

## ДОСЛІДЖЕННЯ НЕЧІТКИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ШВИДКОЗМІННИХ НЕСТАЦІОНАРНИХ ЧАСОВИХ РЯДІВ

*Реалізовано інформаційно-аналітичну систему прогнозу нестационарних швидкозмінних часових рядів на основі нечіткої нейронної мережі Такагі-Сугено-Канга для задачі прогнозування курсу акцій підприємства. Розроблено алгоритм автоматичного формування нечіткої бази правил для нечіткої нейронної мережі. Запропоновано використання чисельних алгоритмів для розрахунку частинних похідних виходу правил по параметрам функції належності. Проаналізовано роботу розробленої мережі при зміні розміру навчальної вибірки та кількості правил для нечіткого логічного виводу.*

**Ключові слова:** нейронні мережі, алгоритм, часові ряди.

*Реализована информационно-аналитическая система прогноза нестационарных быстроизменяющихся временных рядов на основе нечеткой нейронной сети Такаги-Сугено-Канга для задачи прогнозирования курса акций предприятия. Разработан алгоритм автоматического формирования нечеткой базы правил для нечеткой нейронной сети. Предложено использование численных алгоритмов для расчета частных производных выхода правил по параметрам функции принадлежности. Проанализирована работа разработанной сети при изменении размера обучающей выборки и количества правил для нечеткого логического вывода.*

**Ключевые слова:** нейронные сети, алгоритмы, временные ряды.

*The implemented fuzzy neural Takagi-Sugeno-Kang network for prediction of non stationary dynamic time series is researched. There are proposed the algorithm of automatic generation of rules and numerical computing for partial derivations of rules outputs. The high significance of results is developed in prediction dynamics of changes in financial time series.*

**Key words:** neural networks, algorithms, time series.

### Вступ

Оскільки більшість досліджуваних фінансових, економічних або технічних реальних процесів є нестационарними, особливої актуальності набуває вирішення задач моделювання і прогнозування нестационарних процесів. До таких процесів відносяться, наприклад, процеси формування цін на біржах, значні випадкові коливання курсів валют. В останній час бурного розвитку отримали адаптивні методи короткострокового прогнозування. Перевага їх полягає у тому, що це само регульовані моделі, й у разі появи нових даних прогнози оновлюються із мінімальною затримкою в той час як економетрична модель з постійними параметрами буде екстраполувати істотно застарілі залежності. Адаптація до нових даних є перевагою нейронних мереж з їх здатністю до самонавчання. Але звичайні чіткі нейронні мережі перенавчаються дуже повільно, тому в даний час активно розвиваються системи штучного інтелекту, які базуються на нечітких множинах, та системах нечіткого логічного висновку. З іншого боку система з нечіткою логікою має ряд недоліків: база правил, яку задає експерт може бути неповною та суперечливою, функції належності можуть бути неадекватними. Всі ці недоліки відсутні у нечітких нейронних мережах, які, з одного боку, можуть адаптувати параметри функцій належності до нових даних, з іншого боку, навчаються на 2 порядку швидше.

Використання систем нечіткого логічного висновку і побудованих на їх основі нечітких нейронних мереж дозволило вирішувати багато задач: прогнозування, класифікації, кластерного аналізу, розпізнавання образів. Але в відомих нечітких нейронних мережах є певні моменти, які потребують доробки. Так, база правил або формується експертом, або використовується повна база правил, що є надмірним та не завжди відповідає дійсності. Алгоритми навчання нейронної мережі розроблені для Гауссовської функції належності [2] або мають емпіричний характер.

### Задачі дослідження

Метою роботи є удосконалення та дослідження нечітких нейронних мереж в задачах прогнозування нестационарних часових рядів з динамікою, що випадково змінюється.

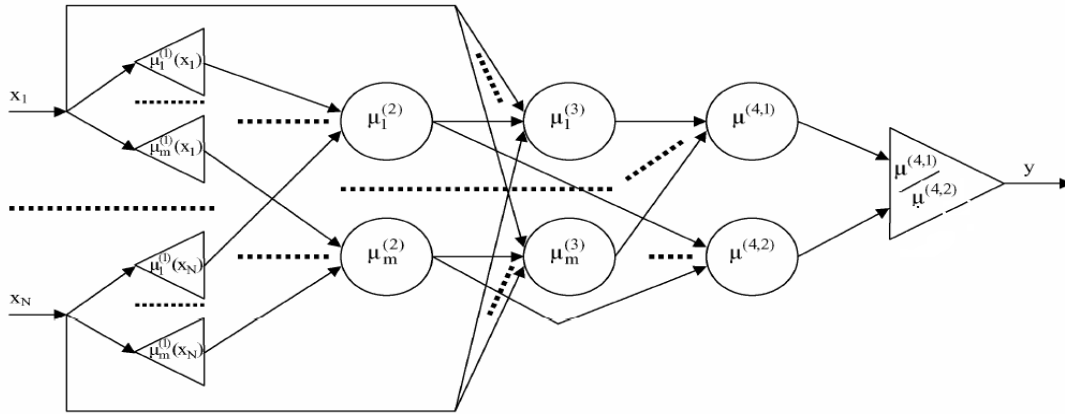
Для цього необхідно розв'язати наступні задачі:

- обрати структуру нейронної мережі для прогнозування часових рядів, які можна використати при різких нестационарних процесах та при випадкових впливах;
- розробити алгоритм автоматичного формування бази правил;
- розробити інформаційно-аналітичну систему прогнозу;
- провести дослідження параметрів мережі в задачі прогнозування.

