

УДК 004.92

Є.Є. Федоров, д-р техн. наук, проф.
В.О. Патрушев, ст. викл.
О.І. Патрушева, ас.Донецький національний технічний університет
fedorovee75@ukr.net,
wa_pat@ukr.net,
patrushevao.op@gmail.com

Метод прогнозу прибутку Інтернет-магазину на основі повнозв'язної довго короткочасної пам'яті, яка навчена за допомогою генетичного алгоритму

У статті розглянуті і проаналізовані існуючі методи прогнозу формування прибутку. Виходячи з основних переваг та недоліків, був розроблений і реалізований нейромережевий метод прогнозу прибутку функціонуючого інтернет-магазину. В основу методу закладена нейронна мережа FC-LSTM, архітектура якої була модифікована введенням тимчасових затримок у вхідному шарі. Навчання FC-LSTM проводилося на основі комбінації як алгоритмів BPTT і RTRL, так і генетичного алгоритму. Для оцінки ефективності запропонованого методу було проведено чисельні дослідження, які доводять ефективність обраної мережі та її архітектури.

Ключові слова: прогноз, інтернет-магазин, штучна нейронна мережа, FC-LSTM, генетичний алгоритм.

DOI: 10.31474/1996-1588-2017-2-25-118-125

Вступ

В даний час однією з найважливіших проблем, з якою стикаються сучасні інтернет магазини є відсутність вбудованих механізмів, які дозволяють виконувати аналіз і подальше прогнозування результатів господарської діяльності, спрямованих на підвищення прибутку. Це призводить до того, що господарська діяльність власника інтернет магазину, спрямована на збільшення прибутку, може бути неефективною. Тому розробка способів прогнозу прибутку інтернет магазину є актуальною.

Метою роботи є розробка методу дослідження процесу формування прибутку інтернет-магазину. Для досягнення мети були поставлені і вирішені наступні завдання:

- провести аналіз існуючих методів прогнозу;
- вибрати штучну нейронну мережу прогнозу;
- визначити структури моделі штучної нейронної мережі;
- вибрати критерій оцінки ефективності нейросетевой моделі прогнозу;
- навчити моделі штучної нейронної мережі;
- виконати чисельні дослідження.

Результати розробки та досліджень

На сьогоднішній день в якості інструменту для прогнозу відомо безліч підходів, серед яких можна виділити:

- регресивні і авторегресійні методи прогнозування;
- методи, засновані на експоненційному згладжуванні;
- методи на базі ланцюгів Маркова;
- на базі класифікаційно-регресійних дерев;
- нейромережеві методи прогнозу.

В роботі [5] наведено порівняльну характеристику перерахованих способів.

Оскільки використання нейронних мереж при прогнозуванні дає відчутну перевагу, яке полягає в тому, що: взаємозв'язки між факторами досліджуються на готових моделях; не потрібні ніякі припущення щодо розподілу факторів; апріорна інформація про фактори може бути відсутньою; вихідні дані можуть сильно корелювати, бути неповними або зашумленими; можливий аналіз систем з високим ступенем не лінійності; швидка розробка моделі; висока адаптивність; можливий аналіз систем з великою кількістю чинників; не вимагається повний перебір всіх можливих моделей; можливий аналіз систем з неоднорідними факторами, в статті буде

використовуватися нейросетевий спосіб прогнозу.

Оскільки фактори, що впливають на отримання прибутку змінюються як в короткостроковій, так і довгостроковій перспективі, то прогноз здійснюється на основі динамічних (з тимчасовою затримкою) штучних нейронних мереж. В якості таких мереж найчастіше використовуються:

- нейромережа Джордона (JNN) [6], яка є динамічною рекурентною двошаровою мережею і побудована на базі MLP.

- нейромережа Елмана (ENN) або проста рекурентна мережа (SRN) [6], яка є динамічною рекурентною двошаровою мережею і побудована на базі MLP.

- рекурентний багатшаровий перцептрон (RMLP) [7], який є динамічною рекурентною багатшаровою мережею і побудований на базі MLP.

- модель нелінійної авторегресії (NAR) [8], яка є динамічною нерекурентною двошаровою мережею і побудована на базі MLP.

- модель нелінійної авторегресії-ковзного середнього (NARMA) [8], яка є динамічною рекурентною двошаровою мережею та побудована на базі MLP.

- повнозв'язна довго короткочасної пам'ять (FC-LSTM) [9, 10], яка є рекурентною двошаровою мережею і побудована на базі MLP.

В табл. 1 приведені порівняльні характеристики динамічних мереж.

Кількість нейронів вхідного шару відповідає кількості контрольованих параметрів. Дослідження проводилися для одного показника – прибуток інтернет-магазину. Це відповідає

одному нейрону вхідного шару. Оскільки FC-LSTM дає найбільшу точність прогнозу, ця мережа вибирається для вирішення завдань прогнозування.

