

УДК 519.8

РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ СЛУЖБИ ТАКСІ

О.Ю. МЕЛЬНИКОВ, А.М. БОБУХ

Донбаська державна машинобудівна академія, м. Краматорськ

Приведено математичну (заснована на нейромережевому підході) і інформаційну (на мові UML) моделі системи для інтелектуального аналізу даних сучасної служби таксі. Описана комп'ютерна реалізація моделей

Служба таксі є невід'ємною частиною життя сьогоденного суспільства. Організація такої служби вимагає централізованого збору і обробки значного об'єму інформації, необхідної для оперативної роботи. Як правило, програмне забезпечення, яке використовується в службах таксі, займається оперативним збором і візуалізацією інформації, а її аналіз і ухвалення рішень покладено на диспетчера. Таким чином, є необхідність підвищити ефективність використання даних, виявити із стандартного їх набору додаткову інформацію і тим самим полегшити процес ухвалення рішення.

Постановка завдання

Найбільш важливими завданнями представляються наступні: прогнозування районів, з яких поступлять виклики в певний час доби, і прогнозування поведінки конкретного працівника служби таксі протягом робочої зміни. Рішення першої задачі дозволить планувати стратегічне розміщення машин по місту залежно від часу доби, і таким чином забезпечувати найскоріше прибуття машини до клієнта, випереджаючи конкуруючі фірми. Рішення другої задачі дозволить оптимізувати дії працівників, а також отримати якесь узагальнене уявлення про їх поведінку за день. Наприклад, якщо спостерігається велике число відвідин станції технічного обслуговування або заправок різними працівниками, особливо в ті години, коли потік клієнтів найбільший – має сенс подумати про модернізацію парку автомобілів, щоб не упускати найбільш прибуткові години роботи.

Результати та їх обговорення

Вихідними даними для моделювання є інформація з баз даних «Замовлення» і «Персонал». Таблиця «Замовлення» містить дані про виклики, що поступили в службу таксі, і про хід їх виконання (дата і час надходження замовлення, точна адреса і район виклику, дані про оператора, диспетчера і водія, які обслуговують це замовлення тощо); таблиця «Персонал» – про дії, що виконуються співробітниками служби за деякий проміжок часу.

Була поставлено й вирішена задача проектування програмної системи, здатної вирішувати таку задачу інтелектуального аналізу даних, як прогнозування. У якості математичного апарата був обраний метод штучних нейронних мереж [1]. Для навчання мережі використовується метод зворотного поширення помилок [2, 3].

Нейронна мережу зворотного поширення складається з декількох шарів нейронів, причому кожен нейрон шару і пов'язаний з кожним нейроном шару $i+1$, тобто мова йде про повнозв'язану нейронну мережу. Приклад повнозв'язаної двошарової нейронної мережі наведено на рис. 1.

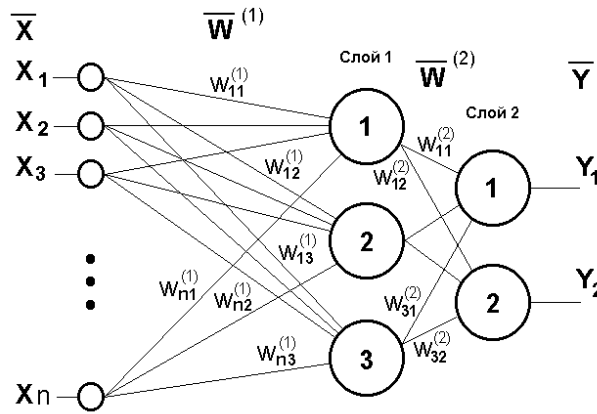


Рис. 1 Двошарова повнзв’язана нейронна сітка

Мережа, зображена на малюнку, має n входів. На них надходять сигнали, що йдуть далі за синапсах на 3 нейрони, які утворюють перший шар. Вихідні сигнали першого шару передаються двом нейронам другого шару. Останні, у свою чергу, видають два вихідних сигнали.

