

ЗАСТОСУВАННЯ НЕЙРОННИХ СІТОК ДО МОДЕЛЮВАННЯ ЕКОНОМІЧНИХ ПРОЦЕСІВ

Поставлена задача и найдено решение моделирования потоков заказов на предприятии связи, а также прогнозирование превышения курса одной валюты над другой путем использования математического аппарата нейронных сетей

To find a solution of modelling string of the calls in telephone firm and prediction the growing cash flow ausrian shilling an australia dollar with used mathematical method "neuron net".

При моделюванні економічних процесів однією з найбільш складних задач є вибір виду функції апроксимації. Це пояснюється тим, що характер залежності одного економічного параметра від іншого є невідомим. Фактично, нам необхідно, знаючи певний набір значень вхідних параметрів, віднести значення вихідного параметра до певного класу чи значення.

Задача в такій постановці відповідає задачам по класифікації (розпізнаванню образів) [1, 3, 7]. В загальному вигляді ця задача має визначається як набір розділяючі функцій, які залежать від вхідних параметрів x

$$C_i = F(x_1, x_2, \dots, x_n) \quad (1)$$

де C_i – функція класифікатора, i – порядковий номер класу ($1 < i < K$), K – кількість класів, на які розбивається початкова вибірка даних, n – кількість вхідних параметрів, на підставі значень яких провадиться класифікація, F – функція класифікації.

При автоматичній класифікації нам треба тільки вказати кількість класів і розбиття буде виконано, частіше всього, на підставі евклідової міри близькості [8]

$$\sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \rightarrow \min, \quad (2)$$

де \bar{x}_i – середнє i -го класу.

Після розрахунку коефіцієнтів функції F на підставі навчальної вибірки, віднесення нового об'єкту до певного класу відбувається шляхом під-

становки значень x_i , що характеризують новий об'єкт, в (1) для кожної розділяючої функції. Яка з C_i розділяючих функцій є більшою, до такого i -го класу і відноситься об'єкт.

Недоліком цієї методики класифікації є складність з вибором виду функції F , оскільки наперед невідомий тип залежності номера класу від вхідних даних. Довільний вибір виду функції не гарантує добру якість розпізнавання.

На наш погляд, для вирішення поставленої задачі найбільш прийнятним є використання такого математичного методу як нейронні сітки [2-5].

Нейронні сітки - це сітки, що складаються зі зв'язаних між собою простих елементів - формальних нейронів. Ядром використовуваних представлень є ідея про те, що нейрони можна моделювати досить простими формулами, а вся складність процесу моделювання визначається зв'язками між нейронами. Кожен зв'язок представляється як зовсім простий елемент, що служить для передачі сигналу.

Навчання нейронної сітки звичайно будується так: існує задачник – набір прикладів із заданими відповідями. Ці приклади пред'являються системі. Нейрони одержують по вхідних зв'язках сигнали – «умови прикладу», перетворюють їх, кілька разів обмінюються перетвореними сигналами і, нарешті, видають відповідь – також набір сигналів. Відхилення від правильної відповіді штрафуються. Навчання складається в мінімізації штрафу як (неявної) функції зв'язків.

Неявне навчання приводить до того, що структура зв'язків стає «незрозумілою» – не існує іншого способу її прочитати, крім як запусити функціонування сітки. Стає складно побудувати зрозумілу людині логічну конструкцію, що відтворює дії сітки. Зате ця методика не вимагає вибору виду розділяючої функції, оскільки використовується невеликий набір типових функцій, якими сітка комбінує

У 1943 році У. Маккаллох і його учень У. Піттс сформулювали основні положення теорії діяльності головного мозку [2, 5]. Ними було розроблено модель нейрона як найпростішого процесорного елемента, що виконує обчислення перехідної функції від скалярного добутку вектора вхідних сигналів x_i і век-

тора вагових коефіцієнтів w_i .

Показана на рис. 1 модель нейрона описується наступною системою рівнянь

$$net = \sum_{j=1}^n x_j w_j$$

$$f(net) = \text{sgn}(net) = \begin{cases} +1, & net > 0, \\ -1, & net < 0. \end{cases} \quad (4)$$

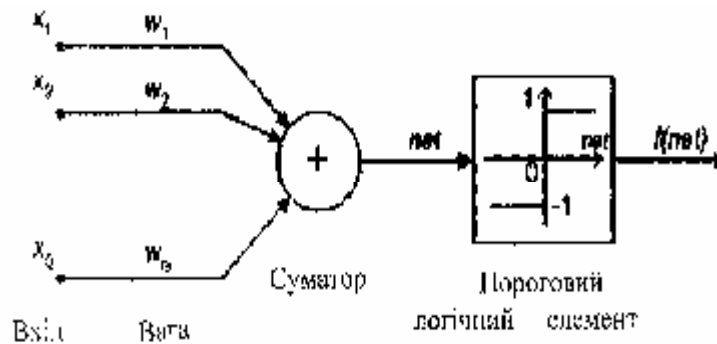


Рис. 1. Модель граничного нейрона МакКаллоха-Піттса

Інколи в (3) додається ще параметр θ – порогової чутливості, яка віднімається від суми зважених сигналів.

