

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Підпис

ВЛАСЕНКО ОЛЕКСАНДР МИКОЛАЙОВИЧ

УДК 004.032.26

**МЕТОДИ ТА МОДЕЛІ ІНТЕЛЕКТУАЛІЗАЦІЇ ПРОЦЕСІВ ОПЕРАТИВНОГО
АНАЛІЗУ РИЗИКІВ НА ОСНОВІ М'ЯКИХ ОБЧИСЛЕНЬ**

05.13.23 – системи та засоби штучного інтелекту

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2021

Дисертацією є рукопис.

Робота виконана в Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник доктор технічних наук, професор
Бодянський Євгеній Володимирович,
Харківський національний університет радіоелектроніки,
професор кафедри штучного інтелекту.

Офіційні опоненти: доктор технічних наук, професор
Литвиненко Володимир Іванович,
Херсонський національний технічний університет
МОН України, завідувач кафедри інформатики та
комп'ютерних наук;

кандидат технічних наук, доцент
Чумаченко Дмитро Ігорович,
доцент кафедри математичного моделювання та штучного
інтелекту Національного аерокосмічного університету
ім. М.Є. Жуковського «Харківський авіаційний інститут».

Захист відбудеться «12» травня 2021 р. о 13-00 годині на засіданні спеціалізованої вченої ради Д 64.052.01 Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Науки, 14.

З дисертацією можна ознайомитись у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Науки, 14.

Автореферат розісланий «09» квітня 2021 р.

Учений секретар
спеціалізованої вченої ради

Підпис

Є.І. Литвинова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. 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 нестаціонарних 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.

Великий внесок у розвиток динамічного інтелектуального аналізу даних на основі штучних нейронних мереж та внесли О.Г. Івахненко, Б. Уїдроу, Т. Кохонен, Г. Голланд, Б. Коско, Є.В. Бодянський, Ю. Шмідхубер, Д.Д. Пелешко, Ян ЛеКун, О.І. Михальов, Л. Ботту, О.Г. Руденко, Й. Бенджіо, Ю.П. Зайченко, Ян Гудфелоу та інші. Фундамент теорії нечіткої логіки та нечітких множин заклали Е. Мамдані, Т. Сугено, Дж. Мендель, Дж. Бездек, В. Педрич, Д. Дюбуа та А. Праде.

До недоліків нейро-фаззі моделей можна віднести проблему великої розмірності у багатьох прикладних задачах.

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, високою точністю, швидкодією 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 інкорпорувала 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 малоїмовірні 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 висновування, [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 висновування, [13] - ієрархічна нечітко модель, [14] – розробка імплементації та імітаційне моделювання нечіткої ієрархічної моделі логічного висновування, [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.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** подано загальну характеристику роботи, обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і задачі дослідження, визначено наукову новизну, практичне значення і впровадження одержаних результатів, розкрито застосовані методи дослідження, зв'язок роботи з НДР, відомості про публікацію і апробацію результатів досліджень, особистий внесок здобувача.

У **першому розділі** проаналізовано стан проблеми побудови методів обробки даних та моделей складних систем та процесів, які використовуються у оперативному аналізі ризиків.

На основі аналізу існуючих методів та прикладів їх застосування зроблений висновок про недоліки існуючих методів та моделей, визначені мета та задачі дослідження роботи.

Другий розділ присвячено побудові адаптивної нейро-фаззи моделі, здатної відтворювати складні нелінійні залежності у вхідних даних та поєднувати велику точність у прогнозуванні з високою швидкістю. Також запропоновано узагальнену модель для декількох вихідних змінних здатну використовувати взаємозалежності між наборами даних для кращої точності.

Запропонована модель є модифікацією класичної моделі ANFIS – адаптивної нейро-фаззи системи логічного висновування, і вона складається з п'яти шарів. На рис. 1 зображена загальна архітектура моделі.

Перший рівень відповідає за фаззифікацію вхідних змінних, представлених n -розмірним вектором $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$. Кожен вхідний параметр обробляється h^ψ функціями належності, отже, загальна кількість функцій належності першого рівня становить $h^\psi \times n$.