В даній статті запропоновано авторський варіант FC-LSTM, який передбачає тимчасові затримки у вхідному шарі.

Таблиця 1 - Результати експериментальних досліджень нейронних мереж

Мережа	RLMP	JNN	ENN (SRN)	NAR	NARMA	FC-LSTM
Критерій						
Наявність зворотного зв'язку	+	+	-	-	+	+
Затримка у вхідному шарі	-	-	-	+	+	+
Точність прогнозу	90	80	85	88	93	95

Структура FC-LSTM з тимчасовими затримками у вхідному шарі зображена на рис. 1. В якості прихованого нейрона використовується блок пам'яті (Memory Block), який містить одну або більше комірок, і пов'язаний з вхідним шлюзом (Input Gate) і шлюзом забування (Forget Gate), які відповідають КІХ-фільтрам. Зв'язки між комірками і шлюзами дозволяють шлюзам враховувати стан осередку. Вихідний сигнал комірки надходить по зворотного зв'язку назад в блок пам'яті.

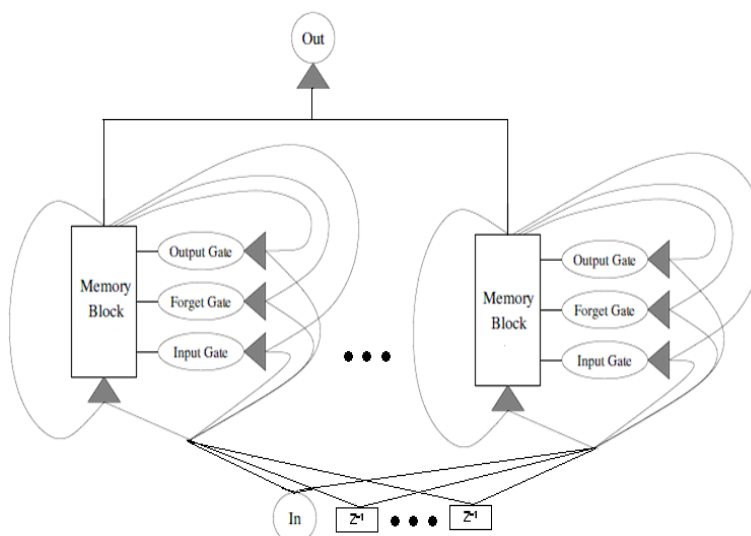


Рисунок 1 - Повнозв'язна довгокороткочасна пам'ять (FC-LSTM) з часовими затримками в вхідному шарі

$$y_j^{gin}(n) = \text{falignl} \left(\sum_{i=0}^{M^{(0)}} w_{ij}^{in-gin} y^{in}(n-(i+1)) + \sum_{v=1}^{S_j} w_{jv}^{s-gin} s_{jv}^c(n-1) \right),$$

$$j \in \overline{1, N^{(1)}},$$

$$y_j^g(n) = \text{falignl} \left(\sum_{i=0}^{M^{(0)}} w_{ij}^{in-g} y^{in}(n-(i+1)) + \sum_{v=1}^{S_j} w_{jv}^{s-g} s_{jv}^c(n-1) \right),$$

$$j \in \overline{1, N^{(1)}},$$

$$net_{jv}^c(n) = \sum_{i=0}^{M^{(0)}} w_{ijv}^{in-s} y^{in}(n-(i+1)) + \sum_{z=1}^{S_j} w_{jzv}^{c-s} y_{jz}^c(n-1)$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$s_{jv}^c(n) = y_j^g(n) s_{jv}^c(n-1) + 2y_j^{gin}(n) h(net_{jv}^c(n))$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$y_j^{gout}(n) = \text{falignl} \left(\sum_{i=0}^{M^{(0)}} w_{ij}^{in-gout} y^{in}(n-(i+1)) + \sum_{v=1}^{S_j} w_{jv}^{s-gout} s_{jv}^c(n-1) \right),$$

$$j \in \overline{1, N^{(1)}},$$

$$y_{jv}^c(n) = y_j^{gout}(n) h(s_{jv}^c(n)),$$

$$v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$y^{out}(n) = f \left(\sum_{i=1}^{N^{(1)}} \sum_{v=1}^{S_i} w_{iv}^{c-out} y_{iv}^c(n) \right).$$