**Нейро-нечітка мережа Такагі-Сугено-Канга**

Для розв'язання задачі прогнозування була обрана нечітка нейронна мережа Такагі-Сугено-Канга, яка дає найкращу точність прогнозування [1,2,3].

Мережа виконує нечітке виведення Сугено із використанням N змінних  $x_j$  та m правил:



**Рис. 1.** Мережа Такагі-Сугено-Канга

Перший шар виконує окрему фазифікацію кожної вхідної змінної  $x_j$ ,  $j = 1, 2, \dots, N$ , визначаючи для кожного k-го правила значення функції належності:

$$\mu_k^{(1)}(x_j) = \frac{1}{1 + \left( \frac{x_j - c_{jk}}{\sigma_{jk}} \right)^{2b_{jk}}}, \quad (1)$$

параметри якої  $c_{jk}$ ,  $\sigma_{jk}$ ,  $b_{jk}$  підлягають адаптації в процесі навчання.

Другий шар виконує агрегування функцій належності вхідних змінних до термів антецедентів нечітких правил, визначаючи приналежність вхідного вектора до k-го правила (рівень активації правила):

$$\mu_k^{(2)} = \prod_{j=1}^N \mu_k^{(1)}(x_j, c_{jk}, \sigma_{jk}, b_{jk}), \quad k = 1, 2, \dots, m. \quad (2)$$

Третій шар генерує значення функцій консеквентів нечітких правил з урахуванням рівнів активації правил:

$$\mu_k^{(3)} = \mu_k^{(2)} \left( w_0^k + \sum_{j=1}^N w_j^k x_j \right). \quad (3)$$

Це параметричний шар, у якому адаптації підлягають лінійні ваги  $w_{kj}$ ,

$$k = 1, 2, \dots, m, \quad j = 1, 2, \dots, N.$$

Четвертий шар агрегує m правил виведення (перший нейрон) і генерує нормалізуючий сигнал (другий нейрон):

$$\mu^{(4,1)} = \sum_{k=1}^m w_k^{(4)} \mu_k^{(3)}, \quad \mu^{(4,2)} = \sum_{k=1}^m \mu_k^{(3)}. \quad (4)$$

П'ятий (вихідний) шар містить один нейрон і здійснює нормалізацію, формуючи вихідний сигнал  $y$ :  $y(x_1, x_2, \dots, x_N) = \mu^{(4,1)} / \mu^{(4,2)}$ .

Мережа Такагі-Сугено-Канга містить тільки два параметричних шари (перший і третій), параметри яких уточнюються в процесі навчання. Параметри першого шару будемо називати нелінійними параметрами, а параметри третього шару – лінійними вагами. Якщо прийняти, що в конкретний момент часу параметри умов

$P_k$ : Якщо  $x_1 \in A_1^{(k)}$  та  $x_2 \in A_2^{(k)}$  та ... та  $x_N \in A_N^{(k)}$ , тоді  $y_k = w_0^k + \sum_{j=1}^N w_j^k x_j$ ,

де  $A_j^k$  – нечіткий терм, до якого має належати j-та вхідна змінна, щоб активізувати k-те правило,  $w_j^k$  – ваговий коефіцієнт,  $k = 1, 2, \dots, m$ .

Мережа Такагі-Сугено-Канга (рис.1) складається з п'яти шарів.

зафіксовані, то функція  $y(x_1, x_2, \dots, x_N)$  є лінійною щодо змінних  $x_1, x_2, \dots, x_N$ .

При наявності N вхідних змінних кожне правило формує N+1 змінних  $w_j^k$  лінійної залежності  $y_k(x_1, x_2, \dots, x_N)$ . При m правилах виведення це дає  $m(N+1)$  лінійних параметрів мережі. У свою чергу, кожна функція належності використовує три параметри  $c_{jk}, \sigma_{jk}, b_{jk}$ , що підлягають адаптації. Якщо прийняти, що кожна змінна  $x_j$ , характеризується власною функцією належності, то при m правилах виведення ми одержимо 3m нелінійних параметрів. У сумі це дає  $m(4N+1)$  лінійних і нелінійних параметрів, значення яких повинні підбиратися в процесі навчання мережі.

На практиці для зменшення кількості параметрів, що адаптуються, оперують меншою кількістю незалежних функцій належності для окремих змінних, керуючись правилами, у яких комбінуються функції належності різних змінних. Якщо прийняти, що кожна змінна  $x_j$  має  $k_j$  різних функцій належності, то максимальна кількість правил, яку можна створити при їхньому комбінуванні, складе:  $m = k_j$ . У такий спосіб сумарна кількість нелінійних параметрів мережі при m правилах висновку зменшується з 3m у загальному випадку до  $3Nm^{1/N}$ . Кількість лінійних параметрів при подібній модифікації залишається без змін, тобто  $m(N+1)$ .