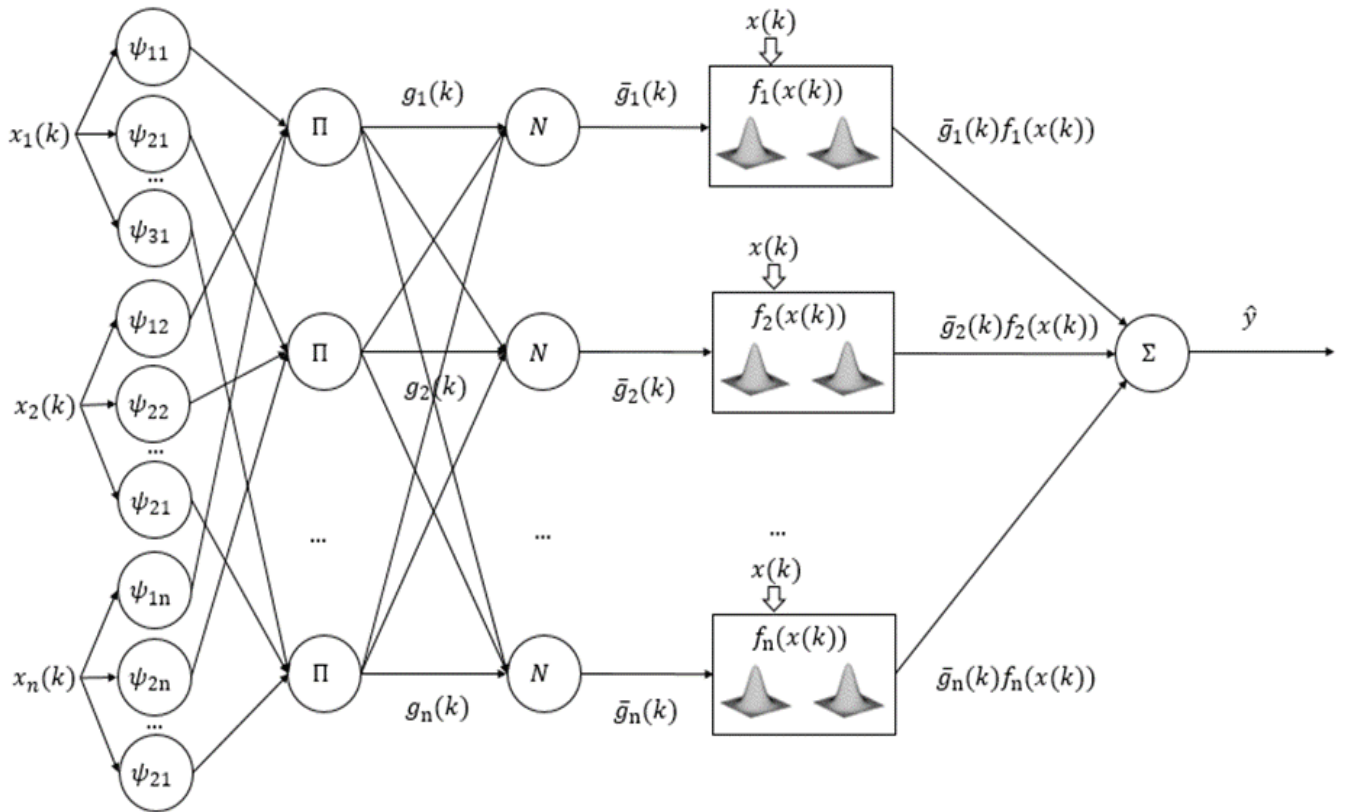


Рисунок 1 – Загальна архітектура запропонованої моделі з одним виходом

У дослідженні використано функцію належності Гауса, яка відноситься до радіально-базисних ядерних функцій має такий вигляд:

$$\psi(x(k)) = \exp\left(-\frac{(x(k) - c)^2}{2\sigma^2(k)}\right),$$

де $x(k)$ – вектор вхідних значень, c_{je}^ψ – центр функції σ_{je} параметр ширини функції. Основними перевагами цієї функції належності є відносно невелика кількість параметрів та стійкість до викидів, яка дозволяє будувати робастні моделі.

Другий рівень відповідає за агрегацію значень функцій фазифікації – виразів антецеденту. Він складається з h^ψ вузлів, що реалізують нечітку Т-норму у вигляді алгебраїчного добутку:

$$g_j(k) = \prod_{i=1}^n \psi_{je}(x_i(k)).$$

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

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

Це необхідно для задоволення критерію розділення Руспіні:

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

Четвертий шар представлений багатовимірними функціями Гауса $\phi_{je}(x(k))$ та їхніми ваговими коефіцієнтами p які використовуються замість стандартних поліномів у моделі ANFIS:

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

де $x(k)$ – вхідний вектор, c_{je}^ϕ – вектор, який репрезентує центр функції Q_{je} – матриці рецепторних полів, $Q_{je} \in S_{++}^n$. У цьому випадку квадратична функція стає квадратною формою цілого вхідного вектору $x(k)$ та узагальнює поняття ширини функції.

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

Таким чином, загальний результат виведення четвертого рівня може бути виражений як:

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

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:

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

У матричній формі результат може бути представлено як:

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

де $x(k)$ вхідний вектор, p -вагові коефіцієнти задані у вигляді вектору $f(x(k))$ вектор значень вузлів консеквентного шару:

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

де h^ψ кількість рецепторних функцій першого рівня та h^ϕ - кількість багатовимірних гаусіанів для кожного нормалізованого результату попереднього шару \bar{g}_1 .

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

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

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

де $p(k)$ – вагові коефіцієнти представлені у вигляді вектору, $y(k)$ – зовнішній навчальний сигнал, $p^T f(x(k))$ - вихідне значення моделі.

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

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

де $y(k)$ – значення зовнішнього навчального сигналу, $\hat{y}(k)$ – значення прогнозоване моделлю та N розмір вибірки.

