

УДК 004.891.3

Притула К. В.,
Симоненко А.В.

ВИБІР МЕТОДІВ ДЛЯ РЕАЛІЗАЦІЇ СИСТЕМИ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ (КУРСА ВАЛЮТ) З ВИКОРИСТАННЯМ НЕЙРОМЕРЕЖ

Національний Технічний Університет України «КПІ»

Розглянуто методи для реалізації системи прогнозування часових рядів. Проведено порівняння з методом групового врахування аргументів. Дану систему можна використовувати для прогнозування курсу валют.

Вступ

Для того, щоб розробити систему прогнозування часового ряду можна використовувати безліч методів, які відрізняються складністю реалізації, рівнем допустимих похибок (достовірності), величиною часових витрат для прогнозування.

Одними з методів, на основі яких може базуватися система прогнозування, є штучні нейронні мережі, а також Метод Групового Врахування Аргументів (рос. МГУА).

МГВА – сімейство індуктивних алгоритмів для моделювання мультипараметричних даних, що оснований на рекурсивном селективном відборі моделей, на основі яких будуються більш складні моделі. Точність моделювання на кожному наступному кроці рекурсії зростає за рахунок ускладнення моделі.

Нейромережі – системи, що мають можливість самонавчання.

В даній статті ми розглянемо прогнозування за допомогою нейромереж. Розглянемо оптимальні методи, на основі яких краще будувати систему прогнозування.

Статистичні моделі прогнозування

Прогнози нейромережі бувають багат шаровими та одно шаровими.

Багат шаровим прогнозом називаються довгостроковий прогноз за декілька кроків, на яких система використовує отримані прогнози як вхідні данні.

Одно шаровим прогнозуванням називають короткостроковий прогноз (на 1 крок). Одно шарові прогнози більш точні,

але вони не дозволяють отримувати довгострокові відносини.

Прогнозування, знаходження прихованих періодичностей в даних, аналіз залежностей, оцінка ризику при прийнятті рішення та інші задачі вирішуються в рамках статистичних моделей.

Теорія стохастичного прогнозування вивчає методи побудови предикторів. Для побудови даної теорії, перш за все, необхідно уточнити сенс наближеної рівності $\Phi(X) \approx Y$. Якщо $\Phi(X)$ використовується для прогнозування величини Y , то одним із розумних заходів розходження між ними є квадратична похибка, але так як величина Y невідома, то для виміру точності предиктора Φ використовується середньоквадратична.

Теоретичні відомості про нейромережі

Основою нейромереж є штучний нейрон, що імітує в першому наближенні властивості біологічного нейрона. На вхід штучного нейрона поступає множина сигналів, які є виходами інших нейронів.

Сила нейромережі напряму залежить від кількості нейронів. Групи нейронів поєднуються в шари, на основі яких можна побудувати одно шарові або багат шарові нейромережі. Багат шарові мережі володіють значно більшими можливостями, але вони можуть привести до збільшення обчислювальної потужності лише в тому випадку, якщо активаційна функція між шарами буде нелінійною.

Активаційна функція застосовується на виході кожного з нейронів для підсилення або послаблення сигналу, що в свою чергу пришвидшує роботу систе-

ми. Позначимо вхідний та вихідний сигнали нейрону як NET та OUT . Активаційна функція $F(NET)$ може бути пороговою бінарною, лінійно обмеженою, функцією гіперболічного тангенсу, або сигмоїдною (S -подібною або логістичною функцією) (1).

$$OUT = \frac{1}{1 + e^{-C \cdot NET}} \quad (1)$$

З виразу для сигмоїда (1) очевидно, що вихідне значення нейрона лежить в діапазоні $[0,1]$. Популярність сигмоїдної функції зумовлюють наступні її властивості:

- здатність підсилювати слабкі сигнали сильніше, ніж великі, і опиратися „насиченню” від потужних сигналів;
 - монотонність і диферентіюваність на всій осі абсцис;
- простий вираз для похідної що дає можливість використовувати широкий спектр оптимізаційних алгоритмів.

$$F'(NET) = C \cdot F(NET) \cdot (1 - F(NET))$$

Навчання системи будемо проводити по методу «Навчання без вчителя» зі зворотним поширенням. Тому що такий вид навчання має ряд переваг, які для поставленого завдання найважливіші, а саме – найменший час навчання та прогнозування системи.

Для алгоритму зворотного поширення потрібно лише, щоб функція була всюди диференційована. Сигмоїд задовольняє цій вимозі. Його додаткова перевага полягає в автоматичному контролі посилення.

Налаштування ваг вихідного шару

Так як при «навчанні зі вчителем» для кожного нейрона вихідного шару задано цільове значення, то налаштування ваг легко здійснюється з використанням модифікованого дельта-правила [1]. Для виходів нейронів прихованого шару не має цільових значень для порівняння, тому навчання ускладнюється.