Завдання мережі полягає в такому відображенні пар даних  $\langle x, y \rangle$ , щоб для вхідного вектора  $x^s$  розрахункове значення вихідної ознаки  $y^{s*} = y(x^s)$  не сильно відрізнялося б від співставленого  $x^s$  фактичного значення цільової ознаки  $y^s$ .

Навчання мережі засновано на мінімізації цільової функції:

$$E = \frac{1}{2} \sum_{s=1}^S (y^{s*} - y^s)^2, \quad (5)$$

де S – кількість навчаючих пар  $\langle x^s, y^s \rangle$ , та є контрольованим.

**Автоматичне формування бази правил**

Так як, база правил або формується експертом і не завжди відповідає дійсності, або використовується повна

база правил, що є надмірним, то було розроблено алгоритм автоматичного формування бази правил.

На першому етапі кожному прикладу з вибірки ставиться у відповідність окреме правило. Для цього для кожного прикладу  $(x_1^{(k)}, x_2^{(k)}, y^{(k)})$ ,  $(k=1..K)$ , визначаються ступені належності заданих значень змінних до відповідних нечітких множин. Після чого кожному навчальному прикладу ставляться у відповідність ті нечіткі множини, ступені відповідності до яких у відповідних значень змінних цього прикладу є максимальними. Сформована таким чином множина правил складає початкову базу правил.

Оскільки початкова сформована база правил ймовірно є надмірною, а також може містити суперечливі правила з однаковою лівою частиною та різною правою. Оптимізацію правил можливо зробити як на основі експертних оцінок правил, так і шляхом адаптації до існуючих навчальних даних. Підрахунок рейтингів проходить наступним чином, всі приклади з навчальної

вибірки  $(x_1^{(k)}, x_2^{(k)} \dots x_m^{(k)}, y^{(k)})$ ,  $(k=1..K)$  подаються кожному правилу. В результаті для кожного правила визначається його рейтинг:

$$r_i = \sum_{k=1}^K \mu_{Ai1}(x_1^{(k)}) * \mu_{Ai2}(x_2^{(k)}) \dots \mu_{Aim}(x_m^{(k)}) * \mu_{Bi}(y^{(k)}), i=1..n \quad (6)$$

Після підрахунку рейтингів з бази правил видаляються правила з найменшим рейтингом. При цьому в першу чергу видалення проводиться по групам правил, що мають однакові передумови та різні виводи, тобто різні функції належності змінної виходу. Такі правила суперечать одне одному, з групи подібних правил потрібно залишити правила з найбільшим рейтингом. Таким чином вирішується суперечливість в базі правил, а також значно зменшується кількість правил. Правила, що залишилися формують вихідну базу правил.