Q_{j0l}^{-1} матриці ініціалізуються як одиничні, і їх навчання можна записати наступним чином:

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

де λ_Q – крок навчання, β_Q – параметр пам'яті, τ_{jl}^Q – вектор значень зворотного розповсюдження помилки по відношенню до Q_{jl}^{-1} .

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

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

де λ_c крок навчання, β_c параметр пам'яті, τ_{jl}^c вектор значень зворотного розповсюдження помилки по відношенню до c_{jl}^ϕ .

Центри функцій належності першого рівня c_{jl}^ψ розподіляються рівновіддалено при ініціалізації, і вони можуть бути налаштовані за аналогічною процедурою:

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

де λ_ψ крок навчання, β_ψ параметр пам'яті, τ_{jl}^ψ вектор значень зворотного розповсюдження помилки по відношенню до c_{jl}^ψ .

Замість обчислення повного градієнта для всього набору даних, цей метод оновлює регульовані параметри моделі після обробки кожного вхідного шаблону. Це дозволяє пришвидшити обробку та використати шум обчислення вхідних градієнтів для досягнення кращих узагальнюючих властивостей. Також це дозволяю уникнути помилок округлення, які виникають, коли доводиться зберігати накопичений градієнт.

Ефект затримки прогнозованого сигналу можна мінімізувати який часто виникає у реальних даних, використовуючи спеціалізований фінансовий критерій у процедурі навчання:

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

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

де $y(k)$ та $y(k-1)$ – поточне та попереднє значення зовнішнього навчального сигналу, $\hat{y}(k)$ – прогнозоване значення та N розмір вибірки для навчання;
 $signal(k)$ функція недиференційована, тому ми використали гіперболічний тангенс як заміну і метод навчання набуває наступного вигляду:

$$\begin{cases} Q_{jl}^{-1}(k+1) = Q_{jl}^{-1}(k) + \lambda_Q \frac{\tau_{jl}^Q(k) \gamma sign(y(k)-y(k-1)) \tanh^2(\gamma(y(k)-y(k-1)))}{\eta_Q(k)}, \\ \eta_Q(k+1) = \beta_Q \eta_Q(k) + Tr(\tau_{jl}^{Q^T} \tau_{jl}^Q) \\ c_{jl}^\phi(k+1) = c_{jl}^\phi(k) + \lambda_c \frac{\tau_{jl}^c(k) \gamma sign(y(k)-y(k-1)) \tanh^2(\gamma(y(k)-y(k-1)))}{\eta_c(k)}, \\ \eta_c(k+1) = \beta_c \eta_c(k) + \tau_{jl}^{c^T} \tau_{jl}^c \end{cases},$$

На рис. 2 зображено узагальнену модель з декількома вихідними змінними.