$$f(s) = \frac{1}{1 + e^{-s}}, \quad h(s) = \tanh(s),$$

де $M^{(0)}$ – затримка у вхідному шарі;
 $N^{(1)}$ – кількість блоків пам'яті; S_j – кількість осередків в j -му блоці пам'яті; w_{ij}^{in-gin} , $w_{ij}^{in-g\phi}$, $w_{ij}^{in-gout}$, – вага зв'язку від i -го вхідного нейрона в момент часу $n - (i + 1)$ до вхідного, вихідного і забування шлюзів j -го блоку пам'яті в момент часу n ; w_{ijv}^{in-s} – вага зв'язку від i -го вхідного нейрона в момент часу $n - (i + 1)$ до v -й осередку j -го блоку пам'яті в момент часу n ; w_{jv}^{s-gin} , $w_{jv}^{s-g\phi}$, w_{jv}^{s-gout} , w_{jv}^{c-gin} , $w_{jv}^{c-g\phi}$, w_{jv}^{c-gout} , – вага зв'язку від v -й осередку j -го блоку пам'яті в момент часу $n - (i + 1)$ до вхідного, вихідного і забування шлюзів j -го блоку пам'яті в момент часу n ; w_{jzv}^{c-s} – вага зв'язку від z -й осередку j -го блоку пам'яті в момент часу $n - (i + 1)$ до v -й осередку j -го блоку пам'яті в момент

часу n ; w_{iv}^{c-out} – вага зв'язку від v -й осередку i -го блоку пам'яті в момент часу $n - (i + 1)$ до v -й осередку i -го блоку пам'яті в момент часу n ; $y_j^{gin}(n)$ – вихід вхідного шлюзу j -го блоку пам'яті; $y_j^{gout}(n)$ – вихід вихідного шлюзу j -го блоку пам'яті; $y_j^{g\phi}(n)$ – вихід шлюзу забування j -го блоку пам'яті; $y_{jv}^c(n)$ – вихід v -й осередку j -го блоку пам'яті; $y^{in}(n)$ – вихід вхідного нейрона; $y^{out}(n)$ – вихід вихідного нейрона; $s_{jv}^c(n)$ – стан v -й осередку j -го блоку пам'яті; f – функція активації нейронів.

Для визначення структури моделі FC-LSTM (визначення кількості блоків пам'яті) був проведений ряд експериментів, результати яких представлені на рис. 2.

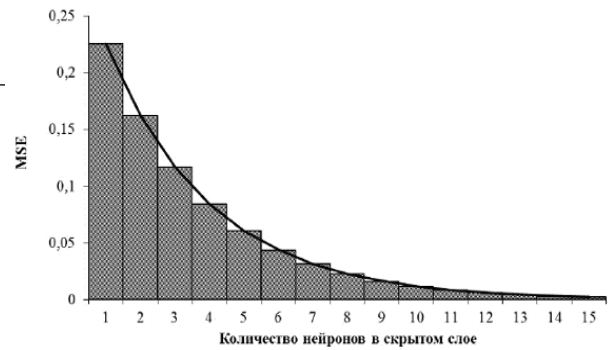


Рисунок 2 – Гістограма залежності значення середньоквадратичної помилки від кількості нейронів у скритому шарі

В якості вхідних даних для навчання FC-LSTM були використані вибірки значень по господарської діяльності існуючого інтернет-магазину за 2016 рік. Критерієм вибору структури моделі мережі було мінімальна середньоквадратична помилка прогнозу. Як видно з рис. 2, зі збільшенням кількості однокомірчастих блоків пам'яті значення помилки зменшується. Для прогнозу продажів і отримання подальшої прибутку досить використовувати 5 однокомірчастих блоків пам'яті, оскільки при подальшому збільшенні кількості однокомірчастих блоків пам'яті, зміна значення помилки є незначним.

В роботі для навчання моделі FC-LSTM обраний критерій адекватності моделі, який означає вибір таких значень параметрів $W = \{w_{ij}^{in-gin}, w_{ij}^{in-g\phi}, w_{ij}^{in-gout}, w_{ijv}^{in-s}, w_{jv}^{s-gin}, w_{jv}^{s-g\phi}, w_{jv}^{s-gout}, w_{jv}^{c-gin}, w_{jv}^{c-g\phi}, w_{jv}^{c-gout}, w_{jzv}^{c-s}, w_{iv}^{c-out}\}$, які доставляють

мінімум середньоквадратичної помилки (різниці виходу за моделлю і бажаного виходу):

$$F = \frac{1}{P} \sum_{p=1}^P (y_p - d_p)^2 \rightarrow \min_W (1)$$

Навчання моделі FC-LSTM підпорядковане критерієм (3), для чого в статті пропонується комбінація алгоритмів зворотного поширення в часі (BPTT) і рекуррентного навчання в реальному часі (RTRL), яка полягає в наступному