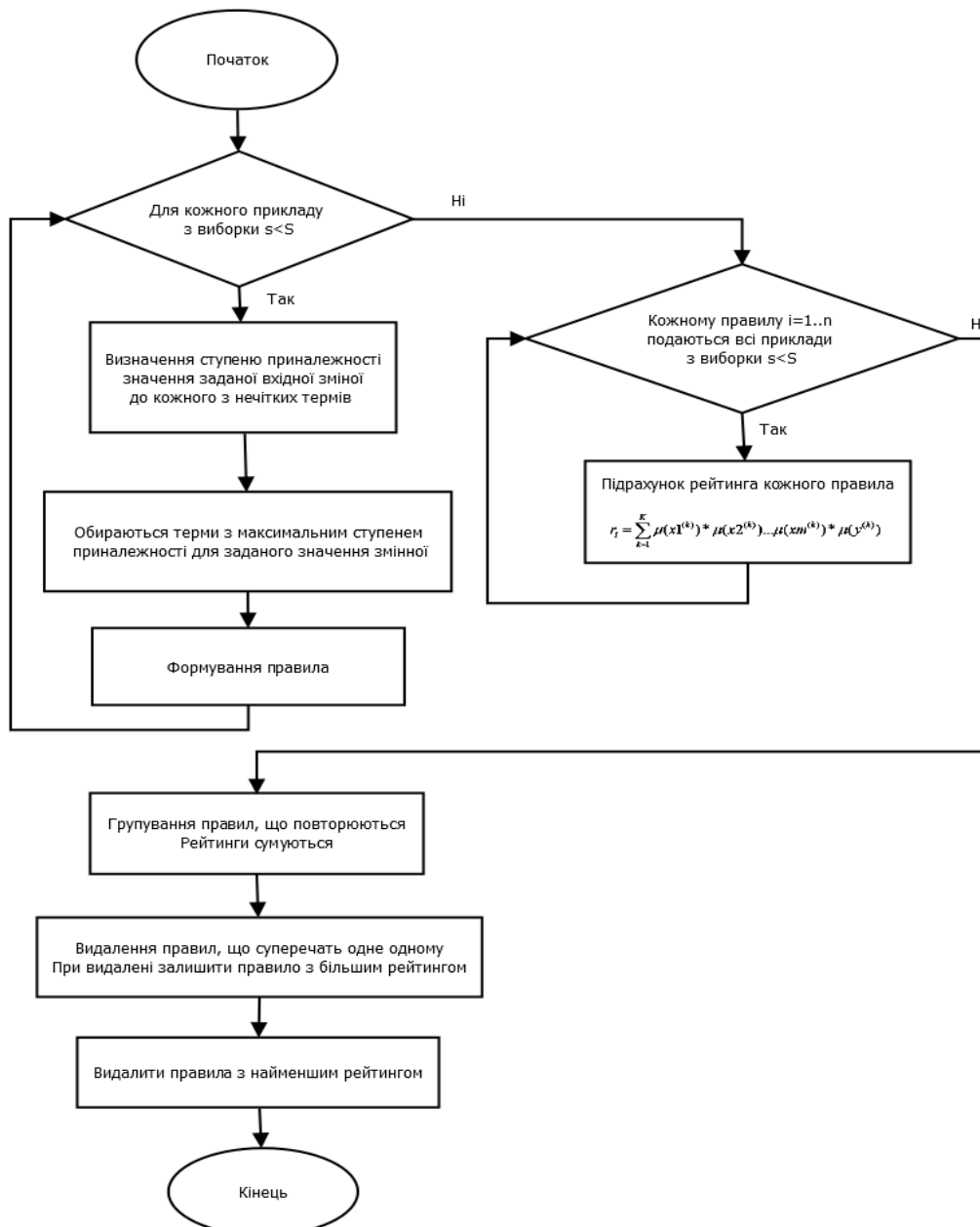


Рис 2. Блок-схема алгоритму генерації бази правил

**Гібридний метод навчання нейронної мережі**

Гібридний метод навчання застосовується для мережі Такагі-Сугено-Канга й інших нейро-нечітких мереж, подібних їй. У гібридному методі підлягаючі адаптації параметри розподіляються на дві групи: одна група складається з лінійних параметрів  $w_j^k$  третього шару, а інша група – з параметрів нелінійних функцій належності першого шару. Уточнення параметрів проводиться в два етапи.

На першому етапі при фіксації визначених значень параметрів функцій належності (у першому циклі – це значення, отримані в результаті ініціалізації) шляхом розв'язку системи лінійних рівнянь розраховуються лінійні параметри  $w_j^k$ . При відомих значеннях функцій належності залежність  $y(x_1, x_2, \dots, x_N)$  можна подати в лінійній формі:

$$y(x_1, x_2, \dots, x_N) = \sum_{k=1}^m w_k' \left( w_0^k + \sum_{j=1}^N w_j^k x_j \right), w_k' = \left( \sum_{p=1}^m \left( \prod_{j=1}^N \mu_p^{(1)}(x_j) \right) \right)^{-1} \prod_{j=1}^N \mu_k^{(1)}(x_j). \tag{7}$$

При  $S$  екземплярах навчаючої вибірки  $\langle x^s, y^s \rangle, s = 1, 2, \dots, S$ , одержимо систему з лінійних рівнянь у матричній формі:

$$Aw = y, \tag{8}$$

де  $w = (w_1, w_2, \dots, w_m)^T, w_k = (w_{k0}, w_{k1}, \dots, w_{kN})^T, y = (y_1, y_2, \dots, y_S)^T, A = (a_1, a_2, \dots, a_S)^T, a_s = (a_{s1}, a_{s2}, \dots, a_{sm})^T$ . Розмірність матриці  $A$  дорівнює  $S(N+1)m$ , при цьому звичайно кількість рядків  $S$  значно більше кількості стовпців  $(N+1)m$ . Розв'язок цієї системи:  $w = A^{-1}y$ . Матрицю  $A$  обертаємо тільки 1 раз, для наступних епох використовуємо адаптивні алгоритми.

На другому етапі після фіксації значень лінійних параметрів  $w_j^k$  розраховуються вихідні сигнали мережі  $y = \{y^{s*}\} (s = 1, 2, \dots, S)$ :  $y = Aw$ , а слідом за ними – вектор помилки  $e = \{e^s\}, e^s = y^{s*} - y^s$ . Сигнали помилок направляються через підключену мережу за напрямком до входу мережі (зворотне поширення) аж до першого шару, де можуть бути розраховані компоненти градієнта цільової функції щодо конкретних параметрів  $c_{jk}, \sigma_{jk}, b_{jk}$ .

Після формування вектора градієнта параметри уточнюються з використанням одного з градієнтних методів навчання. Якщо застосовується найпростіший метод найшвидшого спуску, то відповідні формули адаптації приймають форму:

$$c_{jk}(t+1) = c_{jk}(t) - \alpha_c \frac{\partial E(t)}{\partial c_{jk}}, \sigma_{jk}(t+1) = \sigma_{jk}(t) - \alpha_\sigma \frac{\partial E(t)}{\partial \sigma_{jk}}, b_{jk}(t+1) = b_{jk}(t) - \alpha_b \frac{\partial E(t)}{\partial b_{jk}}, \tag{9}$$

де  $t$  – номер ітерації,  $\alpha_c, \alpha_\sigma, \alpha_b$  – коригувальні прирости (кроки навчання).