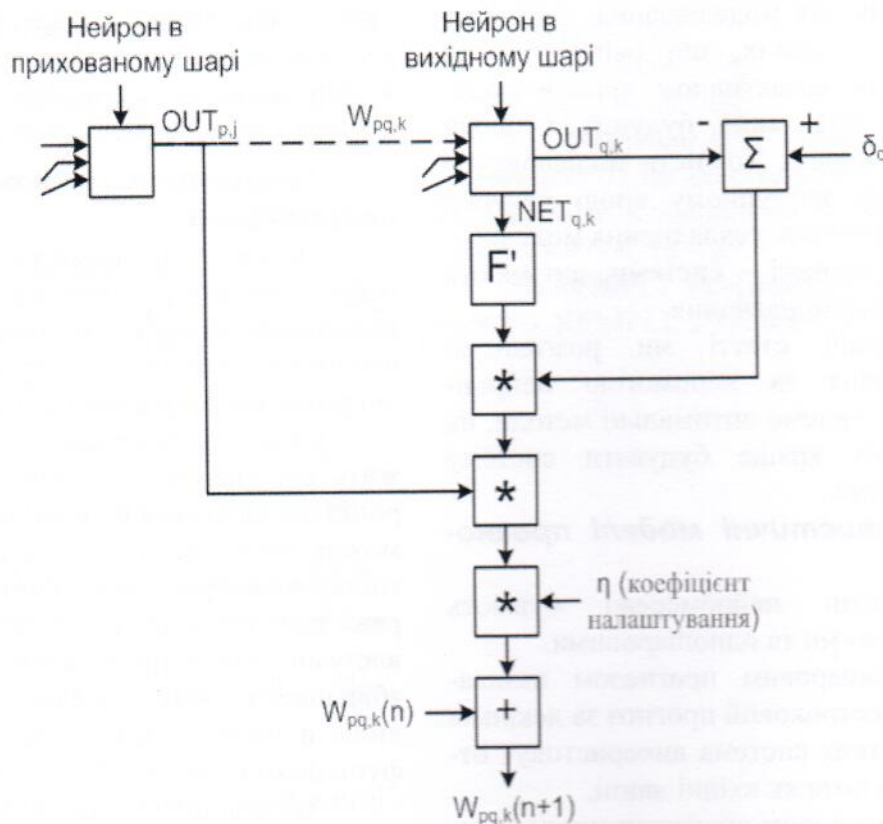


Рис 1. Налаштування ваги вхідного шару

На рис. 1 показаний процес навчання для однієї ваги від нейрона p в прихованому шарі j до нейрона q у вихідному шарі k . Вихід нейрона шару k , виходячи з цільового значення (*Target*), дає сигнал помилки. Він множиться на похідну стискаючої функції $OUT(1-OUT)$, обчислену для цього нейрона шару k , даючи, таким чином, величину δ .

$$\delta = OUT(1-OUT)(Target - OUT)$$

Потім δ множиться на величину OUT нейрона j , з якого виходить розглянута вага. Цей вираз у свою чергу множиться на коефіцієнт швидкості навчання η (зазвичай від 0.01 до 1.0), і результат додається до ваги. Така ж процедура виконується для кожної ваги від нейрона прихованого шару до нейрона у вихідному шарі.

Наступні рівняння ілюструють це обчислення:

$$\Delta W_{pq,k} = \eta * \delta_{q,k} * OUT_{p,j} \quad (2)$$

$$W_{pq,k}(n+1) = W_{pq,k}(n) + \Delta W_{pq,k} \quad (3)$$

де $w_{pq,k}(n)$ - вага від нейрона p в прихованому шарі до нейрона q у вихідному шарі на кроці n ; $\delta_{q,k}$ - величина δ для нейрона q ; $OUT_{p,j}$ - величина OUT для нейрона p .

Налаштування ваг прихованого шару

Під час навчання з'єднуючі ваги функціонують у зворотному порядку, пропускаючи величину δ від вихідного шару до прихованого шару. Величина δ ,

необхідна для нейрона прихованого шару, обчислюється сумою всіх таких виразів і множенням на похідну стискає функції:

$$\delta_{q,k} = OUT_{p,j}(1-OUT_{p,j}) \left[\sum_q \delta_{q,k} W_{pq,k} \right]$$

Коли значення δ отримано, ваги, що живлять перший прихований рівень, можуть бути відкориговані за допомогою рівнянь (2) і (3).

