

УДК 517.977.58:621.3.07

**В.М. Штепа,**  
кандидат технічних наук  
**Н.А. Заєць,**  
кандидат технічних наук, доцент  
**О.В. Ленков**  
**А.С. Шворов**

## МЕТОДИЧНІ ЗАСАДИ ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЗНАЧЕННЯ ВАЖЛИВОСТІ ВХІДНИХ ЕЛЕКТРОННИХ ДОКУМЕНТІВ

*Розглянуто методичні основи побудови та застосування штучних нейронних мереж для визначення важливості вхідних електронних документів (ВЕД). Проаналізовані архітектура, принципи функціонування, алгоритми навчання та можливість використання нейронних мереж для вирішення завдань здобуття нових знань при дослідженні важливості ВЕД.*

**Ключові слова:** нейронна мережа, генетичний алгоритм, оптимальний розв'язок.

*Рассмотрены методические основы построения и применения искусственных нейронных сетей для определения важности входных электронных документов (ВЭД). Проанализированы архитектура, принципы функционирования, алгоритмы обучения и возможность использования нейронных сетей для решения задач получения новых знаний при исследовании важности ВЭД.*

**Ключевые слова:** нейронная сеть, генетический алгоритм, оптимальное решение.

*Methodical bases of the construction and application of artificial neural networks for determination of importance of the front-end electronic documents are considered. Architecture, principles of functioning, algorithms of training and the possibility of use of neural networks for the solution of the problems of the information gaining for the research of an importance of the front-end electronic documents are analyzed.*

**Keywords:** neural network, genetic algorithm, optimal solution.

Як відомо, до складу нейрона (рис. 1) входять помножувачі (синапси), суматор і нелінійний (лінійний) перетворювач [1, 2]. Синапси здійснюють зв'язок між нейронами і перемножують вхідний сигнал на число, яке характеризує силу зв'язку – вагу синапсу [3].

Суматор складає сигнали, що поступають по синаптичних зв'язках від інших нейронів, і зовнішніх вхідних сигналів. Нелінійний (лінійний) перетворювач реалізує функцію одного аргументу – виходу суматора. Ця функція називається “функція активації” або “передаточна функція” нейрона. Нейрон у цілому реалізує скалярну функцію векторного аргументу.

Математична модель нейрона описується співвідношеннями:

$$s = \sum_{i=1}^n \omega_i \cdot x_i + b, \quad (1)$$

де  $\omega_i$  – вага синапсу ( $i=1, \dots, n$ );  $b$  – значення зміщення;  $x_i$  – компонента вхідного вектору (вхідний сигнал) ( $i=1, \dots, n$ );  $s$  – результат додавання;  $y$  – вихідний сигнал нейрона (рис. 1);  $n$  – число входів нейрона;  $f$  – перетворення (функція активації або передаточна функція).

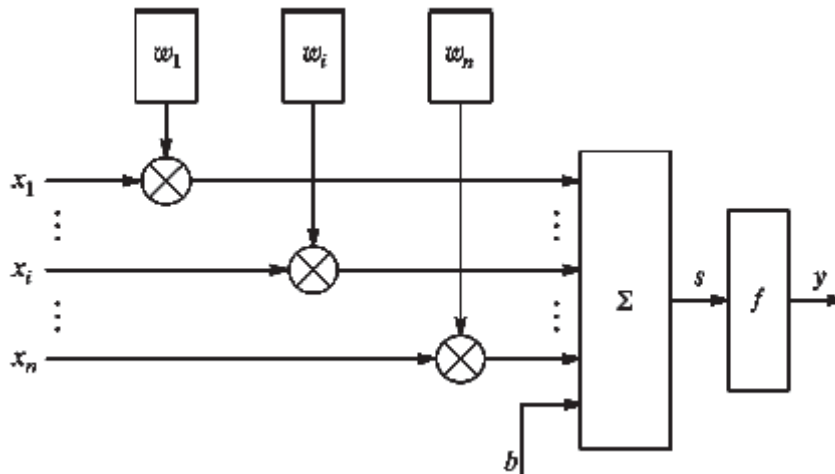


Рис. 1. Структурна схема штучного нейрона

У загальному випадку вхідний сигнал, вагові коефіцієнти і значення зсуву можуть приймати дійсні значення. Вихід ( $y$ ) визначається видом функції активації і може бути як дійсним, так і цілим [4]. У багатьох практичних задачах входи, вага і зсуви можуть приймати лише деякі фіксовані значення.