У загальному випадку завдання навчання нейронної мережі зводиться до знаходження якоїсь функціональної залежності $Y = F(X)$, де X – вхідний, а Y – вихідний вектори. Така задача, при обмеженому наборі вхідних даних, має нескінченну безліч рішень. Для обмеження простору пошуку при навчанні ставиться завдання мінімізації цільової функції помилки нейронної мережі, яка знаходиться за методом найменших квадратів:

$$E(w) = \frac{1}{2} \sum_{j=1}^p (y_j - d_j)^2 \tag{1}$$

де y_j – значення j -го виходу нейронної мережі;

d_j – цільове значення j -го виходу;

p – число нейронів у вихідному шарі;

Розрахункове значення Y визначається за формулою $Y = f(S)$,

де S – стан нейрона, що визначається за формулою:

$$S = \sum_{i=1}^n x_i w_i \tag{2}$$

де n – число входів нейрона;

x_i – значення i -го входу нейрона;

w_i – вага i -го синапсу

Функція f називається активаційний функцією. Ми використовуємо сигмоїдальну функцію:

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{3}$$

Навчання нейронної мережі проводиться методом градієнтного спуску, тобто на кожній ітерації зміна ваги проводиться за формулою

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}} \tag{4}$$

де η – параметр, що визначає швидкість навчання.

Загальний алгоритм навчання нейронної мережі приведений в [3,4].

Проектування інформаційної системи здійснювалося на уніфікованій мові моделювання UML [4]. Можливості системи у вигляді діаграми варіантів використання представлені на рис. 2, її структура у вигляді діаграми класів – на рис. 3.

