



УДК 004.9

ЗГОРТКОВА НЕЙРОМЕРЕЖА ДЛЯ ВИЯВЛЕННЯ ШАХРАЙСЬКИХ ОПЕРАЦІЙ З КРЕДИТНИМИ КАРТКАМИ

Корнієнко О. В.¹, Субботін С. О.²

Національний університет «Запорізька політехніка», Запоріжжя, Україна

ORCID: ¹<https://orcid.org/0000-0003-4812-5382>, ²<https://orcid.org/0000-0001-5814-8268>E-mail: ¹al.korn95@gmail.com, ²subbotin.csit@gmail.com

Copyright © 2018 by author and the journal "Automation technologies and business - processes.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0>

DOI:

Анотація. В роботі вирішено завдання створення математичного забезпечення для побудови моделей кількісних залежностей на основі згорткових нейронних мереж. Запропоновано архітектуру згорткової нейронної мережі, що може використовуватися для даних, в яких вхідні значення не пов'язані між собою. В запропонованій архітектурі в якості першого шару використовується повнозв'язний шар. Завдяки цьому в процесі навчання нейронної мережі між вихідними значеннями нейронів першого шару можуть з'явитись зв'язки, що необхідні для роботи наступних згорткових шарів. Як і в звичайних згорткових нейромережах, згорткові шари можуть чергуватися із шарами підвибірки, але при цьому використовується одновимірна згортка. Після згорткових шарів використовуються повнозв'язні. В якості функції активації останнього шару використовується функція softmax, що дозволяє визначити ймовірності належності розпізнаваного екземпляра до кожного з класів. Вирішено практичне завдання виявлення шахрайських операцій з кредитними картками. Виконано побудову нейромережесевих моделей, їх навчання та тестування на даних за транзакціями протягом 2 діб. Зазвичай кількість шахрайських операцій складає невелику частину від усіх операцій, тому метрика точності (accigasu) не може використовуватися для оцінки якості побудованої моделі. Для цього використано метрику AUPRC, що розраховується як площа під кривою залежності значень precision та recall. Завдяки використанню функції активації softmax на останньому шарі нейромережі, побудувати туку криву набагато простіше, ніж при використанні інших функцій активації. Порівняно результати тестування всіх побудованих моделей. За результатами тестування визначено, що якість запропонованої моделі вища, але на її навчання поребується більше часу.

Abstract. The problem of mathematical support for the construction of models of quantitative dependencies on the basis of convolutional neural networks is solved in the work. A convolutional neural network architecture that can be used for data in which the input values are not related is proposed. The proposed architecture uses the fully connected layer as the first layer. As a result, in the learning of the neural network, the connections necessary to operate the next convolutional layers may appear between the output values of the neurons of the first layer. As with simple convolutional neural networks, convolutional layers may alternate with subsampling layers, but a one-dimensional convolutions is used. After the convolutional layers are used fully connected layers. The softmax function is used as the activation function of the last layer, which allows to determine the probabilities of recognition of a recognized instance to each of the classes. Solved the practical problem of detecting credit card fraud. Neural network models were built, trained and tested on transaction data for 2 days. Usually, the number of fraudulent operations is a small part of all operations, so the accuracy metric cannot be used to evaluate the quality of the model being built. For this purpose, the AUPRC metric calculated as the area under the curve of precision and recall values is used. By using the softmax activation function on the last layer of the neural network, it is much easier to plot this curve than using other activation functions. Test results of all built models are compared. According to the results of testing, the quality of the proposed model is higher, but it takes longer to train it.

Ключові слова: вибірка, навчання, нейронна мережа, згортка, помилка, градієнт, виключення, ймовірність, перекресна ентропія.

Keywords: dataset, training, neural network, convolution, error, gradient, dropout, probability, cross entropy.



Вступ

В наш час кредитні картки є одним з найбільш популярних способів оплати товарів та послуг. Дані карток або самі картки можуть бути викрадені шахраями з метою неправомірного заволодіння грошовими коштами. Отже, для банків важливо мати можливість виявляти шахрайські операції з картками та своєчасно блокувати їх.

Тому актуальною проблемою є побудова математичних моделей, що дозволятимуть визначати правомірність операції з картою за відомими даними клієнта та операції.