Після уточнення нелінійних параметрів знову запускається процес адаптації лінійних параметрів (перший етап) і нелінійних параметрів мережі (другий

етап). Цей цикл повторюється аж до стабілізації всіх параметрів процесу навчання. Остаточний вид формул настроювання параметрів залежить як від використуваного визначення функції помилки на виході мережі, так і від виду функції належності. Так при використанні функцій Гауса як функцій належності та середньоквадратичної помилки, часткові похідні цільової функції мають дуже складний вигляд [2].

Але вираз  $\frac{\partial w_p^i}{\partial c_{jk}}$  у [1] отримано тільки для

гауссовської функції належності і він дуже складний для обчислення. Тому у роботі пропонується його отримання чисельним алгоритмом У алгоритмі кожному параметру функції належності по черзі дається приріст  $\Delta c_{jk} = c_{jk} + \Delta c_{jk}$ . Потім рахується вихід правил при збільшеному параметрі та при звичайних параметрах. Функції належності. Частинна похідна має вигляд

$$\frac{\partial w_p^i}{\partial c_{jk}} = \frac{w_p^i (c_{jk}) - w_p^i (c_{jk})}{\Delta c_{jk}}. \tag{10}$$

Чисельний розрахунок  $\frac{\partial w_p^i}{\partial c_{jk}}$  дав можливість використання будь якої функції належності

Незважаючи на складну структуру наведених у [2] формул, що виражають компоненти вектора градієнта, вони дозволяють аналітично визначити величини, необхідні для уточнення параметрів нечіткої мережі.

При практичній реалізації гібридного методу навчання нечітких мереж домінуючим фактором їхньої адаптації вважається перший етап, на якому ваги  $w_j^k$  підбираються з використанням псевдоінверсії за один крок. Для зрівноважування його впливу другий етап (підбір нелінійних параметрів градієнтним методом) багаторазово повторюється в кожному циклі.

Гібридний метод є одним з найбільш ефективних способів навчання нейро-нечітких мереж. Його головна особливість полягає в розподілі процесу навчання на два відособлених у часі етапи. На кожному етапі уточнюється тільки частина параметрів мережі. Якщо взяти до уваги, що обчислювальна складність кожного методу оптимізації пропорційна (нелінійно) кількості параметрів, то зменшення розмірності задачі оптимізації істотно скорочує кількість математичних операцій і збільшує швидкість виконання методу. Завдяки цьому гібридний метод є значно більш ефективним, ніж звичайний градієнтний метод фронтального типу, відповідно до якого уточнення всіх параметрів мережі робиться паралельно й одночасно.