Функція $f(net)$, яку ще позначають як $OUT(net)$, називається активізуючою. Тому в штучних нейронних мережах використовують інші функції активації, найбільш популярною з яких є так звана логістична уніполярна сигмоїдальна функція (сигмоїд):

$$f(net) = \frac{1}{1 + \lambda^{-1net}}$$

$$f(net) \in [0;1] \quad (4)$$

$\lambda > 0$ – коефіцієнт крутості безупинної функції $f(net)$ біля $net = 0.5$. Функція

симетрична відносно NET=0, OUT=1/2

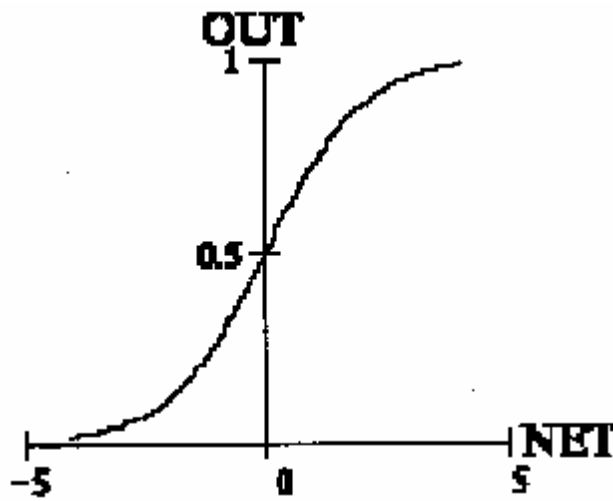


Рис. 2. Вид логістичної функції

Режими навчання з учителем нейронних сіток можуть бути різними, але найбільш ефективним є так зване “дельта-правило” [3, 4]:

1. Початкові ваги можуть бути будь-якими. Корекція провадиться пропорційно величині похідної по даній координаті. Похідна береться від функції активації. Підстроювання j ваги для i нейрона здійснюється за формулою:

$$\Delta w_j = \eta [a_j - f(\text{net}_j)] f'(\text{net}_j) x_j, \quad (5)$$

де $j=1,2,\dots, \eta > 0$ - коефіцієнт навчання, підбирається евристично

2. Помилка при навчанні на k кроці: $E_k = \frac{1}{2} [d_j - f(\text{net}_j)]^2$, (6)

де d_i - очікуваний вихід

3. Загальна помилка при навчанні: $E = \frac{1}{2p} \sum_{j=1}^p [d_j - f(\text{net}_j)]^2$, (7)

де p - число прикладів у навчальній вибірці

4. Похідна від сигмоїди

$$f'(\text{net}) = f(\text{net})[1 - f(\text{net})], \quad (8)$$

де p - число прикладів у навчальній вибірці.

Побудуємо нейронну сітку для двох варіантів економічних процесів. Перший стосується розпізнавання безперервних економічних процесів, а другий – дискретних.

Почнемо зі створення апроксимуючої залежності для даних про кількість замовлень підприємства зв'язку по годинах робочого дня [6].

В якості моделі було обрано одношаровий перцептрон з логістичною функцією активації та з одним нейроном на три входи (рис. 3). На кожен вхід подавалося значення трьох послідовних значень кількості викликів від початку доби. Кожне значення вхідного параметру x бралось з вагою w та з певним значенням порогової чутливості θ . Таким чином, математична модель такого перцептрона мала наступний загальний вигляд

$$OUT = \frac{1}{1 + \lambda^{-(w_1 x_i - q_1 + w_2 x_{i+1} - q_2 + w_3 x_{i+2} - q_3)}} \quad (9)$$

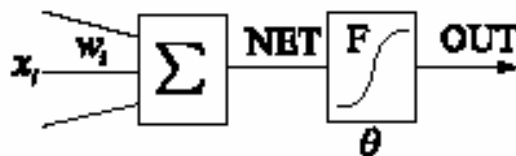


Рис. 3. Загальний вигляд схеми перцептрона з одним нейроном та суматором на вході