Перспективним засобом побудови таких моделей є штучні нейронні мережі. Завдяки їх здатностям до навчання за прецедентами, узагальнення даних та витягу знань з даних нейронні мережі можуть використовуватися для побудови моделей залежностей за кількісними спостереженнями в умовах відсутніх або невідомих аналітичних залежностей [1].

Серед відомих архітектур нейронних мереж найбільш потужними і одночасно доступними для подальшого аналізу та сприйняття людиною є нейронні мережі прямого поширення сигналу. Серед них слід виділити багатoshарові та згорткові нейронні мережі. Багатoshарові нейронні мережі є найбільш простими, але можуть використовуватися з вхідними даними в практично будь-якому вигляді. Згорткові нейронні мережі мають кращі узагальнюючі властивості, та, в більшості випадків, більш високу точність, але можуть використовуватися тільки з даними, в яких вхідні значення пов'язані між собою.

Мета роботи – створення нейромережевої моделі, що поєднує переваги багатoshарових та згорткових нейронних мереж та її застосування для виявлення шахрайських операцій з кредитними картками.

Постановка задачі

Нехай ми маємо вибірку спостережень $x = \langle x^s, y^s \rangle$, де x^s – s -й екземпляр вибірки, $x^s = \{x_j^s\}$, x_j^s – значення j -ї вхідної ознаки, що характеризує s -й екземпляр вибірки, $j = 1, 2, \dots, N$, $s = 1, 2, \dots, S_{all}$, де N – кількість вхідних ознак, що характеризують екземпляри вибірки, S_{all} – кількість екземплярів у вибірці.

Тоді задача побудови моделі залежності $y = f(w, x)$ на основі нейронної мережі із заданою користувачем структурою f буде полягати в знаходженні оптимальних значень набору вагових коефіцієнтів нейронів мережі w , тобто таких, при яких критерій якості моделі – перехресна ентропія [2] E буде мінімальною.

Перехресна ентропія E визначається за формулою [2]:

$$E = \sqrt{\sum_{i=1}^Y -M_i \log R_i}, \quad (1)$$

де R_i – реальне значення вихідного параметра y_i^s , M_i – розрахункове значення параметра Y_i^s , отримане за допомогою моделі, $i = 1, \dots, Y$, Y – кількість вихідних параметрів моделі.

Архітектура нейронної мережі

У багатoshарових нейронних мережах (рис. 1) вхідні ознаки розпізнаваного екземпляра подаються на входи мережі та далі поширюються від вхідного шару ($x_1 \dots x_n$) до вихідного ($y_1 \dots y_m$) [3].

Кожен нейрон на прихованих та вихідному шарах отримує на входах сигнали з виходів нейронів попереднього шару. Вхідний вектор кожного нейрона перетворюється у скаляр за допомогою вагової (дискримінантної) функції, після чого значення вихідного сигналу розраховується за допомогою функції активації [4]. Вихідні значення останнього шару є результатом роботи нейромережі.

Вихідні значення кожного нейрона визначаються за формулою:

$$O_m = \psi(w_{m,0} + \sum_{n=1}^N w_{m,n} \cdot I_n) \quad (2)$$

де $\psi(x)$ – функція активації, w – вагові коефіцієнти, I – вихідні значення нейронів попереднього шару, N – кількість нейронів попереднього шару, $m = 1, 2, \dots, M$, M – кількість нейронів поточного шару.

Згорткова нейронна мережа складається з шарів згортки, субдискретизації та повнозв'язних шарів. Приклад операції згортки зображено на рис. 2. Шари субдискретизації зменшують розміри вхідних карт ознак шляхом вибору максимального або середнього значення з деякої області [5]. Повнозв'язні шари працюють так само, як в багатoshаровій нейронній мережі.