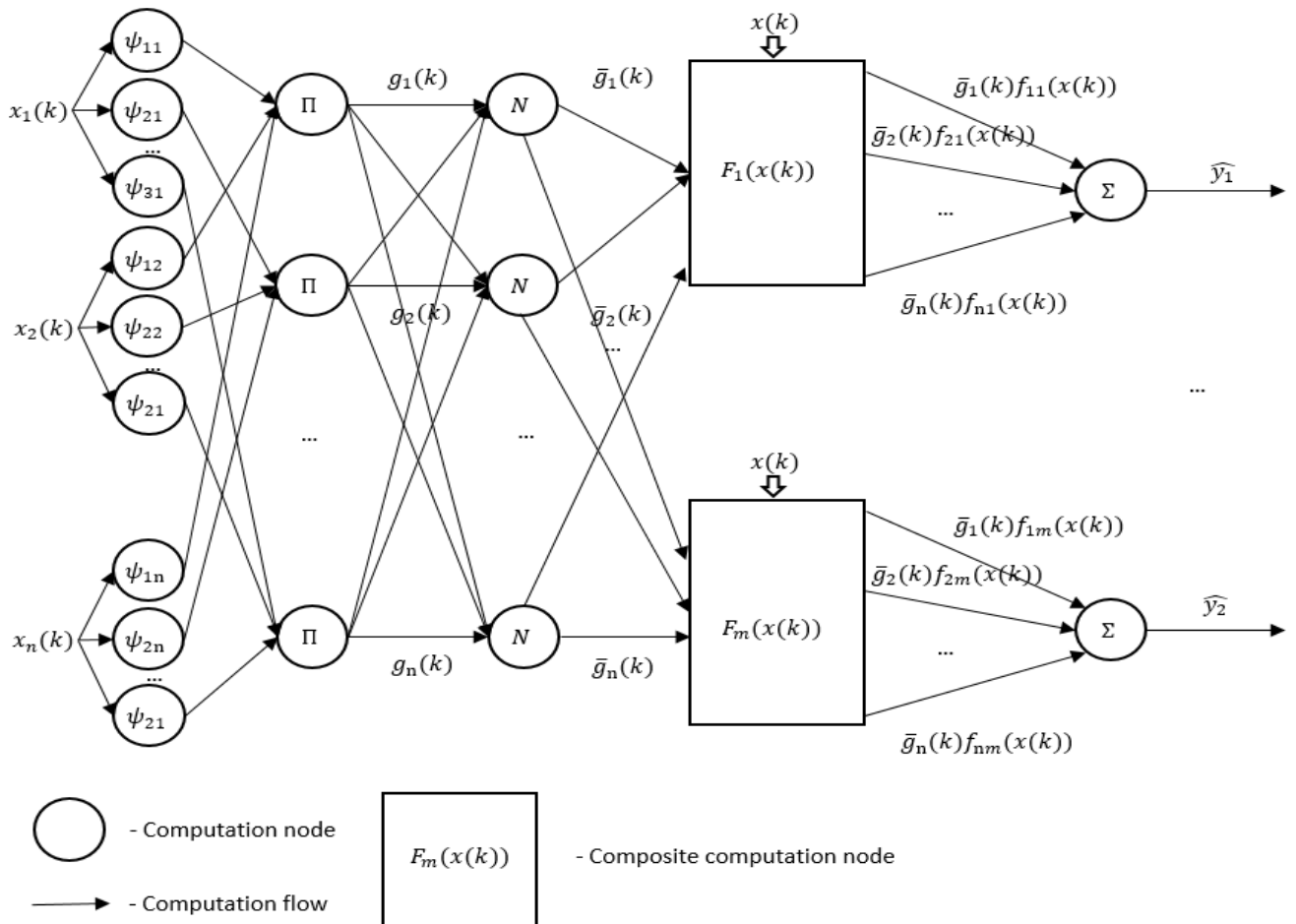


Рисунок 2 – Загальна архітектура моделі з декількома виходами

Ми маємо m вихідних змінних $\hat{y}(k) = (\hat{y}_1(k), \hat{y}_2(k), \dots, \hat{y}_m(k))^T$ замість лише однієї \hat{y} для кожного вхідного шаблону $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ і, відповідно, m вузлів четвертого шару $F = (F_1(k), F_2(k), \dots, F_m(k))^T$.

Запропонований метод узагальнює оптимізацію у просторі вагових коефіцієнтів P і в просторі параметрів багатовимірних гаусіанів для випадку багатьох змінних.

Центри c_{ajl}^ϕ налаштовуються згідно процедури:

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

де λ_c – крок навчання, β_c – параметр пам'яті, τ_{ajl}^c – вектор значень зворотного розповсюдження похибки по відношенню до c_{ajl}^ϕ .

Навчання матриць Q_{ajl} :

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

де λ_Q – крок навчання, β_Q – параметр пам'яті, τ_{ajl}^Q – матриця значень зворотного розповсюдження по відношенню до Q_{ajl} . Вектори η_{ac} та матриці η_{aQ} репрезентують згасаючі середні значення попередніх градієнтів.

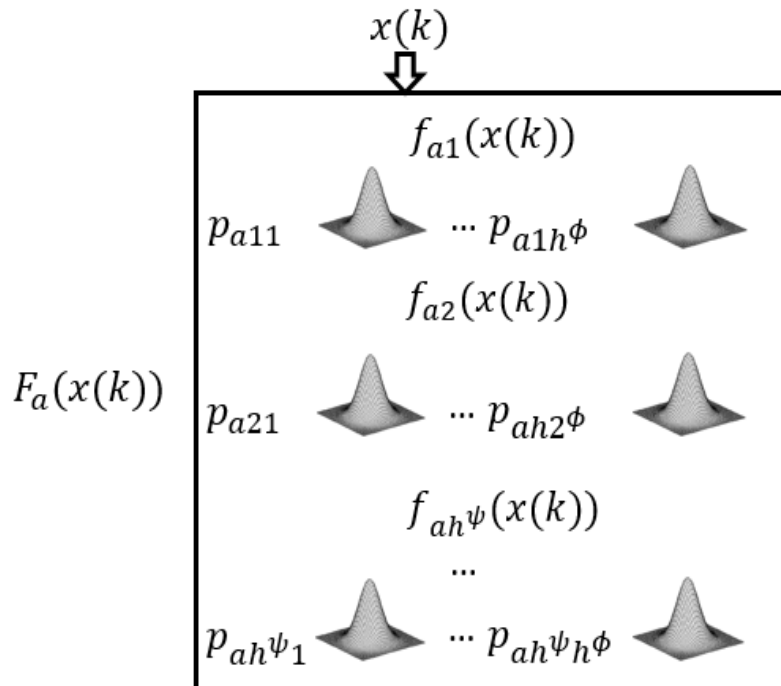


Рисунок 3 – Деталізована структура четвертого (консеквентного) шару

У **третьому розділі** представлено ансамбль нейро-фаззі моделей, який дозволяє зменшити похибку за рахунок зменшення варіативності, спростити вибір

гіперпараметрів та використати можливості сучасних багатоядерних процесорів за рахунок паралельного навчання складових моделей.

Ансамбль складається з моделей M_e . Загальна архітектура представлена на рис. 4.

