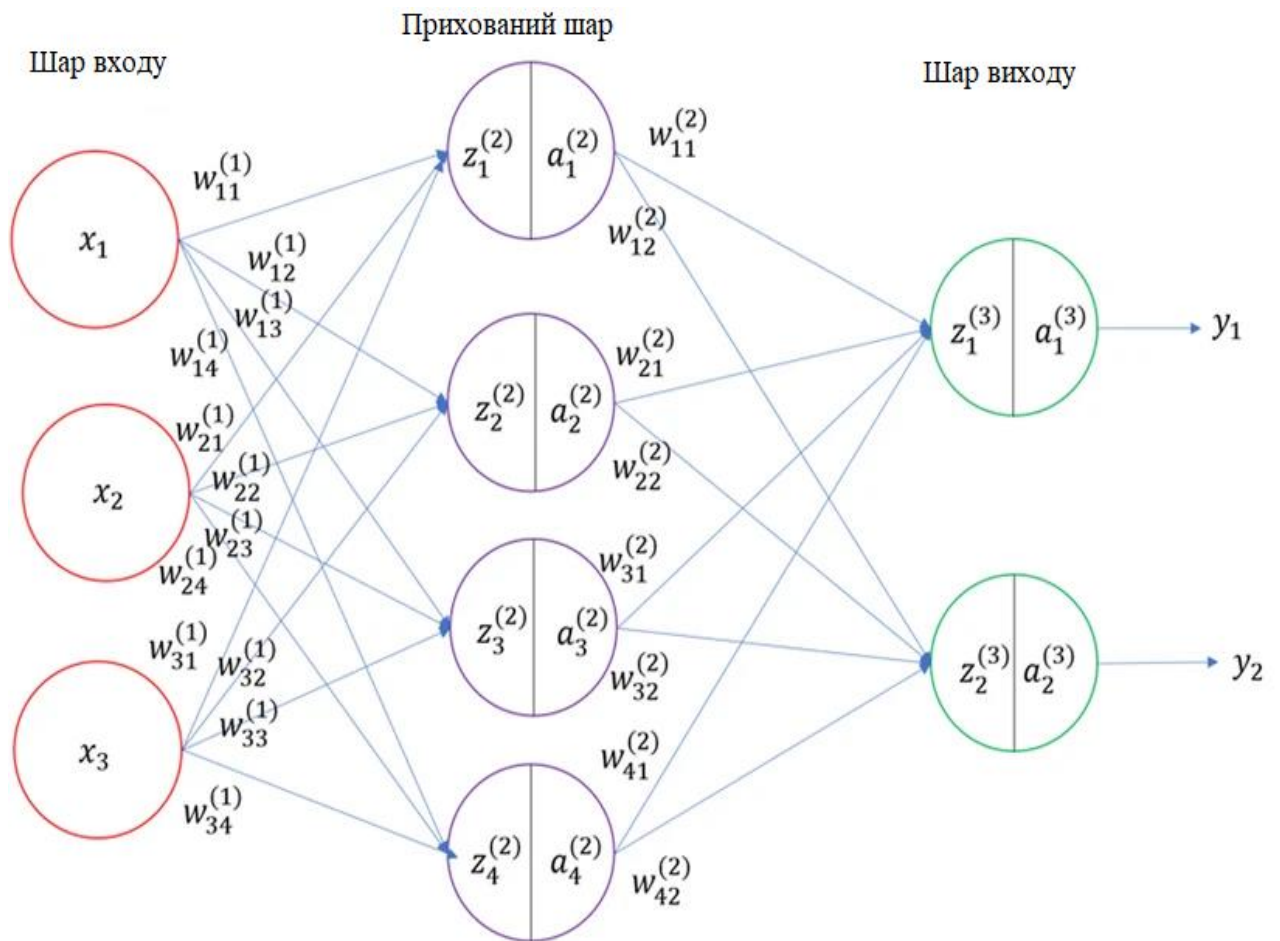


Рисунок 1.5 – Приклад архітектури нейронної мережі з одним прихованим шаром

Багато нейронних моделей випробувано в завданнях прогнозування та аналізу ризиків - дослідження [31] доводить здатність безлічі мереж обробляти складні структури даних, включаючи короткі, дуже часті та сезонні часові ряди. Серед прикладів застосування, прогнозування цінових тенденцій FOREX, використовуючи довгострокову короткочасну пам'ять (LSTM) [32], напрямку щоденної віддачі індексу SPDR S&P 500 ETF за допомогою глибоких нейронних мереж (DNN) [33] та ін [34] [35] [36] [37] [38] [39] [40] [41] [42].

1.2.3 Нейро-фаззі моделі

ANFIS – адаптивна нейро-нечітка система, історично перша та найбільш популярна версія нейро-фаззі систем, яка була розроблена так, щоб функціонально нагадувати нечітку модель типу Такагі-Сугено-Канга. ANFIS складається з п'яти шарів [43]. Приклад архітектури, показаний на рис. 1.6, містить одиничний вихід z з двома вхідними значеннями x, y .

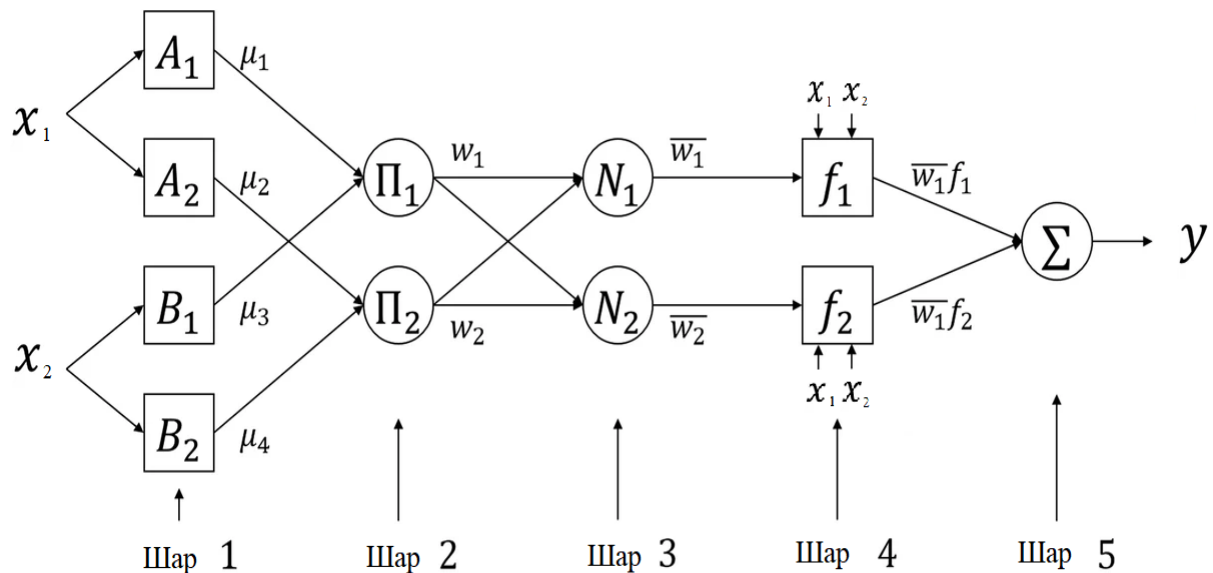


Рисунок 1.6 – Архітектура ANFIS

Для вхідних значень x та y перший рівень виконує фаззіфікацію через функції належності. Можуть бути використані трикутні, трапецієвидні, дзвоновидні функції або функція Гауса [44], наприклад дзвоновидна функція має наступний вигляд.

$$\mu_i(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}}, i = 1, 2, 3, 4 \quad (1.14)$$

Вузли другого рівня відповідають нечітким правилам. Вихід другого рівня є добутком усіх вхідних сигналів, підключених до кожного вузла. Наприклад, на рис. 1.6 вихід першого вузла виражається рівнянням:

$$w_1 = \mu_1 \times \mu_3 \quad (1.15)$$

Третій шар виконує нормалізацію виходу другого рівня. І-й вузол третього шару обчислює відношення і-го вихідного значення другого шару до суми всіх вхідних значень другого шару. Це виражається рівнянням

$$\bar{w}_i = \frac{w_i}{\sum_{j=1}^n w_j}, i = 1,2 \quad (1.16)$$

де n – кількість вузлів у другому шарі.

Четвертий рівень виконує дефазифікацію за допомогою функції які отримують всі вхідні значення. Вихід четвертого шару для і-го вузла визначається як

$$\bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), i = 1,2 \quad (1.17)$$

де $\{p_i, q_i, r_i\}$ називаються параметрами консекветна і налаштовуються за допомогою методів навчання.

Нарешті, вузол п'ятого шару додає всі вихідні значення четвертого шару, щоб сформуванати кінцевий вихідне значення z , як показано в наступному рівнянні:

$$z = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (1.18)$$

Нейро-фаззі системи поєднують в собі сильні сторони ANN та систем нечіткого логічного висновування і були успішно застосовані до різних проблем. [45] [46], завдяки їх універсальним можливостям наближення та надійності в поєднанні з хорошими обчислювальними характеристиками та успадкованими від ANN ефективних процедур навчання.

Серед застосувань у реальних проблемах прогнозування слід виділити, наприклад, прогнозування завантаження електричних мереж [47] [48], у сфері фінансів [49] [50] [51] [52] [53] та інших прикладних областей [54] [55]. Гарсія та ін. [56] застосував гібридну нечітку нейронну мережу для прогнозування цінового напрямку в індексі фондового ринку, який складається з 30 найбільших німецьких компаній.

Їх переваги особливо помітні в швидко мінливих доменах з високим ступенем невизначеності, наприклад, керівництво та планування маршрутів [57] [58], складна логістика [59], змішування сирої нафти [60], управління портфелем людських ресурсів [61] та багато інших. Робота [62] надає короткий огляд застосувань нейро-фаззі моделей, включаючи прогнозування.

Останнім часом нео-фаззі системи [63] [64] [65], які базуються на адитивних моделях нульового порядку, набирають популярності в різних сферах.

У галузі аналізу ризиків фондового ринку приклади успішно застосованих нейро-фаззі моделей включають модель ANFIS із модифікованим алгоритмом навчання Левенберга-Марквардта для прогнозування ціни закриття фондової біржі на день [66] та модель ANFIS, засновану на непрямому підході та протестовану на індексах Тегеранської фондової біржі [52]. У роботі [67] запропонували інтерпретований нейро-фаззі підхід до прогнозування курсу акцій, що застосовується до різних біржових серій.

Моделі нейро-фаззі прогнозування поєднані з генетичними алгоритмами [51], вейвлет-перетворенням [54] [68], методами опорних

векторів [69] [70] та динамічною пам'ятю [50] [53]. Такі комбіновані моделі можуть показати високу ефективність, але вони страждають від додаткових витрат на додаткові процедури кластеризації, шари та гіперпараметри.

Сай та ін. [71] використовували оптимізацію колонії мурашок до навчання нейро-фаззі систем. Деякі методи оптимізації були застосовані до моделей типу 2 [72] [73] [74]. Усі такі моделі отримують вигоду від поєднання переваг різних підходів, але можуть постраждати від зростання параметрів, що підлягають налаштуванню.

Нейро-фаззі моделі загалом вимагають багатьох правил для охоплення складних нелінійних відношень, що відоме як «прокляття розмірності». Для вирішення цієї проблеми для сценаріїв, коли вхідні змінні демонструють певний ступінь взаємозалежності та кореляції, в роботі [75] в нейро-фаззі моделі використовувалась багатовимірна функція Гауса для ефективної обробки корельованих областей вхідного простору, але передбачення ефективність вимірювалася на хаотичних часових рядах, де поведінка може різко відрізнятися від реальних даних. Не так багато уваги було приділено обробці шуму даних.

1.3 Методи математичної оптимізації у машинному навчанні.

Більшість алгоритмів навчання передбачають оптимізацію - завдання мінімізації або максимізації деякої функції $f(x)$ шляхом зміни x . Функція, яку ми хочемо мінімізувати або максимізувати, називається цільовою функцією, або критерієм. Коли ми мінімізуємо його, ми можемо також назвати це функцією витрат, функцією збитків або функцією помилки.

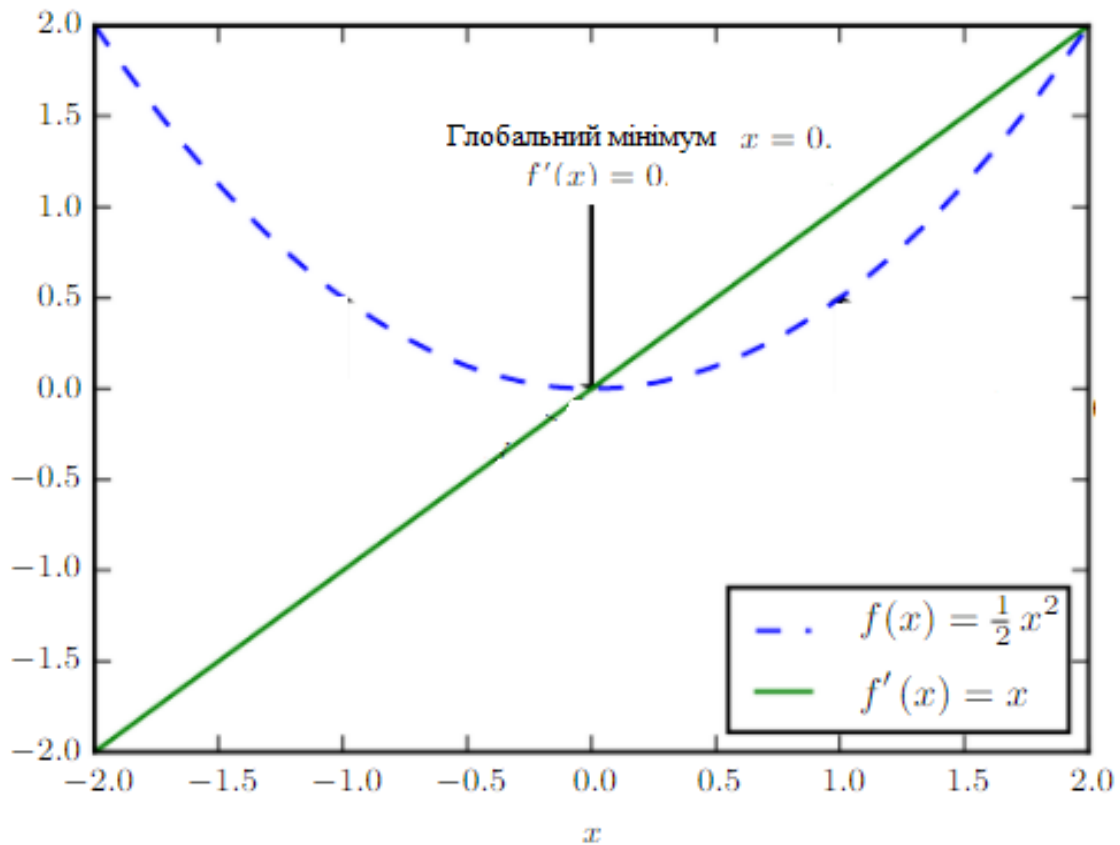


Рисунок 1.7 – Приклад пошуку глобального мінімуму

Точка, яка отримує абсолютне найнижче значення $f(x)$, є глобальним мінімумом. Може бути лише один глобальний мінімум або кілька глобальних мінімумів функції. Також можливе існування локальних мінімумів, які не є оптимальними у глобальному масштабі. У контексті машинного навчання ми оптимізуємо функції, які можуть мати багато локальних мінімумів, які не є оптимальними, і мати багато сідлових точок, оточених дуже плоскими регіонами. Все це ускладнює оптимізацію, особливо коли функція багатовимірною. Тому ми, як правило, погоджуємось визначити результатом оптимізації значення, яке є дуже низьким, але не обов'язково мінімальним у будь-якому формальному розумінні [76].

Ми можемо зменшити значення цільової функції f , рухаючись у напрямку негативного градієнта. Це відомо як метод найкрутішого спуску або градієнтного спуску. Найкрутіший спуск пропонує нову точку як:

$$x' = x - \epsilon \nabla_x f(x), \quad (1.19)$$

де ϵ – крок навчання, позитивний скаляр, що визначає розмір кроку.

Ми можемо обирати крок навчання різними способами. Популярним підходом є встановлення невеликої константи. Іноді ми можемо визначити розмір кроку, при якому похідна напрямку нульова. Інший підхід полягає в оцінці $f(x - \nabla_x f(x))$ для кількох значень ϵ та виборі того, що має найменше значення цільової функції. Ця остання стратегія називається рядковим пошуком.

Алгоритми оптимізації, що використовують лише градієнт, наприклад, градієнтний спуск, називаються алгоритмами оптимізації першого порядку. Алгоритми оптимізації, які також використовують матрицю Гессе, наприклад метод Ньютона, називаються алгоритмами оптимізації другого порядку [77] [78].

Показано, що стохастичні градієнтні методи мають багато переваг у порівнянні з популярними пакетними методами в практичних задачах оптимізації, включаючи програми машинного навчання [79] [80]. Замість обчислення повного градієнта за всіма наборами даних у пакетах, ці методи оновлюють вільні параметри моделі після обробки кожного шаблону введення. Їх головна перевага - властивий розрахунку градієнтів шум. Крім того, вони не страждають від помилок округлення, які виникають, коли доводиться зберігати значення накопиченого градієнту [80], і мають кращі обчислювальні показники. Недоліком є те, що їх неможливо так легко розпаралелювати, як пакетні методи [81] і серйозність такого обмеження залежить від проблеми.

Великою проблемою імплементації методів оптимізації на сучасних цифрових обчислювальних пристроях є те, що потрібно представити дійсні числа за допомогою кінцевого набору бітових шаблонів. Наслідком цього є що

майже для всіх дійсних чисел ми отримуємо певну помилку наближення/округлення, коли представляємо число в комп'ютері. Помилка округлення може викликати проблеми у практичних задачах, особливо коли вона накопичується у довгому ланцюгу операцій, якщо методи не адаптовані для мінімізації накопичення помилок округлення.

Також багато математичних функцій дуже чутливі до зміни значень аргументу близьких до нуля, і коли такі значення округлюються до нуля то методи, які їх використовують дають неправильні результати.

Помилка переповнення відбувається, коли числа з великою величиною апроксимуються як ∞ або $-\infty$. Подальша арифметика зазвичай змінює ці кінцеві значення на нечислові значення.

Також великою проблемою чисельної оптимізації є погана обумовленість, яка відноситься до того, як швидко значення функції змінюється щодо невеликих змін у її аргументах. Функції, які швидко змінюються, коли їхні входи дещо збурені, можуть бути проблематичними для обчислень, оскільки помилки округлення вхідних даних призводять до значних змін у вихідних даних.

1.4 Перетворення сигналів в задачах обробки даних для аналізу ризиків

У багатьох практичних застосуваннях зашумлені дані та сигнали обробляться за допомогою різноманітних перетворень, які розкладають вхідні дані на компоненти з корисними властивостями.

Найбільшу історію мають методи основані на перетворенні Фур'є, яке розкладає функцію вхідних даних на осциляторні функції. Прикладами застосування є роботи [82], де аналіз Фур'є був поєднаний з класифікатором на базі рекурентних нейронних мереж, та [83] у якій він був застосований разом з ARMA.

Розповсюдженими є вейвлет-перетворення розглядають функцію від часу у термінах коливань, локалізованих за часом та/або простором і частотою. Вони часто поєднуються з іншими методами та моделями, наприклад у роботі [84] адаптивні вейвлети були використані для оцінки щільності в моделі типу GARCH. Також вейвлети були поєднані з VaR моделями [85] [86].

Когерентний вейвлет-аналіз здобув популярність у аналізі ризиків у фінансах та економетриці [87] [88] [89]. Приклад практичного застосування когерентного вейвлет аналізу для економічних даних з роботи [89] наведено на рисунку 1.8.

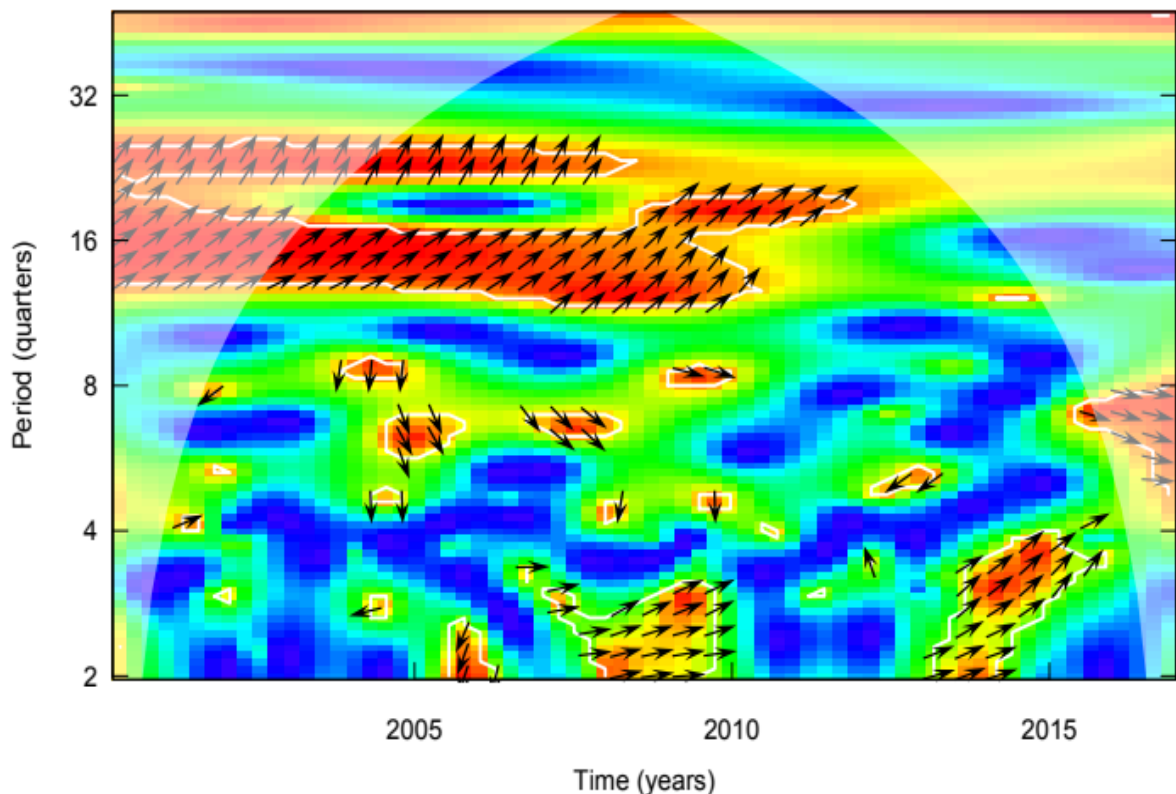


Рисунок 1.8 – Приклад застосування когерентного вейвлет аналізу

Перетворення Гільберта-Хуанга було запропоновано в [90] і застосовано як метод для моделювання нестационарних та нелінійних часових рядів у різних областях. Він складається з емпіричної модової декомпозиції (EMD) та

спектрального перетворення Гільберта. EMD, також відомий як перетворення Хуана, розкладає сигнал на набір простих власних модових функцій (IMF), які є корисними компонентами для подальшого аналізу. В даний час EMD не має надійної теоретичної бази і вважається емпіричним методом, але дослідження ведуться в цьому напрямку, наприклад [91] доводить обґрунтованість та надійність EMD. Робота [92] показує, що це перетворення є конкурентоспроможним як інструмент, що знижує рівень шуму у фінансових даних.

Використання EMD як інструменту декомпозиції та шумозаглушення замість перетворення Фур'є та вейвлетів разом із моделями штучного інтелекту є новою, але перспективною сферою досліджень. EMD у поєднанні з глибокими мережами LSTM використовувались для прогнозування фондового індексу Тайваня [93]. Ансамбль EMD (EEMD) разом з іншими методами декомпозиції був використаний з ANNs та машинами опорних векторів (SVM) для прогнозування фондового ринку [94].

У [95] повний ансамбль емпіричної модової декомпозиції з адаптивним шумом (CEEMDAN) був використаний з LSTM і показав кращі показники порівняно з єдиними мережами LSTM та іншими сучасними моделями.

Серед вартих уваги прикладів гібридних моделей - EEMD - незалежний компонентний аналіз (ICA) у завданні факторного аналізу [96], нейронна мережа EMD - BP та аналіз основних компонентів (PCA) в аналізі фондових ринків Азіатсько-Тихоокеанського регіону [97].

1.5 Висновки до першого розділу

1. Проаналізовано стан проблеми інтелектуалізації процесів оперативного аналізу ризиків.

2. Розглянуто основні принципи м'яких обчислень – штучних нейронних мереж та систем логічного виведення на базі нечіткої логіки.

3. Проведено аналіз існуючих методів прогнозування, діагностики та аналізу в задачах оперативного аналізу ризиків.

4. Показано, що поєднання нейронних мереж і нечіткої логіки може ефективно вирішувати складні практичні завдання, долаючи недоліки кожної з цих технологій в задачах нечіткої кластеризації коротких часових рядів.

5. Сформульовано задачу дослідження.

Список використаних у цьому розділі джерел наведено у повному списку використаних джерел під номерами [1–97].

2 ГІБРИДНА НЕЙРО-ФАЗЗИ МОДЕЛЬ У ЗАДАЧАХ ПРОГНОЗУВАННЯ ТА АНАЛІЗУ

Мета розділу – розробка адаптивної нейро-фаззі моделі, здатної відтворювати складні нелінійні залежності у вхідних даних та поєднувати велику точність з високою швидкістю.

Завдання:

- 1) створити нейро-фаззі модель для задач оперативного аналізу ризиків та методи швидкого навчання для неї;
- 2) узагальнити запропоновану модель та методи навчання до випадку декількох вихідних змінних зі здатністю використовувати взаємозалежності між ними;

2.1 Гібридна п'ятишарова нейро-фаззі модель с багатовимірними гаусіанами у шарі консеквентів та одним виходом.

Запропонована модель є модифікацією класичної моделі ANFIS – адаптивної нейро-фаззі системи логічного висновування, і вона складається з п'яти шарів. На рис. 2.1 зображена загальна архітектура моделі.

Перший рівень відповідає за фаззіфікацію вхідних змінних, представлених n -розмірним вектором $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$. Для задачі прогнозування вхідний вектор означає історичний зріз спостережуваної змінної.

Кожен вхідний параметр – компонент вхідного вектору обробляється h^ψ вузлами - функціями належності, отже, загальна кількість функцій належності першого рівня становить $h^\psi \times n$.

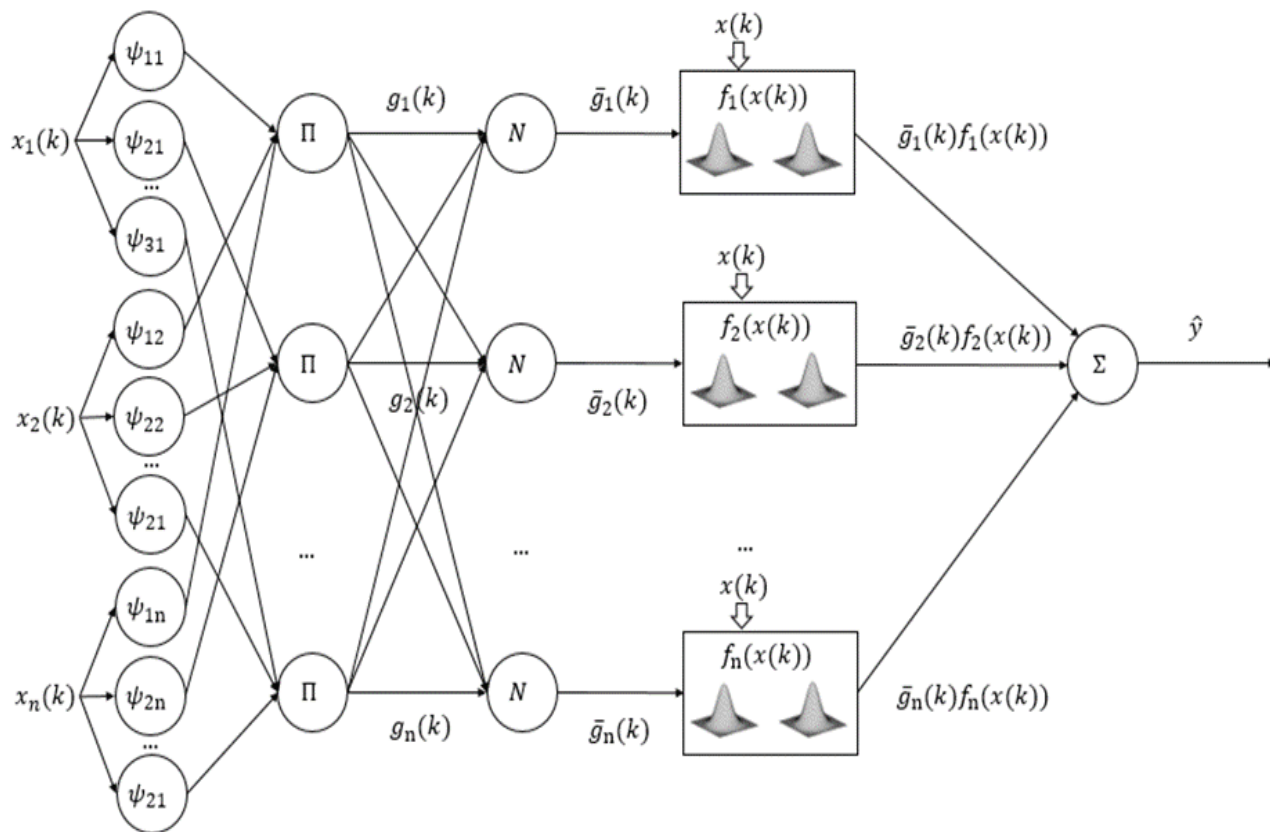


Рисунок 2.1 – Загальна архітектура запропонованої моделі з одним ВИХОДОМ

У дослідженні використано функцію належності Гауса, але у більш загальному випадку можливо її замінити на інші поширені типи функцій належності (наприклад, трикутна або дзвоновидна). Функція належності Гауса має такий вигляд:

$$\psi_{jl}(x(k)) = \exp\left(-\frac{(x(k) - c_{jl}^\psi)^2}{2\sigma_{je}^2}\right), \quad (2.1)$$

де $x(k)$ – вхідне значення;

c_{jl}^ψ – центр функції;

σ_{jl} – параметр ширини функції.

Аргумент експоненціальної функції $-\frac{(x_i(k)-c_{jl}^{\phi})}{2\sigma_{jl}^2(k)}$ - квадратична функція від

x_i . Будучи композицією експоненціальної функції з увігнутою квадратною функцією вона має багато корисних властивостей та відноситься до радіально-базисних ядерних функцій.

Основними перевагами цієї функції належності є відносно невелика кількість параметрів та стійкість до викидів яка дозволяє будувати робастні моделі.

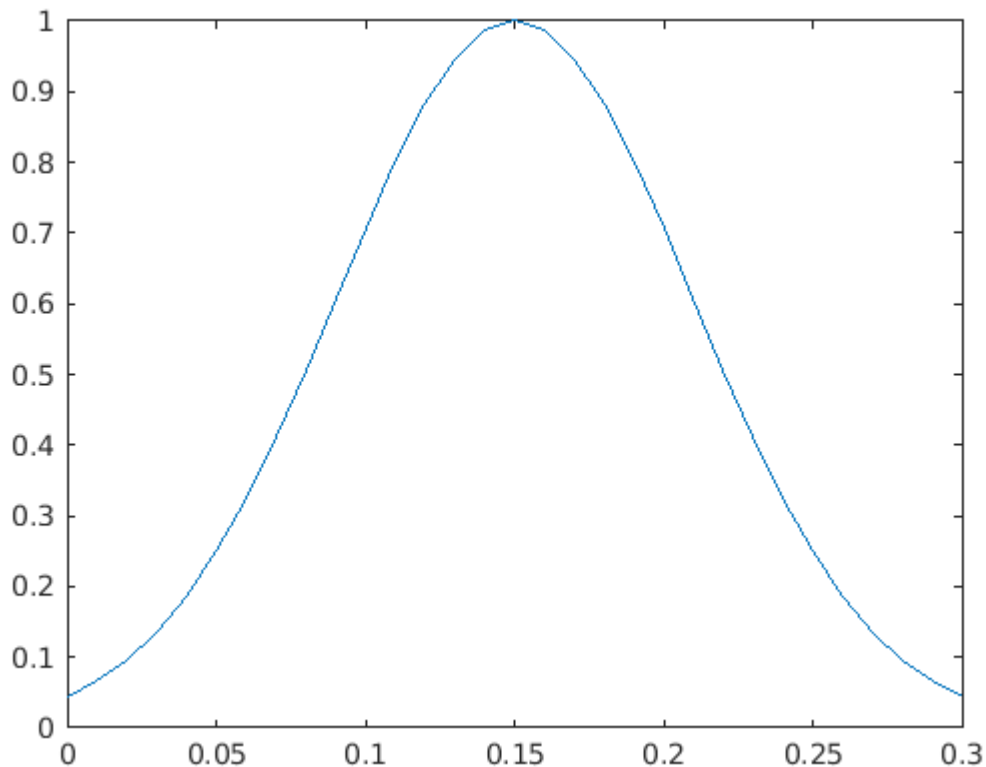


Рисунок 2.2 – Приклад функції належності з c_{jl}^{ϕ} 0.15 та σ_{jl} 0.06

Другий рівень відповідає за агрегацію значень функцій фазифікації – виразів антецеденту. Він складається з h^{ψ} вузлів, що реалізують алгебраїчну добуткову нечітку Т-норму:

$$g_j(k) = \prod_{i=1}^n \psi_{jl}(x_i(k)), \quad (2.2)$$

Третій шар є непараметризованим і відповідає за нормалізацію. Він також складається з h^ψ вузлів, вихід яких обчислюється за наступною формулою:

$$\bar{g}_j(k) = \frac{g_j(k)}{\sum_{j=1}^{h^\psi} g_j(k)} = \frac{\prod_{i=1}^n \psi_{ji}(x_i(k))}{\sum_{i=1}^{h^\psi} \prod_{i=1}^n \psi_{ji}(x_i(k))}, \quad (2.3)$$

Це необхідно для задоволення критерію розділення Руспіні:

$$\sum_{j=1}^{h^\psi} \bar{g}_j(k) = 1, \quad (2.4)$$

Четвертий шар представлений багатовимірними послідовними функціями Гауса $\phi_{je}(x(k))$ та їхніми ваговими коефіцієнтами p . Замість стандартних поліномів моделей TSK / ANFIS використовуються багатовимірні гаусіани:

$$\phi_{je}(x(k)) = \exp\left(-\frac{(x(k)-c_{je}^\phi)^T Q_{je}^{-1}(x(k)-c_{je}^\phi(k))}{2}\right), \quad (2.5)$$

де $x(k)$ вхідний вектор,

c_{je}^ϕ – вектор який представляє центр функції ;

Q_{je} – матриця рецепторних полів, $Q_{je} \in S_{++}^n$.