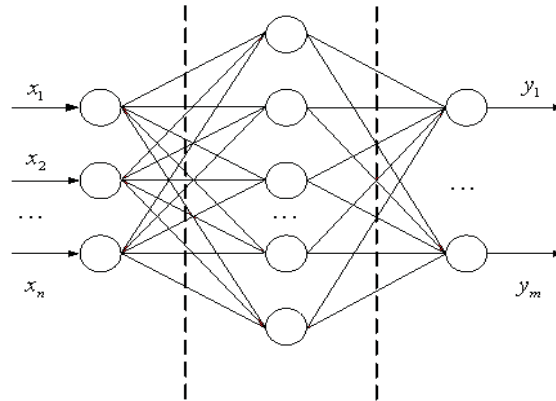


Рис. 1 – Схема багатoshарової нейромережі

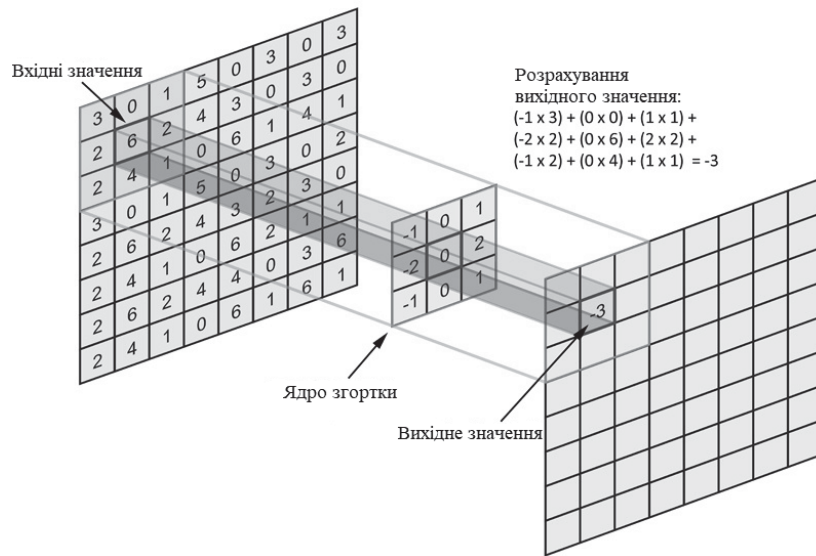


Рис. 2 – Операція згортки

Вихідні значення згорткових шарів обчислюються за формулою:

$$O_{i,j,l} = \psi \left(\sum_{h=0}^H \sum_{m=0}^M \sum_{n=0}^N I_{i+m,j+n,h} \cdot F_{m,n,h,l} + B_{h,l} \right),$$

$$l=0, \dots, L, i=0, \dots, I, j=0, \dots, J,$$

де O – вихідні значення, I_n – вхідні значення, H – кількість вхідних карт ознак, M – ширина ядер згортки, N – висота ядер згортки, L – кількість вихідних карт ознак, I – ширина вихідних карт ознак, J – висота вихідних карт ознак, F – ядра згортки, B – зміщення, $\psi(x)$ – функція активації.

Для можливості використання згорткових нейронних мереж для даних, в яких вхідні значення не мають зв'язків або мають невідомі зв'язки між собою пропонується використовувати повнозв'язний шар перед згортковими. Таким чином в процесі навчання можуть з'явитися зв'язки між вихідними значеннями першого шару, що значно підвищить ефективність наступних згорткових шарів. При цьому слід використовувати одновимірні згорткові шари, тому можна спростити формулу, за якою визначаються їх вихідні значення:

$$O_{i,l} = \psi \left(\sum_{h=0}^H \sum_{m=0}^M I_{i+m,h} \cdot F_{m,h,l} + B_{h,l} \right), \tag{3}$$

$$l=0, \dots, L, i=0, \dots, I,$$

де O – вихідні значення, I_n – вхідні значення, H – кількість вхідних карт ознак, M – ширина ядер згортки, L – кількість вихідних карт ознак, I – ширина вихідних карт ознак, F – ядра згортки, B – зміщення, $\psi(x)$ – функція активації.

Схему запропонованої нейронної мережі з 1 згортковим шаром зображено на рис. 3.