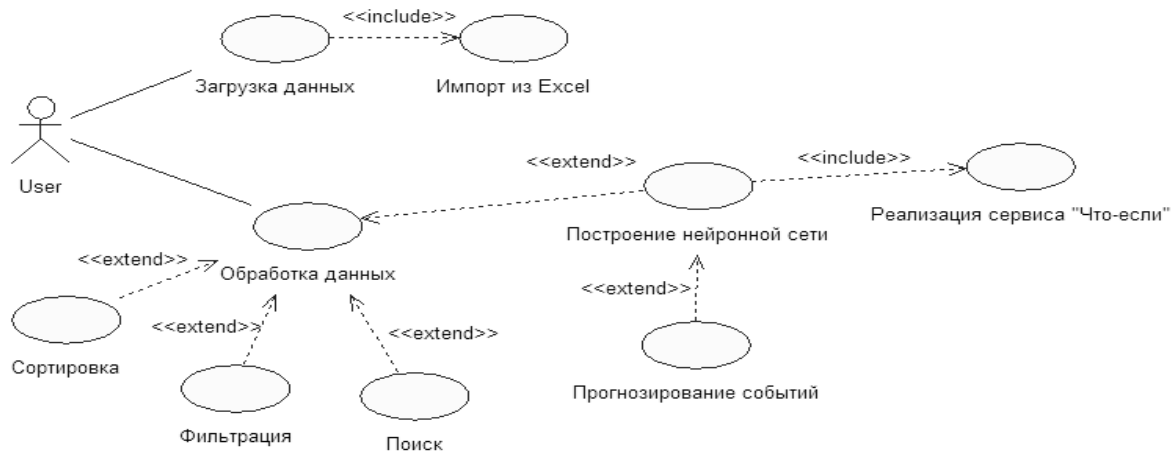


Рис. 2 Діаграма варіантів використання

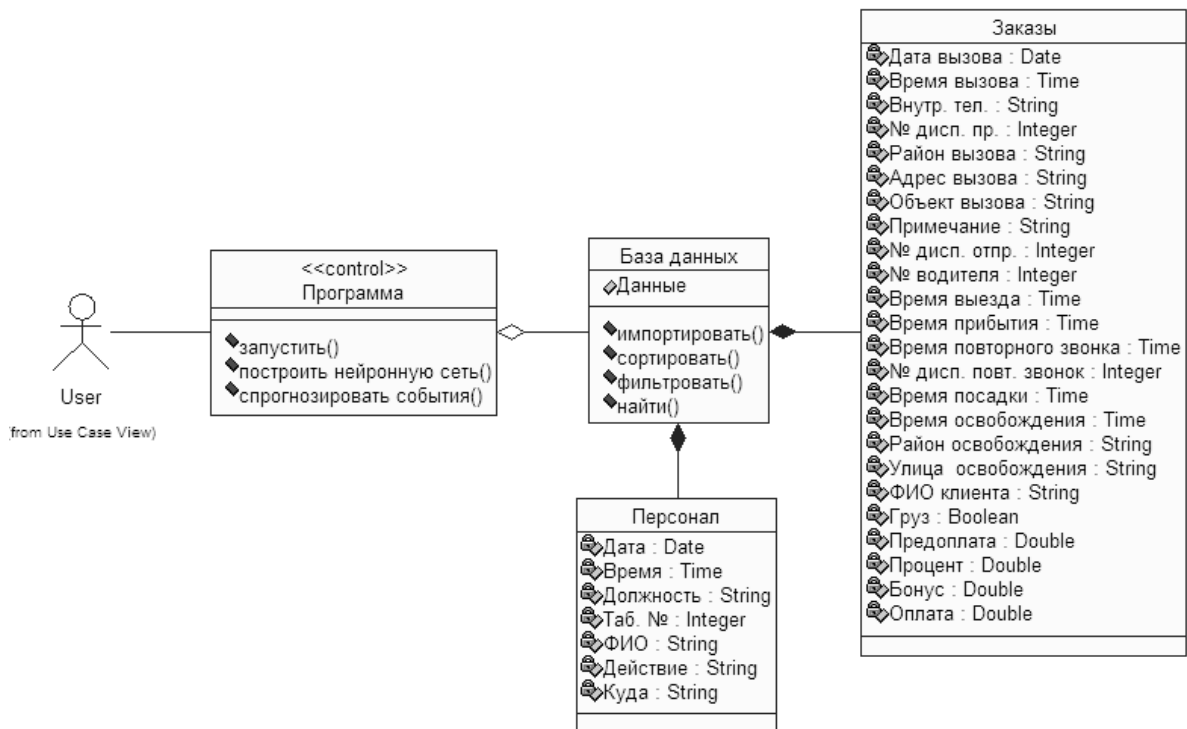


Рис. 3 Діаграма класів

Програмна реалізація спроектованої системи була здійснена в середовищі програмування Borland-Delphi 7. Розроблений додаток на даний момент забезпечує основні можливості, після закінчення навчання нейронної мережі надається доступ тільки до сервісу «Що-якщо».

Розглянемо як приклад рішення задачі прогнозування поведінки працівників протягом зміни. В якості вхідних параметрів у нейронній мережі буде використовуватися табельний номер працівника, час доби і дію (прихід, ухід, переміщення). В якості вихідного поля вибираємо місце дії, тобто куди відправиться (або звідки повернеться) цей працівник у даний час. У мережі буде один прихований шар з чотирма нейронами. У навчальну множину включимо 80% вихідних даних. Описані налаштування і результат роботи програми показані на малюнках 4 – 7.