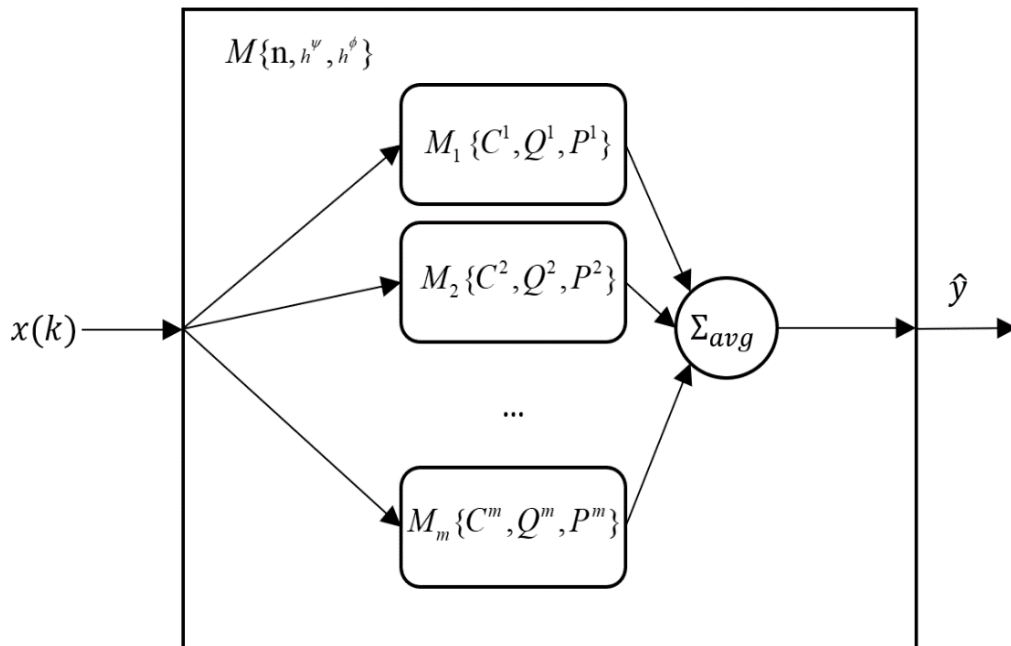


Рисунок 4 – Загальна архітектура запропонованого ансамблю моделей з регульованими параметрами

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.

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

$$\hat{y} = \frac{\sum_{e=1}^m \hat{y}_e}{m},$$

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.

Метод синтезу ансамблю показаний на рис. 5. Його метою є ініціалізація моделей членів перед навчанням і встановлення параметрів навчання. Результати показують, що і гіперпараметри β_c та 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 . Перед навчанням центри c_{ejl}^ϕ розміщуються на однаковій відстані, а матриці ідентичності Q_{ejl} ініціалізуються як одиничні. Після створення моделей ми можемо починати навчати їх паралельно.

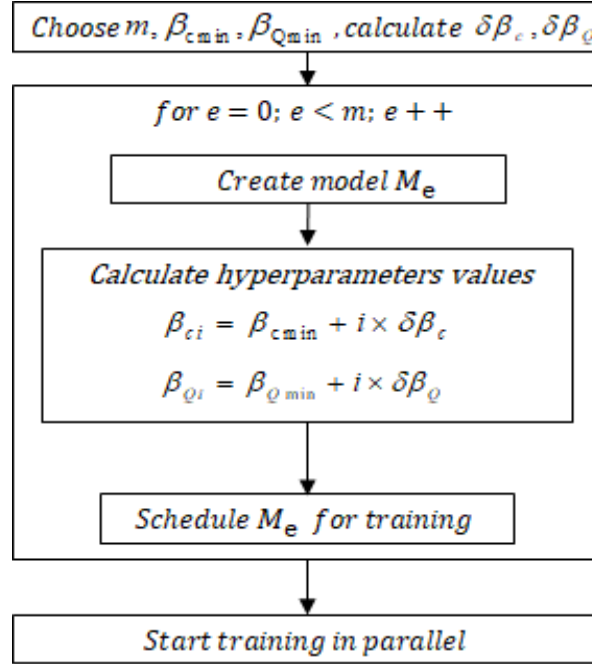


Рисунок 5 – Метод синтезу ансамблю моделей

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

$$\left\{ \begin{array}{l} c_{ejl}^\phi(k+1) = c_{ejl}^\phi(k) + \lambda_c \frac{\tau_{ejl}^c(k)e(k)}{\eta_c(k)} \\ \eta_{ec}(k+1) = \beta_c \eta_{ec}(k) + \tau_{ejl}^c T \tau_{ejl}^c \\ Q_{ejl}(k+1) = Q_{ejl}(k) + \lambda_Q \frac{\tau_{ejl}^Q(k)e(k)}{\eta_{eq}(k)} \\ \eta_{eQ}(k+1) = \beta_{Qe} \eta_{eQ}(k) + Tr \left(\tau_{ejl}^Q T \tau_{ejl}^Q \right) \end{array} \right. ,$$

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.

У **четвертому розділі** описано модифікований метод навчання нейро-фаззи моделей та синтезу ансамблів 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.

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 $m_1(t) = \frac{e_u - e_l}{2}$ та різницю як прообраз власної модової функції $h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t)$.
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:

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

5. Повторити процес та отримати наступні 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. 6.