У цьому випадку компонент квадратичної функції стає квадратною формою всього вхідного вектору $x(k)$ та узагальнює поняття ширини функції.

Багатовимірний гаусіан - це потужний інструмент для представлення даних, які не розподілені рівномірно за основними осями.

Загальний результат виведення четвертого рівня може бути виражений як:

$$f_j(x(k)) = \sum_{e=1}^{h^\phi} p_{je} \phi_{je}(x(k)), \quad (2.6)$$

Де h^ϕ кількість багатовимірних гаусіанів у кожному обчислювальному блоці f_j .

П'ятий рівень є непараметризованим і видає загальний результат моделі як суму вхідних даних:

$$\hat{y} = \sum_{j=1}^{h^\phi} g_j f_j(x(k)), \quad (2.7)$$

У матричній формі результат може бути вираженим як:

$$\hat{y} = p^T f(x(k)), \quad (2.8)$$

де $x(k)$ – вхідний вектор,

p – вагові коефіцієнти задані у вигляді вектору

$f(x(k))$ – вектор значень вузлів консеквентного шару:

$$f(x(k)) = (\bar{g}_1 \phi_{11}(x(k)) \dots \bar{g}_{h^\psi} \phi_{h^\psi h^\phi}(x(k))), \quad (2.9)$$

де h^ψ – кількість рецепторних функцій першого рівня;

h^ϕ – кількість багатовимірних гаусіанів для кожного нормалізованого результату \bar{g}_1 .

Приклад архітектури для двох вхідних значень наведено на рис 2.3:

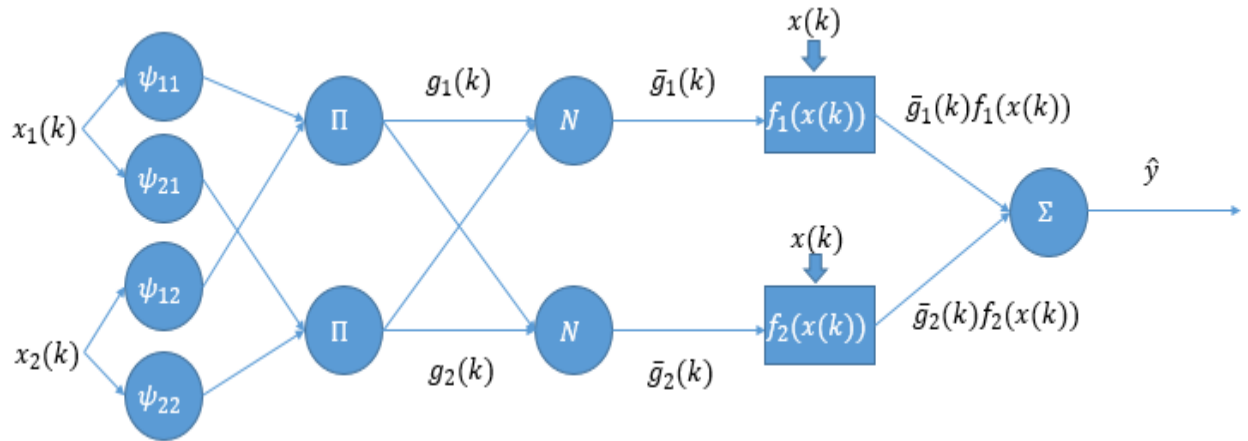


Рисунок 2.3 – Приклад архітектури для двох вхідних значень

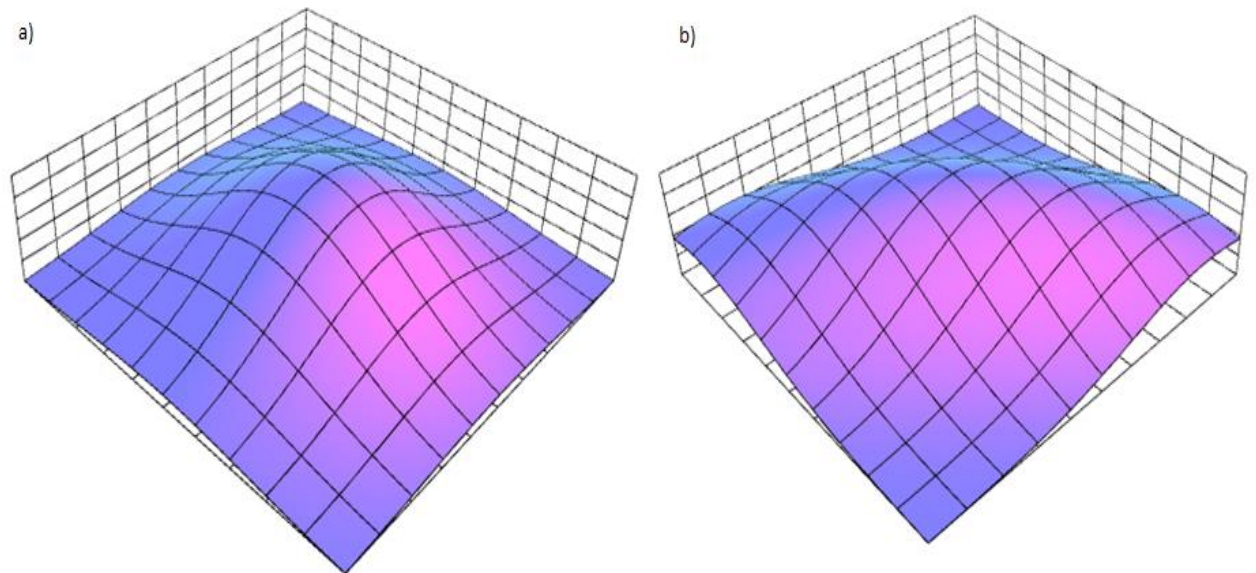


Рисунок 2.4 – Зразки багатовимірних гаусіанів: а) – з діагональною матрицею; б) – з матрицею $\begin{bmatrix} 0.9 & 0.6 \\ 0.7 & 0.9 \end{bmatrix}$.

Процес навчання у запропонованій моделі складається з налаштування вектору ваг p та налаштування параметрів багатовимірних функцій Гауса - центрів c_{jl}^{ϕ} та матриць Q_{jl} . Центри функцій членства першого рівня c_{jl}^{ϕ}

розподіляються рівновіддалено при ініціалізації, і вони не налаштовуються під час навчання.

Вагові коефіцієнти p ініціалізуються випадковим чином у діапазоні $[-0,1; 0,1]$, і їхня оптимізація досягається за допомогою ітераційного проективного методу Качмажа:

$$p(k + 1) = p(k) + \frac{y(k) - p^T f(x(k))}{f^T(x(k)) f(x(k))} f(x(k)), \quad (2.10)$$

де $p(k)$ – вагові коефіцієнти представлені у вигляді вектору;

$y(k)$ – зовнішній навчальний сигнал;

$p^T f(x(k))$ – вихідне значення моделі.

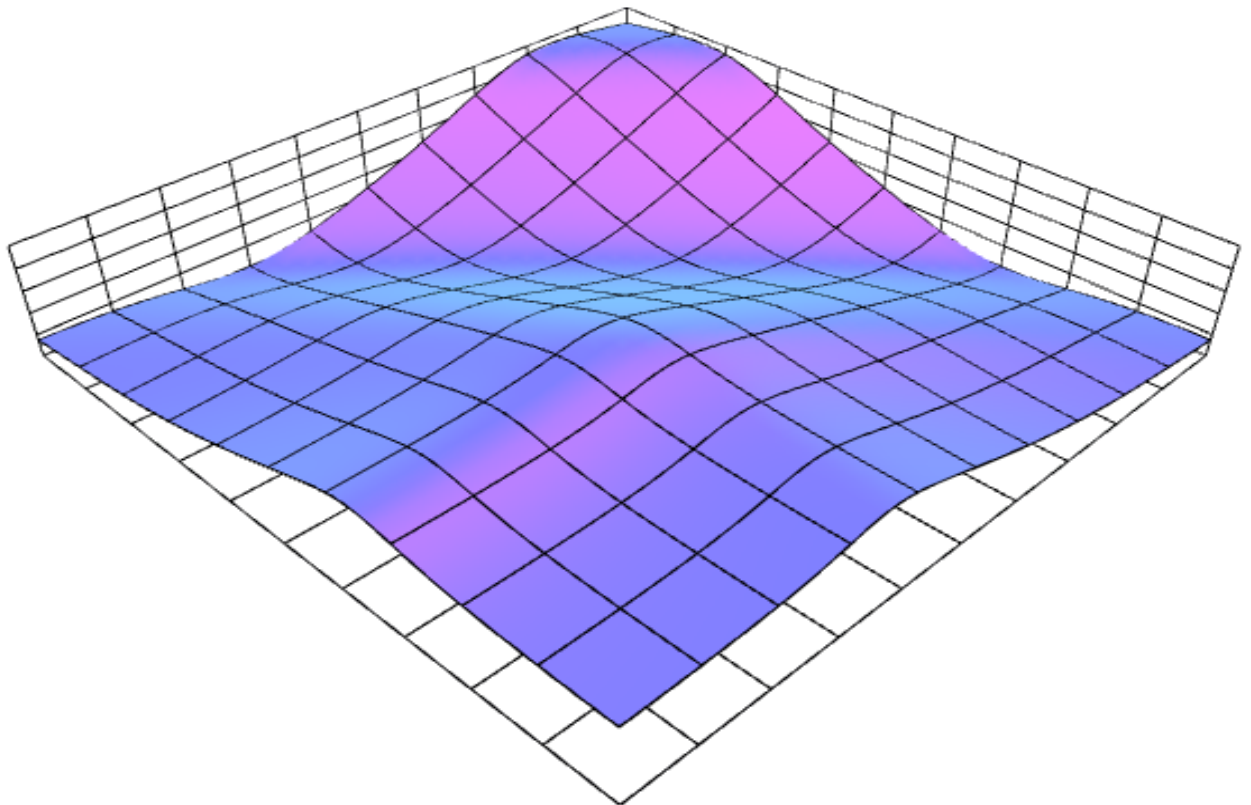


Рисунок 2.5 – Приклад поверхні похибки у просторі вагових коефіцієнтів

Навчання гаусіанів четвертого рівня виконане за допомогою градієнтної процедури зворотного розповсюдження помилки першого порядку на основі стандартного квадратичного критерію:

$$E = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2, \quad (2.11)$$

де $y(k)$ – значення референтного сигналу;

$\hat{y}(k)$ – значення прогнозоване моделлю ;

N – розмір вибірки.

Цей метод навчання є стохастичним, або з точки зору теорії машинного навчання онлайн-методом, бо ми налаштуємо параметри використовуючи градієнт для кожного вхідного шаблону та не обчислюємо весь градієнт навчальної вибірки як у пакетних методах.

Q_{jl}^{-1} матриці ініціалізуються як одиничні, і їх навчання можна записати наступним чином:

$$\begin{cases} Q_{jl}^{-1}(k+1) = Q_{jl}^{-1}(k) + \lambda_Q \frac{\tau_{jl}^Q(k)e(k)}{\eta_Q(k)} \\ \eta_Q(k+1) = \beta_Q \eta_Q(k) + Tr \left(\tau_{jl}^{QT} \tau_{jl}^Q \right) \end{cases}, \quad (2.12)$$

де λ_Q – крок навчання;

β_Q – параметр пам'яті;

τ_{jl}^Q – матриця значень зворотного розповсюдження градієнту похибки по відношенню до Q_{jl}^{-1} ;

η_Q – матриця згасаючих середніх значень попередніх градієнтів.

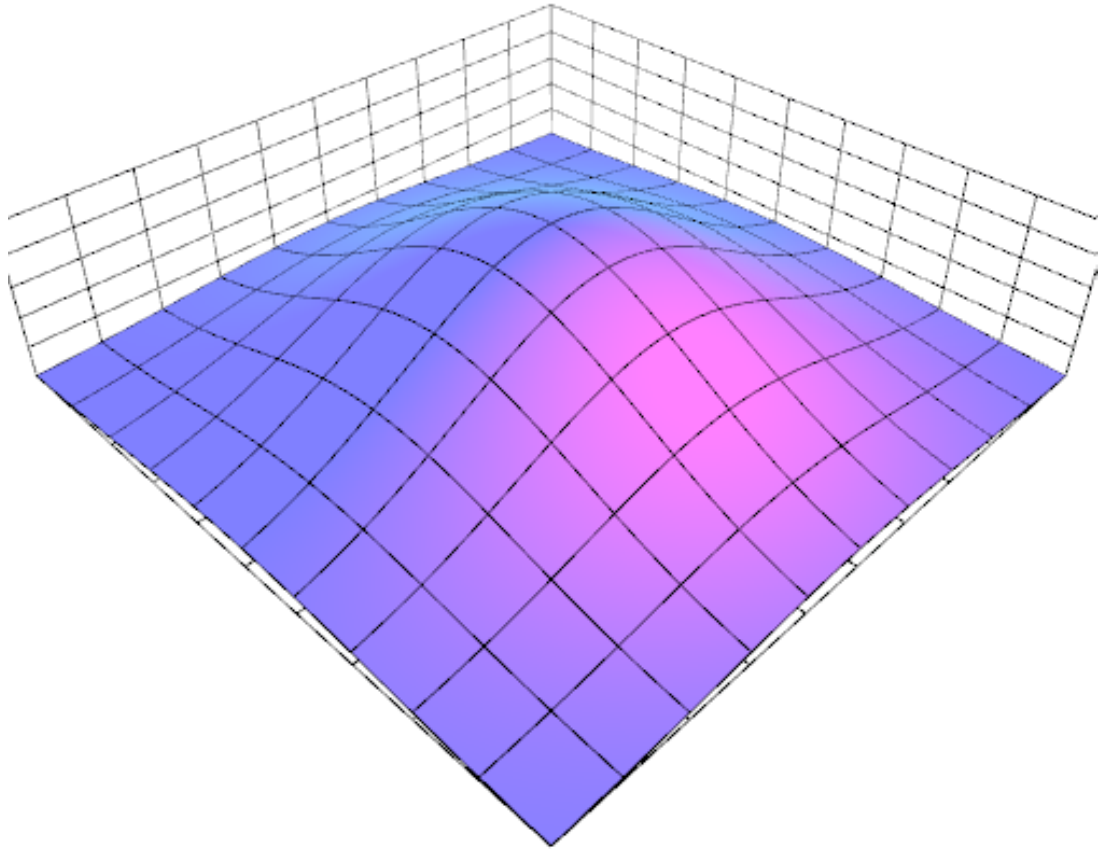


Рисунок 2.6 – Приклад багатовимірного гаусіна після ініціалізації.

Центри c_{jl}^ϕ розміщуються рівновіддалено на кроці ініціалізації, а потім налаштовуються за формулою наведеною нижче.

$$\begin{cases} c_{jl}^\phi(k+1) = c_{jl}^\phi(k) + \lambda_c \frac{\tau_{jl}^c(k)e(k)}{\eta_c(k)} \\ \eta_c(k+1) = \beta_c \eta_c(k) + \tau_{jl}^{cT} \tau_{jl}^c \end{cases}, \quad (2.13)$$

де λ_c – крок навчання;

β_c – параметр пам'яті;

τ_{jl}^c – вектор значень зворотного розповсюдження градієнту похибки по відношенню до c_{jl}^ϕ ;

η_c – вектор згасаючих середніх значень попередніх градієнтів.

На рис. 2.7 зображено приклад двомирної проекції поверхні помилок щодо центрів.

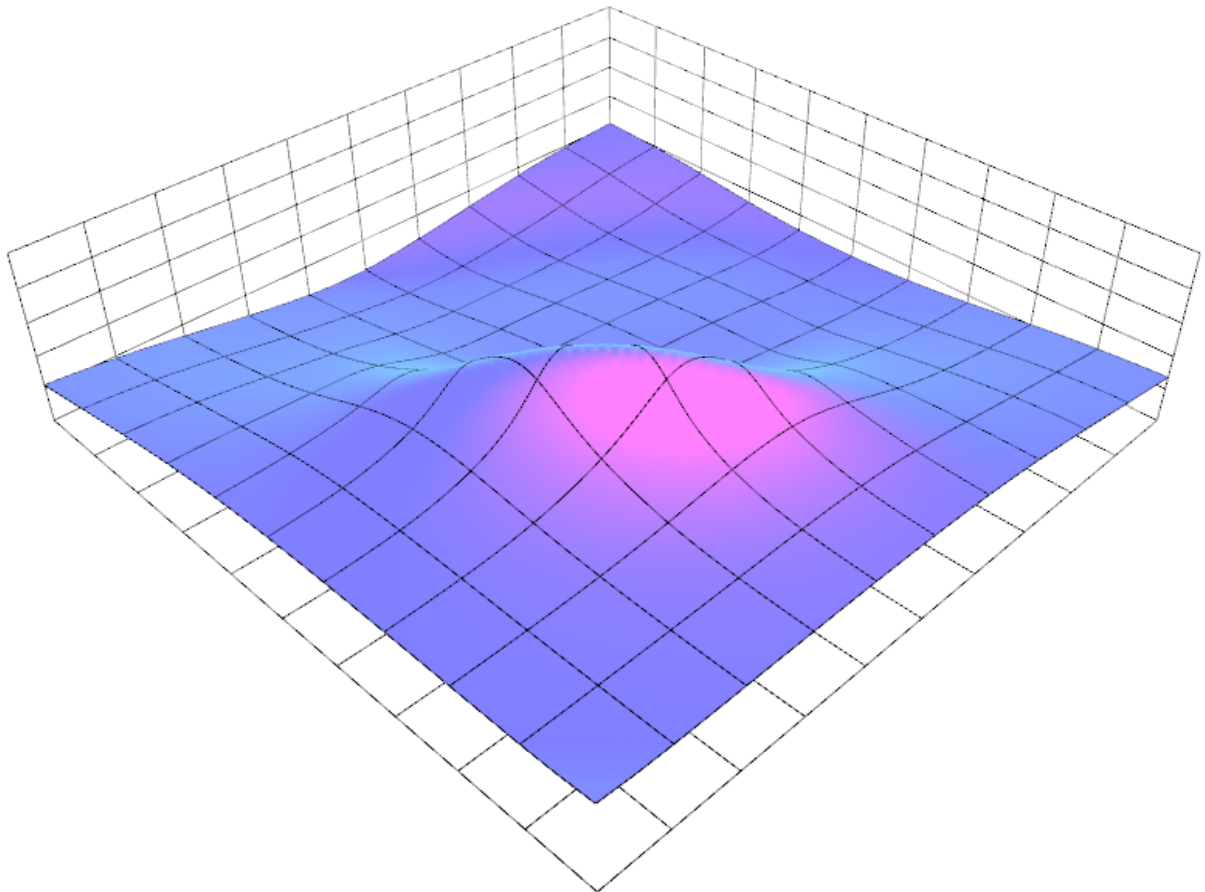


Рисунок 2.7 – Приклад поверхні функції помилки по відношенню до центрів гаусіанів у спрощеному двовимірному випадку

2.1.1 Модифікація з навчанням рецепторів фазифікації

У цьому підрозділі використаний алгоритм зворотного розповсюдження поширено на перший рецепторний шар нейро-фаззі моделі, який відповідає за фазифікацію вхідного вектору $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ та містить $h^\psi \times n$ гаусіанів, які повинні бути оптимізовані.

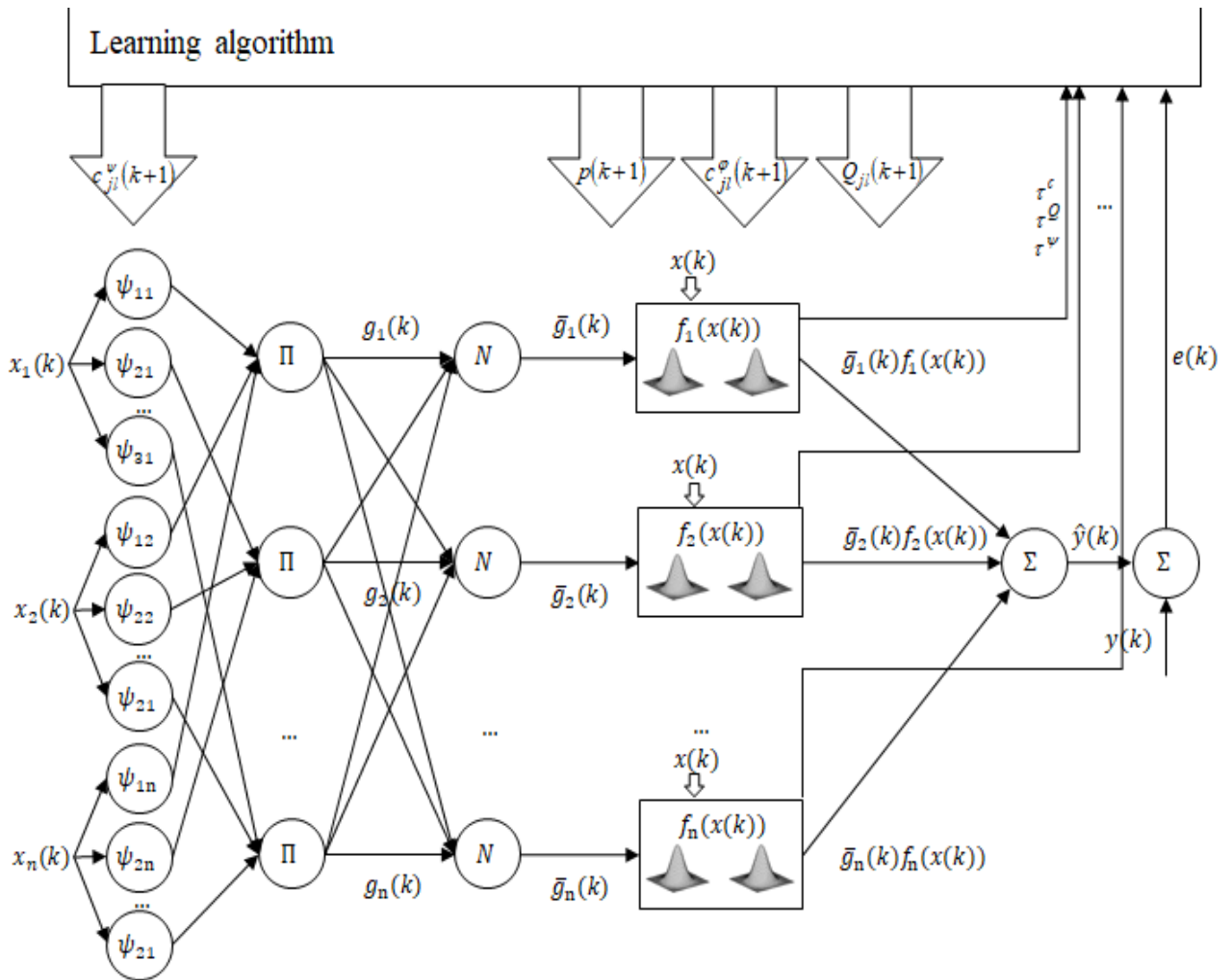


Рисунок 2.8 – Загальна архітектура нейро-фаззі моделі та навчання включаючи перший шар

Рис. 2.8 зображує загальну архітектуру нейро-фаззі моделі з модифікованим алгоритмом навчання, яка тепер включає налаштування функцій належності першого рівня. Показано напрямок зворотнього розповсюдження похибки навчання.

На рис. 2.9 показано приклад з п'ятьма функціями належності для вхідної змінної.

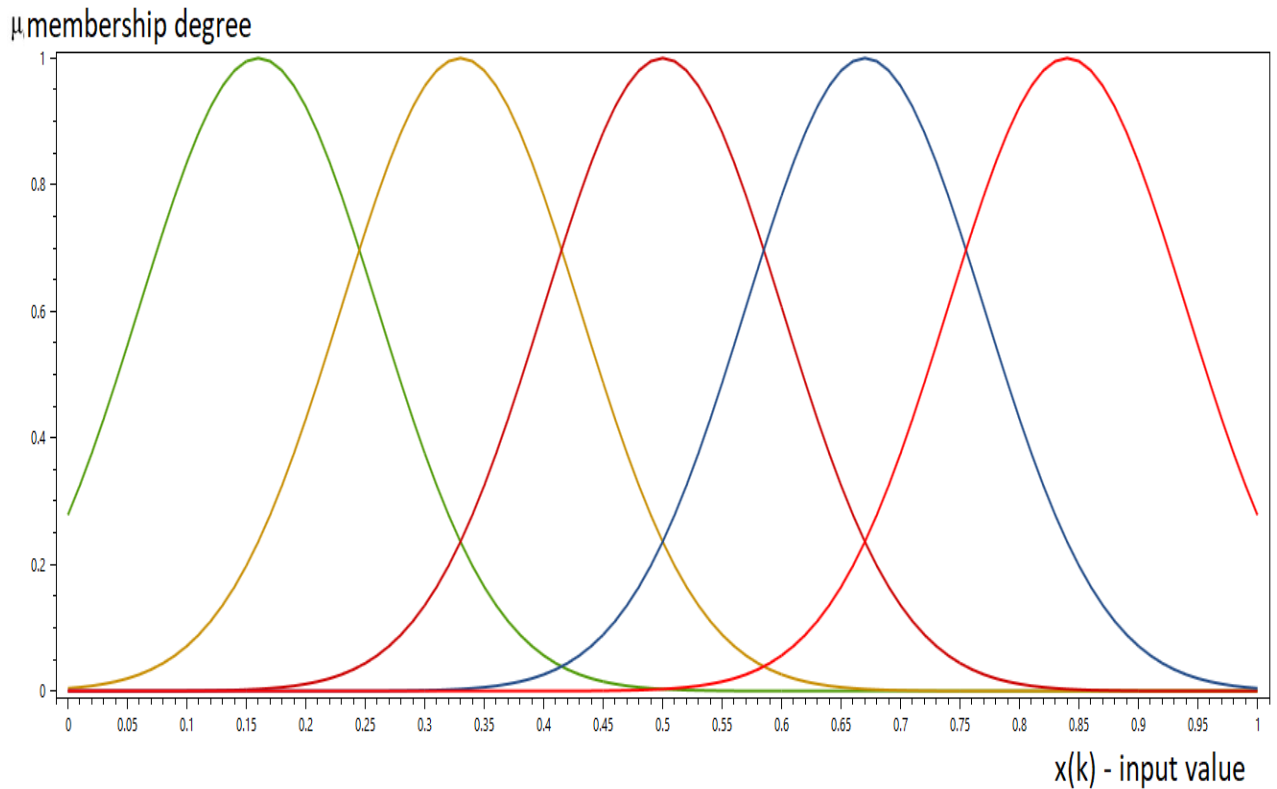


Рисунок 2.9 – приклад функцій належності першого шару

Шляхом диференціації ми досягли наступних формул для налаштування центрів:

$$\begin{cases} c_{jl}^{\psi}(k+1) = c_{jl}^{\psi}(k) - \lambda_{\psi} \frac{\tau_{jl}^{\psi} e(k)}{\eta_{\psi}(k)}, \\ \eta_{\psi}(k+1) = \beta_{\psi} \eta_{\psi}(k) + \tau_{jl}^{\psi 2} \end{cases}, \quad (2.14)$$

де λ_{ψ} – крок навчання;

β_{ψ} – параметр пам'яті;

τ_{jl}^{ψ} – вектор значень зворотного розповсюдження помилки по відношенню до c_{jl}^{ψ} .

2.1.2 Спеціалізований критерій навчання для подолання проблеми затримки у прогнозуванні

Дуже часто у реальних даних виникає проблема затримки прогнозованого сигналу, коли модель здатна відтворити модельовані дані проте з затримкою на декілька кроків.

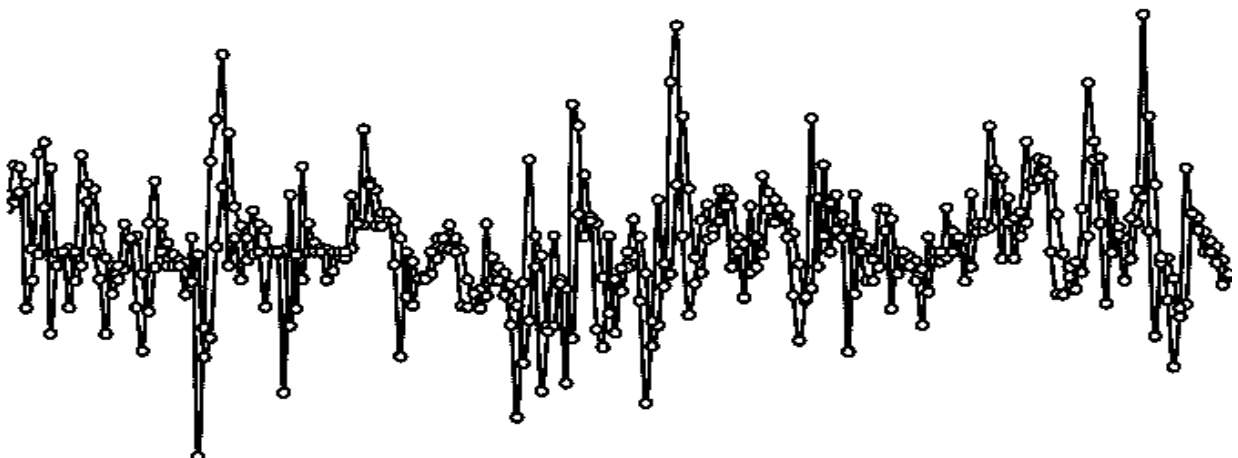


Рисунок 2.10 – Зразок ефекту затримки сигналу

Цей небажаний ефект можна мінімізувати, використовуючи спеціалізований фінансовий критерій від [98] та оновити процедуру навчання відповідно до нього. Критерій має наступний вигляд::

$$Wegstrecke = \frac{\sum_{k=1}^N signal(k)(y(k) - y(k - 1))}{\sum_{k=1}^N |y(k) - y(k - 1)|} \quad (2.15)$$

$$signal(k) = \begin{cases} 1 \text{ if } \hat{y}(k) - y(k - 1) > 0, \\ -1 \text{ if } \hat{y}(k) - y(k - 1) < 0, \\ 0 \text{ in all other cases.} \end{cases}$$

де $y(k)$ – поточне значення зовнішнього навчального сигналу;

$y(k - 1)$ – попереднє значення зовнішнього навчального сигналу;

$\hat{y}(k)$ – прогнозоване значення;

N – розмір вибірки для навчання;

$signal(k)$ функція недиференційована, тому використано гіперболічний тангенс як заміну.

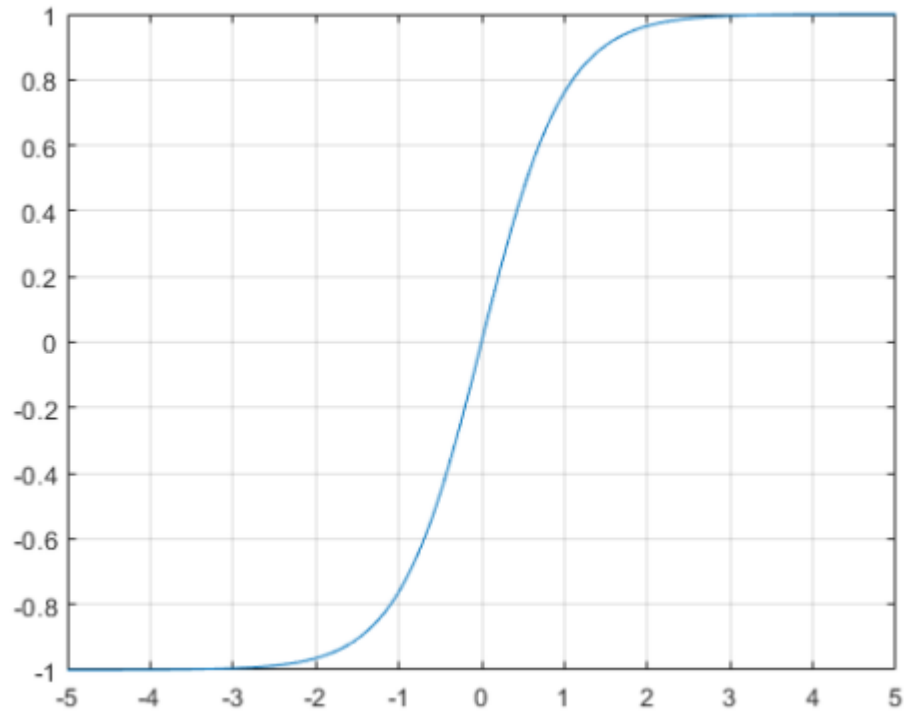


Рисунок 2.11 – Зразок гіперболічного тангенса

Оновлений метод навчання для матриць Q_{jl}^{-1} та центрів c_{jl}^{ϕ} набуває наступної форми:

$$\begin{cases} Q_{jl}^{-1}(k+1) = Q_{jl}^{-1}(k) + \lambda_Q \frac{\tau_{jl}^Q(k) \gamma \text{sign}(y(k) - y(k-1)) \tanh^2(\gamma(y(k) - y(k-1)))}{\eta_Q(k)} \\ \eta_Q(k+1) = \beta_Q \eta_Q(k) + \text{Tr}(\tau_{jl}^{Q^T} \tau_{jl}^Q) \end{cases} \quad (2.16)$$

де λ_Q – крок навчання;

and β_Q – параметр пам'яті;

τ_{jl}^Q – вектор значень зворотного розповсюдження помилки по відношенню до Q_{jl}^{-1} ;

γ – коефіцієнт нахилу функції \tanh .

$$\begin{cases} c_{jl}^\phi(k+1) = c_{jl}^\phi(k) + \lambda_c \frac{\tau_{jl}^c(k) \gamma \operatorname{sign}(y(k)-y(k-1)) \tanh^2(\gamma(y(k)-y(k-1)))}{\eta_c(k)} \\ \eta_c(k+1) = \beta_c \eta_c(k) + \tau_{jl}^{cT} \tau_{jl}^c \end{cases}, (2.17)$$

де λ_c – крок навчання;

and β_c – параметр пам'яті;

τ_{jl}^c – вектор значень зворотного розповсюдження помилки по відношенню до c_{jl}^ϕ ;

Ця зміна мінімізувала вплив зсуву прогнозу на реальні набори даних без негативного впливу на обчислювальні показники. Недоліком є те, що коефіцієнт нахилу функції став додатковим системним гіперпараметром, і його слід вибирати дуже ретельно.