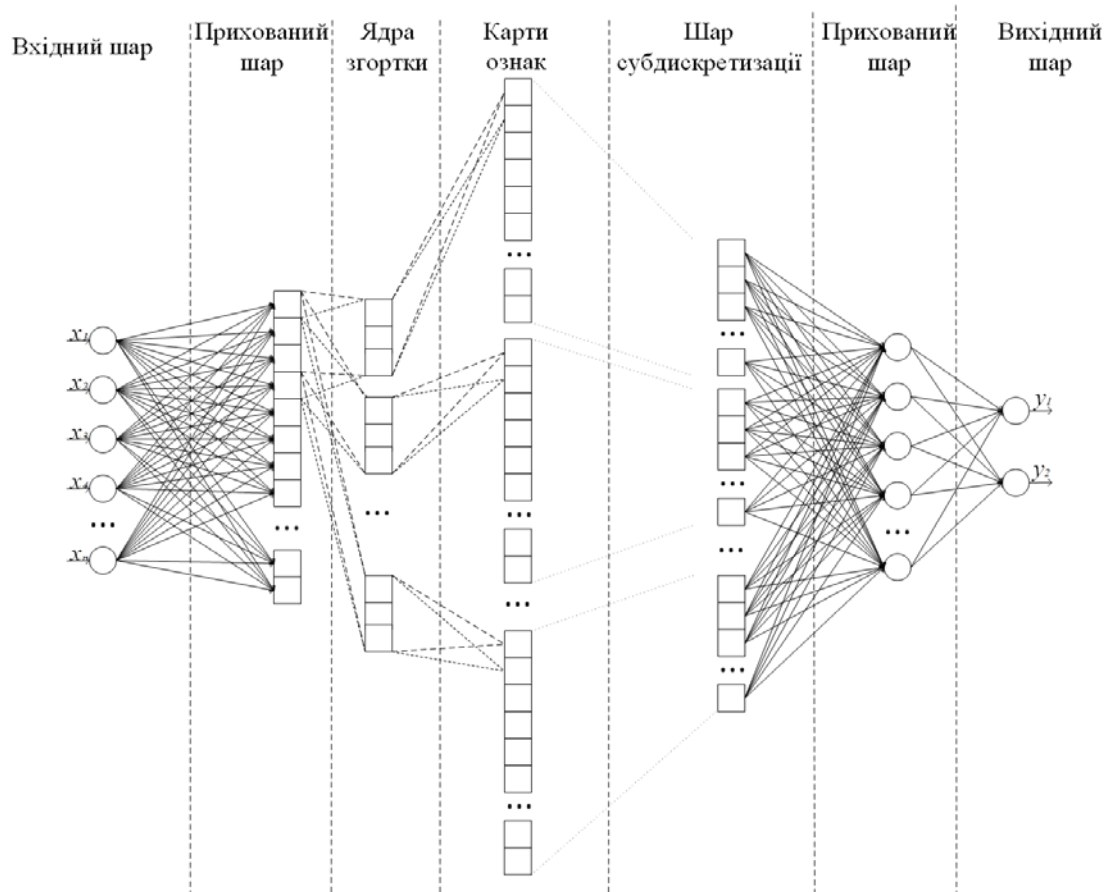


Рис. 3 – Схема нейронної мережі

Така нейронна мережа поєднує переваги багатошарових та згорткових нейромереж. Як і в звичайних згорткових нейромережах кількість згорткових шарів може бути більшою, ніж на схемі. Шари субдискретизації також можуть не використовуватися, але зазвичай вони використовуються після кожного згорткового шару. Вихідні значення запропонованої нейронної мережі з 2 прихованими та 1 згортковим шаром (без шару субдискретизації) для кожного екземпляра вибірки розраховується за формулою:

$$O_j = \psi_4(w_3 \sum_{d=1}^D w_3 \psi_3(w_2 \sum_{c=1}^C w_2 \psi_2(B_k + \sum_{g=1}^G \sum_{h=1}^H F_{h,g} \psi_1(w_1 \sum_{n=1}^N w_1 I_n)))) ,$$

де O – вихідні значення, I – вхідні значення, $j=0..J$ – кількість нейронів вихідного шару, D – кількість нейронів другого прихованого шару, C – загальна кількість значень шару субдискретизації, G – кількість ядер згортки, H – розмір ядер згортки, $m=0..M$ – кількість нейронів першого прихованого шару, N – кількість вхідних нейронів, F – ядра згортки, B – зміщення, w_1, w_2, w_3 – вагові коефіцієнти першого прихованого, другого прихованого та вихідного шарів відповідно, $\psi_1, \psi_2, \psi_3, \psi_4$ – функції активації першого прихованого, згорткового, другого прихованого і вихідного шарів відповідно.