Синаптичні зв'язки з позитивними вагами називають збуджуючими, з негативними – гальмуючими.

Таким чином, нейрон повністю описується своїми вагами  $\omega_i$  і передаточною функцією  $f(s)$ . Одержавши набір чисел (вектор)  $x$ , нейрон видає деяке число  $y$  на виході.

Тобто робота нейромережі полягає у перетворенні вхідного вектора  $X$  у вихідний вектор  $Y$ , причому це перетворення задається вагами мережі.

У процесі функціонування нейронна мережа формує вихідний сигнал  $Y$  відповідно до вхідного сигналу  $X$ , реалізуючи деяку функцію  $Y = G(X)$ . Якщо архітектура мережі задана, то вид функції  $G$  визначається значеннями синаптичних вагів і зсувів мережі.

Нехай рішенням деякої задачі є функція  $Y = F(X)$ , що задана парами вхідних-вихідних даних  $(X^1, Y^1)$ ,  $(X^2, Y^2)$ , ...,  $(X^N, Y^N)$ , для яких  $Y^k = F(X^k)$  ( $k = 1, 2, \dots, N$ ). Відповідно, навчання полягає в пошуці (синтезі) функції  $G$ , близької до  $F$  у значенні деякої функції помилки  $E$ .

Якщо вибрані множини навчальних прикладів – пара  $(X^k, Y^k)$  (де  $k = 1, 2, \dots, N$ ) і спосіб обчислення функції помилки  $E$ , то навчання нейронної мережі перетворюється на задачу багатовимірної оптимізації, що має дуже велику

розмірність при цьому, оскільки функція  $E$  може мати довільний вигляд, у загальному випадку – багатоекстремальна неопукла задача оптимізації.

Для її вирішення можуть бути використані такі (ітераційні) алгоритми:

- локальної оптимізації з обчисленням часткових похідних першого порядку;
- локальної оптимізації з обчисленням часткових похідних першого і другого порядку;
- стохастичні алгоритми оптимізації;
- глобальної оптимізації.

До першої групи належить: градієнтний алгоритм (метод найшвидшого спуску); методи з одновимірною і двовимірною оптимізацією цільової функції у напрямі антиградієнта; метод зв'язаних градієнтів; методи, що враховують напрям антиградієнта на декількох етапах алгоритму.

До другої групи належить: метод Ньютона, квазіньютонівські методи, метод Гауса-Ньютона, метод Льовенберга-Марквардта та інші.

Стохастичними методами є: пошук у випадковому напрямі, імітація відпалу, метод Монте-Карло (чисельний метод статистичних випробувань).

Задачі глобальної оптимізації розв'язуються за допомогою перебору значень змінних, від яких залежить цільова функція (функція помилки  $E$ ).

Класично нейронні мережі застосовуються для: класифікації образів, кластеризації (категоризації), апроксимації, передбачення/прогнозу, оптимізації, управління.

Однак, з урахуванням здатності нейронних мереж адекватно працювати в умовах невизначеності та неповноти інформації [1–4], доцільно використовувати їх і для набуття нових наукових знань щодо досліджуваних ВЕД.