2.2 Узагальнена модель для декількох вихідних значень

У цьому підрозділі представлено узагальнення нейро-фаззі моделі до випадку багатьох вихідних змінних та відповідну зміну процедури навчання. Ця модель має три переваги: (1) досягає кращої точності завдяки використанню репрезентативних здібностей багатовимірних функцій Гауса в четвертому шарі; (2) показує хорошу обчислювальну продуктивність, яка досягається за допомогою процедури стохастичної оптимізації градієнта для налаштування послідовних одиниць шару та ітеративного проєктивного

алгоритму для налаштування їх ваг; та (3) має розумну кількість гіперпараметрів, які можна вибрати.

Основна відмінність від архітектури, описаної у підрозділі 2.1 полягає в тому, що ми маємо t виходів $\hat{y}(k) = (\hat{y}_1(k), \hat{y}_2(k), \dots, \hat{y}_m(k))^T$ замість лише одного \hat{y} для кожного вхідного шаблону $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ і, відповідно, t складених послідовних одиниць шару $F = (F_1(k), F_2(k), \dots, F_m(k))^T$. На рис. 2.12 зображена модифікована архітектура.

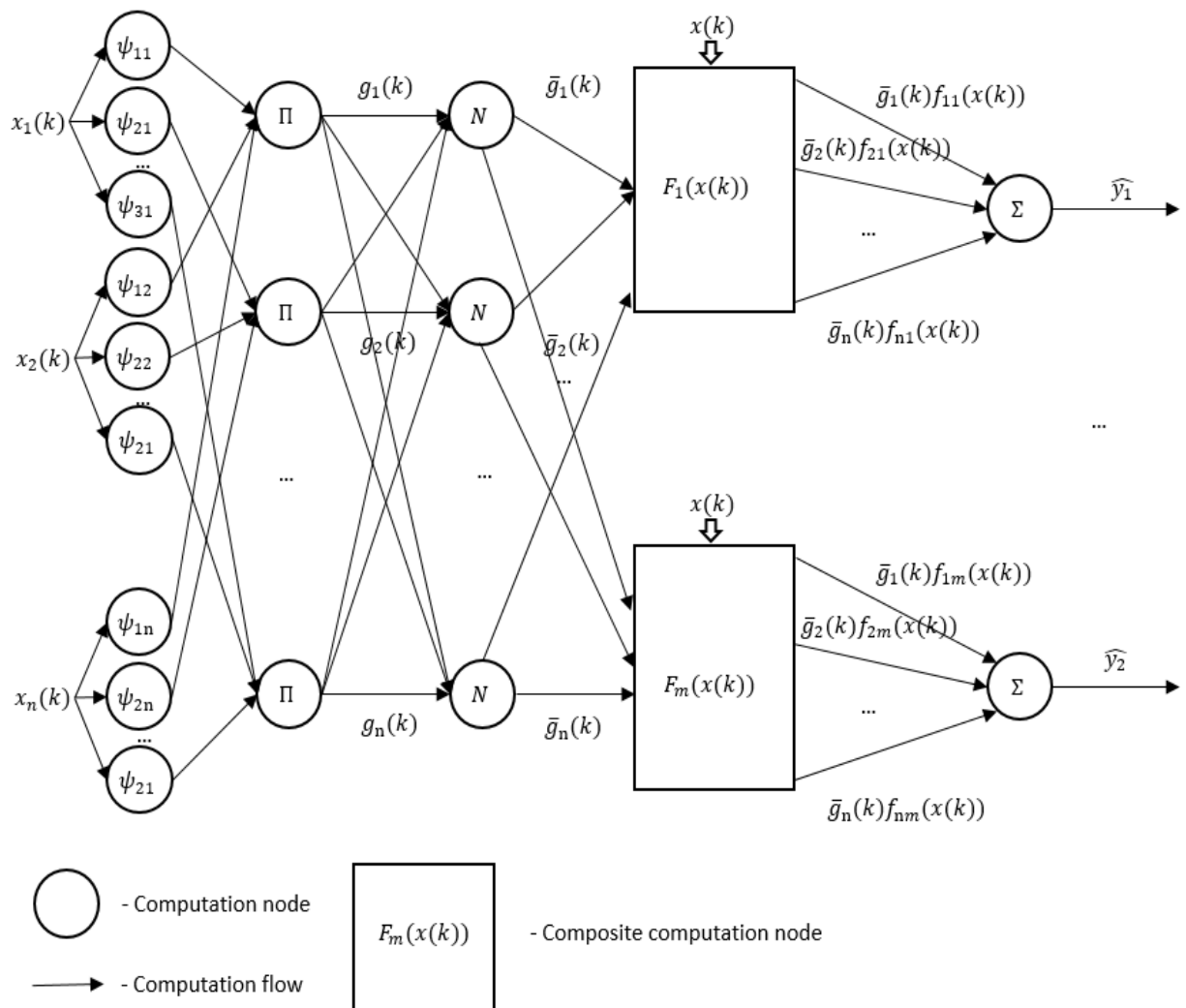


Рисунок 2.12 – Загальна архітектура моделі

Деталізована структура четвертого (консеквентного) шару представлена на рисунку 2.13.

Відповідно загальний результат консеквентного шару:

$$f_{aj}(x(k)) = \sum_{e=1}^{h^\phi} p_{aje} \phi_{aje}(x(k)), \tag{2.18}$$

де h^ϕ кількість багатовимірних гаусіанів для кожного обчислювального вузла f_j .

Вихідне значення моделі:

$$\hat{y}_a = \sum_{j=1}^{h^\phi} g_{aj} f_{aj}(x(k)). \tag{2.19}$$

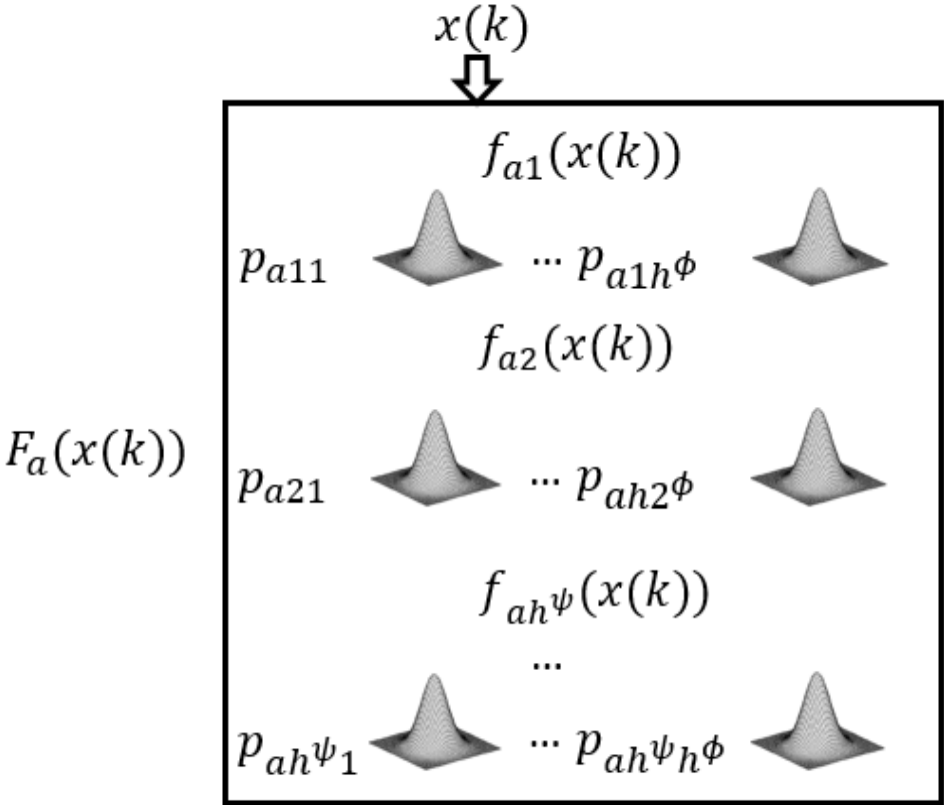


Рисунок 2.13 – Деталізована структура четвертого (консеквентного) шару.

Запропонований метод навчання узагальнює підхід до навчання, запропонований у розділі 2 для багатовимірного випадку. Він включає два одночасних процеси: оптимізацію у просторі вагових коефіцієнтів P і в просторі параметрів багатовимірних гаусіанів.

На етапі ініціалізації центри c_{ajl}^ϕ розміщуються рівновіддалено та матриці Q_{ajl} ініціалізуються як одиничні. Вагові коефіцієнти p ініціалізуються випадковим чином у діапазоні $[-0,1; 0,1]$.

Навчання вагових коефіцієнтів p_a набуває наступної форми:

$$p_a(k+1) = p_a(k) + \frac{y_a(k) - p_a^T f_a(x(k))}{f_a^T(x(k)) f_a(x(k))} f_a(x(k)), \quad (2.20)$$

де $p_a(k)$ – вагові коефіцієнти;

$y_a(k)$ – поточний компонент зовнішнього навчального сигналу;

$p_a^T f_a(x(k))$ – вихідне значення моделі.

Рисунок 2.13 демонструє спрощений двовимірний зріз функції похибки у просторі вагових коефіцієнтів.

Центри c_{ajl}^ϕ налаштовуються за процедурою:

$$\begin{cases} c_{ajl}^\phi(k+1) = c_{ajl}^\phi(k) + \lambda_c \frac{\tau_{ajl}^c e(k)}{\eta_c(k)}, \\ \eta_{ac}(k+1) = \beta_c \eta_{ac}(k) + \tau_{ajl}^c \tau_{ajl}^c \end{cases}, \quad (2.21)$$

де λ_c – крок навчання;

β_c – параметр пам'яті;

τ_{ajl}^c – вектор значень зворотного розповсюдження помилки по відношенню до c_{ajl}^ϕ .

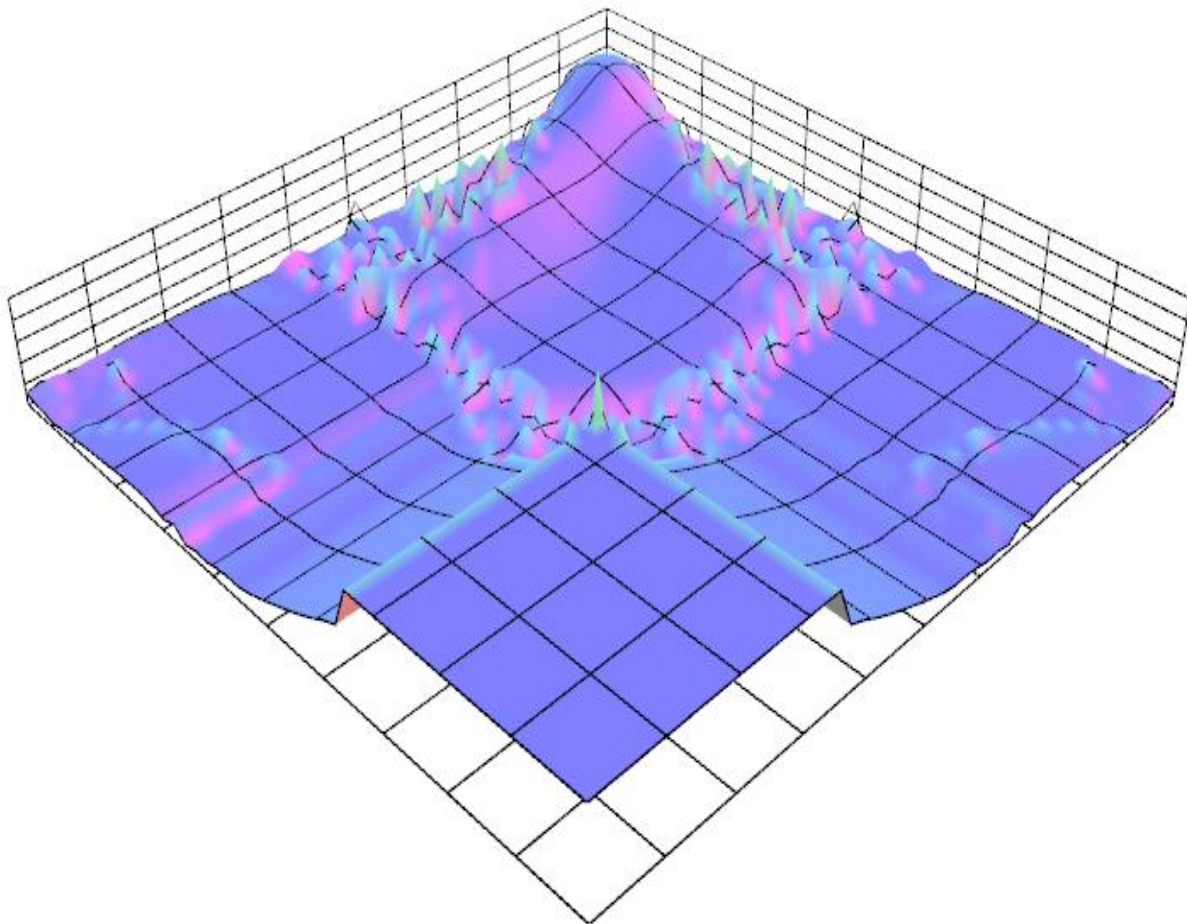


Рисунок 2.14 – Поверхня похибки у просторі вагових коефіцієнтів.

Налаштування матриць Q_{ajl} :

$$\begin{cases} Q_{ajl}(k+1) = Q_{ajl}(k) + \lambda_Q \frac{\tau_{ajl}^Q e(k)}{\eta_{aQ}(k)}, \\ \eta_{aQ}(k+1) = \beta_Q \eta_{aQ}(k) + \text{Tr}(\tau_{ajl}^Q \tau_{ajl}^Q), \end{cases} \quad (2.22)$$

де λ_Q – крок навчання;

β_Q – параметр пам'яті;

τ_{ajl}^Q – вектор значень зворотного розповсюдження помилки по відношенню до Q_{ajl} .

Вектори η_{ac} та матриці η_{aQ} репрезентують згасаючі середні значення попередніх градієнтів

Алгоритм ініціалізується з $\eta_{ac} = \eta_{aQ} = 10,000$.

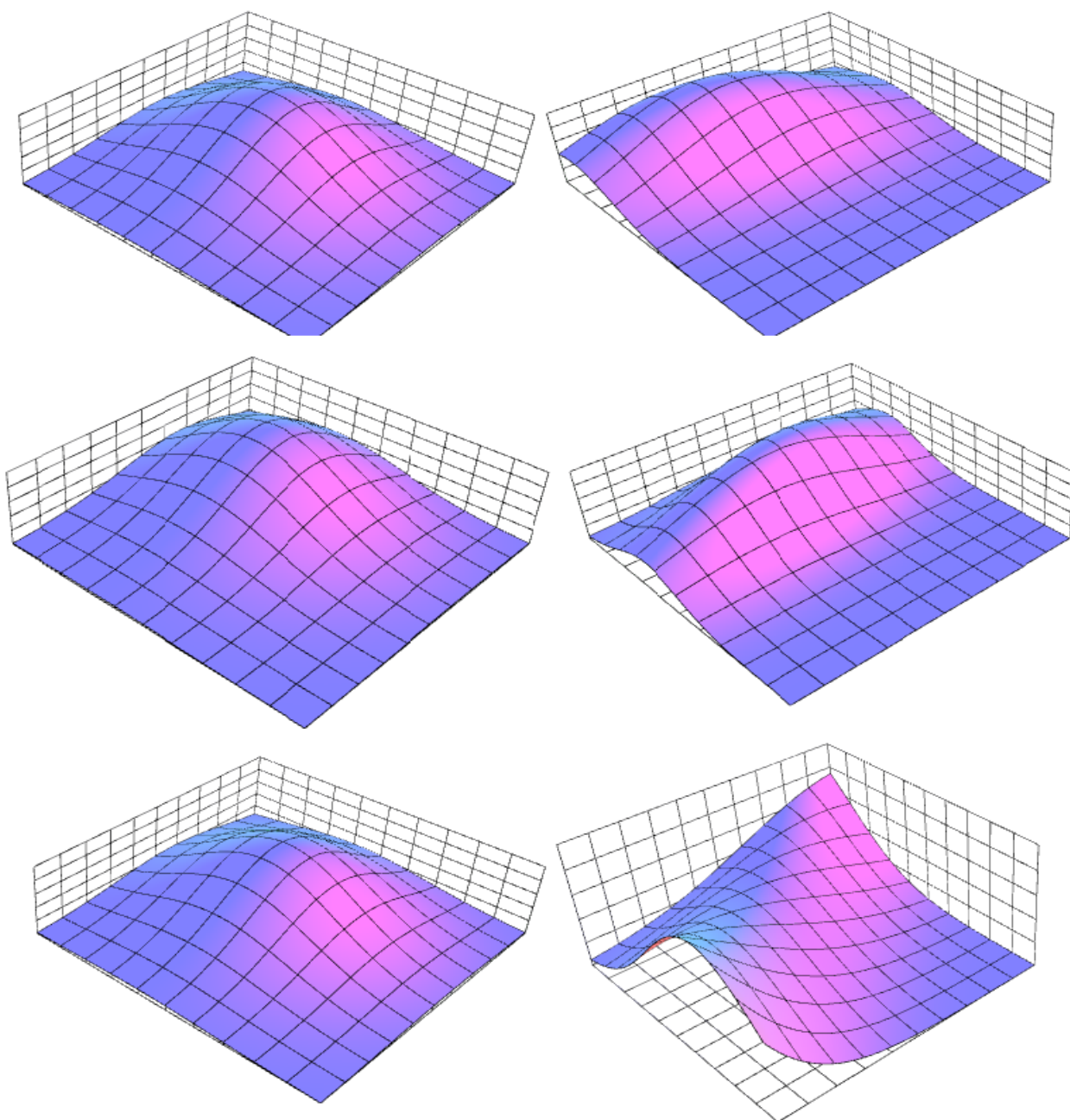


Рисунок 2.15 – приклад гаусіанів до та після навчання – зліва функції після ініціалізації, праворуч – після завершення навчання

Експериментальне моделювання продемонструвало хороші обчислювальні показники та точність прогнозування нашої моделі порівняно з добре відомими моделями штучних нейронних мереж. Аналітична верифікація збіжності та стабільності за допомогою процедури, побудованій на методі Ляпунова, гарантує можливість застосування метода навчання для широкого спектру прикладних задач.

2.3 Висновки до розділу

1. Запропоновано нову п'ятишарову гібридну нейро-фаззі модель з багатовимірними гаусіанами у консеквентному шарі.

2. Запропоновано швидкий метод навчання для запропонованої моделі який поєднує ітеративний проєкційний метод Качмажа для налаштування вагових коефіцієнтів та градієнтний стохастичний метод для оптимізації параметрів функцій четвертого шару.

3. Запропоновану модель та метод навчання розширено для багатовимірного випадку. Модель забезпечує оптимальну ефективність навчання щодо інших шляхом неявної обробки взаємозв'язків між вихідними змінними, які є зустрічаються у даних.

Список використаних у цьому розділі джерел наведено у повному списку використаних джерел під номерами [43] [98].

Основні результати цього розділу опубліковано в роботах [99, 100] [101] [102] [103]

3 АНСАМБЛЬ НЕЙРО-ФАЗЗИ МОДЕЛЕЙ У ЗАДАЧАХ ОПЕРАТИВНОГО АНІЛІЗУ РИЗИКІВ

Мета розділу – створити ансамбль нейро-фаззи моделей для задач оперативного аналізу ризиків, який дозволяє зменшити похибку за рахунок зменшення варіативності, спростити вибір гіперпараметрів та використати можливості сучасних багатоядерних процесорів за рахунок паралельного навчання складових моделей.

Завдання розділу:

1) Розробити ансамбль гібридних нейро-фаззи моделей з багатовимірними гаусіанами у консеквентному шарі, який шляхом усереднення вихідних значень досягає кращої ефективності у задачах прогнозування та аналізу;

2) Створити метод синтезу ансамблю нейро-фаззи моделей, який спрощує підбір гіперпараметрів навчання моделей та дозволяє досягти меншої похибки у процесі навчання, особливо на коротких наборах даних, за рахунок позитивного ефекту від різної динаміки навчання моделей-компонентів викликаної різницею при ініціалізації.

3.1 Архітектура ансамблю нейро-фаззи моделей

Для вирішення проблем чутливості процесу навчання до вибору гіперпараметрів ми пропонуємо модель ансамблю та процедуру ініціалізації для неї. Ансамбль складається з моделей M_e . Загальна архітектура, представлена на рис.3.1.

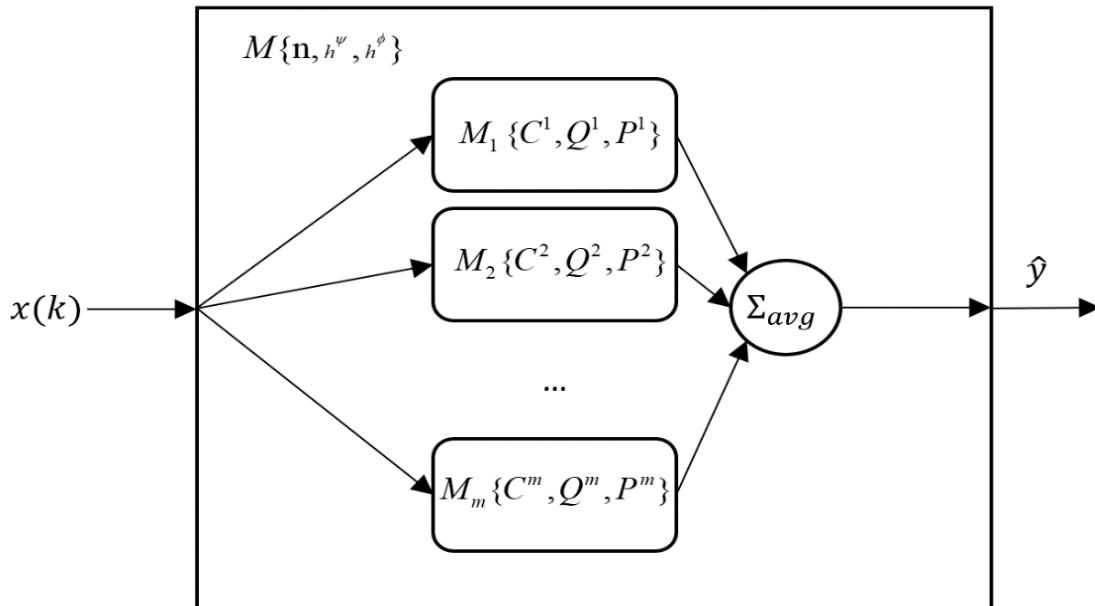


Рисунок 3.1 – Загальна архітектура ансамблю моделей

Кожен вхідний вектор $x(k)$ передається на вхід всім моделям M_e які мають однакову структуру, що визначається трьома параметрами - довжиною вхідного вектору n , кількістю функцій рецепторного шару h^ψ та кількістю функцій четвертого шару h^ϕ (аналогічно кількості поліноміальних членів в ANFIS). Різниця полягає в параметрах четвертого шару C^e, Q^e, P^e , які ініціалізуються випадковим чином, а потім налаштовуються на етапі навчання. Матриця C^e представляє всі вектори-центри багатовимірних гаусіанів, матриця Q^e - їх матриці рецепторних полів і P^e - матриця вагових коефіцієнтів для кожної моделі M_e .

Потім усі результати надсилаються на вузол виводу Σ_{avg} , який обчислює отримане значення \hat{y} як середнє значення:

$$\hat{y} = \frac{\sum_{e=1}^m \hat{y}_e}{m}, \quad (3.1)$$

де – \hat{y}_e вихідний сигнал моделі;

m – кількість моделей-членів.

Усереднення результатів моделей - це найпростіший спосіб поєднання різних моделей та досягнення кращої точності, припускаючи, що різні моделі матимуть частково незалежні помилки, що в свою чергу може призвести до кращих можливостей узагальнення.

Детальна структура моделі-елемента M_e ансамблю зображена на рис. 3.2.

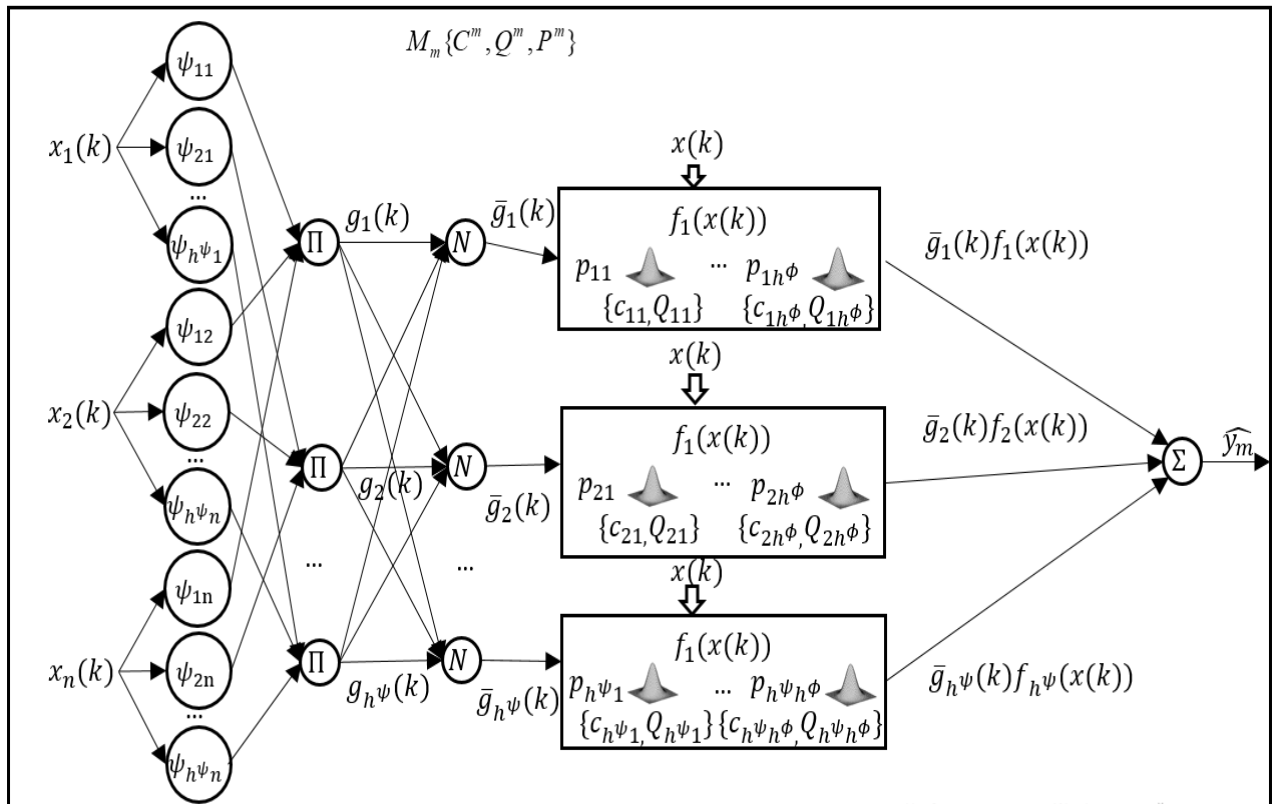


Рисунок 3.2 – Архітектура локальної моделі

Результат \hat{y}_e кожної моделі M_e обчислюється наступним чином:

$$\hat{y}_e = p_e^T f_e(x(k)), \quad (3.2)$$

де $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ – вектор вхідних значень;

p_e – вектор вагових коефіцієнтів (векторне представлення матриці P^e);

$f_e(x(k))$ – вектор нормалізованих значень обчислювальних блоків четвертого шару:

$$f_e(x(k)) = (\bar{g}_{e1}\phi_{e11}(x(k)) \dots \bar{g}_{eh^\psi}\phi_{eh^\psi h^\phi}(x(k))), \quad (3.3)$$

де h^ψ – кількість функцій фаззифкації у першому шарі;

h^ϕ – кількість багатовимірних гаусіанів у четвертому шарі для кожного нормованого виходу \bar{g}_{el} :

$$\bar{g}_{je}(k) = \frac{g_{je}(k)}{\sum_{j=1}^{h^\psi} g_{je}(k)} = \frac{\prod_{i=1}^n \psi_{jle}(x_i(k))}{\sum_{i=1}^{h^\psi} \prod_{i=1}^n \psi_{jle}(x_i(k))}, \quad (3.4)$$

3.2 Алгоритм синтезу ансамблів нейро-фаззі моделей

Алгоритм синтезу показаний на рис. 3.3. Його метою є ініціалізація моделей членів перед навчанням і встановлення параметрів навчання. Результати показують, що і гіперпараметри β_c, β_Q , які є демпінговими параметрами для оптимізації C^e та Q^e відповідно, суттєво впливають на точність прогнозування моделі. Відповідно до цього ми створюємо однакові моделі-члени, але встановлюємо різні значення для β_c і β_Q . Центри розміщені на однаковій відстані, а матриці ідентичності використовуються як сприйнятливі перед навчанням.

Після створення моделей ми можемо починати навчати їх паралельно.

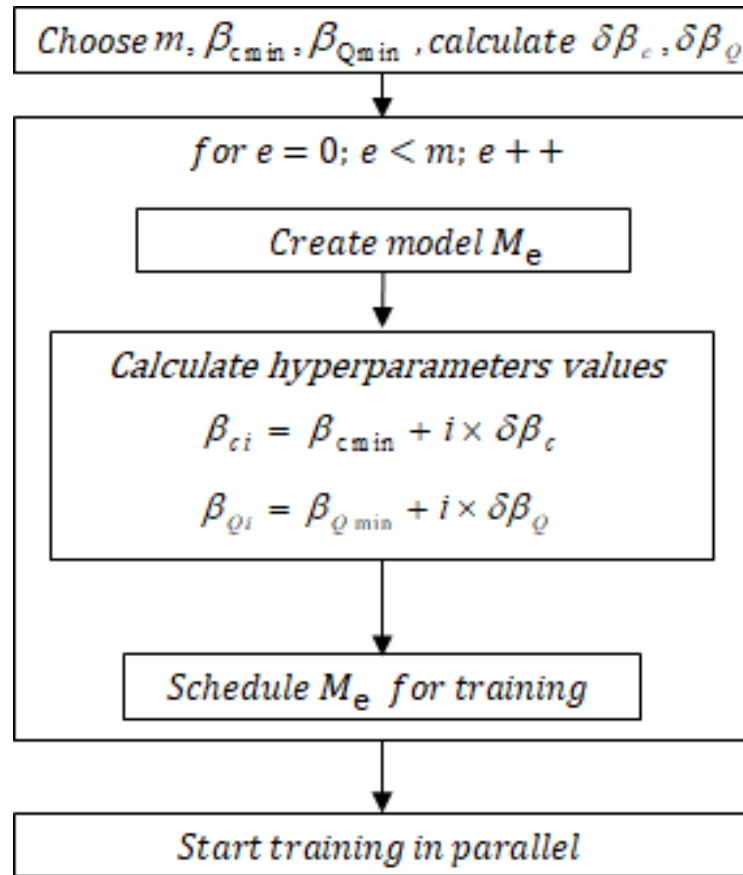


Рисунок 3.3 – Метод синтезу ансамблю моделей.

Загальний навчальний процес зображений на рис. 3.4. Вхідні вектори подаються на входи кожної моделі. Потім результуюче значення $y(k)$ використовується для розрахунку похибки e_e для кожної локальної моделі. Похибка далі розповсюджується для обчислення дельт $\Delta_m\{C^m, Q^m, P^m\}$, які використовуються при оптимізації вільних параметрів C^e, Q^e, P^e .

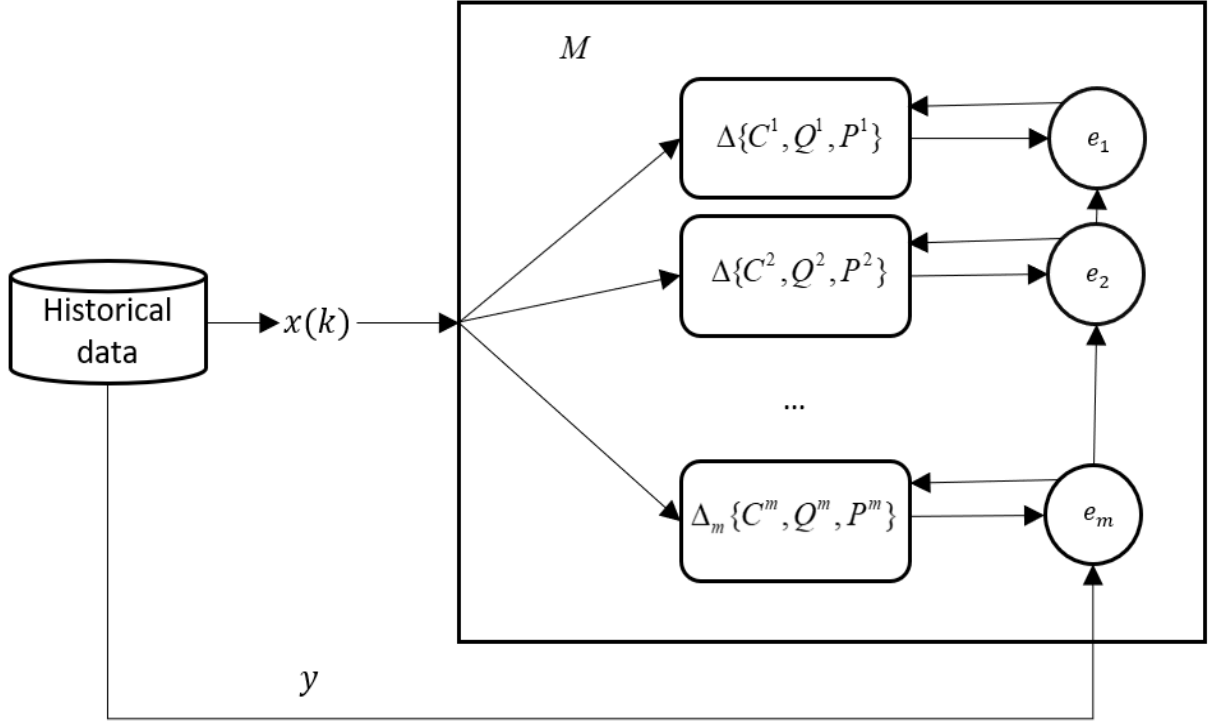


Рисунок 3.4 – Схема навчання ансамблю моделей та розповсюдження похибки.