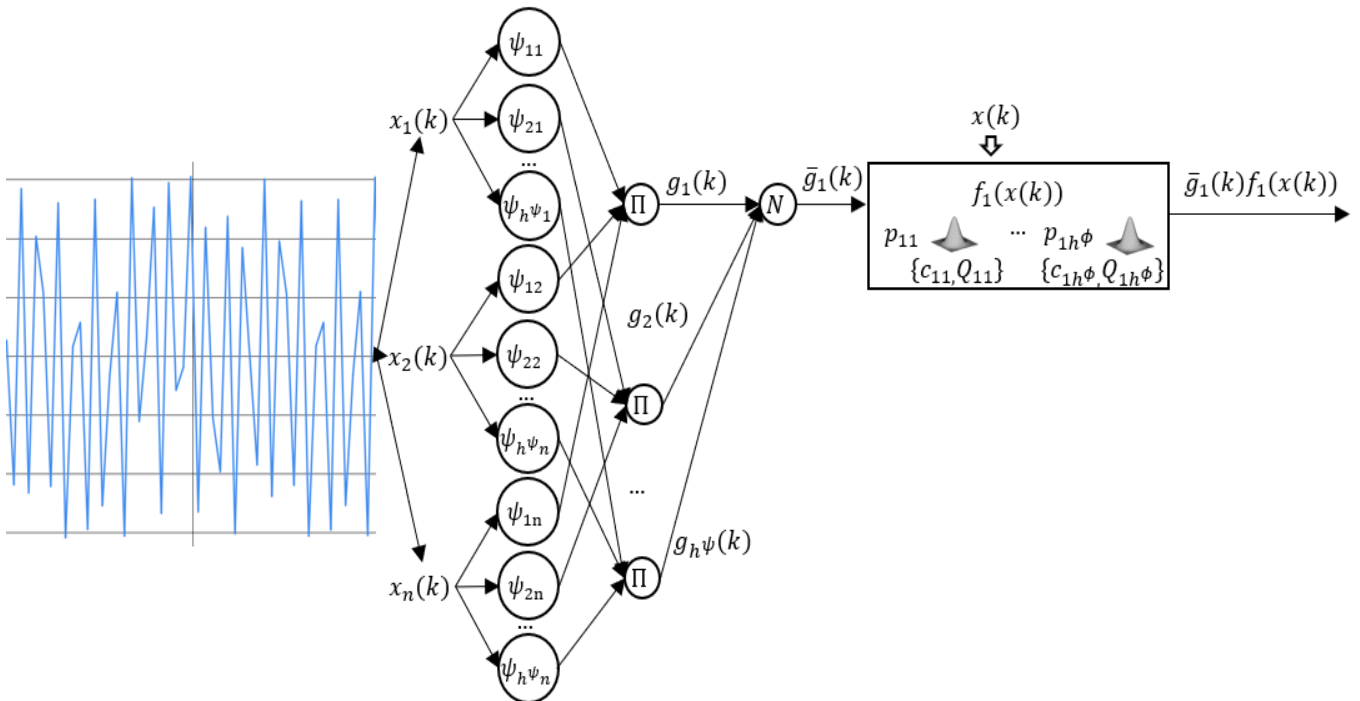


Рисунок 6 – Зразок навчання на власній модовій функції

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

У випадку ансамблів використано результати емпіричної модової декомпозиції для синтезу та навчання ансамблю як показано на рис. 7.