На кожному наступному кроці, підставлялося нове значення x , як $i+2$ -й елемент перцептрона, а i -е значення x відкидалося. Значення OUT порівнювалося зі значенням y з розрахунком погрішності прогнозування вигляду $x' = \frac{(x - m_x)}{s_x}$. На кожному кроці розрахунку провадилося корегування ваги та порогової чутливості за правилом

$$\left. \begin{aligned} \Delta w_{ij} &= e(d_j^s - y_j^s)x_{ij} \\ \Delta \Theta_j &= -e(d_j^s - y_j^s) \end{aligned} \right\} \quad (10)$$

де $d = y$, $y = OUT$, ε – число, яке характеризує швидкість навчання. Було встановлено наступне правило зменшення ε на кожному кроці розрахунку t $\varepsilon' = \varepsilon / 1.5668$, де ε' – нове значення швидкості навчання.

Перед початком навчання перцептрона всі дані були нормовані за наступним правилом:

$$\text{– вхідні значення } x \quad x' = \frac{(x - m_x)}{s_x} \quad (11)$$

, де m_x – середнє значення x , s_x – середнє квадратичне відхилення x .

$$\text{– значення вихідних даних } y: \quad y' = \frac{(x - m_y)}{s_y} + 4, \text{ де } m_y \text{ – середнє значення } y, s_y \text{ – середнє квадратичне відхилення } y.$$

В результаті навчання перцептрона було отримано показане на малюнку співпадіння розрахованих і реальних значень y . При цьому, сама апроксимуюча формула для нормованих значень параметрів, має вигляд

$$OUT = \frac{1}{1 + \lambda^{-(1,58x_i + 0,37x_{i+1} + 1,15x_{i+2} + 1,62)}} \quad (12)$$

На рис. 4 показано графіки навчальної вибірки та кривої, яка реалізується за (12), в якій бралися три попередніх значення кількості викликів. Можна бачити, що точність апроксимації поступово збільшується, оскільки експериментальна і розрахована криві поступово зближуються.

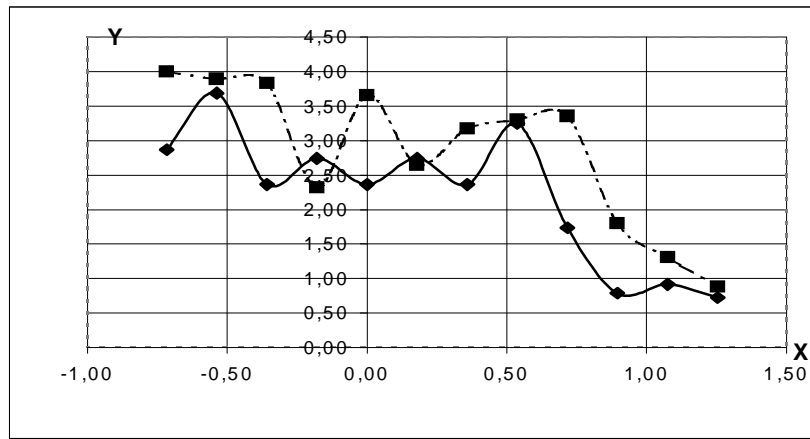


Рис. 4. Графік кількості викликів Y по годинам робочого дня X .

(\blacklozenge – експериментальна крива, \blacksquare – розрахована крива)

Побудуємо тепер прогнозуючу модель зміни курсу австрійського шилінга відносно австралійського долара (дані взяті по відомостям НБУ за період 01.02.2000 по 31.10.2000). В цьому випадку нас цікавить передбачення самого факту збільшення чи зменшення крос-курсу.