Інші алгоритмічні розробки

У роботі [2] описаний метод прискорення навчання для алгоритму зворотного поширення, що збільшує також стійкість процесу. Цей метод, названий імпульсом, полягає в додаванні до корекції ваги члена, пропорційного величині попереднього зміни ваги. Як тільки відбувається корекція, вона «запам'ятовується» і служить для модифікації всіх наступних корекцій. Рівняння корекції модифікуються таким чином:

$$\Delta W_{pq,k}(n+1) = \eta * \delta_{q,k} * OUT_{p,j} + a * \Delta W_{pq,k}(n)$$

$$W_{pq,k}(n+1) = W_{pq,k}(n) + \Delta W_{pq,k}(n+1)$$

де a - коефіцієнт імпульсу, зазвичай встановлюється близько 0,9.

Метод windowing

Задача прогнозування часових рядів дуже схожа по способу реалізації на задачу розпізнавання образів. Для задачі розпізнавання образів існує метод вікон (метод Windowing, Рис. 3), який ідеально підходить для рішення нашої задачі.

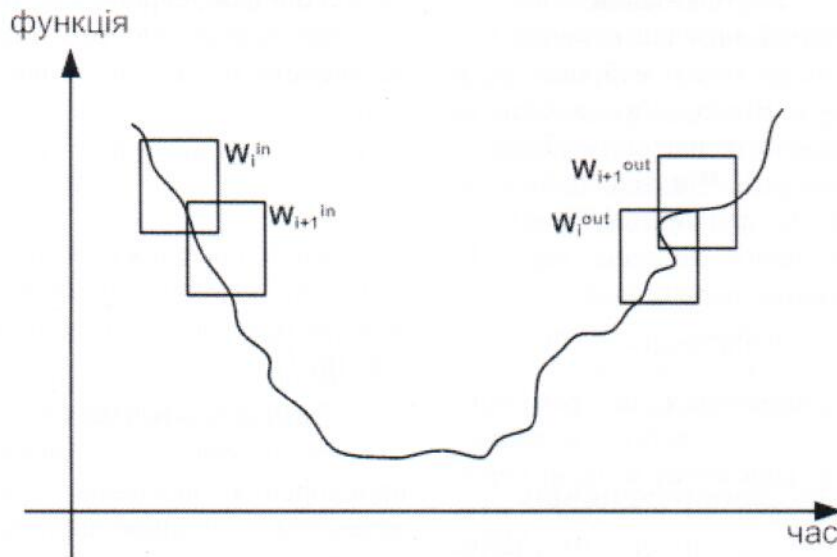


Рис. 3. Метод вікон

Метод Windowing дозволяє виявити закономірності в часовому ряді на основі даних аналізу часового ряду до задачі розпізнавання образів і подальшого її вирішення на нейромережі, описаним вище способом.

Основна ідея методу: вводиться два вікна, одне з яких вхідне (*input*), друге - вихідне (*output*). Ці вікна фіксованого розміру для спостереження даних. Вікна здатні пересуватися з деяким кроком S . У результаті отримуємо деяку послідовність спостережень, яка становить навчальну множину. Вхідному вікну відповідає вхід нейромережі, а вихідному вікну – бажаний образ (ідентифікатор в термінології розпізнавання образів).

Проблеми навчання нейрон-мереж та прогнозування

Тривалий час навчання може бути результатом неоптимального вибору довжини кроку. Невдачі в навчанні зазвичай виникають з двох причин: паралічу мережі і попадання в локальний мінімум.

Зворотне поширення використовує різновид градієнтного спуску, тобто здійснює спуск вниз по поверхні помилки, безперервно налаштовуючи ваги в напрямку до мінімуму. Мережа може потрапити в локальний мінімум (неглибоку долину), коли поруч є набагато більш глибокий мінімум. Статистичні методи навчання можуть допомогти уникнути

цієї пастки, але вони повільні. В [3] запропоновано метод, який об'єднав статистичні методи машини Коші з градієнтним спуском зворотного поширення що призводить до системи, яка знаходить глобальний мінімум, зберігаючи високу швидкість зворотного поширення.

Дуже важливою проблемою прогнозування є так званий горизонт прогнозу. Горизонт прогнозу – це деякий проміжок часу, на який можливо отримати достовірні результати. Після цього проміжку достовірність прогнозу не можливе, так як похибка прогнозування зростає за гіперекспоненціальним законом. Для представленої системи практично було визначено, що горизонт прогнозу рівний 6-8 крокам після початку прогнозування, отже ми маємо можливість точно спрогнозувати курс валют лише на короткий час. Після горизонту прогнозу, нейромережа правильно визначає напрям зміни курсу, але похибка при цьому дуже велика.

Слід згадати теорему Теккенса про глибину занурення прогнозування. Відповідно їй, для будь-якої динамічної системи можна знайти мінімальну кількість значень, на основі яких можна зробити однозначний прогноз на наступне значення. Але так, як система зміни курсу валют підпорядкована хаотичним змінам як самої системи, так і її елементів то визна-

чити оптимальну глибину занурення не вдалося. Але практично було встановлено, що велика кількість вхідних даних дає високу ймовірність того, що горизонт прогнозу буде приблизеним до максимально отриманого, і похибка буде мінімальною.