1. Номер ітерації навчання $n = 2$, ініціалізація за допомогою рівномірного розподілу на інтервалі (0,1) або [-0.5, 0.5] ваг $w_{ij}^{in-gin}(n)$, $w_{ij}^{in-g\phi}(n)$, $w_{ij}^{in-gout}(n)$, $w_{ijv}^{in-s}(n)$, $i \in \overline{0, M^{(0)}}$, $j \in \overline{1, N^{(1)}}$, $v \in \overline{1, S_j}$, $w_{jvj}^{s-gin}(n)$, $w_{jvj}^{s-g\phi}(n)$, $w_{jvj}^{s-gout}(n)$, $w_{jvj}^{c-gin}(n)$, $w_{jvj}^{c-g\phi}(n)$, $w_{jvj}^{c-gout}(n)$, $w_{jzjv}^{c-s}(n)$, $j \in \overline{1, N^{(1)}}$, $z \in \overline{1, S_j}$, $v \in \overline{1, S_j}$, $w_{iv}^{c-out}(n)$, $i \in \overline{1, N^{(1)}}$, $v \in \overline{1, S_j}$, де $M^{(0)}$ – затримка у вхідному шарі, $N^{(1)}$ – кількість блоків пам'яті, S_j – кількість осередків у j -м блоці пам'яті.

2. Задається навчальна множина $\{(x_\mu, d_\mu) | x_\mu \in R, d_\mu \in R\}$, $\mu \in \overline{1, P}$, де x_μ – μ -е вхідне значення, що навчає, d_μ – μ -е вихідне значення, що навчає, P – потужність навчальної множини. Номер поточної пари з навчальної множини $\mu = 2$.

3. Початкове обчислення вихідного сигналу осередку

$$y^{in}(n - (i + 1)) = 0, \quad i \in \overline{1, M^{(0)}},$$

$$s_{jv}^c(n - 1) = 0, \quad y_{jv}^c(n - 1) = 0,$$

$$v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}}.$$

4. Обчислення вихідного сигналу для кожного шару (прямий хід)

$$y^{in}(n - 1) = x_{\mu-1},$$

$$y_j^{gin}(n) = f(\text{net}_j^{gin}(n)), \quad j \in \overline{1, N^{(1)}},$$

$$\text{net}_j^{gin}(n) = \sum_{i=0}^{M^{(0)}} w_{ij}^{in-gin}(n) y^{in}(n - (i + 1)) + \sum_{v=1}^{S_j} w_{jvj}^{s-gin}(n) s_{jv}^c(n - 1) +$$

$$+ \sum_{v=1}^{S_j} w_{jvj}^{c-gin}(n) y_{jv}^c(n - 1)$$

$$y_j^{g\phi}(n) = f(\text{net}_j^{g\phi}(n)) \quad j \in \overline{1, N^{(1)}},$$

$$\text{net}_j^{g\phi}(n) = \sum_{i=0}^{N^{(0)}} w_{ij}^{in-g\phi}(n) y^{in}(n - (i + 1)) + \sum_{v=1}^{S_j} w_{jvj}^{s-g\phi}(n) s_{jv}^c(n - 1) +$$

$$+ \sum_{v=1}^{S_j} w_{jvj}^{c-g\phi}(n) y_{jv}^c(n - 1)$$

$$\text{net}_{jv}^c(n) = \sum_{i=0}^{M^{(0)}} w_{ijv}^{in-s}(n) y^{in}(n - (i + 1))$$

$$+ \sum_{z=1}^{S_j} w_{jzjv}^{c-s}(n) y_{jz}^c(n - 1) \quad v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}},$$

$$s_{jv}^c(n) = y_j^{g\phi}(n) s_{jv}^c(n - 1) + 2 y_j^{gin}(n) h(\text{net}_{jv}^c(n))$$

$$, \quad v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}},$$

$$y_j^{gout}(n) = f(\text{net}_j^{gout}(n)), \quad j \in \overline{1, N^{(1)}},$$

$$\text{net}_j^{gout}(n) = \sum_{i=0}^{M^{(0)}} w_{ij}^{in-gout}(n) y^{in}(n - (i + 1)) + \sum_{v=1}^{S_j} w_{jvj}^{s-gout}(n) s_{jv}^c(n - 1) +$$

$$+ \sum_{v=1}^{S_j} w_{jvj}^{c-gout}(n) y_{jv}^c(n - 1)$$

$$y_{jv}^c(n) = y_j^{gout}(n) h(s_{jv}^c(n)),$$

$$v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}},$$

$$y^{out}(n) = f(\text{net}^{out}(n))$$

$$\text{net}^{out}(n) = \sum_{i=1}^{N^{(1)}} \sum_{v=1}^{S_j} w_{iv}^{c-out}(n) y_{iv}^c(n - 1),$$