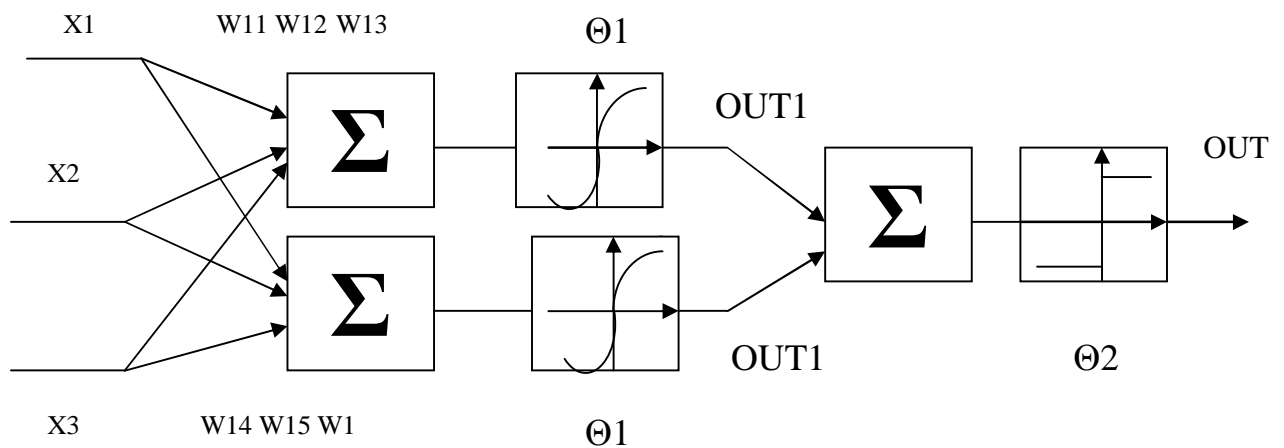


Рис. 5. Схема двошарового перцептрону з трьома входами на кожному нейроні.

Для вирішення цієї задачі було розроблено двошаровий перцептрон (рис. 5), на вхід якого подавалося значення трьох попередніх значеннях крос-курсу австрійського шилінга і робився прогноз вихідного параметра, який являє собою функцію знаку числа

$$OUT = \text{Sign}(NET) = \begin{cases} 1, & NET > 0 \\ 0, & NET = 0 \\ -1, & NET < 0 \end{cases} \quad (13)$$

В якості активуючих функцій першого шару були взяті логістичні (4). Навчання перцептрона виконувалося так. На кожному наступному кроці, підставлялися нові значення x_i як для дієздатних так для недієздатних вузлів. Причому, вихідними значення d_i бралися числа: -1 – для випадку перебільшення курсу австрійського шилінга відносно австралійського долара, +1 – для протилежного випадку дієздатність якого експертами оцінюється як критична, +1 – для випадку спів падіння курсів. Значення OUT порівнювалося зі значенням d_i з розрахунком погрішності прогнозування (7). На кожному кроці розрахунку провадилося корегування ваги та порогової чутливості за правилом [4]

$$\begin{aligned} \Delta w_{ij} &= \varepsilon (d_j^s - y_j^s) x_{ij} \\ \Delta \Theta_j &= -\varepsilon (d_j^s - y_j^s) \end{aligned} \quad (14)$$

де $d = y$, $y = OUT$, ε – число, яке характеризує швидкість навчання. Було встановлено наступне правило зменшення ε на кожному кроці розрахунку t $\varepsilon' = \varepsilon / 1.5668$, де ε' – нове значення швидкості навчання.

Перед початком навчання перцептрона вхідні дані були нормовані за наступним правилом (11)

На рис. 6. представлено схему застосування електронних таблиць Excel для розрахунків такого перцептрона.

	X	W1	X*W1	SUM	TETA1	OUT1	W2	OUT1*W2	SUM	Teta2	OUT
X1	0,3	1	-0,3	-1,1	1	0,10545	2	0,21090802	0,21649	2,8	-1
X2	0,3	2	-0,61	-3,2	21	0,00558	1	0,00557715			
X3	0,1	3	-0,28							Real	1
		4	1,21		Eps	3				Diff	2
1		5	1,51		Sp	1					
		6	-0,46								

Рис. 6. Фрагмент таблиці Excel, яка реалізує перцептрон з рис. 5.

Стрілки на рис. 6 показують порядок передачі даних. Потім вхідні дані повнювалися на один крок, тобто, використовувалися чергові значення вхідних параметрів для іншого вузла. Знаходилася розбіжність поміж значеннями, які видає перцептрон та справжньою оцінкою стану вузла. в порівнянні зі своїм попереднім значенням з корегуванням ваг та порогових функцій. Цей алгоритм показано на рис. 7.

В процесі навчання вже на 9-му кроці перцептрон давав більшість правильних відповідей на наступних значеннях курсу. Тому, процес навчання було припинено, а результуюча функція набула вигляду

$$OUT = Sign \left[\begin{array}{c} 10,37669298 \\ \left[\begin{array}{c} -(-1,443525248X_1 + 1,561436345X_2 \\ + 1,582503396X_3 + 4,683681367) \end{array} \right] + \\ 7,940045065 \\ \left[\begin{array}{c} -(1,556474752X_1 + 4,561436345X_2 + \\ + 4,582503396X_3 + 3,683681367) \end{array} \right] + \\ 2,883681367 \end{array} \right] \quad (15)$$