Процес навчання зосереджений на багатовимірних функціях Гауса четвертого шару та їх вагових коефіцієнтів. Навчання центрів $c_{ejl}^\phi \in C^e$ та матриць коваріації $Q_{ejl} \in Q^e$ набуває наступного вигляду:

$$\begin{cases} c_{ejl}^\phi(k+1) = c_{ejl}^\phi(k) + \lambda_c \frac{\tau_{ejl}^c(k)e(k)}{\eta_c(k)} \\ \eta_{ec}(k+1) = \beta_c \eta_{ec}(k) + \tau_{ejl}^c T \tau_{ejl}^c \\ Q_{ejl}(k+1) = Q_{ejl}(k) + \lambda_Q \frac{\tau_{ejl}^Q(k)e(k)}{\eta_{eq}(k)} \\ \eta_{eq}(k+1) = \beta_{Qe} \eta_{eq}(k) + Tr(\tau_{ejl}^Q T \tau_{ejl}^Q) \end{cases}, \quad (3.4)$$

де λ_c та λ_Q – кроки навчання;

β_{ce} та β_{Qe} – параметри пам'яті для поточної моделі ансамблю,

вектор τ_{ajl}^c та матриця τ_{ejl}^Q – значення градієнту помилки по відношенню до c_{ejl}^ϕ та Q_{ejl} відповідно.

Вектори η_{ec} та матриці η_{eQ} репрезентують згасаючі середні значення попередніх градієнтів. Метод починає своє виконання з $\eta_{ec} = \eta_{eQ} = 10,000$.

На рисунку 3.5 наведено приклади багатовимірних гаусіанів з різних локальних моделей після закінчення навчання - вони були ініціалізовані однаково і навчались на одному і тому ж наборі даних, різниця лише в ініціалізації випадкових ваг та гіперпараметрах, які визначали динаміку навчання.

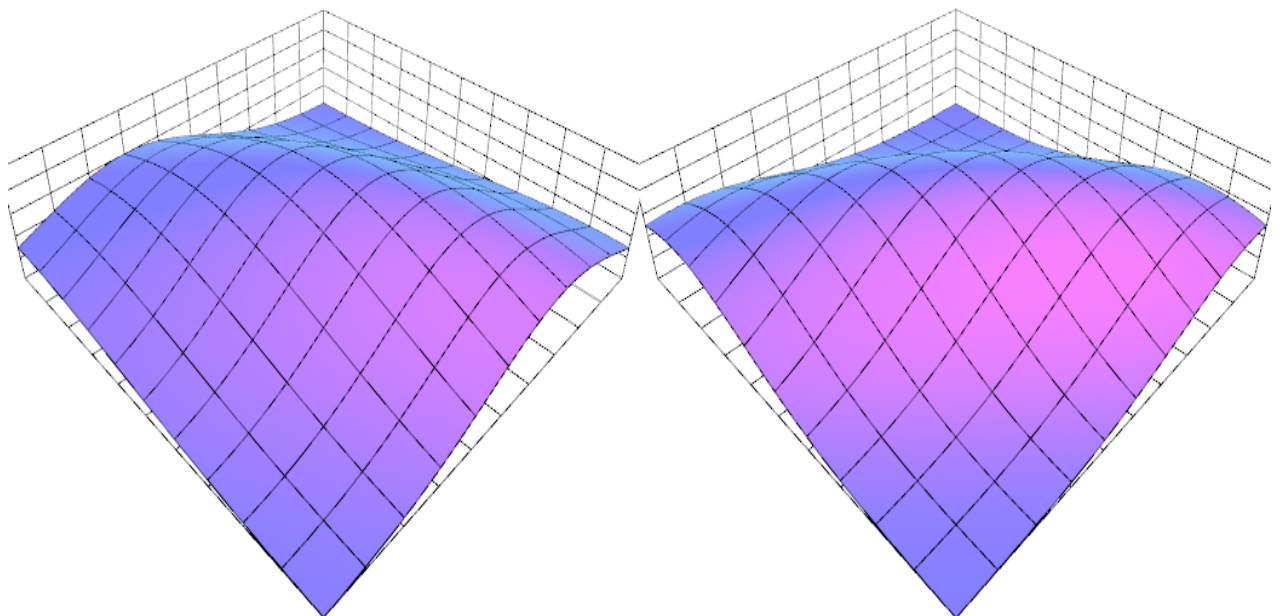


Рисунок 3.5 – Приклади багатовимірних гаусіанів, які були ініціалізовані однаково, але налаштовані процесом навчання з різними гіперпараметрами.

Оптимізація у просторі вагових коефіцієнтів P^e набуває наступної форми:

$$p_e(k+1) = p_e(k) + \frac{e_a}{f_e^T(x(k))f_e(x(k))} f_e(x(k)), \quad (3.5)$$

де $p_a(k)$ – вагові коефіцієнти у векторній формі для спрощення обчислень.

На рис. 3.6 показано процес зменшення помилок під час навчання моделі. Усереднений результат виходу ансамблевою моделі показує як швидший спад помилок, так і кращу стабільність - окремі моделі мають тенденцію до більшої дисперсії.

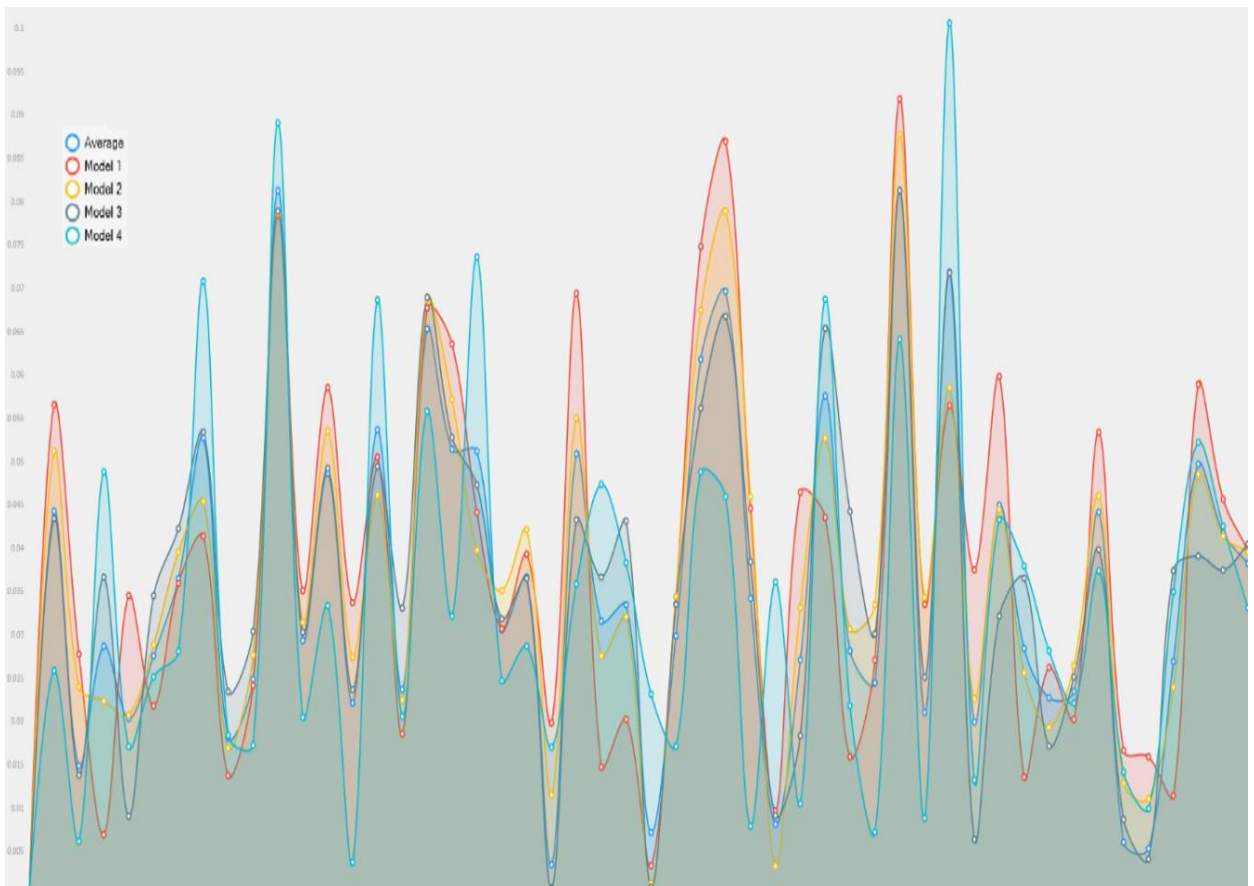


Рисунок 3.6 – Візуалізація згасання помилки ансамблю та його складових моделей

3.3 Висновки до розділу

1. Запропоновано ансамбль нейро-фаззі моделей для задач інтелектуального оперативного аналізу ризиків, який має демонструє меншу похибку ніж окремі моделі.

2. Запропоновано метод синтезу який дозволяє спростити вибір гіперпараметрів та використати можливості сучасних багатоядерних процесорів за рахунок паралельного навчання складових моделей.

Основні результати цього розділу опубліковано в роботі [104].

4 ПОЄДНАННЯ ЕМПІРИЧНОЇ МОДОВОЇ ДЕКОМПОЗИЦІЇ ТА НЕЙРО-ФАЗЗИ МОДЕЛЕЙ, ІЄРАРХІЧНІ МОДЕЛІ

Мета розділу – описано модифікований метод навчання нейро-фаззи моделей та синтезу ансамблів нейро-фаззи моделей з використанням методу емпіричної модової декомпозиції для зниження шумів та попередньої декомпозиції вхідних даних. Також описано використання запропонованих моделей як складових моделей у ієрархічних знання-орієнтованих системах нечіткого логічного висновування та експертних системах побудованих на їхній основі.

Завдання розділу

- 1) Створити метод навчання нейро-фаззи моделі, що використовує властивості емпіричної модової декомпозиції у даних з високим рівнем шумів
- 2) Розробити метод синтезу ансамблю нейро-фаззи моделей, який вирізняється спрощеною процедурою вибору параметрів за рахунок використання декомпозиційних властивостей емпіричної модової декомпозиції
- 3) Створити ієрархічні знання-орієнтовані моделі у задачах оперативного аналізу ризиків, що на відміну від існуючих методів дозволяють опрацьовувати малоймовірні критичні значення ризикових подій

4.1 Метод навчання нейро-фаззи моделей з використанням емпіричної модової декомпозиції

У цьому підрозділі пропонується новий метод, заснований на поєднанні емпіричної модової декомпозиції та нейро-фаззи моделі для досягнення кращої точності.

Точність запропонованої у другому розділі моделі на реальних даних є нижчою ніж на синтетичному наборі даних з доданим шумом. Цей факт

означає, що застосування процедури декомпозиції сигналу та зниження шуму може принести значні переваги.

Запропонована модель складається з трьох високорівневих кроків, зображених на рис. 4.1.

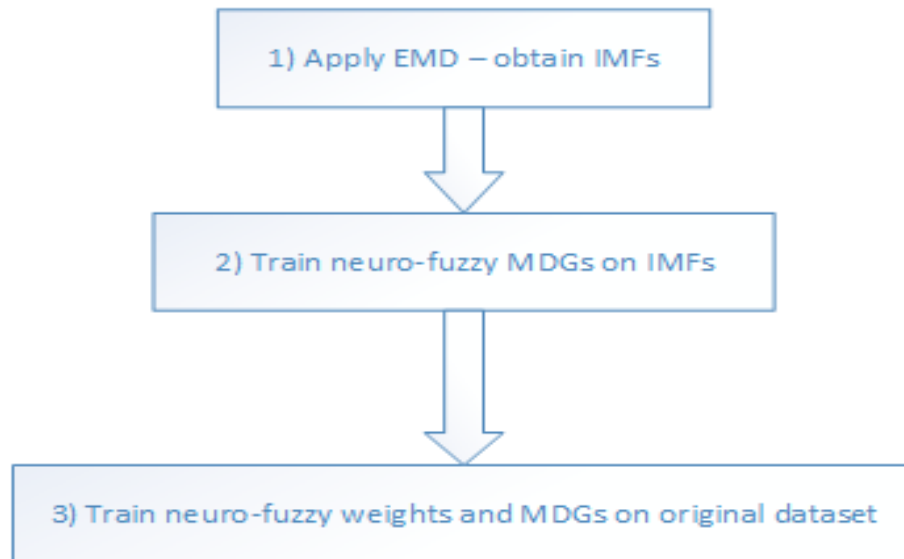


Рисунок 4.1 – Високорівневе представлення процесу навчання з використанням емпіричної модової декомпозиції

На першому кроці ми нормалізуємо набір даних, потім застосовується перетворення використанням методу декомпозиції емпіричних мод для отримання власних модових функцій які представляють прості гармонійні компоненти вхідних даних. Отримання власних модових функцій називається просіюванням і має наступну структуру:

1. Ідентифікація локальних екстремумів у наборі.
2. Інтерполяція кубічними сплайнами для отримання верхньої $e_u(t)$ та нижньої $e_l(t)$ обвідних вибірки.
3. Отримати криву середнього значення

$$m_1(t) = \frac{e_u - e_l}{2} \quad (4.1)$$

та різницю як прообраз власної модової функції

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \quad (4.2)$$

4. Повторити k разів, поки $h_{1k}(t)$ не стане власною модовою функцією $h_{1k}(t) = c_1$. Для визначення кількості кроків просіювання використовується критерій стандартного відхилення:

$$SD_k = \sum_{t=1}^T \frac{|h_{k-1}(t) - h_k(t)|^2}{h_k^2(t)} \quad (4.3)$$

5. Повторити процес та отримати наступні власні модові функції.

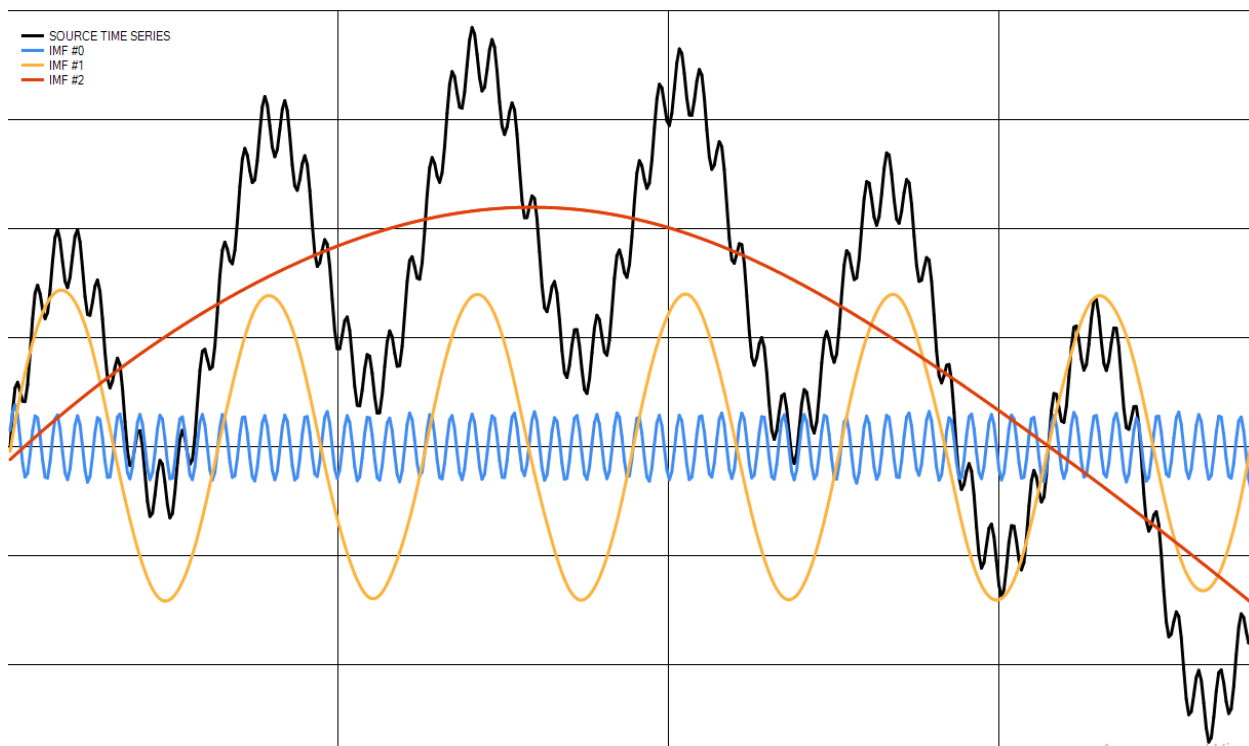


Рисунок 4.2 – Приклад декомпозиції у власні модові функції.

Завершивши процес ми використовуємо отримані власні модові функції для навчання обчислювальних блоків четвертого шару $f_j(x(k))$ функцій рецепторних полів по одному, як у прикладі, показаному на рис. 4.3.

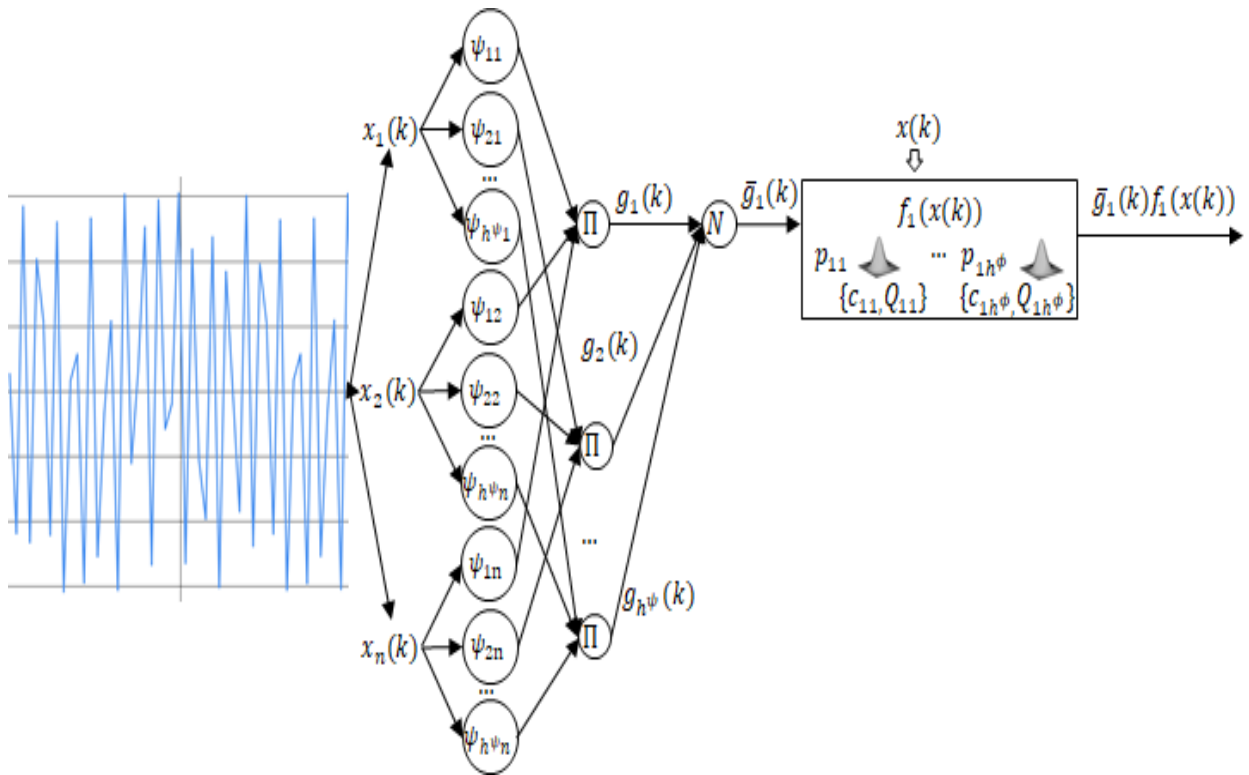


Рисунок 4.3 – Зразок навчання на власній модовій функції

На наступному кроці ми виконуємо навчання вагових коефіцієнтів, центрів c_{jl}^ϕ та матриць Q_{jl}^{-1} на вхідному сигналі:

$$\begin{cases} c_{jl}^\phi(k+1) = c_{jl}^\phi(k) + \lambda_c \frac{\tau_{jl}^c(k)e(k)}{\eta_c(k)} \\ \eta_c(k+1) = \beta_c \eta_c(k) + \tau_{jl}^{cT} \tau_{jl}^c \\ Q_{jl}(k+1) = Q_{jl}(k) + \lambda_Q \frac{\tau_{jl}^Q(k)e(k)}{\eta_Q(k)} \\ \eta_Q(k+1) = \beta_Q \eta_Q(k) + Tr(\tau_{jl}^{QT} \tau_{jl}^Q) \end{cases} \quad (4.4)$$

де λ_c та λ_Q – кроки навчання;

β_c та β_Q – параметри пам'яті;

τ_{jl}^Q та τ_{jl}^c – значення градієнту помилки по відношенню до кожного багатовимірного гаусіану.

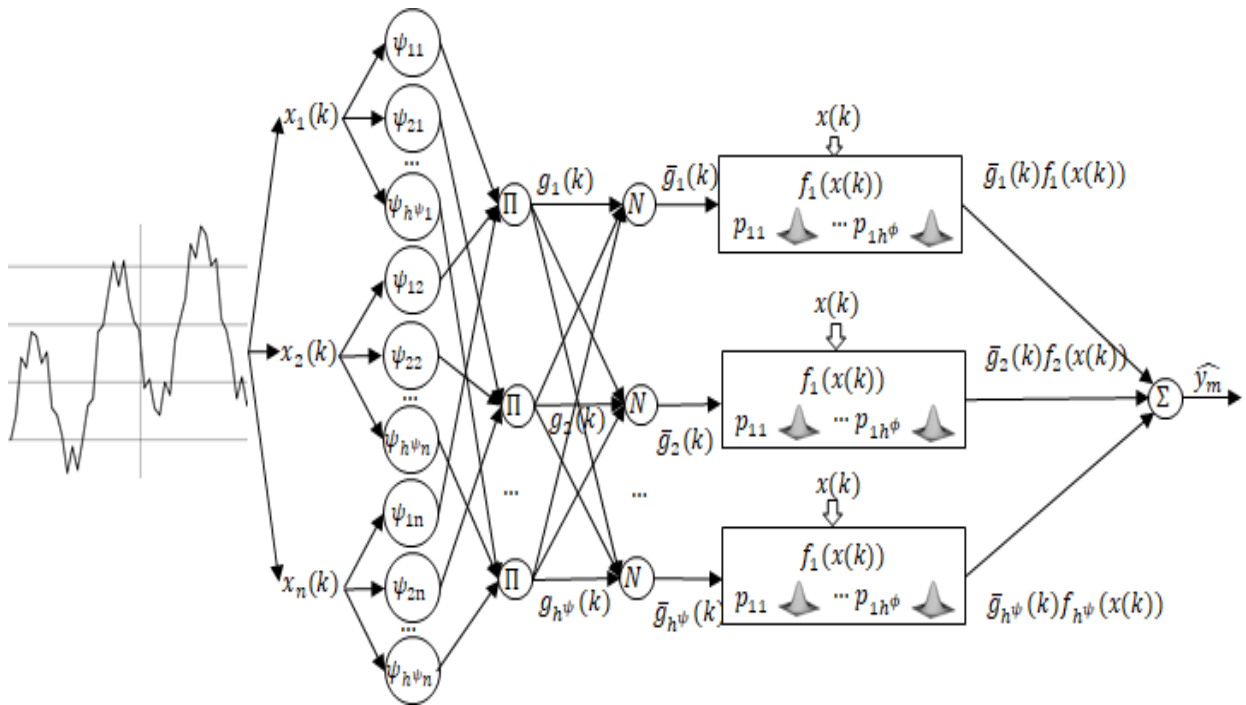


Рисунок 4.4 – Загальна архітектура на третьому кроці з навчанням на оригінальних даних

Різниця з навчанням у розділі 2 полягає в тому, що ми проводимо процедуру навчання окремо для кожного $f_j(x(k))$ та використовуємо відповідну власну модову функцію. Ця частина алгоритму навчання може бути реалізована паралельно.

Потім ми налаштуємо параметри всієї моделі на справжньому навчальному наборі.

Експериментальні результати показують значне покращення порівняно з оригінальною моделлю та її конкурентами.

4.2 Синтез ансамблю нейро-фаззі моделей за допомогою методу емпіричної модової декомпозиції

У випадку ансамблів використано результати емпіричної модової декомпозиції для синтезу та навчання ансамблю нейро-фаззі моделей. Основні етапи побудови запропонованої моделі зображені на рис. 4.5.

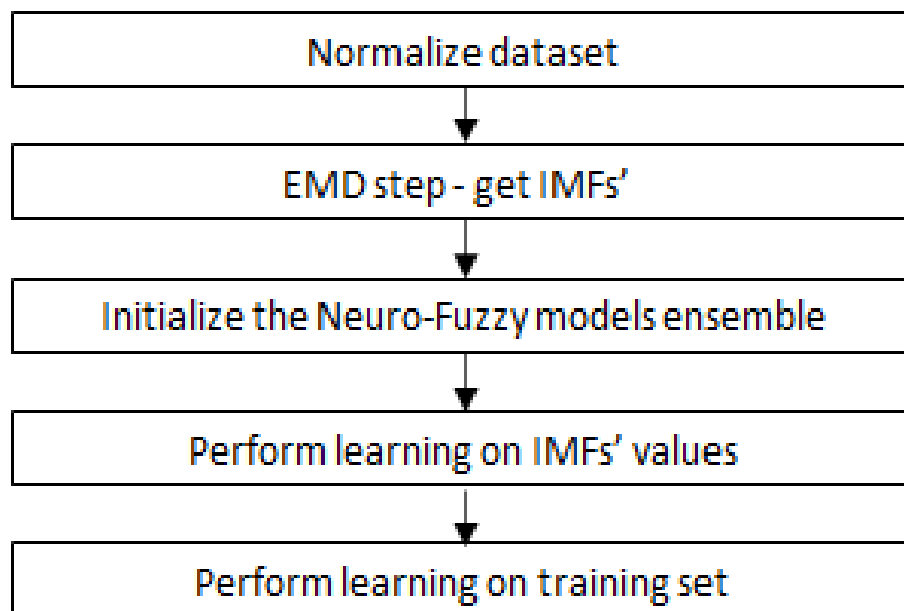


Рисунок 4.5 – Високорівневе представлення методу синтезу та навчання ансамблю

На рис. 4.6 зображена високорівнева структура ансамблю та процесу навчання.

Навчання центрів $c_{ejl}^{\phi} \in C^e$ та матриць коваріації $Q_{ejl} \in Q^e$ набуває наступного вигляду:

$$\begin{cases} c_{ejl}^{\phi}(k+1) = c_{ejl}^{\phi}(k) + \lambda_c \frac{\tau_{ejl}^c(k)e(k)}{\eta_c(k)} \\ \eta_{ec}(k+1) = \beta_c \eta_{ec}(k) + \tau_{ejl}^c T \tau_{ejl}^c \\ Q_{ejl}(k+1) = Q_{ejl}(k) + \lambda_Q \frac{\tau_{ejl}^Q(k)e(k)}{\eta_{eQ}(k)} \\ \eta_{eQ}(k+1) = \beta_{Qe} \eta_{eQ}(k) + Tr(\tau_{ejl}^Q T \tau_{ejl}^Q) \end{cases} \quad (4.5)$$

де λ_c та λ_Q – кроки навчання

β_{ce} та β_{Qe} – параметри пам'яті для поточної моделі ансамблю

вектор τ_{ajl}^c та матриця τ_{ejl}^Q – значення градієнту помилки по відношенню до c_{ejl}^{ϕ} та Q_{ejl} відповідно.

Потім ми налаштуємо параметри всієї моделі на справжньому навчальному наборі.

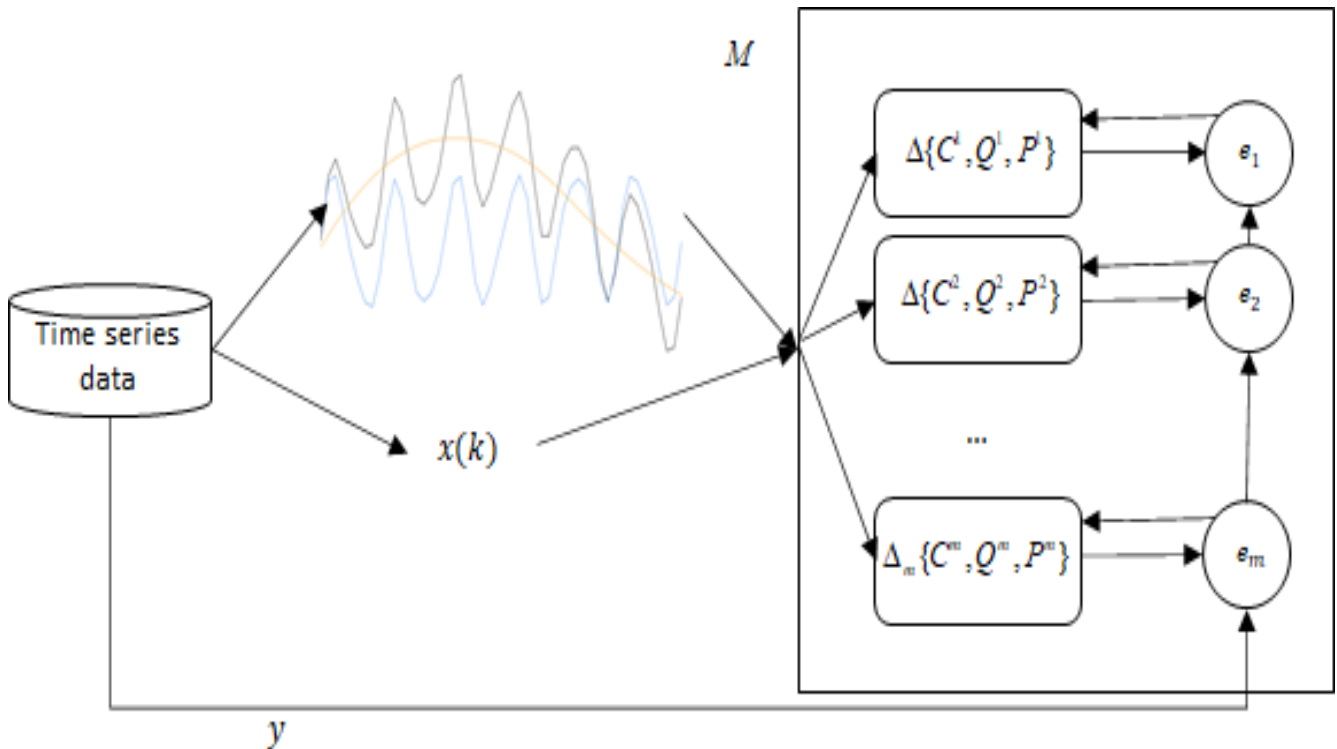


Рисунок 4.6 – Метод синтезу ансамблю моделей з використанням методу емпіричної модової декомпозиції

4.3 Ієрархічні знання-орієнтовані моделі у задачах аналізу ризиків

У багатьох випадках для комплексного вирішення проблеми оперативного аналізу ризиків недостатньо окремих моделей прогнозування та аналізу для окремих наборів даних. Тому виникає потреба у складних знання-орієнтованих методах та моделях які здатні поєднувати різні моделі, відповідальні за різні аспекти системи та експертні знання.

Цей підрозділ присвячено побудові ієрархічних знання-орієнтованих моделей для задач аналізу ризиків, які включають як складові моделі звичайні нечіткі системи логічного висновування, нейро-фаззі моделі та системи нечіткого логічного висновування з ймовірнісною мірою у консеквенті.

Запропонована модель є ієрархічною, оскільки на даний момент це найдієвіший спосіб подолання так званого «прокляття розмірності» - експоненціального характеру залежності між кількістю вхідних параметрів нечіткої моделі і кількістю задіяних правил [29]. Ієрархічні нечіткі моделі, вперше запропоновані в роботі [105], складаються з нечітких моделей менших розмірностей, і виходи моделей одного рівня є входами моделей наступного рівня, що дозволяє значно скоротити кількість правил.

Запропоновану модифікацію моделі ієрархічного нечіткого логічного висновування відрізняє використання локальних локальних моделей двох різних типів - власне нечітких і гібридних нечітко-ймовірнісних. На рис. 4.6 показана одна з можливих структур моделі. Безліч вхідних параметрів (як дискретних, так і безперервних) позначено як X , нечітко-логічні модулі (локальні моделі) - FLU (Fuzzy Local Unit), нечітко-ймовірнісні логічні модулі - як FPLU (Fuzzy-Probabilistic Local Unit). На рис. 4.6 вхідні параметри надходять на входи моделей тільки першого рівня, але це не є обов'язковою характеристикою моделі - вхідні змінні можуть надходити на входи моделей вищих рівнів і комбінуватися з виходами моделей нижніх рівнів.