$$f(s) = \frac{1}{1 + e^{-s}}, \quad h(s) = \tanh(s).$$

5. Обчислення енергії помилки прогнозу нейронної мережі

$$E(n) = \frac{1}{2} e^2(n),$$

$$e(n) = y^{out}(n) - d_{\mu-1}.$$

6. Налаштування синаптичних ваг на основі узагальненого дельта правила (зворотний хід)

$$w_{iv}^{c-out}(n + 1) = w_{iv}^{c-out}(n) - \eta \frac{\partial E(n)}{\partial w_{iv}^{c-out}(n)},$$

$$i \in \overline{1, N^{(1)}}, \quad v \in \overline{1, S_j},$$

$$w_{ij}^{in-gout}(n + 1) = w_{ij}^{in-gout}(n) - \eta \frac{\partial E(n)}{\partial w_{ij}^{in-gout}(n)},$$

$$, \quad i \in \overline{1, N^{(0)}}, \quad j \in \overline{1, N^{(1)}},$$

$$w_{ijv}^{in-s}(n + 1) = w_{ijv}^{in-s}(n) - \eta e_{jv}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{ijv}^{in-s}(n)},$$

$$, \quad i \in \overline{1, N^{(0)}}, \quad v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}},$$

$$w_{jzjv}^{c-s}(n + 1) = w_{jzjv}^{c-s}(n) - \eta \sum_{z=1}^{S_j} e_{jz}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{jzjv}^{c-s}(n)},$$

$$z \in \overline{1, S_j}, \quad v \in \overline{1, S_j}, \quad j \in \overline{1, N^{(1)}},$$

$$w_{ij}^{in-gin}(n+1) = w_{ij}^{in-gin}(n) - \eta \sum_{v=1}^{S_j} e_{jv}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{ij}^{in-gin}(n)},$$

$$i \in \overline{1, N^{(0)}}, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{ij}^{in-g\phi}(n+1) = w_{ij}^{in-g\phi}(n) - \eta \sum_{v=1}^{S_j} e_{jv}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{ij}^{in-g\phi}(n)}$$

$$, i \in \overline{1, N^{(0)}}, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{s-gout}(n+1) = w_{jvj}^{s-gout}(n) - \eta \frac{\partial E(n)}{\partial w_{jvj}^{s-gout}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{s-gin}(n+1) = w_{jvj}^{s-gin}(n) - \eta \sum_{z=1}^{S_j} e_{jz}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{jvj}^{s-gin}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{s-g\phi}(n+1) = w_{jvj}^{s-g\phi}(n) - \eta \sum_{z=1}^{S_j} e_{jz}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{jvj}^{s-g\phi}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{c-gout}(n+1) = w_{jvj}^{c-gout}(n) - \eta \frac{\partial E(n)}{\partial w_{jvj}^{c-gout}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{c-gin}(n+1) = w_{jvj}^{c-gin}(n) - \eta \sum_{z=1}^{S_j} e_{jz}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{jvj}^{c-gin}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

$$w_{jvj}^{c-g\phi}(n+1) = w_{jvj}^{c-g\phi}(n) - \eta \sum_{z=1}^{S_j} e_{jz}^c(n) \frac{\partial s_{jv}^c(n)}{\partial w_{jvj}^{c-g\phi}(n)}$$

$$, v \in \overline{1, S_j}, j \in \overline{1, N^{(1)}},$$

де η – параметр, що визначає швидкість навчання (при великому η навчання відбувається швидше, але збільшується небезпека одержати невірне рішення), $0 < \eta < 1$.

$$\frac{\partial E(n)}{\partial w_{iv}^{c-out}(n)} = y_{iv}^c(n-1) \delta^{out}(n),$$

$$\frac{\partial E(n)}{\partial w_{ij}^{in-gout}(n)} = y_i^{in}(n-1) \delta_j^{gout}(n),$$

$$\frac{\partial E(n)}{\partial w_{jvj}^{s-gout}(n)} = s_{jv}^c(n-1) \delta_j^{gout}(n),$$