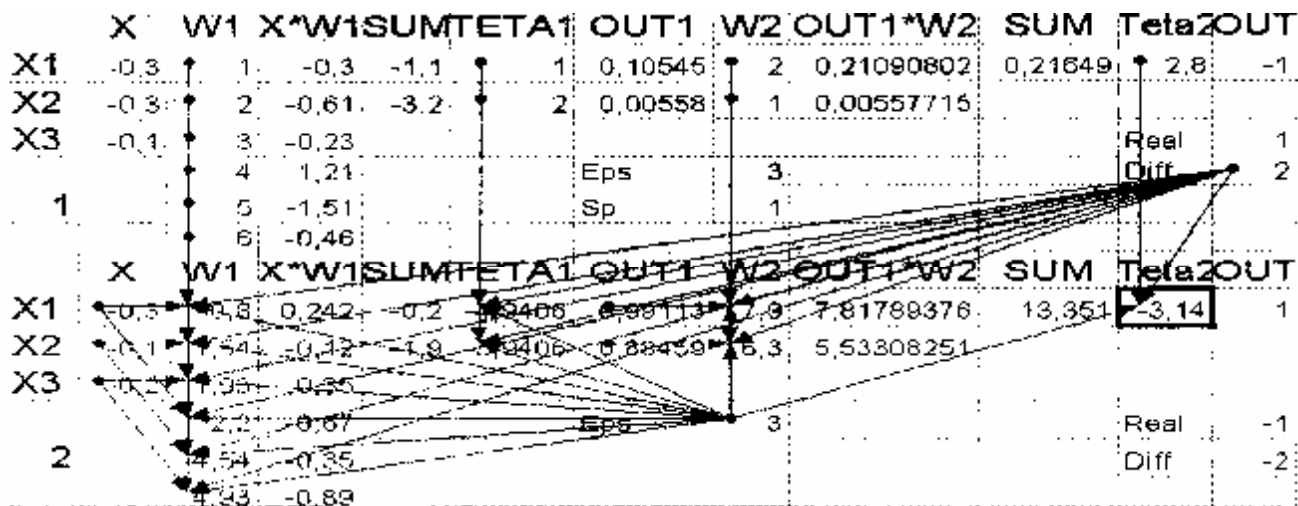


Рис. 7. Схема корегування ваг та порогів на наступному кроці, враховуючи результати з попереднього.

Як видно з наведених прикладів, нейронні сітки мають властивості про-

гнозувати чи класифікувати як безперервні так і дискретні процеси. Змішування таких процесів складає одну проблем при моделюванні економіки, тому ідея перцептрона є, на наш погляд, найбільш привабливою для вирішення поставлених задач.

Збільшення складності модельованих процесів неодмінно має викликати ускладнення і самої схеми перцептрона. Це ускладнення неодмінно викличе появу проблеми в підборі схеми навчання. Тому, перспективою розвитку цього напрямку моделювання нам бачиться в розробці схем навчання для трьохшарових і більше перцептронів.

Потрібно також провести дослідження в напрямку підбору виду активуючих функцій та залежності швидкості сходження процесу навчання від виду цих функцій.

Вирішення цих проблем дозволить створювати моделі складних економічних систем без детального вивчення їх структури та характеру взаємних зв'язків. При цьому, якість прогнозування має бути значно вищою, аніж при застосуванні інших методів моделювання.

Список літератури

1. Акулич И. Л. Математическое программирование в примерах и задачах. - М.: Высшая школа., 1985 г.
2. Горбань А.Н., Россиев Д.А. Нейронні сітки на персональному комп'ютері. – .0Новосибирск: Наука. Сибирская издательская фирма РАН, 1996.
3. Енциклопедія кібернетики. Головна редакція українських радянських енциклопедій. т1 та 2, – Київ, 1973. – 680 с.
4. Минский М., Пайперт С. Перцептроны. - М.: Мир, 1971.
5. Миркес Е.М. Нейрокомп'ютер. Проект стандарта. Новосибирск: Наука, Сибирская издательская фирма РАН, 1998, 337 С
6. Пістунов І.М., Пістунов М.І. Моделювання періодичних процесів в економіці./ Економіка: проблеми теорії та практики. Збірник наукових праць. Випуск 135. – Дніпропетровськ: ДНУ. – 2002. – С.204-207.
7. McKenna C.J., Rees R. Economics: A Mathematical Introduction. Oxford. 1992.
8. Practical Management Science: Spreadsheet Modeling and Applicationsby Wayne L. Winston, S. Christian Albright, Mark Broadie, Chris Albright (Contributor). September 1997.
9. Silberberg E. The Structure of Economics: A Mathematical Analysis. N.Y. 1978.

Рекомендована до публікації д.т.н., проф.. Є.В. Кочура, 25.03.2003