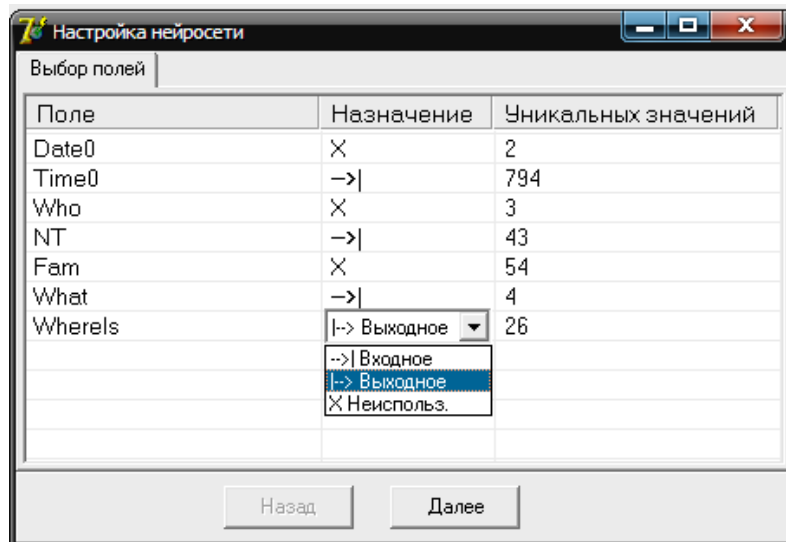


Рис. 4 Вибір вхідних та вихідних полів

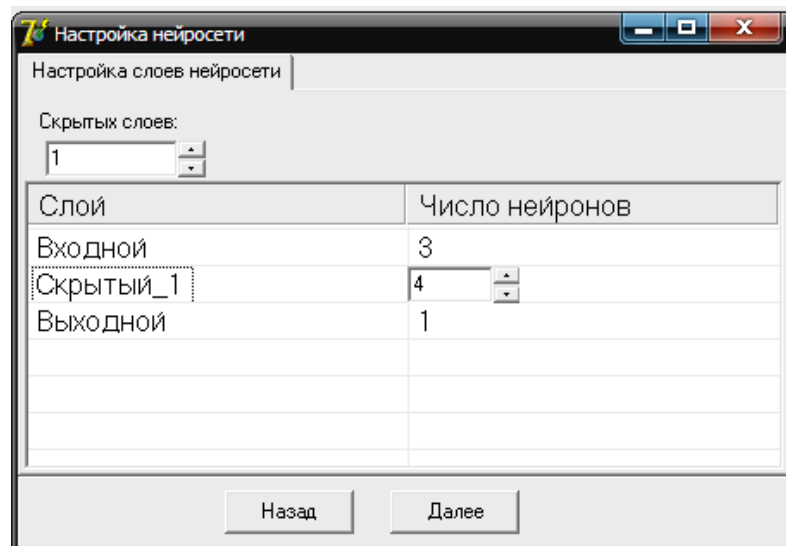


Рис. 5 Налаштування шарів нейронної мережі

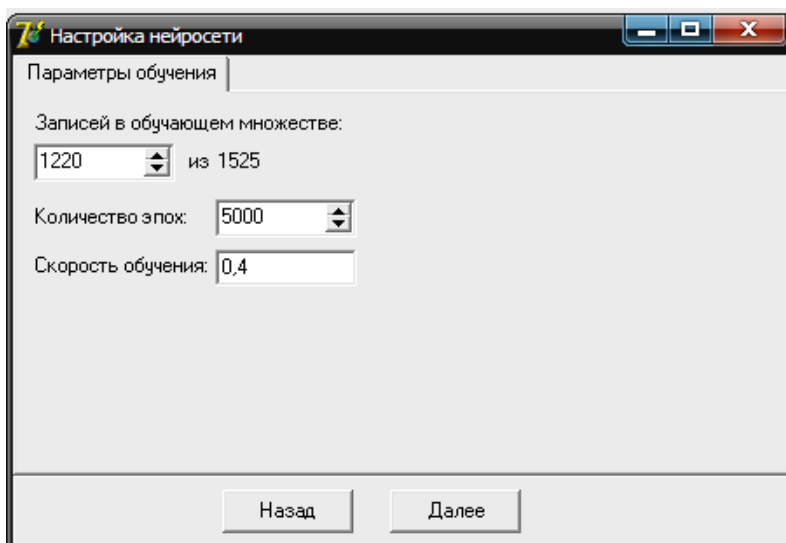


Рис. 6 Налаштування параметрів навчання мережі

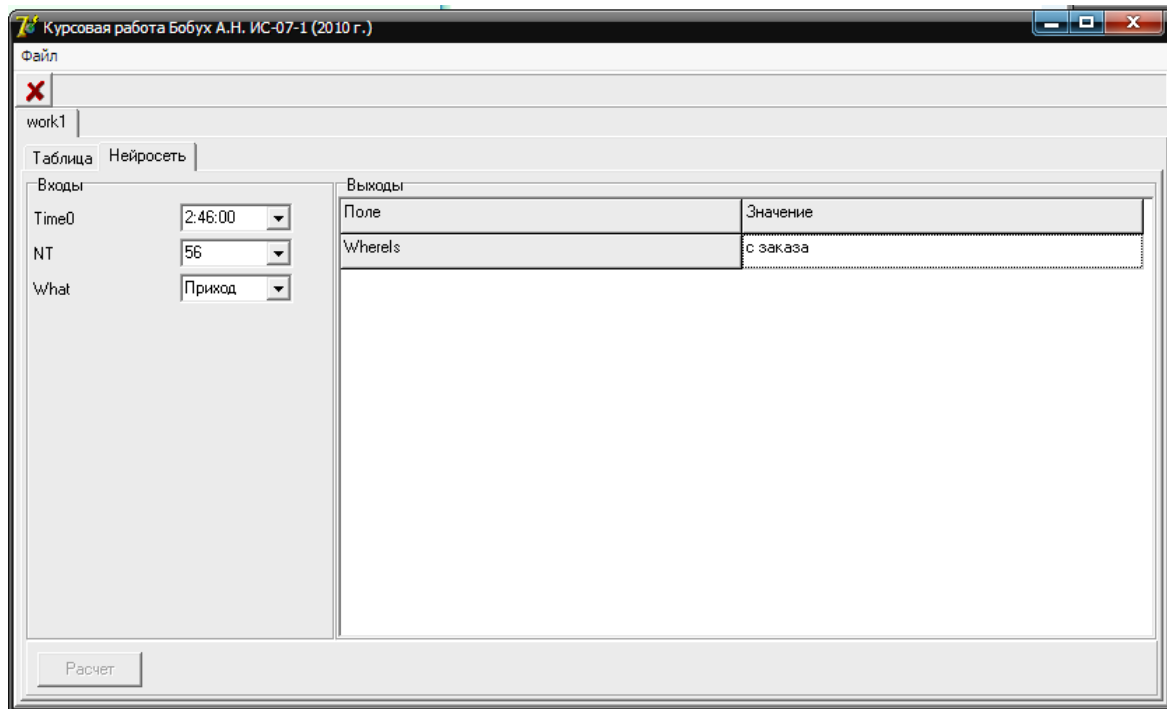


Рис. 7 Сервіс «Що-якщо»

Висновки

У ході роботи була вивчена інформація, що надійшла в деяку службу таксі протягом доби, побудовано математичну та інформаційну моделі системи інтелектуального аналізу даної інформації, здійснено їх комп'ютерну реалізацію.

Створена система дає можливість різними способами проаналізувати інформацію, що надходить, і витягти з неї додаткові відомості, які можна застосувати для підвищення ефективності роботи служби таксі.

ЛІТЕРАТУРА

1. Чубукова И.А. Data Mining: Учебное пособие / И.А. Чубукова. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2006. – 382 с.
2. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотипное. – М.: Горячая линия – Телеком, 2002. – 382 с.
2. Ковалевский С.В., Гитис В.Б. Создание и применение нейронных сетей для решения прикладных задач: Учебно-методическое пособие. – Краматорск: ДГМА, 2005. – 80 с.
3. Мельников А.Ю. Объектно-ориентированный анализ и проектирование информационных систем: Учебное пособие. – Краматорск: ДГМА, 2006.– 184 с.

Надійшла 25.10.2010