Можлива структура правил запропонованої гібридної ієрархічної моделі:

$$\text{Rule } R_q: \text{ If } x \text{ is } A_q \text{ then } \underline{y} = y_{q1} \text{ with } Pr(y_{q1}|A_q) \text{ and } \underline{y} = y_{q2} \text{ with } Pr(y_{q2}|A_q) \text{ and } \dots \underline{y} = y_{qN} \text{ with } Pr(y_{qN}|A_q), \quad (4.6)$$

де $x = (x_1, x_2, \dots, x_m) \in X$ - вхідной вектор розмірності M ,

A_q - лінгвістичне значення антецедент визначається нечіткою функцією належності $\mu_q(x)$,

\underline{y} являє собою стохастичну змінну консеквента, рівну одному зі значень $y_{q1}, y_{q2}, \dots, y_{qN}$.

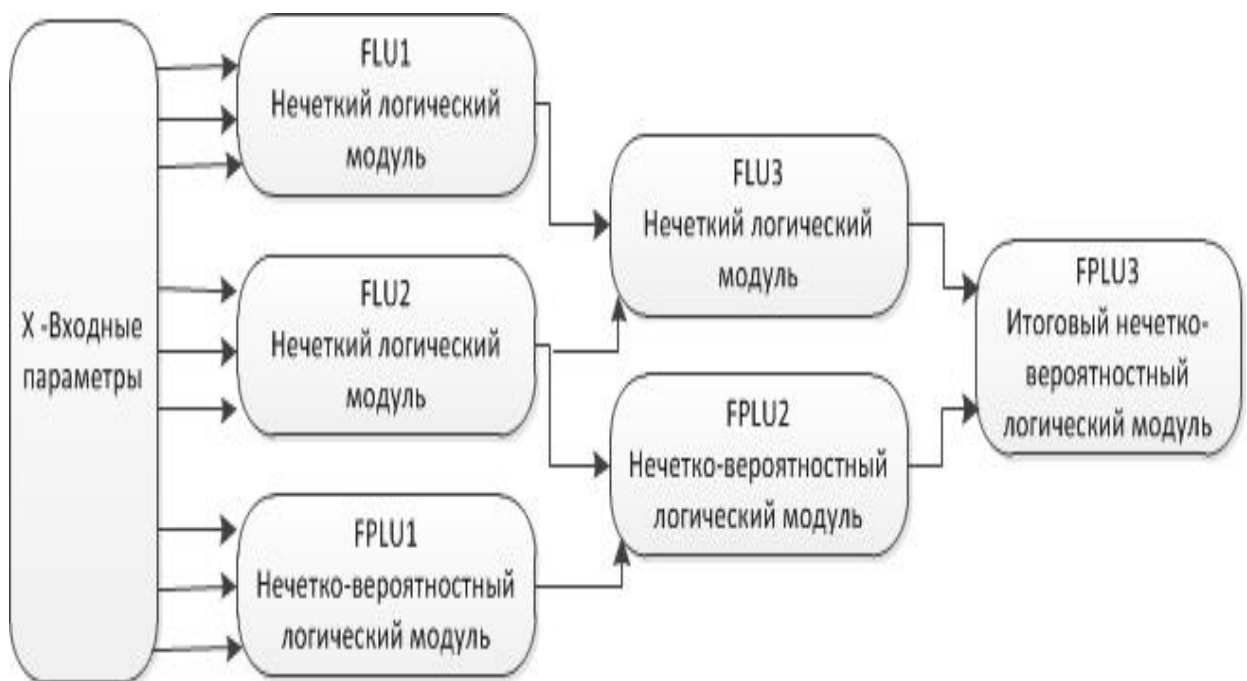


Рисунок 4.7 – Можлива структура запропонованої гібридної ієрархічної моделі

Вибір значення консеквента проводиться у відповідності з умовними ймовірностями $Pr(y_{q1}/A_q), \dots, Pr(y_{qN}/A_q)$. Кожне правило описує імовірність відображення нечіткого антецедента в множину чітких консеквента [9]:

$$Pr(y_j|x) = \sum_{q=1}^Q \varphi_q Pr(y_j | A_q) = \frac{\sum_{q=1}^Q Pr(A_q)\mu_q(x) Pr(y_j|A_q)}{\sum_{q=1}^Q Pr(A_q)\mu_q(x)} \quad (4.7)$$

$$y = E(\underline{y}|x) = \sum_{j=1}^N y_j Pr(y_j|x)$$

Таким чином, якщо ієрархічну систему нечіткого логічного висновування для п'яти вхідних змінних і одного підсумкового виходу можна представити як:

$$f = f_3(f_2(x_1, x_2), f_1(x_3, x_4), x_5) \quad (4.8)$$

Тоді запропонована нами модифікація моделі в загальному матиме вигляд:

$$fp = fp_3(fp_2(x_1, x_2), f_1(x_3, x_4), x_5) \quad (4.9)$$

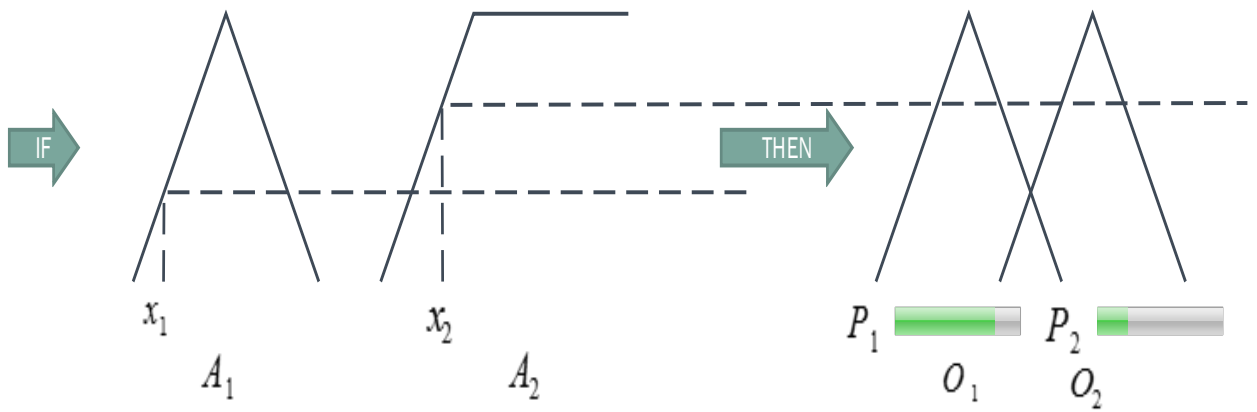


Рисунок 4.8 – Активізація висновувань в нечітко-імовірнісний логічному висновуванні Мамдані з використанням нечіткої диз'юнкції для обчислення значення антецедента

Запропонований метод складається з наступних етапів:

Етап 1. Попередня обробка вхідних значень, включаючи нормалізацію, якщо вона необхідна;

Етап 2. Аналіз моделей на повноту і несуперечність для надійшов набору даних. У разі виявлення проблем - модифікація правил для їх усунення;

Етап 3. Логічне висновування моделей 1 рівня. Моделі першого рівня є МІМО-моделі, правила яких мають наступну структуру:

$$\begin{aligned}
 R_j : & \text{if } x_1 \text{ is } A_{1j} \text{ and } x_2 \text{ is } A_{2j} \text{ and } \dots \text{ and } x_n \text{ is } A_{nj} \\
 & \text{Then } y = c_{1j} \text{ with } p_{1j} \\
 & \text{and } y = c_{2j} \text{ with } p_{2j} \\
 & \text{and } \dots \text{ and } y = c_{nj} \text{ with } p_{nj}, (j = 1, 2, \dots, N)
 \end{aligned}
 \tag{4.10}$$

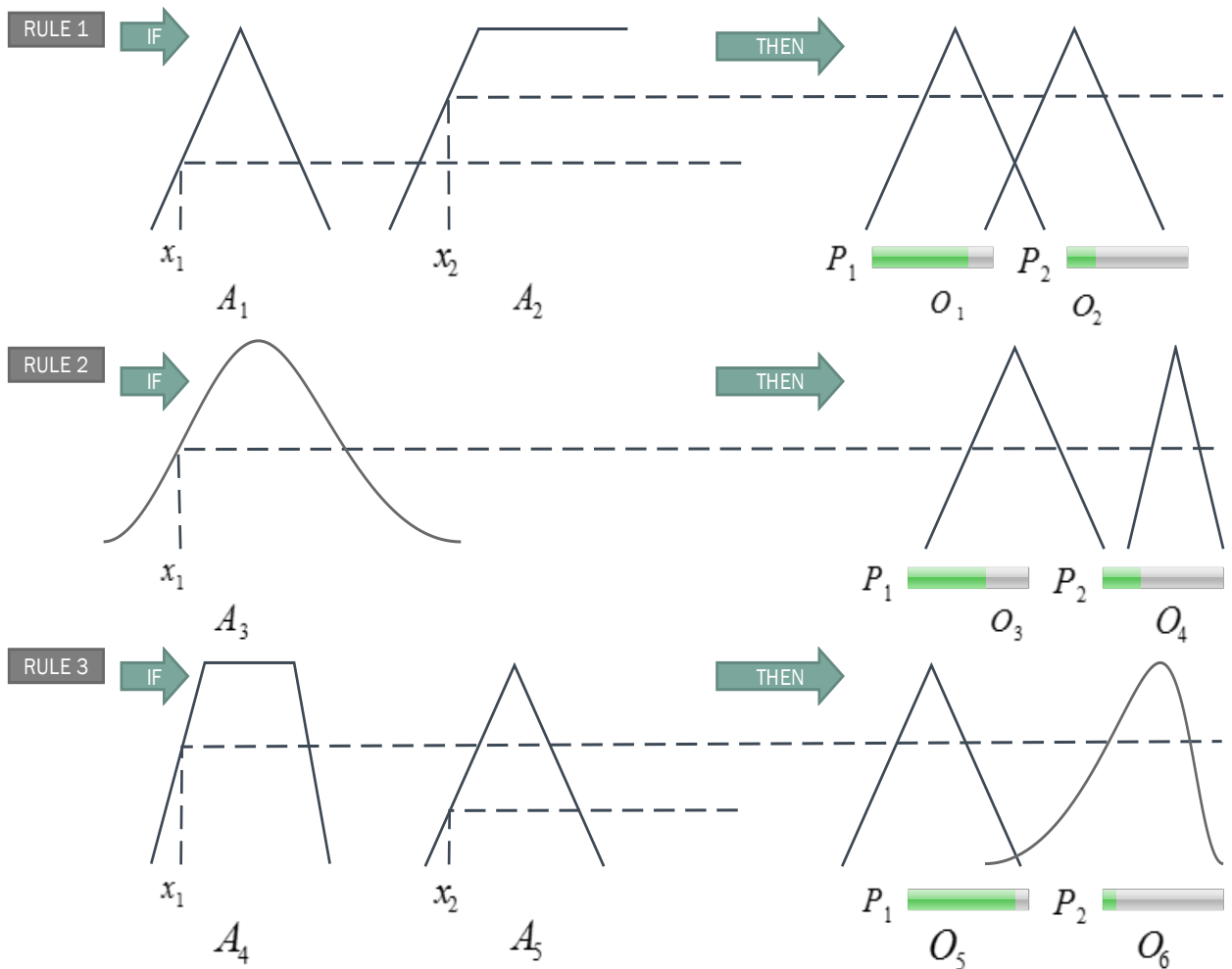


Рисунок 4.9 – Передача виходів локальних моделей першого рівня на локальні моделі другого рівня у вигляді вектора пар значень $\langle y_k, p_k \rangle$,

Підетапи даного етапу:

Фазифікації вхідних первинно оброблених на другому етапі змінних;

Агрегація підумови і активізація заключень за формулою:

$$w_i = T(\mu_1, \mu_2, \dots, \mu_n) \quad (4.11)$$

де T - t - норма нечіткої кон'юнкції;

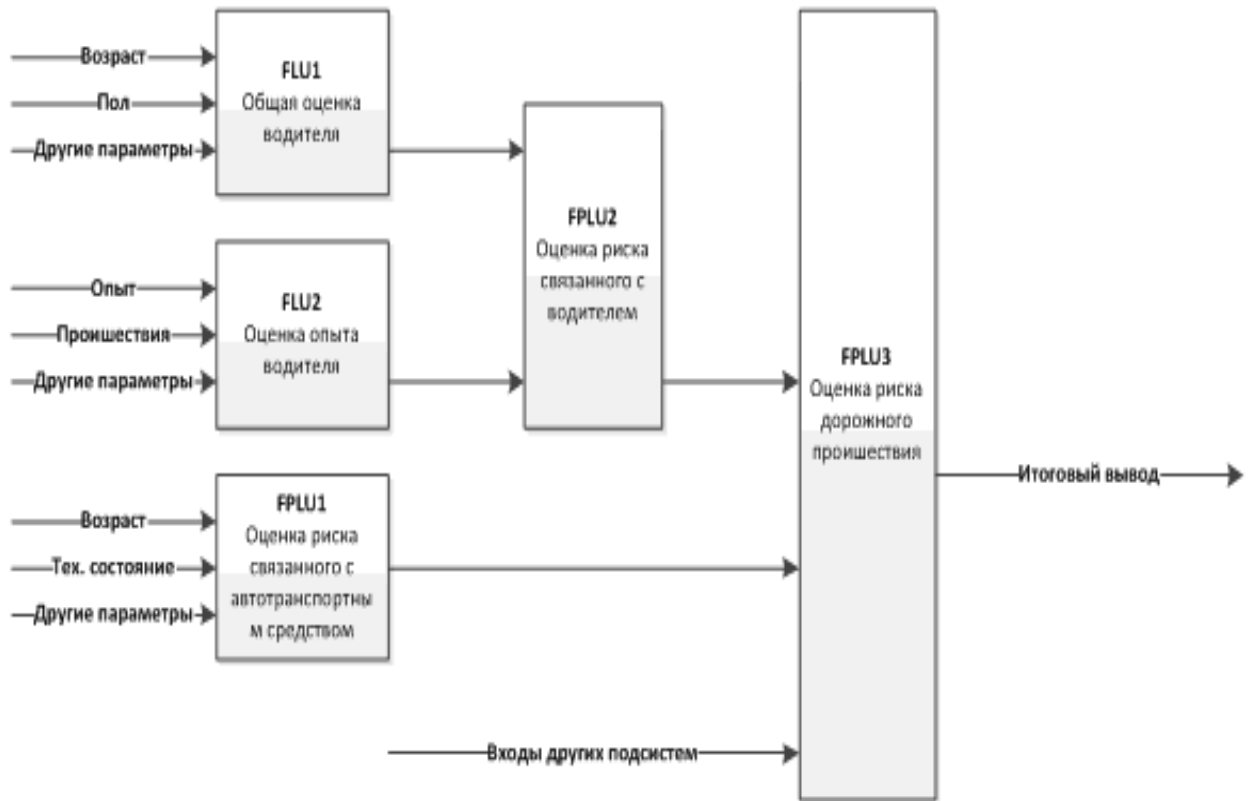


Рисунок 4.10 – Приклад моделі для конкретної прикладної області

Акумуляція для кожної пари виводу моделей першого рівня $\langle y_k, p_k \rangle$, де y_k - чітке значення змінної виводу та p_k - отримана імовірнісна оцінка даного значення. При цьому використовуються наступні формули:

$$y_k = \frac{\sum_{i=1}^n w_i z_i}{\sum_{i=1}^n w_i} \text{ when } p_i \text{ is } P_k \quad (4.12)$$

$$p_k = \frac{\sum_{i=1}^n w_i p_i}{\sum_{i=1}^n w_i} \text{ when } p_i \text{ is } P_k \quad (4.13)$$

де P_k - нечітка підмножина ймовірності.

Етап 4. Передача виходів локальних моделей першого рівня на локальні моделі другого рівня у вигляді пар значень $\langle y_k, p_k \rangle$.

Етап 5. Логічне висновування на моделях другого і наступних рівнів. Правила моделей верхніх рівнів відрізняються від правил локальних моделей першого рівня додаванням в антецедент імовірнісних нечітких множин.

$$\begin{aligned}
&R_j : \text{if } x_1 \text{ is } A_{1j} \text{ and } px_1 \text{ is } PA_{1j} \\
&\text{and } x_2 \text{ is } A_{2j} \text{ and } px_2 \text{ is } PA_{2j} \\
&\text{and...and } x_n \text{ is } A_{nj} \text{ and } px_n \text{ is } PA_{nj} \\
&\text{Then } y = c_{1j} \text{ with } p_{1j} \\
&\text{and } y = c_{2j} \text{ with } p_{2j} \\
&\text{and...and } y = c_{nj} \text{ with } p_{nj},
\end{aligned} \tag{4.14}$$

У розроблюваних нами моделях прикладних предметних областей для представлення експертних оцінок достатнім є наявність трьох нечітких підмножин - «мала ймовірність», «середня ймовірність», «висока ймовірність», для подання оцінок ймовірності, функція приналежності яких трапецієвидна.

Підетапи цього етапу:

- 1) фаззифікації вихідних змінних моделей попереднього рівня і їх ймовірностей;
- 2) Агрегація підумови з ймовірнісним компонентом і активізація заключень;
- 3) Акумуляція для кожної вихідної пари за тими ж правилами, що і для моделей першого рівня.

Етап 6. Передача виходів локальних моделей на входи локальних моделей наступного рівня або на входи зовнішніх систем.

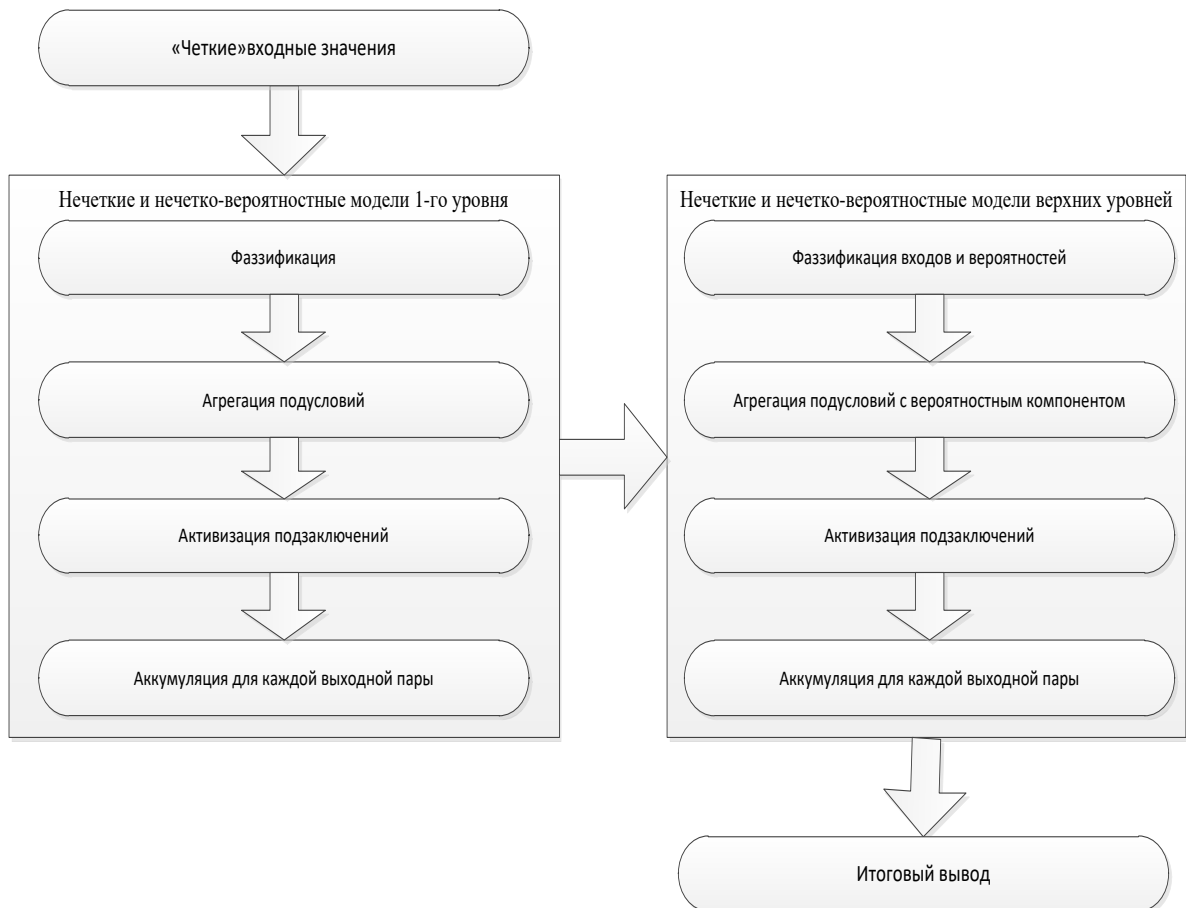


Рисунок 4.11 – Узагальнена схема запропонованої модифікації методу нечітко-імовірнісного логічного висновування

Приклад нечітко-імовірнісного правила з локальної моделі автоматизованого тестування (перший рівень) у системі аналізу ризиків розробки програмного забезпечення для керуванням фінансовим портфоліо:

$$\text{If } x \text{ is } A \text{ then } y = f_1(x) \text{ with } 0.9 \text{ and } y = f_2(x) \text{ with } 0.1 \quad (4.15)$$

де x – кількість тестів для модуля

y – кількість непокритих тестами ситуацій

$f_1(x)$ – лінійна залежність

$f_2(x)$ – експоненціальна залежність

Приклад такого ж правила без використання імовірнісного елемента визначається одне значення консеквенту для кожного правила:

$$\text{If } x \text{ is } A \text{ then } y = f_1(x), \quad (4.16)$$

де $f_1(x)$ – лінійна залежність.

Нечітко-імовірнісне правило дозволяє представити в базі знань менш ймовірний рівень залежності, а запропонована модифікація моделі і методу - коректно обробити це значення в вищих рівнях ієрархії.

Базова архітектура розробленої системи підтримки прийняття рішень заснована складається з поточного контексту прийняття рішень, бази даних, бази знань, механізму умовиводу та користувальницького інтерфейсу для представлення результатів. Загальна структура представлена на рис. 4.11.

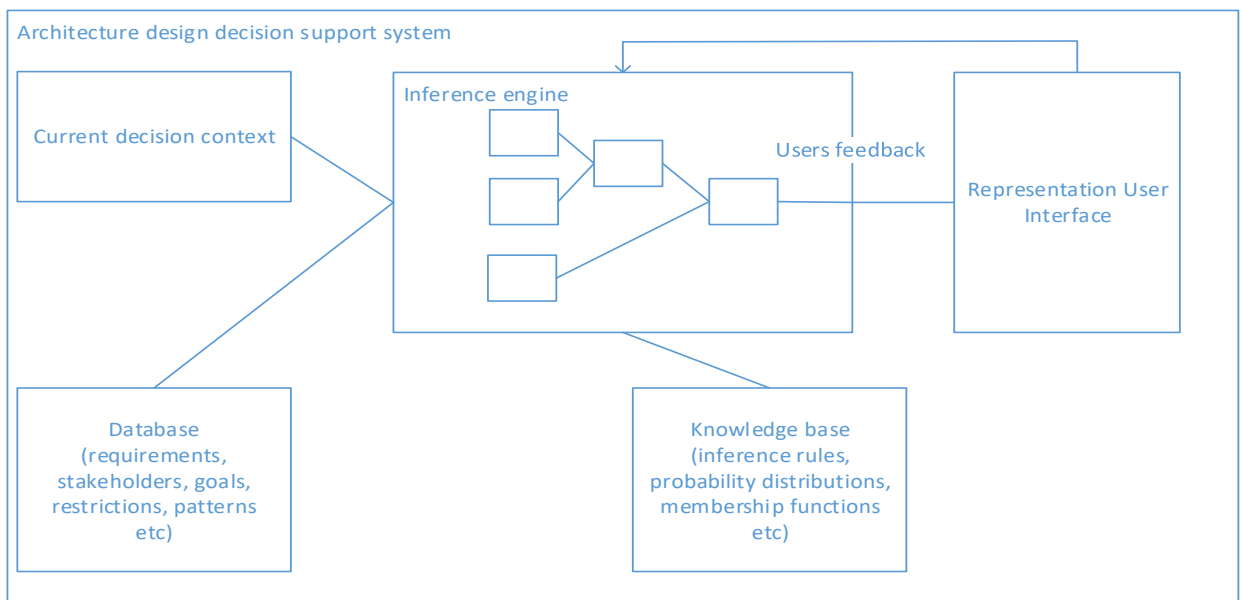


Рисунок 4.12 – загальна архітектура системи підтримки прийняття рішень

4.4 Висновки до розділу

1. Запропоновано метод навчання нейро-фаззі моделі, що використовує шумознижуючі властивості емпіричної модової декомпозиції.

2. Створено метод синтезу ансамблю нейро-фаззі моделей, що, за рахунок декомпозиційних властивостей емпіричної модової декомпозиції дозволяє спростити процедуру вибору параметрів ансамблю.

3. Розвинено та розширено метод побудови ієрархічних знання-орієнтованих нечітких систем шляхом включення оцінок ймовірностей у процес логічного висновування та використання нейро-фаззі моделей як компонентів ієрархічної структури, що на відміну від існуючих методів дозволяє опрацьовувати малоймовірні критичні значення у задачах оперативного аналізу ризиків.

Список використаних у цьому розділі джерел наведено у повному списку використаних джерел під номерами [105] [29].

Основні результати цього розділу опубліковано в роботах [106] [107] [108] [109] [110] [111].

5 ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ ТА РОЗВ'ЯЗАННЯ ПРАКТИЧНИХ ЗАДАЧ

Мета розділу – провести обчислювальні експерименти на основі синтетичних тестових та реальних даних . Розробити програмні модулі, що можуть бути застосованими у практичних впровадженнях. Зокрема реалізувати модуль, що підтвердив свою ефективність у задачах оперативного аналізу ризиків у управлінні фінансовим портфоліо.

Завдання:

- 1) провести імітаційне моделювання методів навчання нейро-фаззі моделей, порівняти їхню точність та швидкодію з розповсюдженими сучасними методами та моделями;
- 2) провести імітаційне моделювання методів навчання ансамблів нейро-фаззі моделей, порівняти їх практичні характеристики з одиничними моделями
- 3) провести аналіз результатів впровадження розроблених методів та моделей у задачах оперативного аналізу ризиків у управлінні фінансовим портфоліо

Програмна бібліотека машинного навчання Accord.NET [112] була використана для реалізації еталонних штучних нейронних мереж.

Для запропонованих нейро-фаззі моделей було створене спеціальне програмне забезпечення на базі імплементацій лінійних операцій у програмному пакеті математичних операцій Math.NET Numerics [113].

Експерименти проводились на комп'ютері з процесором Intel® Core(TM) Core i7-7700 (Intel, Santa Clara, California, U.S.) та 32 гігабайтами оперативної пам'яті.

5.1 Моделювання нейро-фаззі моделі з багатовимірними гаусіанами у консеквентному шарі

Правильне та неупереджене порівняння різних моделей прогнозування є складним завданням [31].

Запропонована модель та алгоритм навчання показали хорошу ефективність у реальних наборах даних фондового ринку – щоденні повернення акцій Cisco.

Таблиця 2

Приклад оригінальних на нормалізованих даних акцій Cisco.

| Cisco | Normalized |
|--------------|-------------------|
| -0.05382 | 0.50755 |
| 0.01827 | 0.63085 |
| -0.00909 | 0.58405 |
| 0.00909 | 0.61515 |
| -0.02752 | 0.55253 |
| -0.01878 | 0.56748 |
| 0.01878 | 0.63172 |
| 0.08895 | 0.75173 |
| -0.00855 | 0.58498 |
| 0.01702 | 0.62871 |

Ми порівняли продуктивність та точність прогнозування з ANN на основі функцій біполярної сигмоїдної активації та еластичного алгоритму навчання зворотного поширення (resilient backpropagation learning algorithm - RPROP) [114].

Для оцінки точності прогнозування використовувались критерії RMSE та MAPE:

$$RMSE = \sqrt{\frac{\sum_{k=1}^N (y(k) - \hat{y}(k))^2}{N}}$$

(5.1)

$$MAPE = \frac{100}{N} \sum_{k=1}^N \left| \frac{\hat{y}(k) - y(k)}{\hat{y}(k)} \right|$$

де $y(k)$ – значення зовнішнього навчального сигналу,

$\hat{y}(k)$ – прогнозоване значення

N – розмір вибірки для навчання.

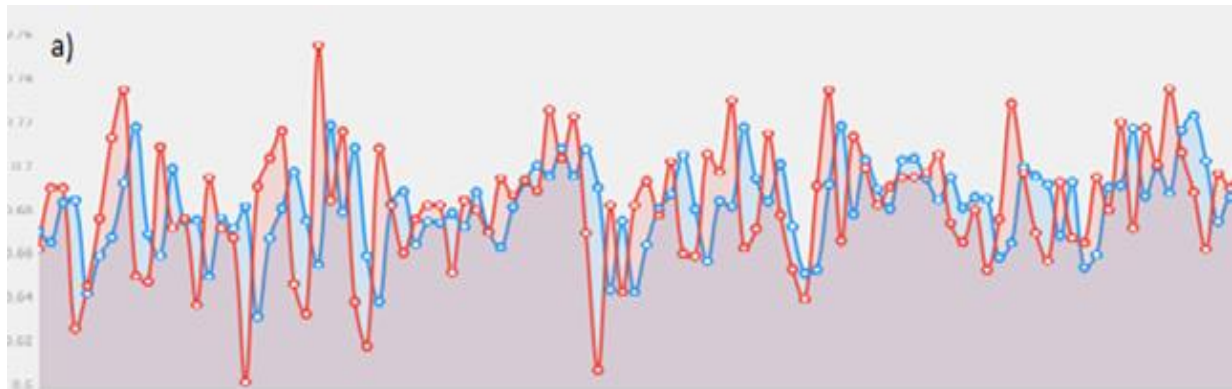


Рисунок 5.1 – Графік експериментів на даних щоденних повернень акцій Cisco запропонованої моделі - навчання

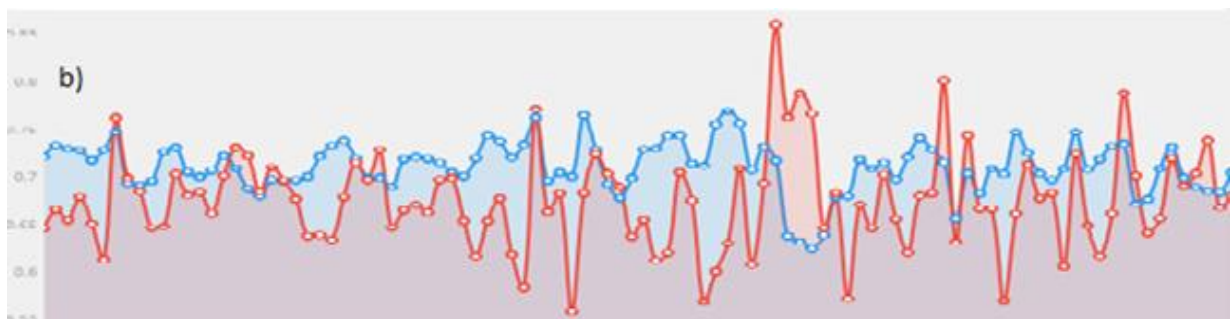


Рисунок 5.2 – Графік експериментів на даних щоденних повернень акцій Cisco запропонованої моделі - Верифікація

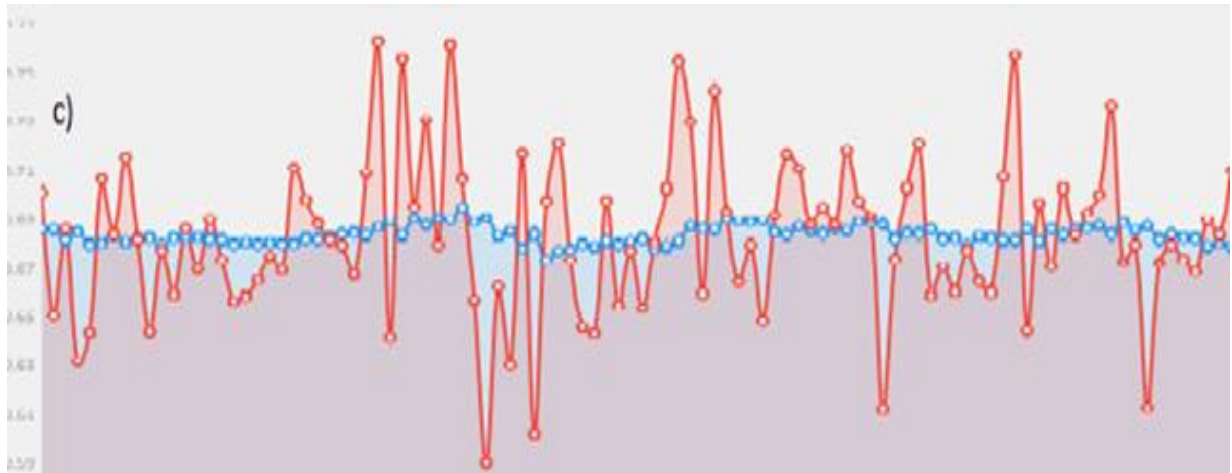


Рисунок 5.3 – Графік експериментів на даних щоденних повернень акцій Cisco, штучна нейронна мережа на біполярних сигмоїдах - навчання

Найкращі результати для представленої нейро-фаззі моделі були досягнуті з $h^{\phi} = 2$, $\lambda_c = 0.87$, $\beta_c = 1.01$, $\lambda_Q = 0.81$, $\beta_Q = 1.02$.

Біполярну сигмоподібну мережу навчали зі значенням альфа 0,4, коефіцієнтом навчання 0,67 та 100 епохами.