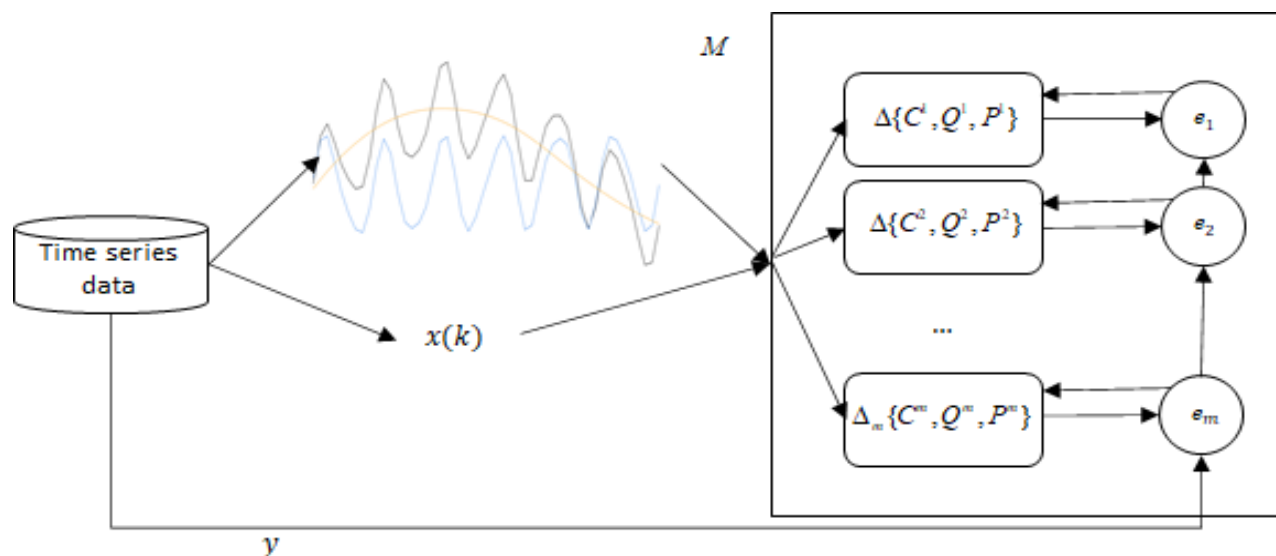


Рисунок 7 – Метод синтезу ансамблю моделей з використанням методу емпіричної модової декомпозиції

У п'ятому розділі описаний процес імітаційного моделювання розроблених методів та моделей. Розв'язані практичні задачі з управління ризиками у фінансових портфоліо на базі моделей прогнозування фондових індексів та управління ризиками архітектурного проектування програмного забезпечення.

Запропоновані методи та моделі були порівняні у різних задачах і наборах даних з відомими широко використовуваними штучними нейронними мережами та методами навчання – багаторівневим перцептроном на основі біполярної сигмоїдної функції активації з еластичним методом зворотного поширення помилки та методом Левенберга-Маркварта, обмеженою машиною Больцмана, опорно-векторними мережами та іншими.

В якості даних для експериментів і перевірки гіпотез були використані різноманітні синтетичні та реальні набори даних. Приклад результатів експерименту на часових рядах згенерованими рівняннями Макі-Гласса з додаванням білого шуму наведено у таблиці 1. Процес навчання проілюстровано на рис. 8.

Таблиця 1 – Порівняння результатів прогнозування в умовах зашумлених даних

Модель	Mackey-Glass time series		
	Час, ms	RMSE, %	MAPE, %
a) Запропонована мадель з навчанням першого шару $h^\phi = 4, \beta_c = 0.78, \beta_Q = 0.91$	2610	1.481	5.98
b) Запропонована мадель без навчання першого шару $h^\phi = 4, \beta_c = 0.75, \beta_Q = 0.91$	2373	3.01	7.64
c) Bipolar Sigmoid Network RBPR 50 epoch	7097	6.315	11.78
d) Bipolar Sigmoid Network RBPR 300 epoch	37442	2.72	7.33