Для розрахування значень шару субдискретизації необхідно кожен карту ознак з попереднього шару розділити на частини розміром n , після чого обрати максимальне значення з кожної частини. Обрані значення будуть новими картами ознак розміром m/n , m – розмір карт ознак попереднього шару.

Навчання нейромережі із контрастуванням та видаленням нейронів

Навчання нейронних мереж прямого поширення, як правило, проводиться на основі градієнтного методу [6] та техніки зворотного поширення помилки [7]. В якості функції помилки доцільно використовувати перехресну ентропію E , що обчислюється за формулою (1).

Вектор R визначається в залежності від класу екземпляра і функції активації нейронів вихідного шару нейромережі. Значення R_i , що відповідає класу екземпляра, повинне бути близьким до максимального значення функції активації, а всі інші – близькими до мінімального. Якщо область значень функції активації необмежена, доцільно використовувати значення 1 та -1 або 1 та 0.

Метод навчання нейронної мережі полягає у послідовному виконанні таких етапів.



Етап ініціалізації. Задати навчальну вибірку даних $\langle x, y \rangle$ та структуру нейромережі. Ваговим коефіцієнтам нейронів задати випадкові значення з діапазону $[-0,1; 0,1]$. Задати: кількість циклів (epoch) навчання $Epochs=0$, максимальну кількість циклів (epoch) навчання $MaxEpochs$, крок навчання α [2].

Етап балансування вибірки. В задачі виявлення шахрайських операцій з кредитними картками кількість екземплярів одного класу може значно перевищувати кількість екземплярів іншого. Навчання нейронної мережі на такій вибірці приводить до ситуації, коли незалежно від вхідних значень прогнозований клас буде завжди однаковим. Отже необхідно видалити існуючі екземпляри або додати нові, щоб кількість екземплярів обох класів була однаковою. Видалення екземплярів може привести до недостатньої кількості навчальних даних, а додавання даних шляхом їх копіювання може привести до перенавчання моделі. Пропонується використовувати додавання даних шляхом копіювання та додавання випадкових шумів. Для додавання шумів необхідно помножити кожне значення в нових екземплярах на коефіцієнт $K=1-P$, де P – випадкове значення з діапазону $[-0,01; 0,01]$.

Етап нормування вибірки. Вибірка даних нормується шляхом відображення значень ознак на інтервал $[0; 1]$. Нові значення визначаються за формулою:

$$x_i^s = \frac{x_i^s - x_i^{\min}}{x_i^{\max} - x_i^{\min}},$$

де x_i^{\max} , x_i^{\min} – максимальне та мінімальне значення i -ї вхідної ознаки серед всіх екземплярів вибірки, відповідно.

Етап корекції ваг. Послідовно у циклі для $Epochs=0, 1, \dots, MaxEpochs$, де $MaxEpochs$ – максимально допустима задана кількість epoch навчання, виконується навчання мережі на основі екземплярів вибірки шляхом виконання етапу навчання мережі на основі екземпляра та коректується крок навчання α .

Етап навчання мережі на основі екземплярів вибірки. Послідовно на входи нейромережі подається i -й екземпляр навчальної вибірки, $i=1,2,\dots,S_{tr}$, де S_{tr} – кількість екземплярів навчальної вибірки. Для поточного екземпляра розраховуються вихідні значення нейронної мережі за формулами (2), (3). В процесі розрахування важливо зберегти вихідні значення нейронів всіх шарів. Після цього обчислюється перехресна ентропія E .

Далі обчислюються градієнти помилок нейронів вихідного шару:

$$G_i = O_i - R_i, \quad i=0, \dots, I,$$

де I – кількість вихідних нейронів, O – вихідний вектор нейронної мережі, R – бажаний вихідний вектор, G – градієнти помилок вихідного шару.

Після цього обчислюються градієнти для нейронів повнозв'язних шарів (крім першого):

$$G_i = \sum_{j=0}^J Gr_j \cdot w_{i,j} \cdot (1 - O_i^2),$$

$$i=0, \dots, I, j=0, \dots, J,$$

де I – кількість нейронів поточного шару, J – кількість нейронів наступного шару, Gr – градієнти помилок наступного шару, G – градієнти помилок поточного шару, O – вихідний вектор поточного шару, w – вагові коефіцієнти нейронів наступного шару.