$$\frac{\partial E(n)}{\partial w_{jvj}^{c-gout}(n)} = y_{jv}^c(n-1) \delta_j^{gout}(n),$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{jv}^{in-s}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{jv}^{in-s}(n-1)} y_j^s(n) + & n > 2, \\ + 2y^{in}(n-(j+1))y_j^{gin}(n)h'(net_{jv}^c(n)), & \\ 2y^{in}(n-(j+1))y_j^{gin}(n)h'(net_{jv}^c(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{izjv}^{c-s}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{izjv}^{c-s}(n-1)} y_j^{g\phi}(n) + 2y_{iz}^c(n-1)y_j^{gin}(n)h'(net_{jv}^c(n)), & n > 2 \\ 2y_{iz}^c(n-1)y_j^{gin}(n)h'(net_{jv}^c(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{ij}^{in-g\phi}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{ij}^{in-g\phi}(n-1)} y_j^{\phi}(n) + & n > 2, \\ + y^{in}(n-(j+1))s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & \\ y^{in}(n-(j+1))s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{jzj}^{c-g\phi}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{jzj}^{c-g\phi}(n-1)} y_j^{\phi}(n) + & n > 2, \\ + y_{jz}^c(n-1)s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & \\ y_{jz}^c(n-1)s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{jzj}^{s-g\phi}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{jzj}^{s-g\phi}(n-1)} y_j^{\phi}(n) + & n > 2, \\ + s_{jz}^c(n-1)s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & \\ s_{jz}^c(n-1)s_{jv}^c(n-1)f'(net_j^{\phi}(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{ij}^{in-gin}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{ij}^{in-gin}(n-1)} y_j^{g\phi}(n) + & n > 2 \\ + y^{in}(n-(j+1))g(net_{jv}^c(n))f'(net_j^{gin}(n)), & \\ y^{in}(n-(j+1))g(net_{jv}^c(n))f'(net_j^{gin}(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{jzj}^{c-gin}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{jzj}^{c-gin}(n-1)} y_j^{g\phi}(n) + & n > 2 \\ + y_{jz}^c(n-1)g(net_{jv}^c(n))f'(net_j^{gin}(n)), & \\ y_{jz}^c(n-1)g(net_{jv}^c(n))f'(net_j^{gin}(n)), & n = 2 \end{cases}$$

$$\frac{\partial s_{jv}^c(n)}{\partial w_{jzj}^{s-gin}(n)} = \begin{cases} \frac{\partial s_{jv}^c(n-1)}{\partial w_{jzj}^{s-gin}(n-1)} y_j^{g\phi}(n) + & n > 2 \\ + s_{jz}^c(n-1)g(net_{jv}^c(n))f'(net_j^{gin}(n)), & \\ s_{jz}^c(n-1)g(net_{jv}^c(n))f'(net_j^{gin}(n)), & n = 2 \end{cases}$$

$$e_{jv}^c(n) = y_j^{gout}(n)h'(s_{jv}^c(n))w_{jv}^{c-out}(n)\delta^{out}(n)$$

$$\delta^{out}(n) = f'(net^{out}(n))(y^{out}(n) - d_{\mu})$$

$$\delta_j^{gout}(n) = f'(net_j^{gout}(n)) \sum_{v=1}^{S_j} h(s_{jv}^c(n))w_{jv}^{c-out}(n)\delta^{out}(n)$$

7. Перевірка умови завершення

Якщо $(n-1) \bmod P > 0$, тоді

$\mu = \mu + 1, n = n + 1$, перехід до п.4.

Якщо $(n-1) \bmod P = 0$ та

$\frac{1}{P} \sum_{s=1}^P E(n-P+s) > \varepsilon$, тоді $n = n + 1$, перехід

до п.2.

Якщо $(n-1) \bmod P = 0$ та
 $\frac{1}{P} \sum_{s=1}^P E(n-P+s) < \varepsilon$, тоді завершення.

Альтернативою алгоритму зворотного поширення є генетичний алгоритм. В якості особини в генетичному алгоритмі використовується вектор ваг W , а в якості фітнес-функції використовується критерій (1).

Авторами досліджувалися три типи генетичного алгоритму (таблиця 2) [13, 14].

Таблиця 2 – Структура генетичних алгоритмів трьох типів

Оператори генетичного алгоритму	Генетичний алгоритм		
	Тип 1	Тип 2	Тип 3
Репродукція	Рівномірний відбір	Лінійно підпорядкований відбір	Комбінація рівномірного та лінійно підпорядкованого відбору
Кроссингвер	Рівномірний, відбір особин – схрещування кращих особин з гіршими	Рівномірний, відбір особин – схрещування кращих особин з кращими	Рівномірний, відбір особин – комбінація схрещування кращих особин з гіршими та кращих особин з кращими
Мутація	однорідна з високою ймовірністю	однорідна з низкою ймовірністю	Неоднорідна
Редукція	рівномірна схема	селекційна схема	комбінація рівномірної та селекційної схеми

Перший тип досліджує весь простір пошуку і не є спрямованим. Для цього типу можуть бути втрачені кращі рішення. Він вимагає значного часу пошуку. Другий тип є

спрямованим. Для цього типу можливе попадання в локальний оптимум. Третій тип є комбінованим, тобто поєднує спрямованість пошуку з дослідженням всього простору пошуку. Для вибору найбільш ефективного алгоритму навчання було проведено чисельне дослідження, результати якого представлені в таблиці 3. Для оцінки ефективності були обрані такі критерії, як кількість ітерацій і точність прогнозу.