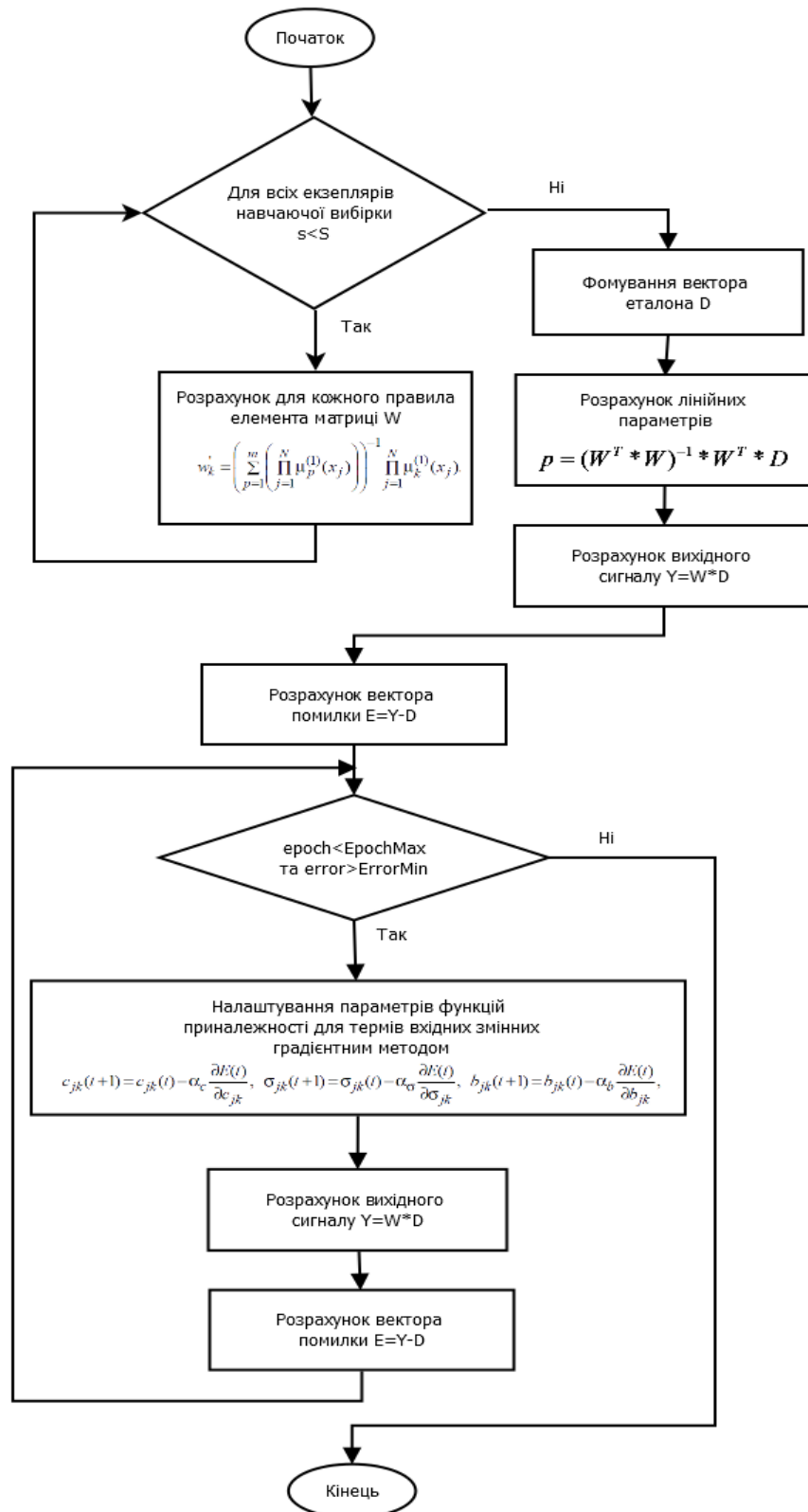


Рис. 3. Блок-схема гібридного алгоритму навчання мережі

Для дослідження задач прогнозування нестационарних часових рядів була розроблена інформаційна аналітична система у середовищі Java. Дані на фондових біржах змінюються щохвилини, тому важливо мати останні актуальні ціни на акції підприємства. Більшість мережевих порталів надає актуальну інформацію про динаміку зміни котирувань акцій в реальному часі. Дані надаються зазвичай у текстовому вигляді або у формі

таблиць. Для отримання актуальних даних обрано інтернет портал <http://trade.bo.ua/ticker>. Він містить всю історію змін цін акцій підприємства «Укрнафта». Програма кожного разу при запуску та за розкладом отримує оновлені дані з порталу, вони зберігаються на жорсткому диску користувача та можуть бути використані для подальшої роботи та прогнозу.

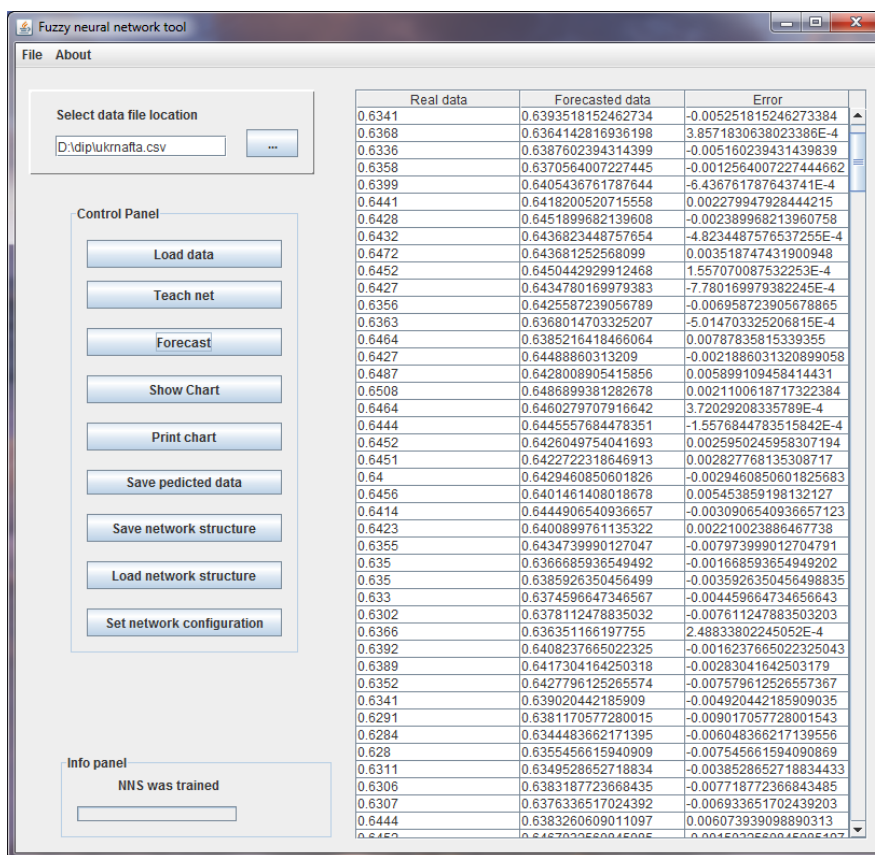


Рис. 4. Головне вікно програми

Розглядалися часові ряди курсів акцій підприємства «Укрнафта» за 2010-2011 роки. Дослідження проводилося за допомогою розробленого програмного забезпечення. Результати дослідження порівнювалися з результатами, отриманими за допомогою адаптивної нейро-нечіткої мережі ANFIS. Для порівняння обрано структуру мережі ANFIS з чотирма входами та трьома

нечіткими термами. Кількість точок в навчальній вибірці було взято 250, в перевірочній вибірці також 250.