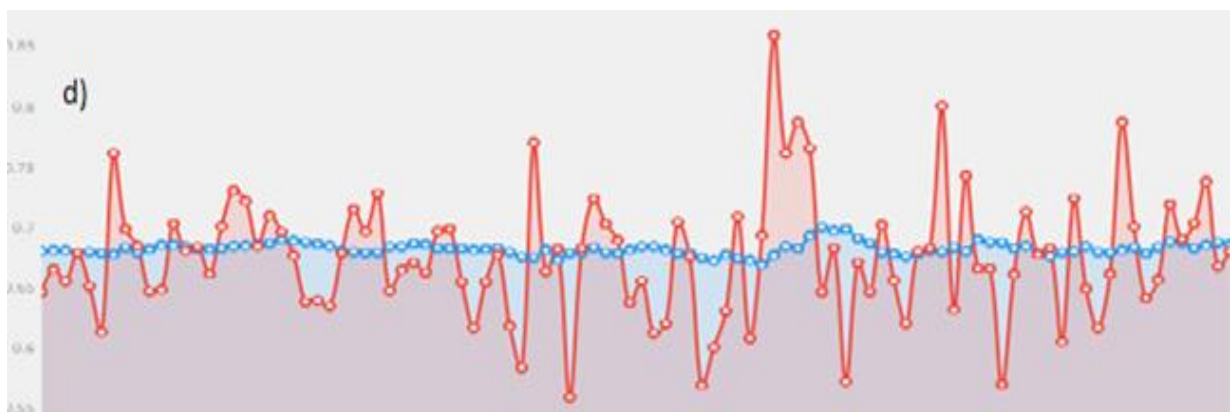


Рисунок 5.4 – Графік експериментів на даних щоденних повернень акцій Cisco штучна нейронна мережа на біполярних сигмоїдах – верифікація

Результати, наведені нижче у таблиці 3, доводять що запропонована модель має нижчу похибку та кращу швидкодію.

Таблиця 3

Результати експериментів

| Модель | Cisco daily log returns dataset results | | |
|--|---|---------|----------|
| | Час виконання, ms | RMSE, % | SMAPE, % |
| Запропонована модель | 442 | 3.237 | 3.53 |
| штучна нейронна мережа на біполярних сигмоїдах | 2668 | 3.31 | 3.6 |

5.2 Експериментальне моделювання навчання функцій належності першого шару

В якості еталону було обрано завдання прогнозування часових рядів Маккі-Гласса. Це стандартний тест для обчислювальних моделей. Точки часових рядів генеруються наступним нелінійним диференціальним рівнянням затримки часу [115]:

$$x(t) = \frac{0.2x(t - \tau)}{1 + x^{10}(t - \tau)} - 0.1x(t) \quad (5.2)$$

Це просте рівняння може породжувати складну динаміку, включаючи хаос.

Ми порівняли обчислювальні показники та точність прогнозування з ANN на основі функцій біполярної сигмоїдної активації та еластичного алгоритму навчання зворотного поширення.

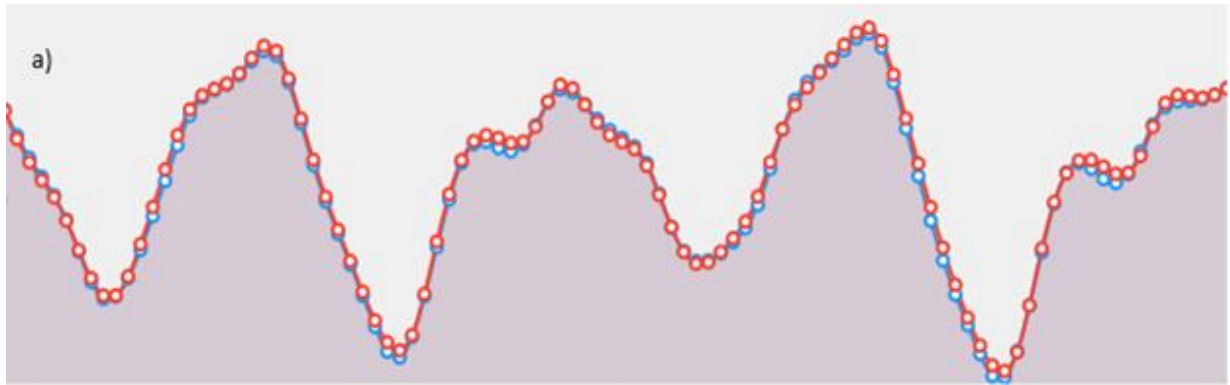


Рисунок 5.5 – Експерименти на даних часового ряду Маккі-Гласса - нейро-фаззі модель із навчанням першого рівня;

Ми порівняли точність прогнозування та обчислювальні показники з ANN за допомогою біполярних функцій активації сигмоїдної системи. Для навчання моделі ANN ми використовували алгоритм Левенберга-Марквардта [116] та еластичний алгоритм навчання зворотного розповсюдження як популярні методи пакетної оптимізації.

Найкращих результатів для нейро-фаззі моделі було досягнуто з $h^{\phi} = 2$, $h^{\psi} = 3$, $\lambda_c = 0.92$, $\beta_c = 0.945$, $\lambda_Q = 0.94$, $\beta_Q = 0.98$, $\lambda_{\psi} = 1$, $\beta_{\psi} = 0.98$.

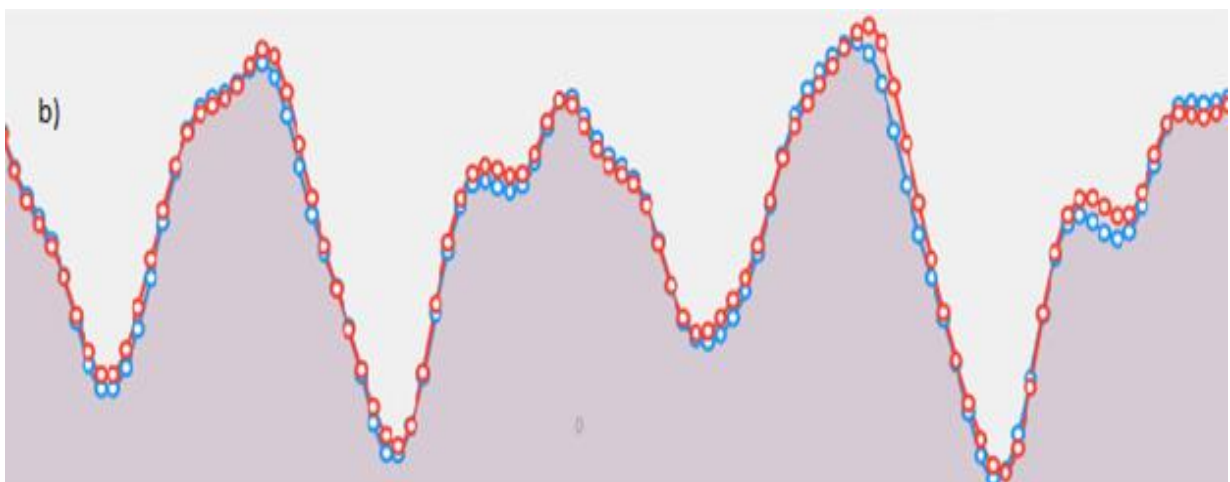


Рисунок 5.6 – Експерименти на даних часового ряду Маккі-Гласса - нейро-фаззі модель без навчанням першого рівня

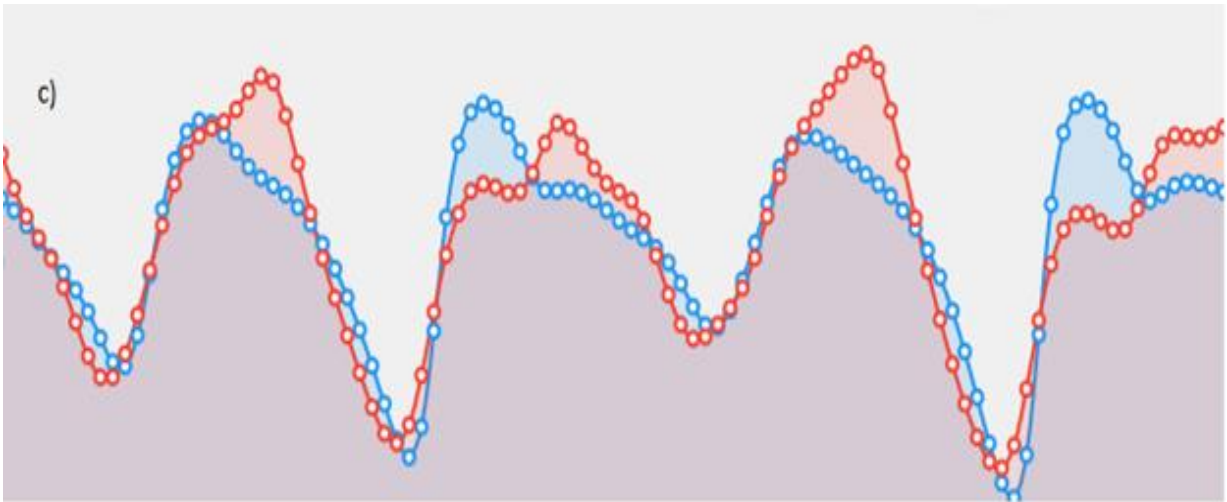


Рисунок 5.7 – Експерименти на даних часового ряду Маккі-Гласса – штучна нейронна мережа – 50 епох навчання

Експерименти показали, що біполярній сигмоподібній мережі потрібна величезна кількість епох, щоб досягти гарної точності, і вона стрімко падає в обчислювальній ефективності.

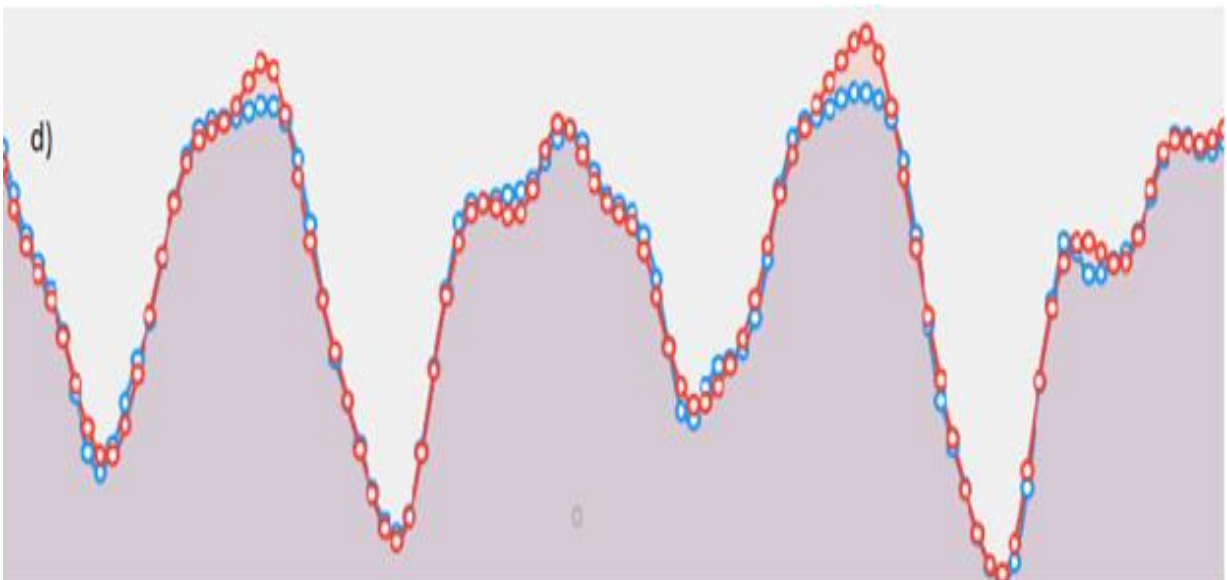


Рисунок 5.8 – Експерименти на даних часового ряду Маккі-Гласса – штучна нейронна мережа – 300 епох навчання

Нейро-фаззі мережа показала найкращу точність з налаштуванням першого шару. Втрачена обчислювальна продуктивність незначна.

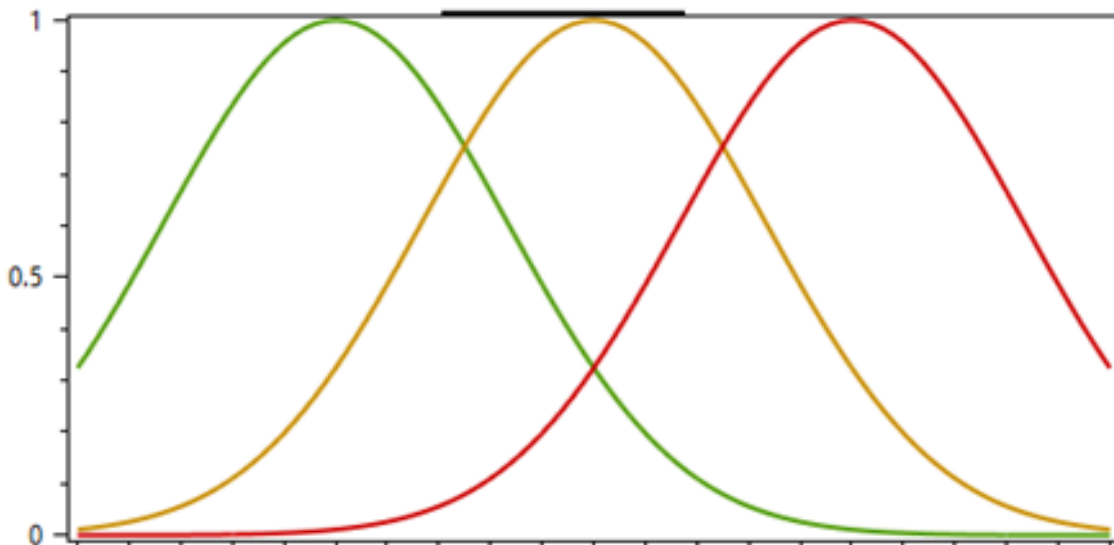


Рисунок 5.9 – Приклад налаштування функцій належності першого шару – ініційовані функції з центрами 0.25, 0.5, 0.75;

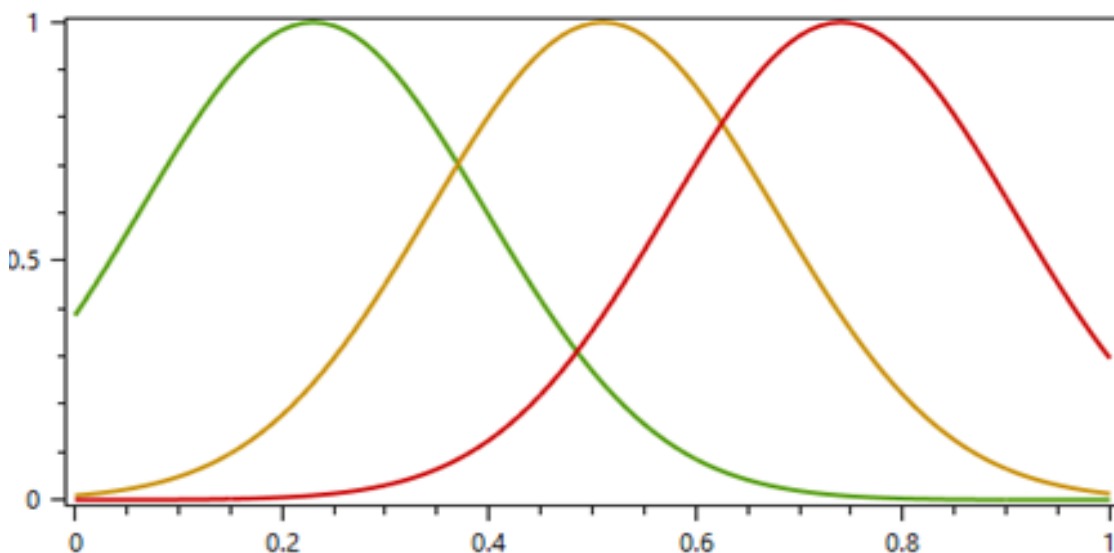


Рисунок 5.10 – Приклад налаштування функцій належності першого шару – налаштовані функції після завершення навчання з центрами 0.2392391, 0.5066654, 0.7482878.

Результати показані на рис.5.11 та таблиці 3.

Таблиця 4

Порівняння результатів прогнозування в умовах зашумлених даних

| Модель | Mackey-Glass time series | | |
|--|--------------------------|-----------------|-----------------|
| | Час, <i>ms</i> | <i>RMSE</i> , % | <i>MAPE</i> , % |
| Запропонована модель з навчанням першого шару $h^\phi = 4, \beta_c = 0.78, \beta_Q = 0.91$ | 2610 | 1.481 | 5.98 |
| Запропонована модель без навчання першого шару $h^\phi = 4, \beta_c = 0.75, \beta_Q = 0.91$ | 2373 | 3.01 | 7.64 |
| Bipolar Sigmoid Network RBPR 50 епох | 7097 | 6.315 | 11.78 |
| Bipolar Sigmoid Network RBPR 300 епох | 37442 | 2.72 | 7.33 |

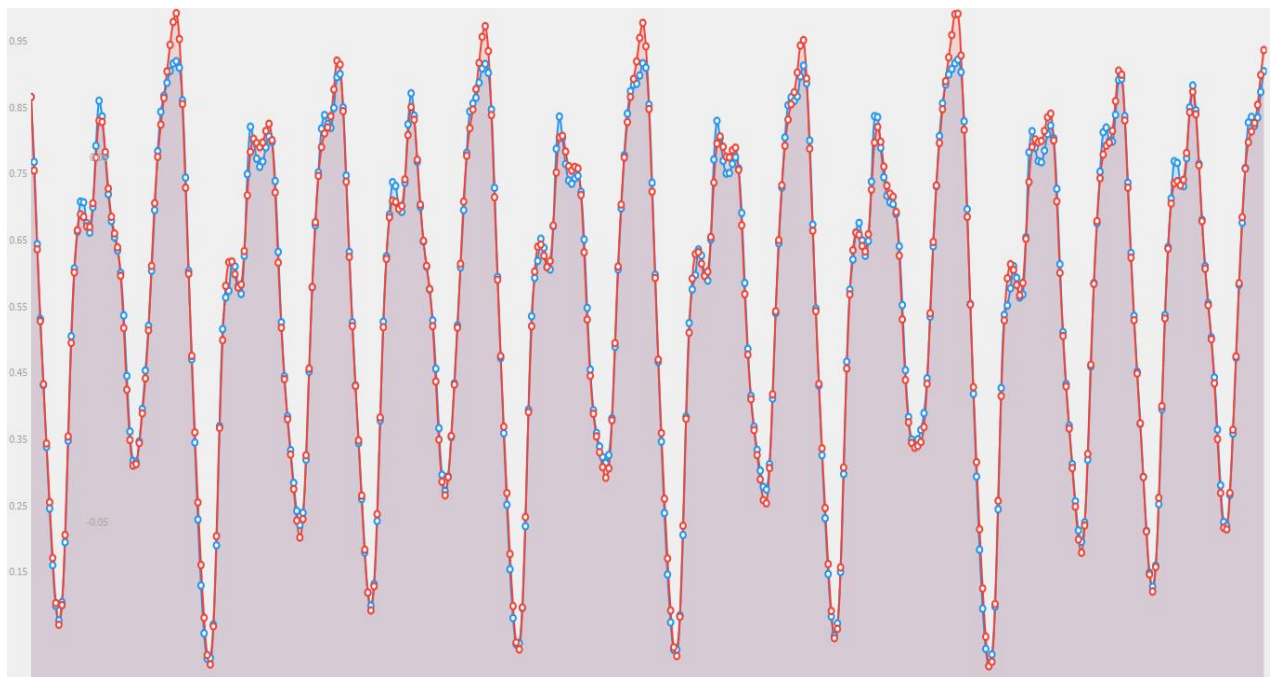


Рисунок 5.11 – Загальний навчальний графік Mackey-Glass time-series.

5.3 Експериментальне моделювання нейро-фаззі моделі з декількома вихідними змінними.

Для того, щоб перевірити ефективність нашої моделі, ми використовували історію доходності акцій ІВМ та індексу S&P 500 з січня 1926 по грудень 1999 року, в якості набору даних було 888 щомісячних записів. Відповідно, набір даних мав два стовпці з числовими значеннями: ІВМ представляє історію доходності акцій ІВМ, а стовпець SP містить значення індексу S&P 500. Ці два набори даних містять двовимірний одночасно корельований часовий ряд [61]. Ми вибрали цей набір даних як основу завдяки його добре вивченим статистичним властивостям та реальному характеру. Оригінальний набір даних можна знайти в [117] і таблиця 1 містить випадково вибраний блок із 10 рядків.

Таблиця 5

Приклад записів запасів ІВМ та набору даних індексу S&P 500

| ІВМ | SP |
|------------|-----------|
| -8.24798 | -4.06866 |
| 6.67236 | -4.81089 |
| 4.73701 | 7.43094 |
| 0.70948 | 2.61842 |
| 3.81336 | 7.52652 |
| 0.00000 | 0.29955 |
| -4.91377 | -5.46357 |
| 8.11467 | 0.75415 |
| -4.96210 | -3.05211 |
| 8.11467 | 0.75415 |

Дані були нормалізовані, щоб правильно порівняти різні моделі. На рисунку 1 показано перші 300 записів у нормалізованій версії набору даних.

$$SMAPE = \frac{2}{N} \sum_{k=1}^N \frac{|\hat{y}(k) - y(k)|}{\hat{y}(k) + y(k)} \quad (5.3)$$

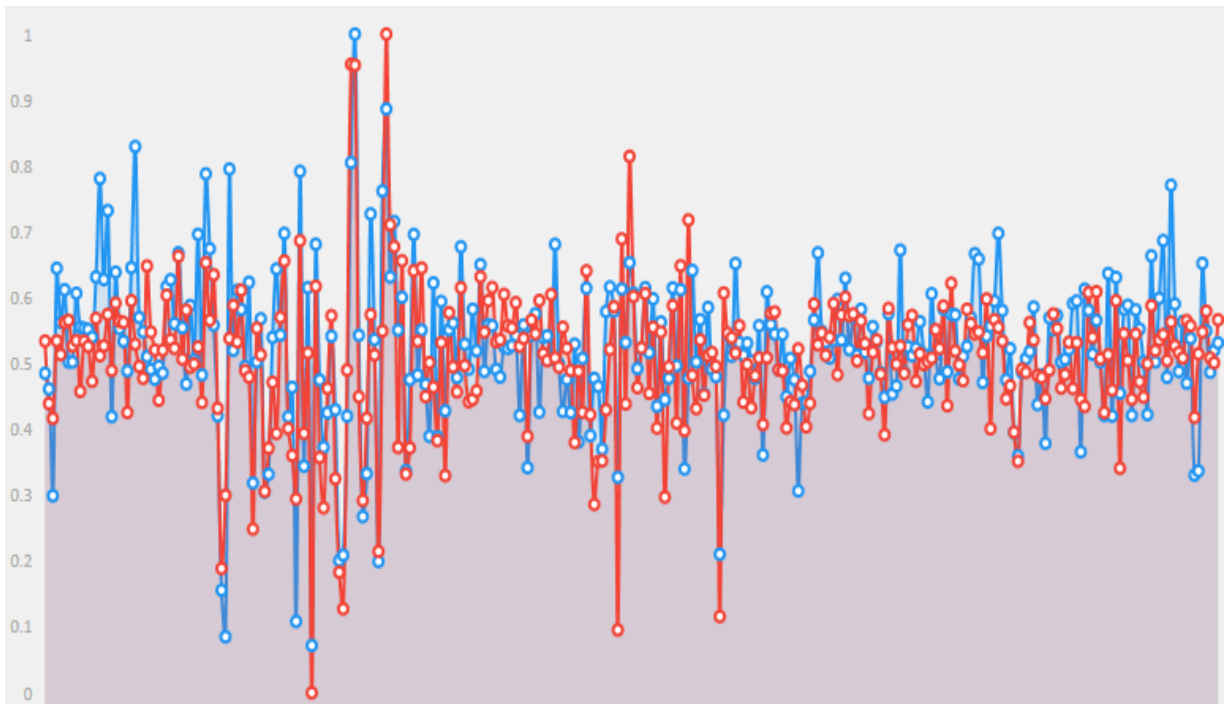


Рисунок 5.12 – Перші 300 записів нормалізованих значень доходності акцій IBM та індексу S&P 500

Біполярні сигмоподібні мережі були побудовані з сигмоїдів з альфа-значенням 0,4 для всіх випадків, швидкістю навчання 0,67 і 50 епох використовували для пружного зворотного розмноження. Реалізація Левенберга-Марквардта була навчена 10 епох. Нейро-фаззі модель мала такі параметри $h^\psi = 3$, $\lambda_c = 1$, та $\lambda_Q = 1$.

Оригінальний набір даних був розділений на навчальний набір із 660 записів та 228 записів у наборі перевірки. Вхідні вектори для двовимірного

випадку склалися з однакової кількості послідовних чисел від обох компонентів набору даних.

Ми також порівняли характеристики всіх моделей в одновимірному корпусі, що відображено в таблиці 5.

Таблиця 6

Результати експериментів

| Модель | IBM Stock Daily Log Returns and S&P 500 Index Dataset Results | | | | |
|--|---|--------------------|---------------------|------------------------|-------------------------|
| | Час виконання (ms) | IBM stock RMSE (%) | IBM stock SMAPE (%) | S&P 500 index RMSE (%) | S&P 500 index SMAPE (%) |
| Proposed model $h^\phi = 1, \beta_c = 0.89, \beta_Q = 0.98$ | 64 | 8.938 | 10.049 | 4.443 | 4.594 |
| $h^\phi = 1, \beta_c = 1.19, \beta_Q = 0.97$ | 39 | 12.956 | 20.548 | - | - |
| $h^\phi = 1, \beta_c = 0.88, \beta_Q = 0.93$ | 38 | - | - | 6.097 | 9.005 |
| Proposed model $h^\phi = 2, \beta_c = 0.85, \beta_Q = 0.93$ | 120 | 8.878 | 10.104 | 4.356 | 4.532 |

| | | | | | |
|--|-----|--------|--------|-------|-------|
| $h^\phi = 2, \beta_c = 0.82, \beta_Q = 0.91$ | 63 | 12.865 | 20.138 | - | - |
| $h^\phi = 2, \beta_c = 0.75, \beta_Q = 0.92$ | 65 | - | - | 6.056 | 8.677 |
| Proposed model $h^\phi = 4, \beta_c = 0.78, \beta_Q = 0.91$ | 257 | 8.866 | 10.013 | 4.294 | 4.463 |
| $h^\phi = 4, \beta_c = 0.75, \beta_Q = 0.91$ | 126 | 12.693 | 20.133 | - | - |
| $h^\phi = 4, \beta_c = 0.74, \beta_Q = 0.93$ | 110 | - | - | 6.113 | 8.775 |

Проведені експерименти показали, що ANN зі стійким зворотним розповсюдженням потребує значної кількості епох для досягнення хорошої точності, значно знижуючи обчислювальні показники.

Навчання Левенберга-Марквардта вимагає менше епох, але кожна епоха коштує дорожче з точки зору обчислень.

Запропонована модель демонструвала кращі результати роботи у всіх випадках - і обчислювальний час, і точність були значно кращими, ніж у конкурентів. Крім того, результати в таблиці 2 показують, що точність покращується із більшим значенням h^ψ , але мінусом є довший час виконання.

Багатовимірна нейро-фаззі модель вимагає менше обчислювальних ресурсів, ніж дві незалежні одновимірні моделі, і має значно кращу точність.

Таблиця 7

Результати експериментів

| Модель | IBM Stock Daily Log Returns and S&P 500 Index Dataset Results | | | | |
|--|---|--------------------|---------------------|------------------------|-------------------------|
| | Час виконання (ms) | IBM stock RMSE (%) | IBM stock SMAPE (%) | S&P 500 index RMSE (%) | S&P 500 index SMAPE (%) |
| Bipolar Sigmoid Network RBPR | 847 | 8.91 | 10.156 | 4.66 | 4.96 |
| | 514 | 13.162 | 20.876 | | |
| | 545 | - | - | 6.308 | 9.139 |
| Bipolar Sigmoid Network Levenberg-Marquart | 302 | 8.940 | 10.07 | 4.542 | 4.755 |
| | 313 | 14.544 | 22.664 | - | - |
| | 297 | - | - | 7.025 | 11.019 |

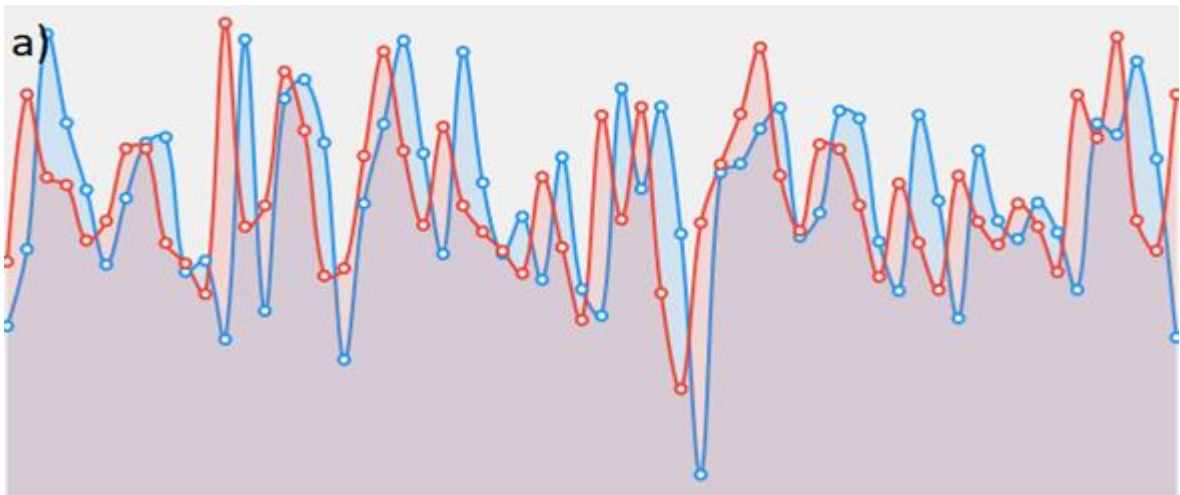


Рисунок 5.13 – Процес навчання нейро-фаззі моделі з двома виходами;

Візуалізація навчального процесу представлена на малюнках 9 і 10; Повернення акцій IBM зображено для багатовимірних та одновимірних випадків.

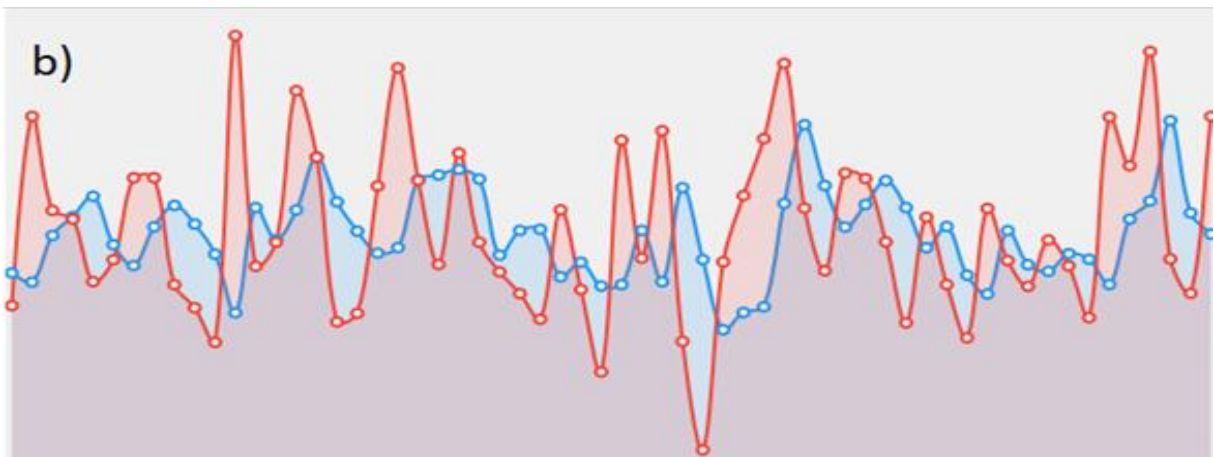


Рисунок 5.14 – Процес навчання нейро-фаззи моделі з одним виходом;

На рисунку 5.15 показано суттєве покращення в графіку прогнозування, отриманому з використанням багатовимірної нейро-фаззи моделі. І навпаки, на малюнку 5.16 не показано такого ступеня поліпшення для біполярної сигмовидної штучної нейронної моделі.