Коли будуть розраховані градієнти повнозв'язних шарів, необхідно розрахувати градієнти для згорткових шарів:

$$G_{i,j+k} = \sum_{m=0}^M Gr_{m,j} \cdot F_{i,m,k} \cdot (1 - O_{i,j+k}^2),$$

$$i=0, \dots, I, j=0, \dots, J, k=0, \dots, K,$$

де I – кількість карт ознак поточного шару, J – ширина карт ознак наступного шару, $ш$ – ширина ядер згортки, M – кількість карт ознак наступного шару, Gr – градієнти помилок наступного шару, G – градієнти помилок поточного шару, O – вихідні карти ознак поточного шару, F – вагові коефіцієнти нейронів (ядер згортки) наступного шару.

Далі визначаються градієнти для першого повнозв'язного шару за формулою (4).

Після розрахування градієнтів необхідно виконати корекцію вагових коефіцієнтів. Для згорткових шарів корекція виконується за формулою:

$$F_{i,j,m} = F_{i,j,m} - \sum_{p=0}^P O_{i,p+m} \cdot G_{j,p} \cdot \alpha,$$

$$i=0, \dots, I, j=0, \dots, J, m=0, \dots, M,$$

де I – кількість карт ознак поточного шару, J – кількість карт ознак наступного шару, M – ширина ядер згортки, P – ширина карт ознак наступного шару, G – градієнти помилок поточного шару, O – вихідні карти ознак поточного шару, F – вагові коефіцієнти нейронів (ядер згортки) поточного шару.

Для повнозв'язних шарів корекція вагових коефіцієнтів виконується за формулою:



$$w_{i,j} = w_{i,j} - O_i \cdot G_j \cdot \alpha,$$

$$i=0, \dots, I, j=0, \dots, J,$$

де I – кількість нейронів поточного шару, J – кількість нейронів наступного шару, G – градієнти помилок, O – вихідний вектор поточного шару, w – вагові коефіцієнти нейронів поточного шару.

Етап контрастування нейронної мережі. Контрастування полягає у видаленні частини найменш значимих зв'язків в нейронній мережі. Видалення слабких зв'язків приведе до збільшення узагальнюючих властивостей нейромережі [8]. Для видалення зв'язку достатньо для відповідного вагового коефіцієнта задати нульове значення. Отже, для проведення контрастування нейронної мережі необхідно для всіх вагових коефіцієнтів $|w_{j,k}^i| < P$, де $i=1..I-1$, $j=1..J$, $k=1..K$, задати нульове значення. I – кількість шарів нейронної мережі, J – кількість нейронів i шару, K – кількість нейронів $i+1$ шару, P – порогове значення для вагових коефіцієнтів.

Одним з методів оптимізації процесу навчання нейронних мереж є виключення нейронів [9]. Він полягає у тому, що в процесі навчання деяка частина нейронів видаляється з мережі. На кожній ітерації навчання виключаються випадкові нейрони. Кожен нейрон може бути виключеним з ймовірністю P . Ймовірність виключення може бути різною для кожного шару нейронної мережі. На рис. 3 зображено звичайну нейронну мережу, а на рис. 4 – нейронну мережу після виключення нейронів.

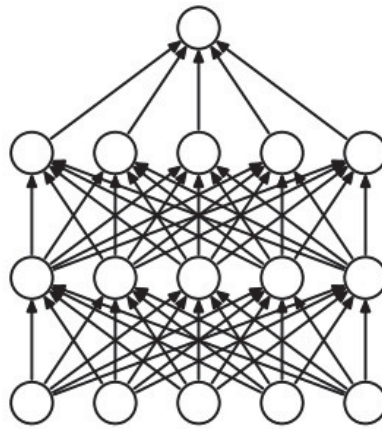


Рис. 4 – Звичайна нейронна мережа

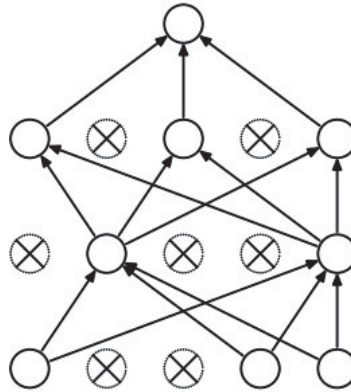


Рис. 5 – Нейронна мережа після виключення нейронів

Виключення нейронів проводиться тільки при навчанні моделі. При реалізації методу достатньо прирівняти до 0 вихідні значення тих нейронів, що повинні бути виключені на ітерації в процесі навчання. Для згорткових шарів слід видаляти карти ознак.