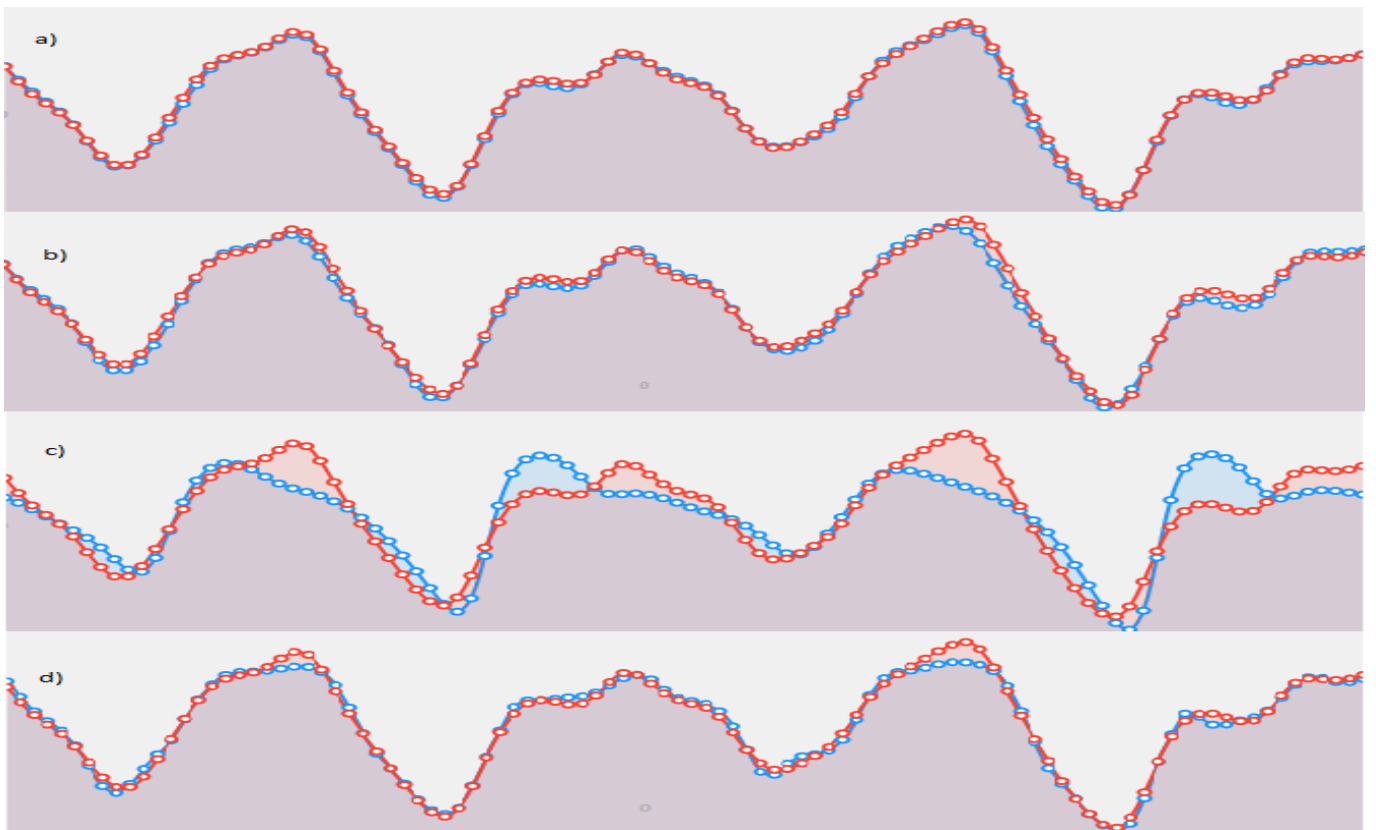


Рисунок 8 – Експериментальні результати у задачі прогнозування часових рядів

У **висновках** сформульовано наукові та практичні результати, що одержані у дисертаційній роботі.

У **додатку** наведено акти про впровадження результатів дисертаційної роботи.

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 нестаціонарних 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;
- розроблено ансамбль нейро-фаззі моделей та метод його синтезу для зменшення похибки та спрощення процедури селекції гіперпараметрів;
- розроблено метод навчання, який використовує метод емпіричної модової декомпозиції для зниження шумів;

- 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 портфоліо. 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 «Нейромеревеві методи обчислювального інтелекту» та «Штучні нейронні мережі: архітектура, навчання, застосування».

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ 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.

АНОТАЦІЯ

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

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

На сьогоднішній час є актуальними задачі розробки нових методів та моделей аналізу даних в умовах нестаціонарності та високочастотної динаміки. Такі моделі повинні відрізнятися невеликим розміром та обчислювальною ефективністю.

Предметом дослідження є методи та моделі інтелектуального аналізу даних на базі м'яких обчислень у задачах оперативного аналізу ризиків.

Розглянуто існуючі методи та підходи в задачах прогнозування, ідентифікації та аналізу, їх недоліки та можливості для покращення.

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.

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 даних.

Набув подальшого розвитку метод побудови ієрархічних знання-орієнтованих нечітких систем, 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.

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

АННОТАЦІЯ

Власенко А.Н. Методы и модели интеллектуализации процессов оперативного анализа рисков на основе мягких вычислений – на правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2021.

Предметом исследования являются методы и модели интеллектуального анализа данных на базе мягких вычислений в задачах оперативного анализа рисков.

Рассмотрены существующие методы и подходы в задачах прогнозирования, идентификации и анализа, их недостатки и возможности для улучшения.

Предложена пятислойная гибридная нейро-фаззі модель с многомерными гауссианами в консеквентном слое и быстрый метод ее обучения с вариантами на базе квадратичной ошибки и специализированного критерия обучения, которая характеризуется высокой точностью, скоростью обработки данных и быстрым обучением, что достигается снижением количества настраиваемых параметров модели.

Предложен ансамбль гибридных нейро-фаззи моделей с многомерными гауссианами в консеквентном слое и метод его синтеза, который характеризуется уменьшенной ошибкой обучения, улучшенными обобщающими возможностями и упрощенной процедурой подбора гиперпараметров.

Усовершенствован метод обучения нейро-фаззи моделей и их ансамблей путем использования метода декомпозиции эмпирических мод в качестве инструмента снижения шума.

Получил дальнейшее развитие метод построения иерархических знание-ориентированных нечетких систем путем включения оценок вероятностей в процесс логического вывода и использования нейро-фаззи моделей как компонентов иерархической структуры, который позволяет обрабатывать маловероятные критические значения в задачах оперативного анализа рисков.

Проведен ряд имитационных экспериментов на основе тестовых и реальных данных, результаты которых подтверждают целесообразность применения предложенного подхода для решения задач интеллектуализации процессов оперативного анализа рисков. На базе разработанных методов решена практическая задача оперативного управления финансовым портфолио.

Ключевые слова: нейро-фаззи модели, градиентные методы обучения, метод декомпозиции эмпирических мод, ансамбли нейро-фаззи моделей

ABSTRACT

Vlasenko O.M. Soft Computing methods and models of risk analysis processes intellectualization. – The manuscript.

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.

This research is devoted to the soft computing methods and models of data mining applications in the tasks of operational risk analysis.

A novel five-layer hybrid neuro-fuzzy model with multidimensional Gaussians in the consequent layer and a fast training method with variants based on standard quadratic error and a specialized criterion are proposed.

An ensemble of hybrid neuro-fuzzy models with multidimensional Gaussians in a consequent layer and a method of its synthesis are proposed.

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

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.

Key words: neuro-fuzzy models, gradient optimization, empirical mode decomposition, neuro-fuzzy ensemble models.