Виконавши навчання мереж та провівши прогнозування, можна зробити висновок, що обидві мережі TSK та ANFIS показали гарні результати. На рисунках зображено результати прогнозу за допомогою мережі TSK та мережі ANFIS.

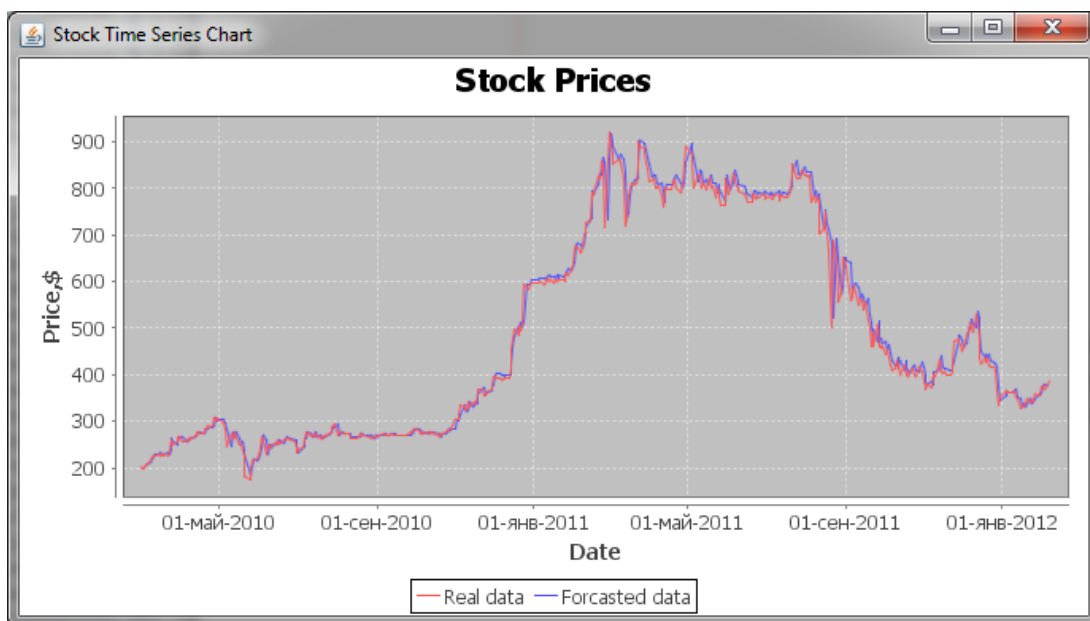


Рис. 5. Результати прогнозу за допомогою мережі TSK

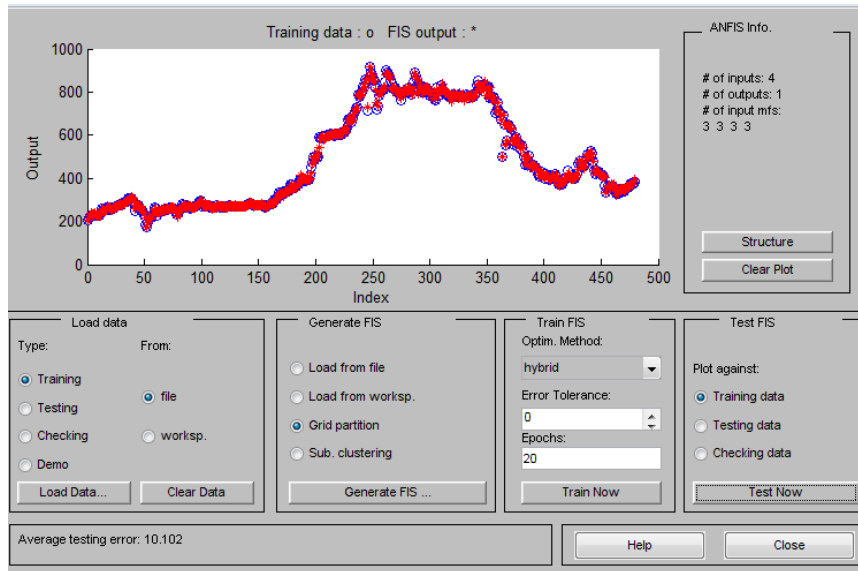


Рис. 6. Результати прогнозу за допомогою мережі ANFIS

Критерієм оцінки якості прогнозу було обрано середньоквадратичну помилку СКП

$$СКП = \frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i)^2 \quad (10)$$

де  $\hat{x}_i$  – прогноз на  $i$ -ий момент часу, зроблений в момент часу  $(i - 1)$ .

В таблиці 1 наведено значення середньоквадратичної помилки при прогнозуванні для навчальної та перевірконої вибірки мереж TSK та ANFIS.