При тестуванні та під час роботи моделі необхідно використовувати всі нейрони. При цьому вихідні значення нейронів на шарах, де використовувалося виключення, необхідно помножити на коефіцієнт $Q = 1 - P$, де P – ймовірність виключення нейронів на шарі.

Виключення нейронів дозволяє уникнути перенавчання моделі [10], що в більшості випадків приводить до збільшення точності моделі на тестових даних.

Експерименти та результати



Для навчання нейромережових моделей було використано вибірку даних за транзакціями, що проводилися протягом 2 діб у вересні 2013 року [11]. Через банківську таємницю авторами даних не розголошується назва банку та назви входних параметрів у вибірці. Серед відомих параметрів надані лише сума транзакції та час проведення (кількість секунд після проведення першої транзакції). До всіх інших даних застосовано PCA трансформації [12], що не зменшує якість вибірки, але дозволяє не розкривати банківську таємницю. Вибірка складається з 284807 екземплярів, серед яких 492 шахрайські операції. Фрагмент вибірки наведено в табл. 1.

Таблиця 1 – Фрагмент вибірки даних

<i>Time</i>	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>V7</i>	<i>V8</i>	<i>V9</i>	<i>V10</i>	<i>V11</i>	<i>V12</i>	<i>Amount</i>	<i>Class</i>
0	-1,36	-0,07	2,54	1,38	-0,34	0,46	0,24	0,1	0,36	0,09	-0,55	-0,62	149,62	0
0	1,19	0,27	0,17	0,45	0,06	-0,08	-0,08	0,09	-0,26	-0,17	1,61	1,07	2,69	0
1	-1,36	-1,34	1,77	0,38	-0,5	1,8	0,79	0,25	-1,51	0,21	0,62	0,07	378,66	0
1	-0,97	-0,19	1,79	-0,86	-0,01	1,25	0,24	0,38	-1,39	-0,05	-0,23	0,18	123,5	0
2	-1,16	0,88	1,55	0,4	-0,41	0,1	0,59	-0,27	0,82	0,75	-0,82	0,54	69,99	0
2	-0,43	0,96	1,14	-0,17	0,42	-0,03	0,48	0,26	-0,57	-0,37	1,34	0,36	3,67	0
4	1,23	0,14	0,05	1,2	0,19	0,27	-0,01	0,08	0,46	-0,1	-1,42	-0,15	4,99	0
7	-0,64	1,42	1,07	-0,49	0,95	0,43	1,12	-3,81	0,62	1,25	-0,62	0,29	40,8	0
7	-0,89	0,29	-0,11	-0,27	2,67	3,72	0,37	0,85	-0,39	-0,41	-0,71	-0,11	93,2	0
9	-0,34	1,12	1,04	-0,22	0,5	-0,25	0,65	0,07	-0,74	-0,37	1,02	0,84	3,68	0
10	1,45	-1,18	0,91	-1,38	-1,97	-0,63	-1,42	0,05	-1,72	1,63	1,2	-0,67	7,8	0
10	0,38	0,62	-0,87	-0,09	2,92	3,32	0,47	0,54	-0,56	0,31	-0,26	-0,33	9,99	0
10	1,25	-1,22	0,38	-1,23	-1,49	-0,75	-0,69	-0,23	-2,09	1,32	0,23	-0,24	121,5	0
11	1,07	0,29	0,83	2,71	-0,18	0,34	-0,1	0,12	-0,22	0,46	-0,77	0,32	27,5	0
12	-2,79	-0,33	1,64	1,77	-0,14	0,81	-0,42	-1,91	0,76	1,15	0,84	0,79	58,8	0
12	-0,75	0,35	2,06	-1,47	-1,16	-0,08	-0,61	0	-0,44	0,75	-0,79	-0,77	15,99	0
12	1,1	-0,04	1,27	1,29	-0,74	0,29	-0,59	0,19	0,78	-0,27	-0,45	0,94	12,99	0
13	-0,44	0,92	0,92	-0,73	0,92	-0,13	0,71	0,09	-0,67	-0,74	0,32	0,28	0,89	0
14	-5,4	-5,45	1,19	1,74	3,05	-1,76	-1,56	0,16	1,23	0,35	0,92	0,97	46,8	0
15	1,49	-1,03	0,45	-1,44	-1,56	-0,72	-1,08	-0,05	-1,98	1,64	1,08	-0,63	5	0
16	0,69	-1,36	1,03	0,83	-1,19	1,31	-0,88	0,45	-0,45	0,57	1,02	1,3	231,71	0
17	0,96	0,33	-0,17	2,11	1,13	1,7	0,11	0,52	-1,19	0,72	1,69	0,41	34,09	0
18	1,17	0,5	-0,07	2,26	0,43	0,09	0,24	0,14	-0,99	0,92	0,74	-0,53	2,28	0
18	0,25	0,28	1,19	-0,09	-1,31	-0,15	-0,95	-1,62	1,54	-0,83	-0,58	0,52	22,75	0
22	-1,95	-0,04	-0,41	-1,01	2,94	2,96	-0,06	0,86	0,05	0,57	-0,08	-0,22	0,89	0