Згідно з результатами порівняння, генетичний алгоритм типу 3 для задачі навчання нейронної мережі є найбільш підходящим для процедури навчання

Таблиця 3 - Порівняльна характеристика алгоритмів навчання нейронної мережі

Алгоритм	BPTT + RTRL	Генетичний алгоритм		
		Тип 1	Тип 2	Тип 3
Характеристика				
Кількість ітерацій	2000	320	260	200
Точність прогноза (СКО)	$236 \cdot 10^{-3}$	$360 \cdot 10^{-4}$	$360 \cdot 10^{-4}$	$300 \cdot 10^{-4}$

Висновки

1. Для вирішення проблеми підвищення точності прогнозу отримання прибутку інтернет-магазином були досліджені існуючі методи прогнозу прибутку підприємством. Дані дослідження показали, що на сьогоднішній день найбільш ефективним є використання штучних нейронних мереж.

2. Для підвищення якості прогнозу була обрана найбільш ефективна динамічна нейромережа FC-LSTM, модифікована (введенням тимчасових затримок у вхідному шарі) і ідентифікована структура її моделі. Проведені експерименти показали, що при десяти прихованих нейронах (однорічних блоків пам'яті), значення середньоквадратичної помилки істотно не змінюється, і обрана мережа дає результати прогнозу з мінімальним відхиленням.

3. В якості алгоритму навчання була запропонована комбінація алгоритмів зворотного поширення в часі (BPTT) і рекурентного навчання в реальному часі (RTRL), а також був використаний генетичний алгоритм. Це дозволило забезпечити високу швидкість і точність прогнозу.

4. Запропонований підхід може використовуватися в різних інтелектуальних системах прогнозу.

Список літератури

1. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. . . . кандидата технических наук: 05.13.18 – математическое моделирование, численные методы и комплексы программ / Чучуева Ирина Александровна. – Москва, 2012. – 153 с.
2. Бодянский Е.В. Искусственные нейронные сети: архитектуры, обучение, применения / Е.В. Бодянский, О.Г. Руденко. – Харьков: Телетех, 2004. – 159 с.
3. Haykin S. Neural networks / S. Haykin. – NY: Pearson Education, 1999. – p. 823.
4. Галушкин А.И. Теория нейронных сетей / А.И. Галушкин. – М.: ИПРЖР, 2000. – 416 с.
5. Каллан Р. Основные концепции нейронных сетей / Р. Каллан. – М.: Издательский дом «Вильямс», 2001. – 288 с.
6. Sivanandam S.N. Introduction to Neural Networks using Matlab 6.0 / S.N. Sivanandam, S. Sumathi, S.N. Deera – New Delhi: The McGraw-Hill Comp., Inc., 2006. – 660 p.
7. Галушкин А.И. Нейронные сети: история развития теории: Учеб. пособие для вузов. / А.И. Галушкин, Я.З. Цыпкин. – М.: ИПРЖР, 2001. – 840 с.
8. Осовский С. Нейронные сети для обработки информации / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
9. Комарцова Л.Г. Нейрокомпьютеры / Л.Г. Комарцова, А.В. Максимов. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2002. – 320 с.
10. Федоров Е.Е. Искусственные нейронные сети: монография / Е.Е. Федоров. – Красноармейск: ДВНЗ «ДонНТУ», 2016. – 338 с.
11. Hochreiter S. Long short-term memory / S. Hochreiter, J. Schmidhuber // Technical Report FKI-207-95, Fakultat fur Informatik, Technische Universitat Munchen.
12. Gers F. Long Short-Term Memory in Recurrent Neural Networks / F. Gers // PhD thesis, Ecole Polytechnique Federale de Lausanne.
13. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская. – М.: Горячая линия – Телеком, 2006. – 452 с.
14. Скобцов Ю.А. Метаэвристики: [монография] / Ю.А. Скобцов, Е.Е. Федоров. – Донецк: Ноулидж, 2013. – 426 с.