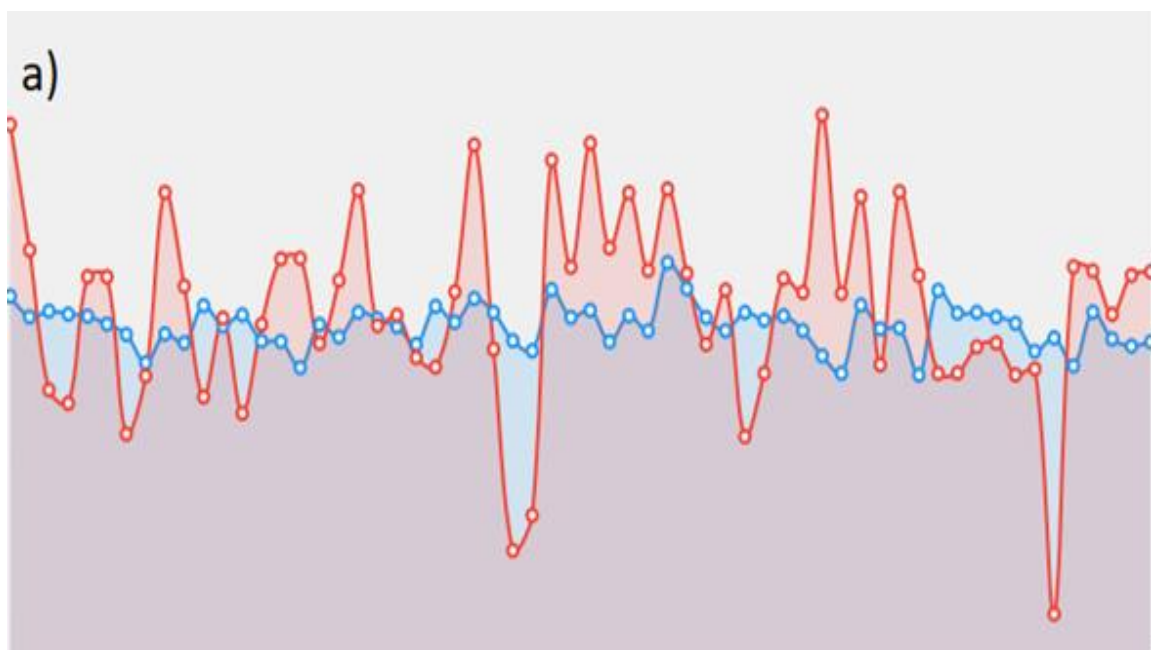


Рисунок 5.15 – Процес навчання штучної нейронної мережі моделі з двома виходами;

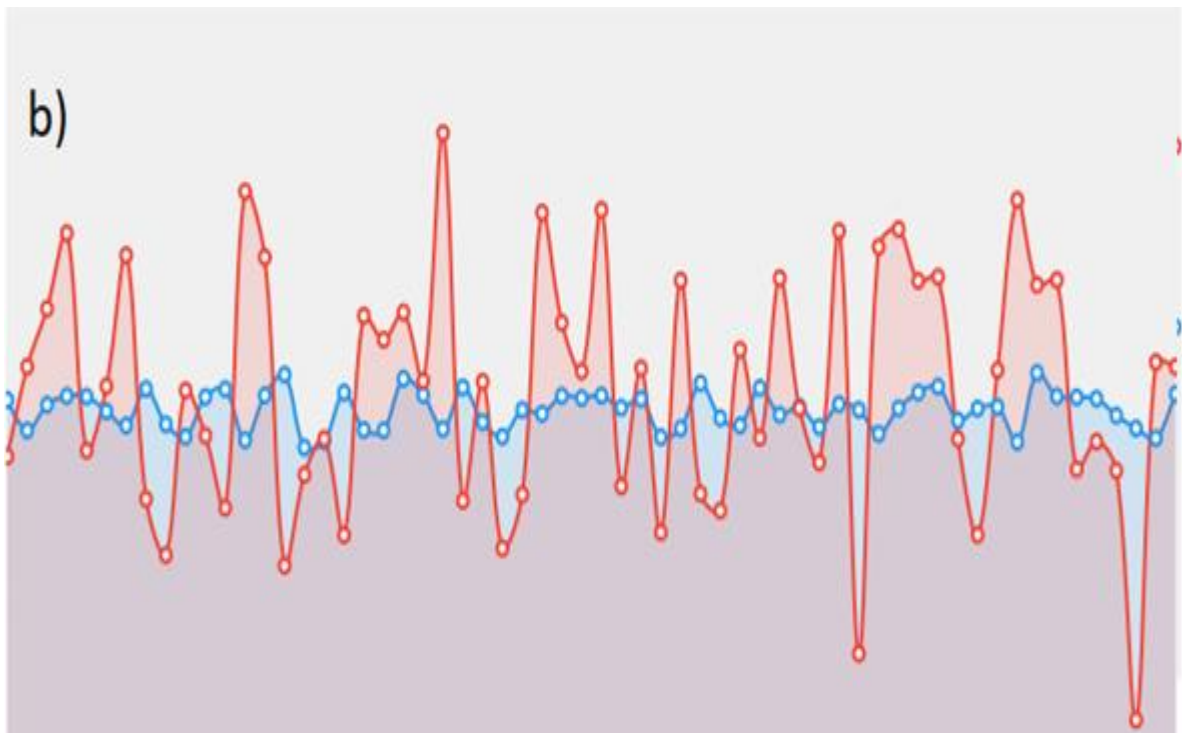


Рисунок 5.16 – Процес навчання штучної нейронної мережі моделі з одним виходом

5.4 Експериментальне моделювання ансамблів моделей

Для перевірки точності нашої моделі та обчислювальних характеристик ми використовували наступні набори даних:

Акції Cisco з 6329 записами.

Акції Alcoa з 2528 записами.

Акції American Express з 2528 записами.

Фонд Діснея з 2528 записами.

Фонд Діснея з 2528 записами.

Набір даних має один стовпець із числовими значеннями. Ми вибрали цей набір даних як базовий каркас завдяки його добре вивченим властивостям та реальному характеру. Оригінальні набори даних можна знайти в Tsay [4].

Кожен набір даних ми розділили на набір валідації із 800 записів та навчальний набір з рештою.

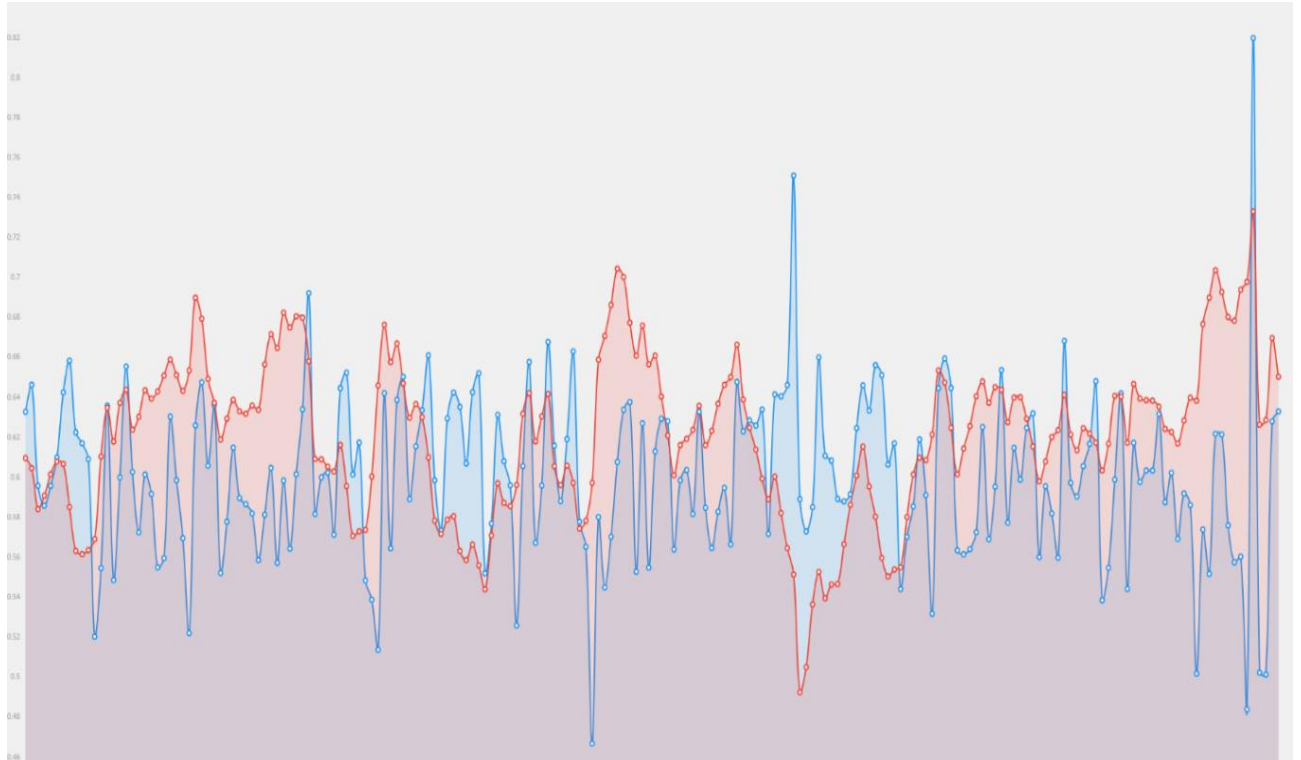


Рисунок 5.17 – Графік прогнозування ансамблевої моделі

На додаток до запропонованої моделі, ми провели тести з одиничними нейро-фаззі моделями та конкуруючими моделями:

Біполярна сигмоподібна нейронна мережа.

Машини опорних векторів з послідовною мінімальною оптимізацією.

Обмежені машини Больцмана як приклад стохастичної нейронної мережі. Ми також використовували еластичний алгоритм навчання зворотного розповсюдження для їх навчання.

Візуалізація навчального процесу представлена на малюнках 8 і 9. На малюнку 8 показано покращення в графіку прогнозування, отриманому шляхом усереднення.

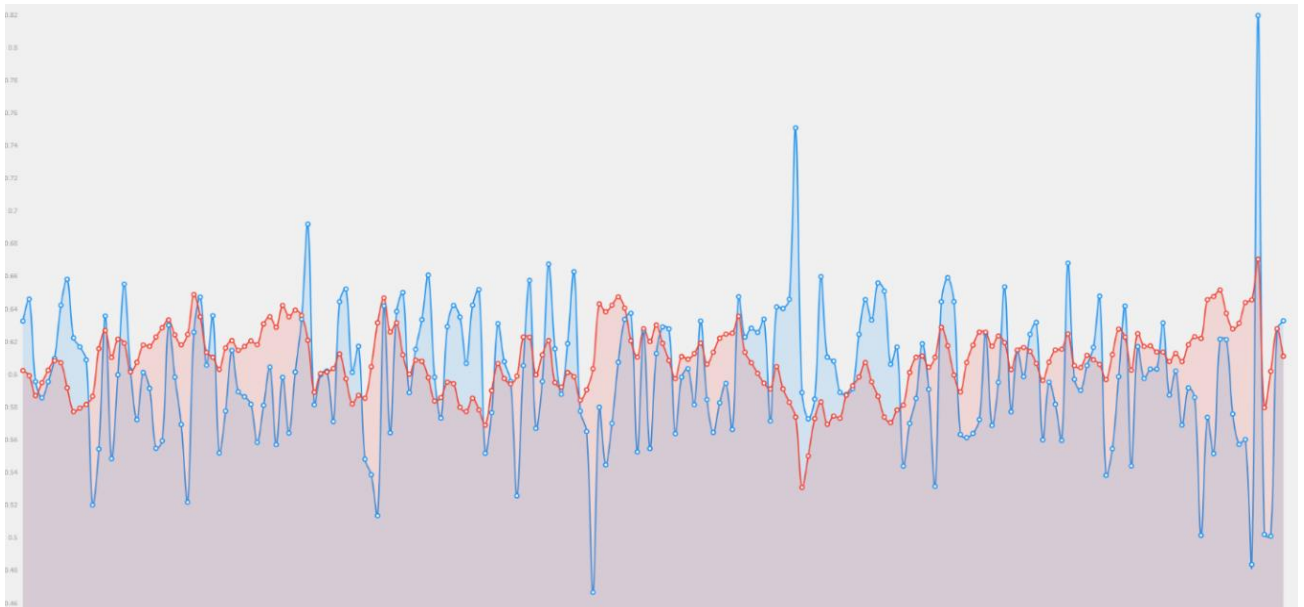


Рисунок 5.18 – Графік прогнозування ансамблевої моделі

Запропонована модель, як видно з таблицях 8-11, дозволяє досягти кращої точності завдяки невеликій кількості моделей-членів та при допустимих витратах на обчислювальні ресурси.

Таблиця 8

Результати експериментів

| Модель | Cisco Stock Daily Log Returns | | |
|---|-------------------------------|----------|-----------|
| | Час виконання (ms) | RMSE (%) | SMAPE (%) |
| Proposed model $m = 4, h^\phi = 2$ | 196 | 4.007 | 5.08246 |
| $m = 6, h^\phi = 2$ | 279 | 4.007 | 5.09458 |
| $m = 8, h^\phi = 2$ | 361 | 4.104 | 5.16344 |
| $m = 12, h^\phi = 2$ | 417 | 4.058 | 5.14296 |
| Single model $h^\phi = 2, \beta_c = 0.65, \beta_Q = 0.91$ | 92 | 4.012 | 5.14493 |
| Bipolar Sigmoid Network Resilient BackProp | 307 | 4.022 | 5.14132 |
| Bipolar Sigmoid Network Levenberg-Marquart | 1377 | 4.054 | 5.16275 |
| Support Vector Machine | 10800 | 4.021 | 5.13844 |
| Restricted Boltzmann Machine | 210 | 4.018 | 5.14531 |

Таблиця 9

Результати експериментів

| Модель | Alcoa Stock Daily Log Returns | | |
|---|-------------------------------|----------|-----------|
| | Час виконання (ms) | RMSE (%) | SMAPE (%) |
| Proposed model $m = 4, h^\phi = 2$ | 60 | 9.017 | 15.299 |
| $m = 6, h^\phi = 2$ | 77 | 9.032 | 15.385 |
| $m = 8, h^\phi = 2$ | 102 | 9.452 | 16.025 |
| $m = 12, h^\phi = 2$ | 145 | 9.578 | 16.245 |
| Single model $h^\phi = 2, \beta_c = 0.65, \beta_Q = 0.91$ | 27 | 9.111 | 15.377 |
| Bipolar Sigmoid Network Resilient BackProp | 192 | 9.894 | 16.667 |
| Bipolar Sigmoid Network Levenberg-Marquart | 472 | 9.896 | 16.711 |
| Support Vector Machine | 1136 | 9.910 | 16.634 |
| Restricted Boltzmann Machine | 72 | 9.889 | 16.663 |

Як видно з результатів, найкращі результати досягнуто при невеликій кількості одиничних моделей.

Таблиця 10

Результати експериментів

| Модель | American Express Stock Daily Log Returns | | |
|---|--|----------|-----------|
| | Час виконання (ms) | RMSE (%) | SMAPE (%) |
| Proposed model $m = 4, h^\phi = 2$ | 102 | 10.028 | 16.880 |
| $m = 6, h^\phi = 2$ | 154 | 10.001 | 16.998 |
| $m = 8, h^\phi = 2$ | 189 | 10.044 | 17.022 |
| $m = 12, h^\phi = 2$ | 211 | 10.045 | 17.025 |
| Single model $h^\phi = 2, \beta_c = 0.65, \beta_Q = 0.91$ | 58 | 10.022 | 16.917 |

| | | | |
|--|------|--------|--------|
| Bipolar Sigmoid Network Resilient BackProp | 188 | 10.06 | 17.043 |
| Bipolar Sigmoid Network Levenberg-Marquart | 458 | 0.999 | 16.885 |
| Support Vector Machine | 1584 | 10.141 | 17.858 |
| Restricted Boltzmann Machine | 87 | 10.054 | 17.038 |

Таблиця 11

Результати експериментів

| Модель | Disney Stock Daily Log Returns | | |
|---|--------------------------------|----------|-----------|
| | Час виконання (ms) | RMSE (%) | SMAPE (%) |
| Proposed model $m = 4, h^\phi = 2$ | 41 | 8.90 | 12.908 |
| $m = 6, h^\phi = 2$ | 52 | 8.902 | 12.952 |
| $m = 8, h^\phi = 2$ | 102 | 8.997 | 12.984 |
| $m = 12, h^\phi = 2$ | 145 | 9.009 | 12.989 |
| Single model $h^\phi = 2, \beta_c = 0.65, \beta_Q = 0.91$ | 27 | 8.960 | 12.983 |
| Bipolar Sigmoid Network Resilient BackProp | 166 | 8.991 | 12.953 |
| Bipolar Sigmoid Network Levenberg-Marquart | 468 | 9.024 | 12.985 |
| Support Vector Machine | 825 | 8.997 | 12.949 |
| Restricted Boltzmann Machine | 65 | 8.989 | 12.939 |

5.4 Експериментальні результати застосування емпіричної модової декомпозиції у навчанні нейро-фаззі моделей

Запропонована гібридна модель показала покращення продуктивності порівняно з оригінальною одиничною нейро-фіззі моделлю в завданні прогнозування фондового ринку на щоденних Alcoa та American Express із 2528 записами. Графік навчального процесу представлений на малюнку 5.

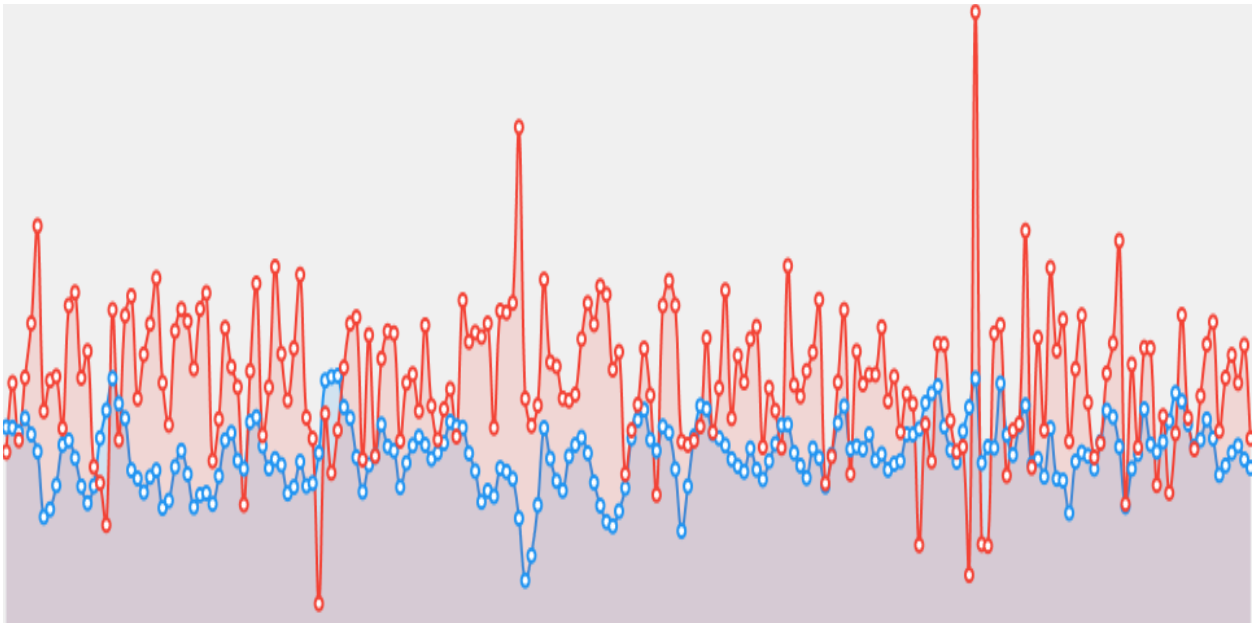


Рисунок 5.19 – Візуалізація процесу навчання на наборі даних Alcoa та American Express

Оптимальні результати були отримані з наступними параметрами $h^{\phi} = 1$, $\lambda_c = 0.85$, $\beta_c = 0.9$, $\lambda_Q = 0.85$, $\beta_Q = 0.95$. Навчання конкуруючої біполярної сигмоїдної моделі проводилось із значенням альфа 0,37, швидкістю навчання 0,6 та 100 епох.

Результати, представлені в таблиці нижче, показують, що введена модель має кращу точність прогнозування, ніж оригінальна нейро-фаззі модель, але вимагає більше ресурсів.

Таблиця 12

| Модель | Назва | | |
|------------------------------------|-----------------------------------|---------|----------|
| | Daily log returns dataset results | | |
| | Час виконання, ms | RMSE, % | SMAPE, % |
| Proposed model | 889 | 2.98 | 3.13 |
| Simple MDG based neuro-fuzzy model | 450 | 3.17 | 3.45 |

| Модель | Daily log returns dataset results | | |
|-------------------------|-----------------------------------|---------|----------|
| | Час виконання, ms | RMSE, % | SMAPE, % |
| Bipolar Sigmoid Network | 2512 | 3.31 | 3.65 |

Ефективність запропонованої моделі ансамблю перевірено даних IBM та Cisco з 2528 записами.

Програмні експерименти на реальних часових рядах показали хорошу точність прогнозування запропонованої моделі, яку порівнювали зі штучними нейронними мережами, одиночною нейро-фаззі моделлю та нейро-фаззі ансамблем. Поєднання шумозаглушення, виконуваного EMD, та ансамблевої структури з варіацією у виборі оригінальних гіперпараметрів призводить до швидшого зменшення помилок при навчанні за допомогою стохастичного градієнтного спуску без шкоди для його узагальнюючих здібностей.

Єдиним недоліком є дещо більші обчислювальні витрати в порівнянні з простою нейро-фаззі моделлю та нейро-фаззі ансамблем, але обчислювальна продуктивність все ще залишається дуже конкурентоспроможною, і невелика її втрата може бути виправдана перевагами, які надає модель.

Застосування EMD значно спрощує синтез ансамблю та вибір гіперпараметрів, а також покращує точність прогнозування, забезпечуючи нейро-фаззі модель сигналами для навчання.

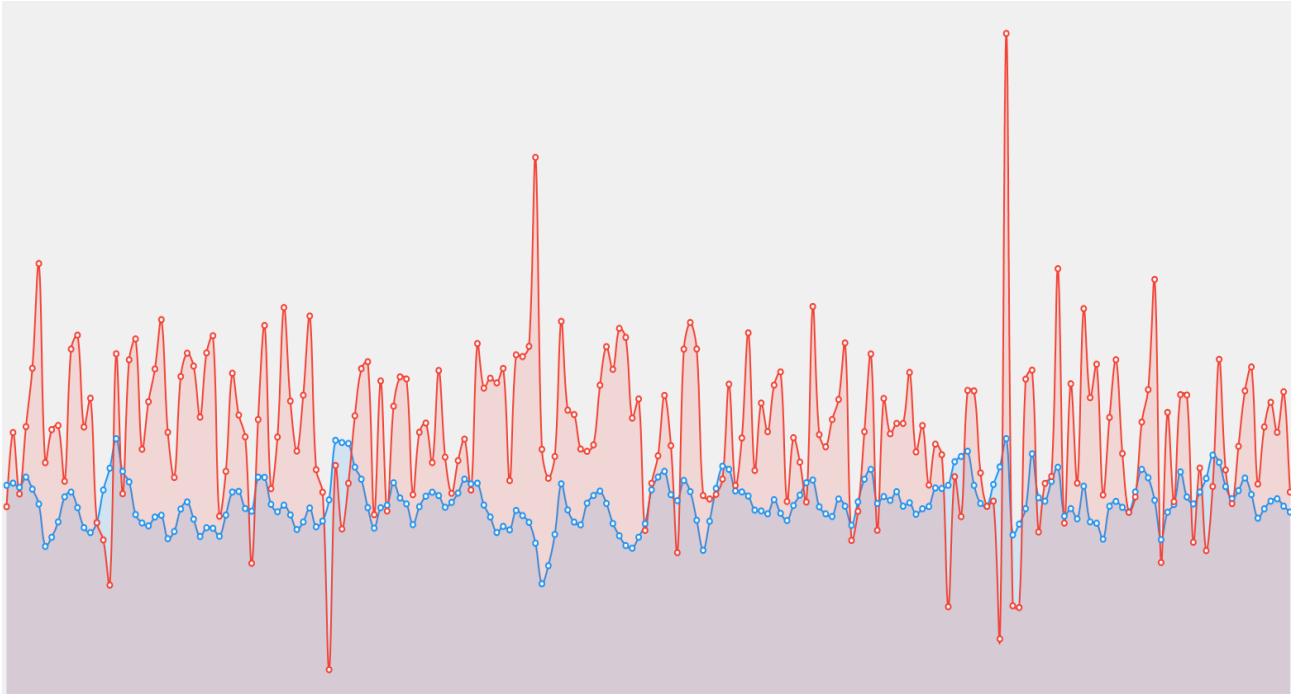


Рисунок 5.20 – Візуалізація процесу навчання на наборі даних Cisco

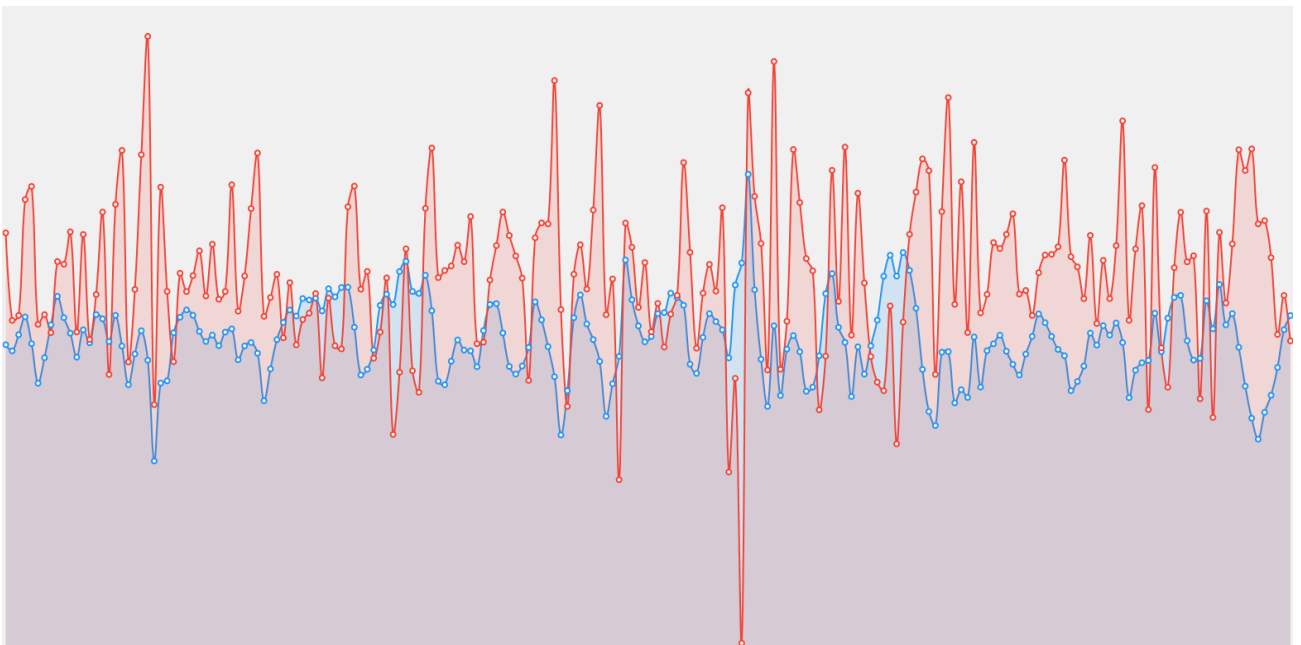


Рисунок 5.21 – Візуалізація процесу навчання на наборі даних IBM.

Навчання конкуруючої біполярної сигмоїдної моделі проводилось із значенням альфа 0,4, швидкістю навчання 0,55 та 80 епох.

Таблиця 13

Результати експериментів

| Модель | Час виконання, ms | RMSE, % | SMAPE, % |
|--------------------------------------|--------------------------|----------------|-----------------|
| Proposed model | 970 | 2.67 | 2.98 |
| Simple MDG based neuro-fuzzy model | 495 | 3.12 | 3.41 |
| Ensemble MDG based neuro-fuzzy model | 900 | 3.01 | 3.09 |
| Bipolar Sigmoid Network | 2215 | 3.32 | 3.67 |

Таблиця 14

Результати експериментів

| Модель | Час виконання, ms | RMSE, % | SMAPE, % |
|--------------------------------------|--------------------------|----------------|-----------------|
| Proposed model | 1005 | 2.88 | 3.11 |
| Simple MDG based neuro-fuzzy model | 467 | 3.19 | 3.42 |
| Ensemble MDG based neuro-fuzzy model | 940 | 3.12 | 3.29 |
| Bipolar Sigmoid Network | 2645 | 3.29 | 3.63 |

5.5 Ієрархчна інтелектуальна система підтримки прийняття рішень у оперативному аналізі ризиків.

У даному підрозділі описано розроблене програмне забезпечення, що реалізує запропоновані і досліджені в дисертації методи, і яке було впроваджено у ТОВ «САЙТОСС» при розробці програмного забезпечення для управління ризиками інвестиційного портфоліо (акт впровадження від 14.01.2021)

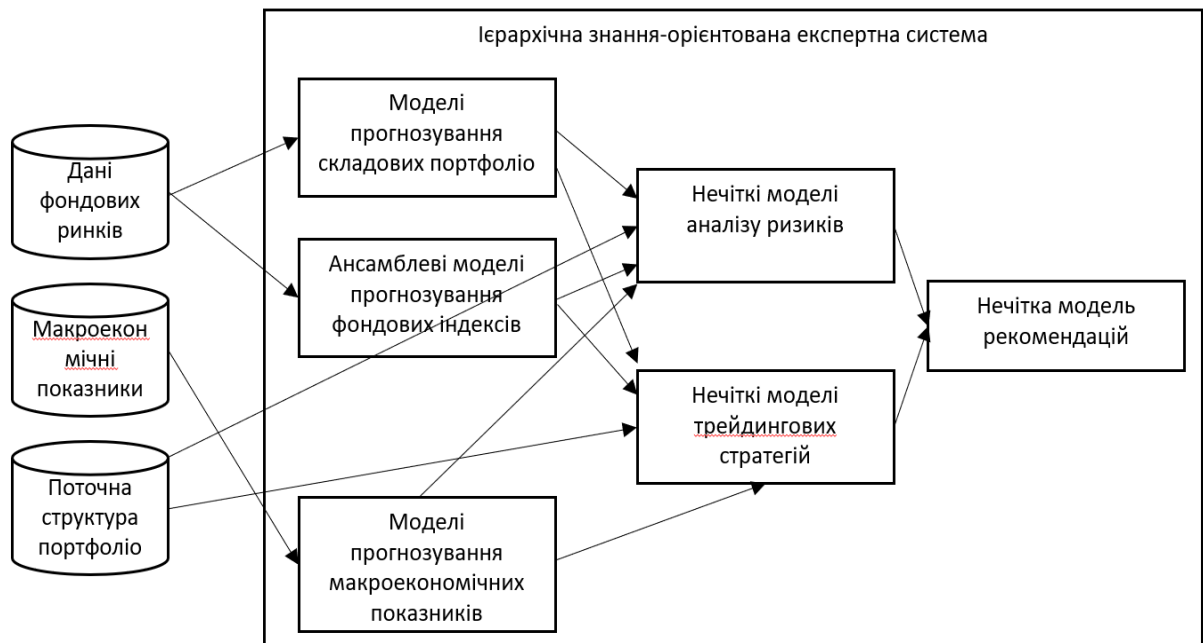


Рисунок 5.22 – Загальна архітектура експертної системи управління ризиками.

Загальну архітектуру експертної системи управління ризиками управління фінансовим портфоліо, на якій представлено ключові підсистеми, представлено на рис. 5.22.

Ефективність запропонованої системи було порівняно з існуючими програмними системами, результати представлено у таблиці 15.

Таблиця 15

Результати експериментів з експертною системою

| Модель | Час виконання, ms | RMSE, % | SMAPE, % |
|---|-------------------|---------|----------|
| Запропонована модель | 5662 | 3.55 | 3.98 |
| Існуюча модель на основі моделі Марковиця | 9854 | 5.29 | 6.41 |
| Існуюча модель на основі моделі Квazi-Шарпа | 10542 | 6.12 | 7.89 |

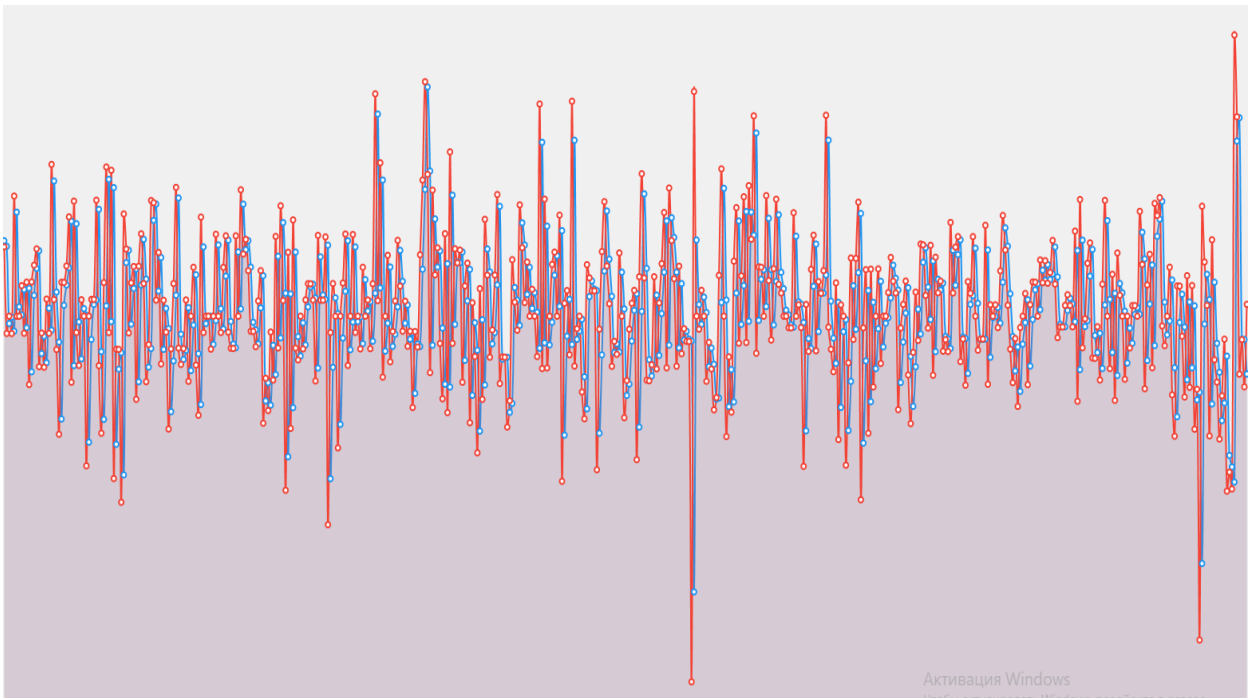


Рисунок 5.23 – Візуалізація процесу прогнозування на компоненті портфолію.

Запропонована система показала набагато кращу точність при прийнятті рішень та менший час обробки даних.