Приклад прогнозування

Для прикладу прогнозування візьмемо курс гривні до євро з 2000 по 2007 роки. Нехай крок прогнозу – середній курс кожного місяця.

Налаштування системи:

- кількість входів мережі – 5;
- кількість шарів – 3;

- кількість нейронів в кожному шарі – 20, 20, 1;
- довжина часового ряду – 62;
- кількість прогнозних значень – 43;
- похибка навчання – 0,01;
- крок навчання – 0,05.

При навчанні системи можна спостерігати зміну похибки прогнозу від 11 до 0,01. По закінченні отримуємо короткий звіт по навчанню (рис. 4).

Результати прогнозу отримуємо в вигляді графіку курсу (рис. 5), а також таблиці.

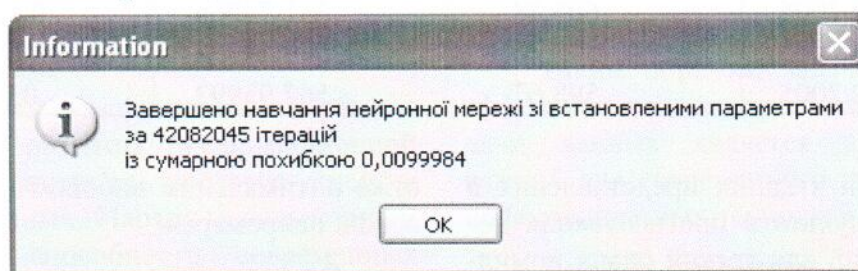


Рис. 4 Кінець навчання



Рис. 5 Графік прогнозу курсу

На рис. 5 блакитною лінією позначені реальні дані курсу за вказаний проміжок. Прогнозування почалося з 01.10.2004, що відповідає 63 точці на графіку (початок зеленого інтервалу). Система виділила зеленою лінією інтервал с 8 точок (період 01.10.2004-01.05.2005), так

як на даному інтервалі ми маємо досить достовірний прогноз. Цей прогноз ілюструє горизонт прогнозу. Далі (червоний інтервал) – прогноз досить не точний, зі збільшенням похибки з кожним наступним кроком, але напрям зміни курсу визначений вірно.

Таблиця 1. Достовірний прогноз системи

Дата	Реальні дані	Прогноз	Похибка
1.10.2004	661,99	661,97988	1,52901E-05
1.11.2004	688,01	672,54622	0,022476094
1.12.2004	710,65	674,14596	0,051367113
1.1.2005	694,68	698,33178	0,005256776
1.2.2005	689,07	680,56223	0,012346749
1.3.2005	697,89	657,08632	0,058467203
1.4.2005	671,44	707,58794	0,053836446
1.5.2005	641,48	641,62722	0,000229496
1.6.2005	615,37	637,40423	0,035806469
1.7.2005	608,28	578,49632	0,048963762
1.8.2005	620,71	593,28392	0,044185017
1.9.2005	619,31	626,93453	0,01231133
1.10.2005	606,78	531,72066	0,123701074
1.11.2005	595,76	599,48394	0,006250744
1.12.2005	598,67	567,93992	0,05133058

Виходячи з даних представлених в таблиці 1, результати прогнозування досить достовірні, але деякий ризик помилки все ж таки існує.

Висновок

Розглянувши основні методи, а також беручи до уваги данні про МГВА, зрозуміло, що систему прогнозування часових рядів доцільніше створювати на основі нейромережі з багат шаровим прогнозом, з навчанням «з вчителем» на основі зворотнього поширення, і використати метод Windowing для підвищення збіжності навчання і прогнозів, і зменшення похибок.

Систему на основі МГВА використовувати недоцільно, так як цей метод рекурсивний та ітераційний, що суттєво збільшує кількість операцій для точних розрахунків, а також їх час.

Система на основі запропонованих методів має великі переваги в порівнянні з іншими типами систем для даної задачі. Зокрема система має покращені властивості швидкості збіжності похибки, а також зменшений час навчання та прогнозування системи.

Беручи до уваги, що завданням є прогнозування курсу валют, час розрахунків є дуже важливим критерієм,

отже оптимальним вибором є система на основі нейромереж.

Список літератури

1. Ф. Уоссерман Нейрокомпьютерная техника: теория и практика. – М.: Мир. 1992 – 592 с.
2. Rumelhart D. E., Hinton G. E., Williams R. J. 1986. Learning internal representations by error propagation. In Parallel distributed processing, vol. 1, pp. 318 – 62. Cambridge, MA: MIT Press.
3. Wasserman P. D. 1988a. Combined backpropagation / Cauchy machine. Proceedings of the International Neural Network Society. New York: Pergamon Press. – 556 p.