References

1. Chuchueva, I.A. (2012) "Model of forecasting of time series in a sample of maximum likeness: dis. . . . candidate of technical sciences: 05.13.18 - mathematical modeling, numerical methods and complexes of programs" [Model' prognozirovaniya vremennykh ryadov po vyborke maksimal'nogo podobiya: dis. . . . kandidata tekhnicheskikh nauk], Moscow, 153 p.
2. Bodiansky, E.V., Rudenko, O.G., (2004) "Artificial Neural Networks: Architecture, Training, Applications", Kharkiv, Teletech, 159 p.
3. Haykin, S. (1999) "Neural networks" [Neyronnyye seti], NY: Pearson Education, 1999, 823 p.
4. Galushkin, A.I. (2000), "Theory of Neural Networks" [Teoriya neyronnykh setey], Moscow, IPRHP, 416 pp.
5. Callan, R. (2001), "Basic Concepts of Neural Networks" [Osnovnyye kontseptsii neyronnykh setey], Moscow, Williams Publishing House, 288 pp.
6. Sivanandam, S.N. (2006) "Introduction to Neural Networks using Matlab 6.0", New Delhi: The McGraw-Hill Comp., Inc., 660 p.
7. Galushkin, A.I., Tsyppkin, Ya.Z. (2001) "Neural Networks: History of Theory Development: Textbook. allowance for high schools" [Neyronnyye seti: istoriya razvitiya teorii: Ucheb. posobiye dlya vuzov], Moscow, IPRHR, 840 p.
8. Osovsky, S. (2002), "Neural Networks for Information Processing", [Neyronnyye seti dlya obrabotki informatsii], Moscow, Finance and Statistics, 344 pp.
9. Komartsova, L.G., Maximov, A.V. (2002), "Neurocomputers" [Neyrokomputeri], Moscow, Izv-MGTU them. N.E. Bauman, 320 p.
10. Fedorov, E. E. (2016) "Artificial Neural Networks: Monograph" [Iskusstvennyye neyronnyye seti: monografiya], Krasnoarmeysk: State University "DonNTU", 338 p.
11. Hochreiter, S., Schmidhuber, J. (1995), "Long short-term memory", Technical Report FKI-207-95, Fakultat fur Informatics, Technische Universitat Munchen.
12. Gers, F. "Long Short-Term Memory in Recurrent Neural Networks", PhD thesis, Ecole Polytechnique Federale de Lausanne.
13. Rutkovskaya, D. (2006), "Neural networks, genetic algorithms and fuzzy systems", [Neyronnyye seti, geneticheskiye algoritmy i nechetkiye sistemy], Moscow, Hotline - Telecom, 452 p.
14. Skobtsov, Yu.A., Fedorov, E.E. (2013) "Meta-heuristics: [monograph]" [Metaevristiki: monografiya], Donetsk, Knowledge, 426 p.

Надійшла до редакції 12.11.1017

E.E.FEDOROV, V.A. PATRUSHEV, O.I. PATRUSHEVA

Donetsk National Technical University (Ukraine)

METHOD FOR THE PROMOTION OF INTERNET ACCOUNT OF THE SHOP ON THE BASIS OF A COMPLETE LONG-TERM SHORT-MEMORY MEMORY, TRAINING BY A GENETIC ALGORITHM

At present, one of the most important problems encountered by modern online stores is the lack of built-in mechanisms that allow for analysis and subsequent forecasting of the results of economic activities aimed at increasing profits. This leads to the fact that the business activities of the owner of the online store, aimed at increasing profits, may be ineffective. Therefore, the development of ways to forecast the profit of an online store is relevant. The purpose of the work is to develop a method for studying the process of generating the profit of an online store. To achieve the goal, the following tasks were set and solved: to analyze existing methods of forecasting; select an artificial neural network of forecast; determine the structure of the artificial neural network model; choose a criterion for evaluating the effectiveness of the neural network forecast model; train models of artificial neural network; carry out numerical studies. Since the use of artificial neural networks in predicting gives a tangible advantage, which is that: the relationship between factors is explored on ready-made models; no assumptions about the distribution of factors are required; A priori information about the factors may not be available; the source data can strongly correlate, be incomplete or noisy; analysis of systems with a high degree of non-linearity is possible; rapid development of the model; high adaptability; analysis of systems with a large number of factors is possible; do not require a complete overview of all possible models; analysis of systems with inhomogeneous factors is possible, the neural network method of forecasting was used in the article. Since the factors that affect profitability change both in the short and long term, the forecast is based on dynamic (with a temporary delay) artificial neural networks. Since the FC-LSTM neural network provides the highest prediction accuracy, this network is chosen to solve the online store profit forecasting tasks. This article proposes an authoritative version of FC-LSTM, which provides for time delays in the input layer. The performed experiments showed that with ten hidden neurons (single-cell memory blocks), the value of the root mean square error does not change significantly, and the author's version of FC-LSTM gives the results of the forecast with a minimum deviation. As a learning algorithm, a combination of BPTT and RTRL was proposed as well as a genetic algorithm (the choice of its operators was preferred: the reproduction is a combination of equally probable and linearly ordered selection; cross-over - equally probable, selection of individuals - a combination of crossings of the best individuals with the worst and best individuals with the best: mutation - heterogeneous: reduction - a combination of equilibrium and breeding schemes s) This allowed for high speed and accuracy of the forecast. The proposed approach can be used in various intelligent forecasting systems

Keywords: *forecast, online store, artificial neural network, FC-LSTM, genetic algorithm.*