Методику такого застосування нейромереж апробуємо для розпізнавання важливості ключових слів ВЕД [5]. Відповідно до цієї методики, вхідними параметрами мережі будуть: щільність потоку ключового слова ВЕД ( $Z_1$ ); вага ключового слова ВЕД ( $Z_2$ ); частота ключового слова на сторінці ( $Z_3$ ); частота появи ключового слова у ВЕД ( $Z_4$ ); коефіцієнти, що враховують розташування ключового слова ( $X_1$ ): у назві сторінки (title), у мета-тегах, місці сторінки, в тексті малюнка, у гіперпосиланні; оформлення ключового слова: стиль; розмір шрифту; жирність шрифту ( $X_2$ ); редизайн документа ( $X_3$ ); поправочний коефіцієнт ( $X_4$ ).

На виході мережа генерувала важливість ключових слів (ранжир) згідно з експертно поділеними ступенями важливості в умовних одиницях (у.о.).

Набори навчальних, контрольних та перевірочних інформаційних блоків містили по 500 прикладів.

Для побудови математичної моделі застосували лінійну нейронну мережу і відповідним аналітичним виразом:

$$Y = X * W + B, \quad (2)$$

де  $Y$  – вихід мережі;  $W$  – ваговий коефіцієнт;  $X$  – вхід;  $B$  – зміщення.

У демоверсії програмного середовища Statistica 6.1 синтезуємо лінійну нейронну мережу оцінки важливості ВЕД (рис. 2): навчальна помилка (середньоквадратичне значення помилок за всіма виходами) – 0,103727 у.о, контрольна помилка – 0,15143 у.о., перевірочна помилка – 0,1236 у.о.

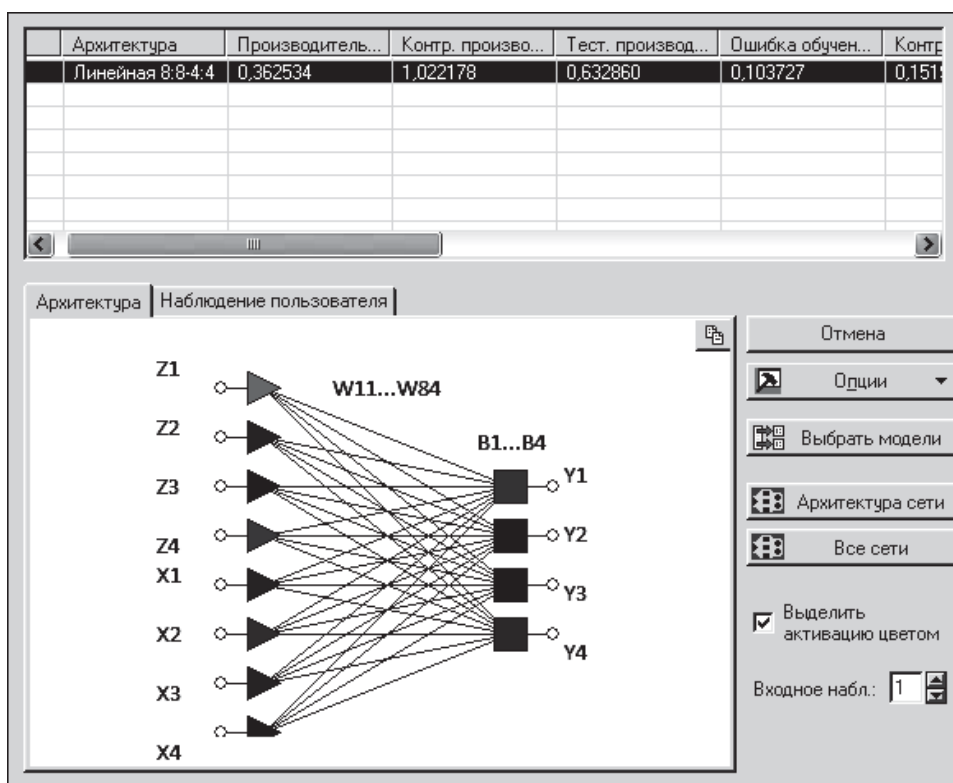


Рис. 2. Архітектура нейронної мережі:  
 $W_{11}...W_{84}$  – вагові коефіцієнти (всього їх – 32; перша цифра – номер вхідного параметру, друга – вихідного);  $B_1...B_4$  – зміщення.