Таблиця 1

Отримані значення СКП

| Назва нечіткої нейронної мережі | СКП для навчальної вибірки | СКП для перевірконої вибірки |
|---------------------------------|----------------------------|------------------------------|
| TSK                             | 0.05                       | 0.08                         |
| ANFIS                           | 0.248                      | 0.343                        |

З таблиці 1 видно, що реалізована нейро-нечітка мережа TSK показала кращі результати у порівнянні з ANFIS, похибка при прогнозуванні менша.

**Дослідження чутливості мережі TSK до довжини навчальної вибірки**

У всіх експериментах варіювалася довжина навчальної вибірки, вона становила 50, 100, 200, 300 і 400 точок. На рисунку 7 наведено графік залежності СКП від довжини навчальної вибірки.

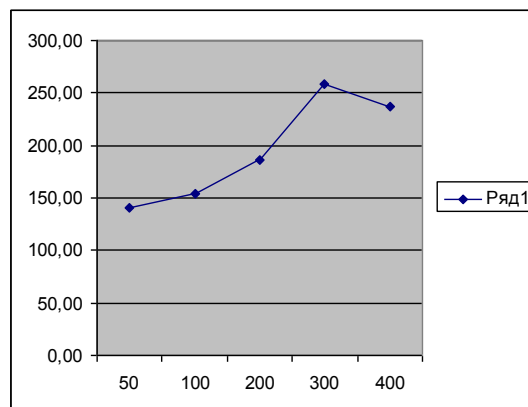


Рис. 7. Залежність СКП ( $\times 10^{-4}$ ) при прогнозуванні на перевірконій вибірці від зміни розміру навчальної вибірки

На рисунку 7 зображено графік критеріїв СКП на перевірконій вибірці. Як видно з графіку, при збільшенні навчальної вибірки збільшуються значення СКП, вона досягає свого максимального значення при довжині навчальної вибірки в 400 точок. Далі графіки починають спадати. Можна зробити висновок, що оптимальне

значення довжини навчальної вибірки становить 50 точок.

**Дослідження чутливості мережі TSK до кількості правил**

На рисунку 8. зображена залежність критерію СКП при прогнозуванні на перевірконій вибірці.

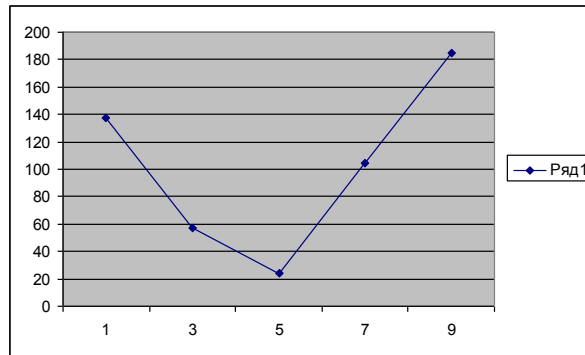


Рис. 8. Залежність СКП ( $\times 10^{-4}$ ) при прогнозуванні на перевіірочній вибірці від зміни кількості правил

Як видно з цього графіку, при збільшенні кількості правил спочатку досягається мінімум при кількості правил  $M \approx 5$ , а потім помилка збільшується.

Отже, можна зробити висновок, що нечіткі нейронні мережі є потужним інструментом для вирішення задач прогнозування динамічних нестационарних процесів. Переваги цих мереж полягають у швидкому навчанні, гнучкості використання та точності прогнозу.

**Висновки:** реалізовано інформаційно-аналітичну систему прогнозу нестационарних швидкозмінних часових рядів на основі нечіткої нейронної мережі Такагі-Сугено-Канга для задачі прогнозування курсу акцій підприємства;

– розроблено алгоритм автоматичного формування нечіткої бази правил для нечіткої нейронної мережі;

– запропоновано використання чисельних алгоритмів для розрахунку частинних похідних виходу правил по параметрам функції належності;

– проведено порівняння розробленої мережі Такагі-Сугено-Канга та нечіткої мережі ANFIS. Встановлено, що точність прогнозу за допомогою нечіткої мережі TSK краща у порівнянні з мережею ANFIS.

– проаналізовано роботу розробленої мережі при зміні розміру навчальної вибірки та кількості правил для нечіткого логічного виводу. Встановлено, що достатній обсяг навчальної вибірки до 100 прикладів а бази правил 5-6.

## ЛІТЕРАТУРА

1. Takagi T., Sugeno M. Fuzzy identification of system and its application for modeling and control. IEEE Trans. Systems, Man and Cybernetics. 1995 V.15 N1P.116-132.
2. Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах.-К., „Слово», 2008.- 333 с.
3. Бодянский С.В., Винокурова О.А.. Интеллектуальная обработка данных на основе гибридной вейвлет-нейро-фаззи системы на адаптивных W - нейронах. // Наукові праці: Науково- методичний журнал. – Миколаїв: Збірка Чорноморського державного університету ім П.Могили 2009.Вип 104.Т.117. –С88-98.

© Кравець І. О., Тімошин А. Ю., 2012

Дата надходження статті до редколегії 03.05.2012 р.

**КРАВЕЦЬ І. О.** – Чорноморський державний університет імені Петра Могили, м. Миколаїв, Україна.

**ТІМОШИН А. Ю.** – Чорноморський державний університет імені Петра Могили, м. Миколаїв, Україна.