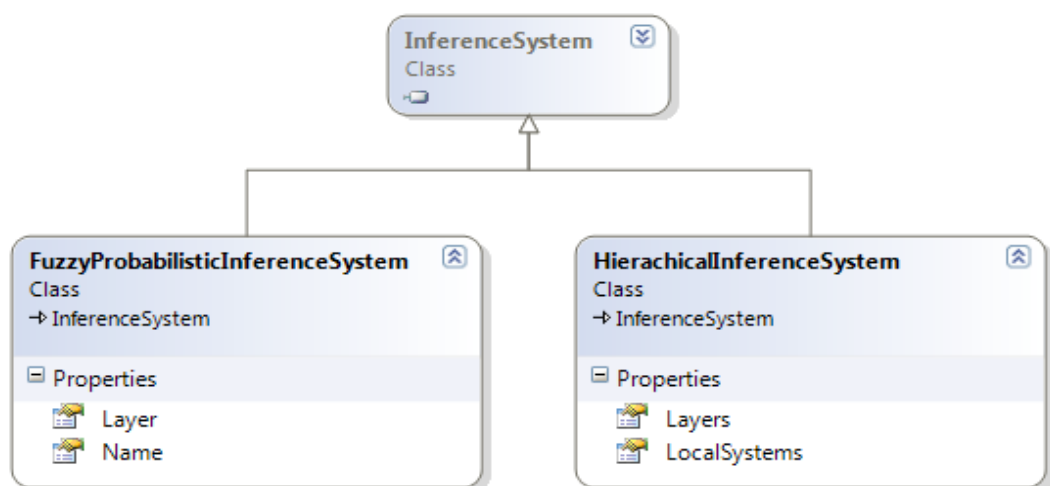


Рисунок 5.24 – Діаграма класів програмної реалізації нечітких систем висновування

5.4 Висновки до розділу

1. Створено програмні реалізації запропонованих методів та моделей та систему для проведення порівняльного імітаційного моделювання.

2. Експерименти довели високу точність та обчислювальну швидкість запропонованих методів.

3. Розглянуто та доведено ефективність розроблених методів та моделей у задачах оперативного аналізу ризиків у управлінні фінансовим портфоліо

Список використаних у цьому розділі джерел наведено у повному списку використаних джерел під номерами [4] [112] [114] [113].

Основні результати цього розділу опубліковано в роботах [93 – 96], [101, 102].

ВИСНОВКИ

У дисертаційній роботі представлені результати, які є відповідно до поставленої мети є рішенням актуальної науково-практичної задачі розробки нових методів та моделей аналізу даних в умовах нестационарних даних з високочастотною динамікою, та які відрізняються невеликим розміром та обчислювальною ефективністю. Внаслідок виконання роботи були отримані нові наукові та практичні результати. Проведені дослідження дозволили зробити наступні висновки:

– проведено аналіз існуючих методів та підходів в задачах прогнозування, ідентифікації та аналізу ризиків в умовах неповних, нелінійних даних зі складною динамікою, показано недоліки існуючих методів;

– розроблено нейро-фаззі модель з багатовимірними гаусіанами у консеквентному шарі та онлайн навчанням з варіантами на базі квадратичної похибки та спеціалізованого критерію для задач прогнозування, ідентифікації та діагностики, що характеризується високою точністю, швидкістю обробки та обчислювальною швидкодією у процесі навчання за рахунок зменшення налаштованих параметрів моделі;

– розроблено ансамбль гібридних нейро-фаззі моделей з багатовимірними гаусіанами у консеквентному шарі та метод його синтезу, що характеризується зменшеною похибкою навчання, покращеними узагальнюючими можливостями та спрощеною процедурою селекції гіперпараметрів;

– розроблено метод навчання нейро-фаззі моделей та їх ансамблів застосуванням емпіричної модової декомпозиції, що відрізняється від аналогів можливістю зниження рівня шуму у випадку високодинамічних даних;

– створено ієрархічну модель підтримки прийняття рішень, яка би інкорпоровала запропоновані моделі;

– проведено експериментальні дослідження розроблених методів на основі реальних та тестових даних;

– розроблені програмні реалізації було використано для ряду практичних задач. Зокрема реалізовано модуль для управління ризиками інвестиційного портфолію. Результати досліджень впроваджені у ТОВ «САЙТОСС» (акт впровадження від 14.01.2021). Окремі положення, висновки та рекомендації дисертаційної роботи використано в освітньому процесі у Харківському національному університеті радіоелектроніки на кафедрі штучного інтелекту з дисциплін «Нейромережеві методи обчислювального інтелекту» та «Штучні нейронні мережі: архітектура, навчання, застосування».

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] S. Kaplan and B. J. Garrick, "On The Quantitative Definition of Risk," *Risk Analysis– Society for Risk Analysis*, vol. 1, pp. 11-27, 1981.
- [2] G. E. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time series analysis: forecasting and control*, John Wiley & Sons, 2015.
- [3] M. S. Paolella, "ARMA Model Identification," in *Linear Models and Time-Series Analysis: Regression, ANOVA, ARMA and GARCH*, John Wiley & Sons, Ltd., 2019.
- [4] R. S. Tsay, *Analysis of Financial Time Series*, John Wiley & Sons, 2015.
- [5] Z. Malki, E. Atlam and A. Ewis, "ARIMA models for predicting the end of COVID-19 pandemic and the risk of second rebound," *Neural Comput & Applic*, vol. 33, p. 2929–2948, 2021.
- [6] M. L. Challa, V. Malepati and S. N. Kolusu, "S&P BSE Sensex and S&P BSE IT return forecasting using ARIMA," *Financ Innov*, vol. 6, no. 47, 2020.
- [7] R. Jiao, Y. Zhang, C. Di, B. Wu, J. Shu, Y. Liu and S. Kuang, "Prediction of the Income and Expenditure Risk of Social Medical Insurance Fund Based on ARIMA Model.", 2020.
- [8] J. F. Commandeur and S. J. Koopman, *Introduction to State Space Time Series Analysis.*, Oxford University Press., 2007.
- [9] A. C. Logue, "The Efficient Market Hypothesis and Its Critics.," *CFA Digest*, vol. 33, p. 40–41, 2003.
- [10] D. Wen, W. Gang-Jin, M. Chaoqun and W. Yudong, "Risk spillovers between oil and stock markets: A VAR for VaR analysis," *Energy Economics*, vol. 80, pp. 524-535, 2019.
- [11] N. G. Emenogu, M. O. Adenomon and N. O. Nweze, "On the volatility of daily stock returns of Total Nigeria Plc: evidence from GARCH models, value-at-risk and backtesting," *Financ Innov*, vol. 6, no. 18, 2020.

- [12] H. Yu, F. Khan and B. Veitch, "A Flexible Hierarchical Bayesian Modeling Technique for Risk Analysis of Major Accidents," *Risk Analysis*, vol. 37, pp. 1668-1682, 2017.
- [13] H. Wang, X. Deng, Z. Zhang and W. Jiang, "A New Failure Mode and Effects Analysis Method Based on Dempster–Shafer Theory by Integrating Evidential Network," *IEEE Access*, vol. 7, pp. 79579-79591, 2019.
- [14] J. Rathman, C. Yang and H. Zhou, "Dempster-Shafer theory for combining in silico evidence and estimating uncertainty in chemical risk assessment," *Computational Toxicology*, vol. 6, pp. 16-34, 2018.
- [15] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338-353, 1965.
- [16] D. Dubois and H. Prade, "Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets.," *Fuzzy Sets Syst*, vol. 192, p. 3–24, 2012.
- [17] D. Dubois and H. Prade, "The legacy of 50 years of fuzzy sets: A. discussion.," *Fuzzy Sets Syst*, vol. 281, p. 21–31, 2015.
- [18] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1-13, 1975.
- [19] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE transactions on systems, man, and cybernetics*, pp. 116-132, 1985.
- [20] A. Sheta, "Software Effort Estimation and Stock Market Prediction Using Takagi-Sugeno Fuzzy Models," *IEEE International Conference on Fuzzy Systems*, pp. P.171-178, July 16-21 2006.
- [21] P. C. Chang and C. H. Liu, "A TSK type fuzzy rule based system for stock price prediction," *Expert Syst. Appl.*, vol. 34, p. 135–144, 2008.
- [22] J. v. d. Berg, U. Kaymak and W.-M. v. d. Bergh, "Financial markets analysis by using a probabilistic fuzzy modeling approach," *International Journal of Approximate Reasoning*, vol. 35, no. 12, pp. 291-305, 2004.

- [23] K. Rudnik and A. Walaszek-Babiszewska, "Probabilistic-fuzzy knowledge-based system for managerial applications," *Management and Production Engineering Review*, vol. 3, no. 1, p. 49–61, 2012.
- [24] R. Galankashi, M. Rafiei and F. Ghezelbash, "Portfolio selection: a fuzzy-ANP approach," *Financ Innov*, vol. 6, no. 17, 2020.
- [25] S. Moradi and F. Mokhatab Rafiei, "A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks," *Financ Innov*, vol. 5, no. 15, 2019.
- [26] S. Rizvi, J. Mitchell and A. Razaque, "A fuzzy inference system (FIS) to evaluate the security readiness of cloud service providers," *Cloud Comp*, vol. 9, no. 42, 2020.
- [27] S. Lauguico, R. Concepcion II, J. Alejandrino, D. Macasaet, R. R. Tobias, A. Bandala and E. Dadios, "A fuzzy logic-based stock market trading algorithm using bollinger bands.," *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, pp. 1-6.
- [28] M. Khalilzadeh, L. Katoueizadeh and E. Zavadskas, "Risk identification and prioritization in banking projects of payment service provider companies: an empirical study," *Front. Bus. Res. China*, vol. 14, no. 15, 2020.
- [29] G. Tsekouras, H. Sarimveis, E. Kavakli and G. Bafa, "A hierarchical fuzzy-clustering approach to fuzzy modeling," *Fuzzy Sets and Systems*, vol. 107, p. 245–266, 2005.
- [30] R. Rojas, *Neural networks: a systematic introduction*, Springer Science & Business Media, 2013.
- [31] S. F. Crone, H. Michele and K. Nikolopoulos, "Advances in forecasting with neural networks: Empirical evidence from the NN3 competition on time series prediction," *International Journal of Forecasting*, vol. 27, no. 3, pp. 635-660, 2011.

- [32] M. Dobrovolny, I. Soukal, K. C. Lim, A. Selamat and O. Krejcar, "Forecasting of forex price trend using recurrent neural network-long short-term memory," 2020.
- [33] X. Zhong and D. Enke, "Predicting the daily return direction of the stock market using hybrid machine learning algorithms," *Financial Innovation*, vol. 5, no. 1, p. 4, 2019.
- [34] J. Park, S. Hossain, J. Oh, H. Yoo and H. Kim, "Assessment of risk potential due to underground box structure installation employing ANN model and field experimental approaches," *Journal of Performance of Constructed Facilities*, vol. 34, no. 4, 2020.
- [35] H. Chu, W. Wu, Q. Wang, R. Nathan and J. Wei, "An ANN-based emulation modelling framework for flood inundation modelling: Application, challenges and future directions.," *Environmental Modelling & Software*, vol. 124.
- [36] X. Zhong and S. Zhou, "Risk analysis method of bank microfinance based on multiple genetic artificial neural networks," *Neural Computing and Applications*, vol. 1, no. 11.
- [37] R. Hasanpour, J. Rostami, J. Schmitt, Y. Ozcelik and B. Sohrabian, "Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 12, no. 1, pp. 21-31, 2020.
- [38] L. Zhao and D. Feng, "Deep neural networks for survival analysis using pseudo values," *IEEE journal of biomedical and health informatics*, vol. 24, no. 11, pp. 3308-3314, 2020.
- [39] R. Ahmadi and S. Shybt, "Study of artificial neural networks in information security risk assessment," *Journal of Management and Accounting Studies*, vol. 8, no. 2, 2020.
- [40] N. Nezamoddini, A. Gholami and F. Aqlan, "A risk-based optimization framework for integrated supply chains using genetic algorithm and artificial neural network," *International Journal of Production Economics*, vol. 225.

- [41] E. Sariev and G. Germano, "Bayesian regularized artificial neural networks for the estimation of the probability of default.," *Quantitative Finance*, vol. 20, no. 2, pp. 311-328, 2020.
- [42] P. M. Addo, D. Guegan and B. Hassani, "Credit risk analysis using machine and deep learning models," *Risks*, vol. 6, no. 2, p. 38, 2018.
- [43] J. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 3, pp. 665-685, 1993.
- [44] L. Rutkowski and K. Cpalka, "Flexible neuro-fuzzy systems," *IEEE transactions on neural networks*, vol. 14, no. 3, pp. 554-574, 2003.
- [45] S. Rajab, "A review on the applications of neuro-fuzzy systems in business," *Artificial Intelligence Review*, vol. 19, no. 4, pp. 481-510, 2018.
- [46] P. V. de Campos Souza, "Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature," *Applied Soft Computing*, p. 106275, 2020.
- [47] Y. V. Bodyanskiy, S. Popov and T. Rybalchenko, "Multilayer neuro-fuzzy network for short term electric load forecasting.," *Computer Science–Theory and Applications*, pp. 339-348, 2008.
- [48] P. Otto, Y. V. Bodyanskiy and V. Kolodyazhniy, "A new learning algorithm for a forecasting neuro-fuzzy network.," *Integrated Computer-Aided Engineering*, vol. 10, no. 4, pp. 399-409, 2003.
- [49] Y. Bodyanskiy and S. Popov, "Neural network approach to forecasting of quasiperiodic financial time series.," *European Journal of Operational Research*, vol. 175, no. 3, pp. 1357-1366, 2006.
- [50] A. K. Parida, "Times Series Forecasting using Chebyshev Functions based Locally Recurrent neuro-Fuzzy Information System.," *International Journal Of Computational Intelligence Systems*, vol. 10, no. 1, pp. 375-393, 2017.
- [51] E. Hadavandi, H. Shavandi and A. Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 800-808, 2010.

- [52] A. Esfahanipour and W. Aghamiri, "Adapted neuro-fuzzy inference system on indirect approach TSK fuzzy rule base for stock market analysis," *Expert Systems with Applications*, vol. 37, no. 7, pp. 4742-4748, 2010.
- [53] G. S. Atsalakis and K. P. Valavanis, "Forecasting stock market short-term trends using a neuro-fuzzy based methodology," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10696-10707, 2009.
- [54] Y. Bodyanskiy, I. Pliss and O. Vynokurova, "Adaptive wavelet-neuro-fuzzy network in the forecasting and emulation tasks.," *Int. journal on information theory and applications*, vol. 15, no. 1, pp. 47-55, 2008.
- [55] A. S. Weigend, B. A. Huberman and D. A. Rumelhart, "Predicting the future: A connectionist approach," *International journal of neural systems*, vol. 1, no. 03, pp. 193-209, 1990.
- [56] F. García, "Hybrid fuzzy neural network to predict price direction in the German DAX-30 index," *Technological and Economic Development of Economy*, vol. 24, no. 6, pp. 2161-2178, 2018.
- [57] D. Pamučar, S. Ljubojević, D. Kostadinović and B. Đorović, "Cost and risk aggregation in multi-objective route planning for hazardous materials transportation—A neuro-fuzzy and artificial bee colony approach.," *Expert Systems with Applications*, vol. 65, pp. 1-15, 2016.
- [58] D. Pamučar and G. Cirovic, "Vehicle Route Selection with an Adaptive Neuro Fuzzy Inference System in Uncertainty Conditions.," *Decis. Mak. Appl. Manag. Eng.*, vol. 1, pp. 13-37, 2008.
- [59] S. Sremac, "ANFIS model for determining the economic order quantity.," *Decision Making: Applications in Management and Engineering*, vol. 1, no. 2, pp. 81-92, 2018.
- [60] D. J. J. Rubio, "USNFIS: Uniform stable neuro fuzzy inference system," *Neurocomputing*, vol. 262, pp. 57-66, 2017.
- [61] V. Lukovac, D. Pamučar, M. Popović and B. Đorović, "Portfolio model for analyzing human resources: An approach based on neuro-fuzzy modeling and

- the simulated annealing algorithm," *Expert Systems with Applications*, vol. 90, pp. 318-331, 2017.
- [62] S. Kar, S. Das and P. K. Ghosh, "Applications of neuro fuzzy systems: A brief review and future outline," *Applied Soft Computing*, vol. 15, pp. 243-259, 2014.
- [63] J. Mendes, F. Souza, R. Araújo and S. Rastegar, "Neo-fuzzy neuron learning using backfitting algorithm," *Neural Computing and Applications*, vol. 31, no. 8, pp. 3609-3618, 2019.
- [64] A. M. Silva, W. Caminhas, A. Lemos and F. Gomide, "A fast learning algorithm for evolving neo-fuzzy neuron," *Applied Soft Computing*, vol. 14, pp. 194-209, 2014.
- [65] Y. V. Bodyanskiy, O. K. Tyshchenko and D. S. Kopaliani, "Adaptive learning of an evolving cascade neo-fuzzy system in data stream mining tasks," *Evolving Systems*, vol. 7, no. 2, pp. 107-116, 2016.
- [66] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," *2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, pp. 1-4, 2016.
- [67] S. Rajab, "An interpretable neuro-fuzzy approach to stock price forecasting," *Soft Computing*, vol. 23, no. 3, pp. 921-936, 2019.
- [68] S. K. Chandar, "Fusion model of wavelet transform and adaptive neuro fuzzy inference system for stock market prediction," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-9, 2019.
- [69] D. Y. Chiu, "Dynamically exploring internal mechanism of stock market by fuzzy-based support vector machines with high dimension input space and genetic algorithm".
- [70] J. A. González, "Fuzzy GA-SVR for Mexican Stock Exchange's Financial Time Series Forecast with Online Parameter Tuning".

- [71] Q. Cai, D. Zhang, W. Zheng and S. C. Leung, "A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression," *Knowledge-Based Systems*, 2015.
- [72] R. S. Lee, "Chaotic Type-2 Transient-Fuzzy Deep Neuro-Oscillatory Network (CT2TFDNN) for Worldwide Financial Prediction," *IEEE Transactions on Fuzzy Systems*, 2019.
- [73] M. Pulido, "Optimization of Ensemble Neural Networks with Type-1 and Type-2 Fuzzy Integration for Prediction of the Taiwan Stock Exchange," *InRecent Developments and the New Direction in Soft-Computing Foundations and Applications*, pp. 151-164, 2018.
- [74] D. Bhattacharya and A. Konar, "Self-adaptive type-1/type-2 hybrid fuzzy reasoning techniques for two-factored stock index time-series prediction," *Soft Computing*, 2017.
- [75] M. M. Ebadzadeh and S. B. Armin, "CFNN: Correlated fuzzy neural network," *Neurocomputing*, vol. 148, pp. 430-444, 2015.
- [76] I. Goodfellow, Y. Bengio, A. Courville and Y. Bengio, *Deep learning*, Cambridge: MIT press, 2016.
- [77] J. Nocedal and S. Wright, "Conjugate gradient methods," *Numerical optimization*, pp. 101-134, 2006.
- [78] J. Nocedal and S. Wright, "Large-scale unconstrained optimization," *Numerical Optimization*, pp. 164-192, 2006.
- [79] Y. A. LeCun, L. Bottou, G. B. Orr and K. R. Müller, "Efficient BackProp," in *Neural networks: Tricks of the trade*, Berlin, Heidelberg, Springer, 2012, pp. 9-48.
- [80] S. Wiesler, A. Richard, R. Schlüter and H. Ney, "A critical evaluation of stochastic algorithms for convex optimization," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6955-6959, 2013.
- [81] L. Bottou, F. E. Curtis and J. Nocedal, "Optimization Methods for Large-Scale Machine Learning," *Siam Review*, vol. 60, no. 2, pp. 223-311, 2018.

- [82] B. Radojičić and S. Kredatus, "The impact of stock market price Fourier transform analysis on the Gated Recurrent Unit classifier model," *Expert Systems with Applications*, vol. 159, 2020.
- [83] P. Talbot, "Correlated synthetic time series generation for energy system simulations using Fourier and ARMA signal processing," *International Journal of Energy Research*, vol. 44, no. 10, pp. 8144-8155, 2020.
- [84] K. Cao and J. Wei, "Adaptive wavelet estimations for the derivative of a density in GARCH-type model," *J Inequal Appl*, vol. 106, 2019.
- [85] E. Bouri, S. Shahzad, D. Roubaud, L. Kristoufek and B. Lucey, "Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis," *The Quarterly Review of Economics and Finance*, vol. 77, pp. 156-164, 2020.
- [86] W. Mensi, "Global financial crisis and co-movements between oil prices and sector stock markets in Saudi Arabia: A VaR based wavelet," *Borsa Istanbul Review*, vol. 19, no. 1, pp. 24-28, 2019.
- [87] X. Qiao, H. Zhu and L. Hau, "Time-frequency co-movement of cryptocurrency return and volatility: Evidence from wavelet coherence analysis," *International Review of Financial Analysis*, vol. 71, 2020.
- [88] S. Kang, R. McIver and J. Hernandez, "Co-movements between Bitcoin and Gold: A wavelet coherence analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 536, 2019.
- [89] B. Václav, L. Pfeifer and D. Kolcunov, "Are the Risk Weights of Banks in the Czech Republic Procyclical?," *Evidence from Wavelet Analysis*, 217.
- [90] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, p. 903–995, 1998.

- [91] H. Ge, G. Chen, H. Yu, H. Chen and F. An, "Theoretical Analysis of Empirical Mode Decomposition," *Symmetry*, vol. 10, no. 11, p. 623, 2018.
- [92] Y. Li, H. Han and Y. Li, "A New HHT-Based Denoising Algorithm for Financial Time Series Data Mining," in *IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 2019.
- [93] S. L. Lin and H. W. Huang, "Improving Deep Learning for Forecasting Accuracy in Financial Data," *Discrete Dynamics in Nature and Society*, pp. 1-12, 2020.
- [94] D. Jothimani and A. Başar, "Stock Index Forecasting Using Time Series Decomposition-Based and Machine Learning Models," *Lecture Notes in Computer Science Artificial Intelligence XXXVI*, p. 283–292, 2019.
- [95] J. Cao, Z. Li and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and its Applications*, vol. 519, p. 127–139, 2019.
- [96] L. Xian, K. He, C. Wang and K. K. Lai, "Factor analysis of financial time series using EEMD-ICA based approach," *Sustainable Futures*, vol. 2, p. 100003, 2020.
- [97] Z. Chengzhao, P. Heping, M. Yu and H. Xun, "Analysis of Asia Pacific stock markets with a novel multiscale model," *Physica A: Statistical Mechanics and its Applications*, vol. 534, p. 120939, 2019.
- [98] Е. В. Бодянский, Е. А. Винокурова, Н. С. Ламонова and И. П. Плисс, "Алгоритм обучения вэйвлет-нейрона на основе комбинированного критерия," *Радіоелектроніка, інформатика, управління.*, vol. №2, no. 14, pp. 83-89, 2005.
- [99] Y. Bodyanskiy, O. Vlasenko, O. Vynokurova and I. Pliss, "A hybrid neuro-fuzzy network for financial time-series prediction," *Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту»*, 2017.

- [100] A. Vlasenko and Y. Bodyanskiy, "An approach to overcome prediction delay problem in hybrid neuro-fuzzy networks for financial time-series forecasting," *Матеріали VI Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології»*, 2017.
- [101] A. Vlasenko, O. Vynokurova, N. Vlasenko and M. Peleshko, "A Hybrid Neuro-Fuzzy Model for Stock Market Time-Series Prediction," *IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, pp. 352-355, 2018.
- [102] A. Vlasenko, N. Vlasenko, O. Vynokurova and Y. Bodyanskiy, "An Enhancement of a Learning Procedure in Neuro-Fuzzy Model," *IEEE First International Conference on System Analysis & Intelligent Computing (SAIC)*, pp. 1-4, 2018.
- [103] A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko, "A novel neuro-fuzzy model for multivariate time-series prediction," *Data*, vol. 3, no. 4, p. 62, 2018.
- [104] A. Vlasenko, N. Vlasenko, O. Vynokurova, Y. Bodyanskiy and D. Peleshko, "A Novel ensemble neuro-fuzzy model for financial time series forecasting," *Data*, vol. 4, no. 3, p. 126, 2019.
- [105] G. Raju, J. Zhou and R. Kisner, "Hierarchical fuzzy control," *International Journal of Control*, vol. 54, p. 1201–1216, 1991.
- [106] A. Vlasenko, Y. Rashkevych, N. Vlasenko, D. Peleshko and O. Vynokurova, "A Hybrid EMD - Neuro-Fuzzy Model for Financial Time Series Analysis," *IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, pp. 112-115, 2020.
- [107] A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko, "An Empirical Mode Decomposition Based Method to Synthesize Ensemble Multidimensional Gaussian Neuro-Fuzzy Models in Financial Forecasting," *Communications in Computer and Information Science*, vol. 1158, pp. 140-149, 2020.

- [108] А. Н. Власенко and Е. И. Кучеренко, "Иерархическая нечетко-вероятностная модель в задачах управления рисками," *Системы обработки информации*, vol. 1, pp. 145-149, 2013.
- [109] А. Н. Власенко and Е. И. Кучеренко, "Применение нечеткой кластеризации в задачах структурной идентификации нечетких моделей Такаги-Сугено-Канга," *ISDMCI'2012*, pp. 328-330, 2012.
- [110] А. Н. Власенко and Е. И. Кучеренко, "Метод логического вывода в иерархических нечетко-вероятностных моделях сложных систем," *Бионика интеллекта*, vol. 2, no. 81, p. 47–52, 2013.
- [111] А. Н. Власенко and Е. И. Кучеренко, "Иерархические системы нечеткого вывода Такаги-Сугено в задачах управления рисками," *матер. XVI междунар. молод. форума «радиоэлектроника и молодежь в XXI веке»*, pp. 69-70, 2012.
- [112] C. R. Souza, "The Accord.NET Framework," [Online]. Available: <http://accord-framework.net>.
- [113] "Math.NET Numerics," [Online]. Available: <https://numerics.mathdotnet.com>.
- [114] M. Riedmiller and B. Heinrich, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *IEEE international conference on neural networks*, 1993.
- [115] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, pp. 287-289, 1977.
- [116] C. Kanzow, "Erratum to “Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints”.
- [117] "Monthly log returns of IBM stock and the S&P 500 index dataset," [Online].
- [118] А. Н. Власенко and Е. И. Кучеренко, "Архитектура иерархической системы логического вывода с нечетко-вероятностной базой знаний," *матер. XI международной научно-практической конференции "Математическое обеспечение интеллектуальных систем"*, pp. 42-43, 2013.

- [119] А. Н. Власенко and Е. И. Кучеренко, "Реализация иерархической нечетко-вероятностной модели логического вывода на базе библиотеки Aforge.Net.," *матер. XVII междунар. молод. форума «радиоэлектроника и молодежь в XXI веке»*, 2013.
- [120] А. Н. Власенко and Е. И. Кучеренко, "Реляционная база нечетко-вероятностных знаний для логического вывода в задачах анализа рисков," *Матер. XVII междунар. молод. форума «радиоэлектроника и молодежь в XXI веке»*, 2014.

ДОДАТОК А

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Список публікацій здобувача, в яких опубліковані основні наукові результати дисертації:

1. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "A novel neuro-fuzzy model for multivariate time-series prediction", *Data*, 3(4), 62, 2018. (Входить до міжнародних науково-метричних баз SCOPUS та Web of Science)
2. A. Vlasenko, N. Vlasenko, O. Vynokurova, Y. Bodyanskiy and D. Peleshko. "A Novel ensemble neuro-fuzzy model for financial time series forecasting", *Data*, 4(3), 126, 2019. (Входить до міжнародних науково-метричних баз SCOPUS та Web of Science)
3. A. Vlasenko, N. Vlasenko, O. Vynokurova and D. Peleshko. "An Empirical Mode Decomposition Based Method to Synthesize Ensemble Multidimensional Gaussian Neuro-Fuzzy Models in Financial Forecasting", *Communications in Computer and Information Science*, volume 1158, pp. 140-149, 2020. Springer, Cham. (Входить до міжнародної науково-метричної бази SCOPUS)
4. А. Власенко, Е. Кучеренко, «Иерархическая нечетко-вероятностная модель в задачах управления рисками», *Системи обробки інформації*, Вип. 1, С. 145-149, 2013. (Входить до міжнародної науково-метричної бази Index Copernicus International)
5. А. Власенко, Е. Кучеренко. «Метод логического вывода в иерархических нечетко-вероятностных моделях сложных систем», *Бионика интеллекта*, Вып. 2 (81), С. 47–52, 2013.

Результати, які засвідчують апробацію матеріалів дисертації:

6. A. Vlasenko, O. Vynokurova, N. Vlasenko and M. Peleshko, "A Hybrid Neuro-Fuzzy Model for Stock Market Time-Series Prediction", in *2018 IEEE*

Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2018. (Входить до міжнародних науково-метричних баз SCOPUS та Web of Science)

7. A. Vlasenko, Y. Rashkevych, N. Vlasenko, D. Peleshko and O. Vynokurova, "A Hybrid EMD - Neuro-Fuzzy Model for Financial Time Series Analysis", in *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, Ukraine, 2020. (Входить до міжнародної науково-метричної бази SCOPUS)

8. A. Vlasenko, N. Vlasenko, O. Vynokurova and Y. Bodyanskiy, "An Enhancement of a Learning Procedure in Neuro-Fuzzy Model", in *2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC)*, Kyiv, Ukraine, 2018. (Входить до міжнародних науково-метричних баз SCOPUS та Web of Science)

9. A. Vlasenko, Y. Bodyanskiy, «An approach to overcome prediction delay problem in hybrid neuro-fuzzy networks for financial time-series forecasting», в *Матеріали VI Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології»*, Одеса, 2017.

10. Y. Bodyanskiy, A. Vlasenko, O. Vynokurova and I. Pliss, "A hybrid neuro-fuzzy network for financial time-series prediction", в *Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту»*, Залізний порт, Україна, 2017.

11. А. Власенко, Е. Кучеренко, «Применение нечеткой кластеризации в задачах структурной идентификации нечетких моделей Такаги-Сугено-Канга» в *Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту»*, Євпаторія, Україна, 2012.

12. А. Власенко, Е. Кучеренко, «Архитектура иерархической системы логического вывода с нечетко-вероятностной базой знаний», в *Материалы XI*

международной научно-практической конференции "Математическое обеспечение интеллектуальных систем" Днепропетровск, 2013.

13. А. Власенко, Е. Кучеренко, «Иерархические системы нечеткого вывода Такаги-Сугено в задачах управления рисками», в *Материалы XVI международного молодежного форума «Радиоэлектроника и молодежь в XXI веке»*, Харьков, 2012.

14. А. Власенко, Е. Кучеренко, «Реализация иерархической нечетко-вероятностной модели логического вывода на базе библиотеки Aforge.Net», в *Материалы XVII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке»*, Харьков, 2013.

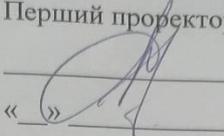
15. А. Власенко, Е. Кучеренко, «Реляционная база нечетко-вероятностных знаний для логического вывода в задачах анализа рисков», в *Материалы XVIII международного молодежного форума «Радиоэлектроника и молодежь в XXI веке»*, Харьков, 2014.

ДОДАТОК В

ДОКУМЕНТИ, ЩО ПІДТВЕРДЖУЮТЬ ВПРОВАДЖЕННЯ

Затверджую

Перший проректор ХНУРЕ

 проф. Рубан І.В.

« » _____

Акт

щодо впровадження у навчальний процес кафедри штучного інтелекту Харківського національного університету радіоелектроніки результатів дисертаційної роботи Власенка Олександра Миколайовича на тему «Методи та моделі інтелектуалізації процесів оперативного аналізу ризиків на основі м'яких обчислень».

Комісія у складі

Голови комісії

завідувача кафедри штучного інтелекту, д.т.н.,
проф. Філатова В.О.

Членів комісії

професора кафедри штучного інтелекту, д.т.н.,
проф. Бодянського Є.В.


професора кафедри штучного інтелекту, д.т.н.,
проф. Петрова К.Е.

доцента кафедри штучного інтелекту, к.т.н.,
Дейнеко А.О.

розглянули результати наукових досліджень Власенка О.М. та прийняли рішення щодо впровадження їх у навчальний процес кафедри ШІ ХНУРЕ при викладанні дисциплін «Нейромережеві методи обчислювального інтелекту» та «Штучні нейронні мережі: архітектура, навчання, застосування».


Отримані в дисертаційній роботі нові наукові результати, що стосуються розробки моделей та методів динамічного інтелектуального аналізу даних на основі штучних нейронних мереж, дозволяють підвищити науковий рівень вказаних дисциплін та доповнити їх прикладами практичного застосування.

Голова комісії




В.О.Філатов


Члени комісії



Є.В. Бодянський



К.Е. Петров



А.О. Дейнеко

ЗАТВЕРДЖУЮ



Директор ТОВ «САЙТОСС»

Луців В.В.

«14 січня 2021 р.»

Акт

про впровадження результатів дисертаційної роботи
**МЕТОДИ ТА МОДЕЛІ ІНТЕЛЕКТУАЛІЗАЦІЇ ПРОЦЕСІВ ОПЕРАТИВНОГО АНАЛІЗУ
 РИЗИКІВ НА ОСНОВІ М'ЯКИХ ОБЧИСЛЕНЬ**
 на здобуття
 наукового ступеня кандидата технічних наук
ВЛАСЕНКА ОЛЕКСАНДРА МИКОЛАЙОВИЧА

Комісія у складі:

Голова:

Луців В.В. – директор;

Члени комісії:

Харкевич О.М. – менеджер;

Матікайнен Т.О. – менеджер;

склала даний акт про те, що у ТОВ «САЙТОСС» був впроваджений програмний модуль, у якому реалізовано п'ятишарову гібридну нейро-фаззі модель із багатовимірними гаусіанами у консеквентному шарі та швидкий алгоритм її навчання. Розроблений модуль орієнтовано на задачі прогнозування та ідентифікації ризиків, він вирізняється здатністю виявляти та обробляти складні нелінійні залежності у вхідних даних при компактній моделі та високій обчислювальній швидкодії. Запропонований модуль дає можливість ефективно оцінювати допустимі сценарії зміни для цін компонентів інвестиційного портфелю, оперативно виявляти ризики інвестиційних втрат та адаптувати структуру портфелю відповідно до них.

Результати впровадження довели, що розроблені Власенком О.М. методи та моделі прогнозування та ідентифікації мають переваги у прикладному застосуванні над існуючими підходами інтелектуального аналізу даних та підтримки прийняття рішень.

Голова комісії:

В.В. Луців

Члени комісії:

О.М. Харкевич

Т.О. Матікайнен