Тоді виводимо аналітичні вирази для кожного із 4-х виходів нейронної мережі:

$$Y_1 = W_{11} * Z_1 + W_{21} * Z_2 + W_{31} * Z_3 + W_{41} * Z_4 + W_{51} * X_1 + W_{61} * X_2 + W_{71} * X_3 + W_{81} * X_4 + B_1; \quad (3)$$

$$Y_2 = W_{12} * Z_1 + W_{22} * Z_2 + W_{32} * Z_3 + W_{42} * Z_4 + W_{52} * X_1 + W_{62} * X_2 + W_{72} * X_3 + W_{82} * X_4 + B_2; \quad (4)$$

$$Y_3 = W_{13} * Z_1 + W_{23} * Z_2 + W_{33} * Z_3 + W_{43} * Z_4 + W_{53} * X_1 + W_{63} * X_2 + W_{73} * X_3 + W_{83} * X_4 + B_3; \quad (5)$$

$$Y_4 = W_{14} * Z_1 + W_{24} * Z_2 + W_{34} * Z_3 + W_{44} * Z_4 + W_{54} * X_1 + W_{64} * X_2 + W_{74} * X_3 + W_{84} * X_4 + B_4. \quad (6)$$

Запишемо вираз цільової функції для оптимізаційної задачі в аналітичному вигляді:

$$F = \sqrt{(Y_1 - R_1)^2} + \sqrt{(Y_2 - R_2)^2} + \sqrt{(Y_3 - R_3)^2} + \sqrt{(Y_4 - R_4)^2} \rightarrow 0 \quad (7)$$

де  $R_1 \dots R_4$  – регламентні значення вихідних величин.

Наприклад, якщо при розпізнаванні важливості слів потрібно корегувати лише параметрами  $X_1, X_2, X_3, X_4$ . Тоді, підставивши у вираз цільової функції  $F$  (7): значення входів  $Z_1 \dots Z_4$ , регламентні значення  $R_1 \dots R_4$ , вагові коефіцієнти

$W_{11} \dots W_{84}$ , зміщення  $B_1 \dots B_4$ , шукатимемо з використанням оптимізаційних мінімумів  $F(7)$  у залежності від комбінації  $X_1 \dots X_4$ .

**Висновок.** Запропоновані методичні засади отримання знань з використанням архітектури нейронних мереж у процесі проведеного дослідження продемонстрували науково-практичну цінність та можуть бути застосовані при побудові інформаційно-аналітичних структур розпізнавання важливості тексту ВЕД на основі систем обчислюваного інтелекту.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы : пер с польск. / Д. Рутковская, М. Пилиньский, Л. Рутковский. – М. : Горячая линия – Телеком, 2004. – 452 с.
2. Круглов В.В. Искусственные нейронные сети. Теория и практика / Круглов В.В. – М. : Горячая линия – Телеком, 2002. – 382 с.
3. Рідкокаша А.А. Основи систем штучного інтелекту : навч. посіб. / А.А. Рідкокаша, К.К. Голдер. – Черкаси : ВІДЛУННЯ-ПЛЮС, 2002. – 240 с.
4. Гунченко Ю.О. Архітектура систем управління біотехнічними об'єктами / Ю.О. Гунченко, В.П. Лисенко, С.А. Шворов, В.М. Штепа // Сучасна спеціальна техніка. – К. : МВС ДНДІ. – 2012. – № 2 (29). – С. 33–40.
5. Melin P. Voice Recognition with Neural Networks, Type-2 Fuzzy Logic and Genetic Algorithms. Engineering Letters / Melin P., Urias J., Solano D., Soto M., Lopez M., Castillo O.

Отримано 19.09.2013