Через те, що вибірка незбалансована, використання точності класифікації як метрики якості моделі неможливе. Пропонується використовувати метрику AUPRC [13], що визначає площу під кривою залежності між значеннями precision та recall [14]. Для побудови кривої залежності та розрахування значення AUPRC необхідно визначити вихідні значення моделі для всіх екземплярів тестової вибірки та відсортувати їх за спаданням ймовірності належності транзакції до класу шахрайських. Використання функції активації softmax [2] для нейронів вихідного шару дозволяє визначати ймовірність за вихідним значенням одного нейрона. Після сортування необхідно розрахувати значення precision та recall для всіх випадків, коли $n=0..N$ перших екземплярів віднесено до класу шахрайських транзакцій. N – кількість екземплярів у тестовій вибірці. Розрахування значень precision та recall виконується за формулами:

$$precision = \frac{TP}{TP + FP},$$

$$recall = \frac{TP}{TP + FN},$$

де TP – кількість екземплярів, що відносяться до класу шахрайських операцій і були правильно розпізнані, FP – кількість екземплярів, що не відносяться до класу шахрайських операцій і не були правильно розпізнані, FN – кількість екземплярів, що відносяться до класу шахрайських операцій і не були правильно розпізнані.

Апроксимоване значення AUPRC можна визначити за формулою:

$$AUPRC = \sum_{n=0}^N precision_n \cdot \Delta recall,$$

де N – кількість екземплярів у тестовій вибірці.

В ході проведення експериментів будувалися три нейромоделі, перші дві – багатошарові, третя – згорткова. Параметри мереж задавалися наступним чином.

Перша мережа: кількість нейронів прихованого шару – 50, функція активації – ReLU, ймовірність виключення нейронів – 15%, функція активації вихідного шару – softmax.

Друга мережа: кількість нейронів першого прихованого шару – 40, функція активації – ReLU, ймовірність виключення нейронів – 15%, кількість нейронів другого прихованого шару – 20, функція активації – ReLU, ймовірність виключення нейронів – 15%, функція активації вихідного шару – softmax.

Третя мережа: кількість нейронів першого прихованого шару – 25, функція активації – ReLU, ймовірність виключення нейронів – 15%, кількість карт ознак згорткового шару – 10, ширина ядер згортки – 6, функція активації – ReLU, ймовірність виключення карт ознак – 15%, ширина ядер шару субдискретизації – 2, кількість нейронів другого прихованого шару – 15, функція активації – ReLU, ймовірність виключення нейронів – 15%, функція активації вихідного шару – softmax.

Для всіх моделей було використано однакові значення кількості епох та кроку навчання. Обсяг оперативної пам'яті для всіх моделей під час навчання не перевищував 1 МБ.

На рис. 6 зображено криві precision-recall для всіх моделей.

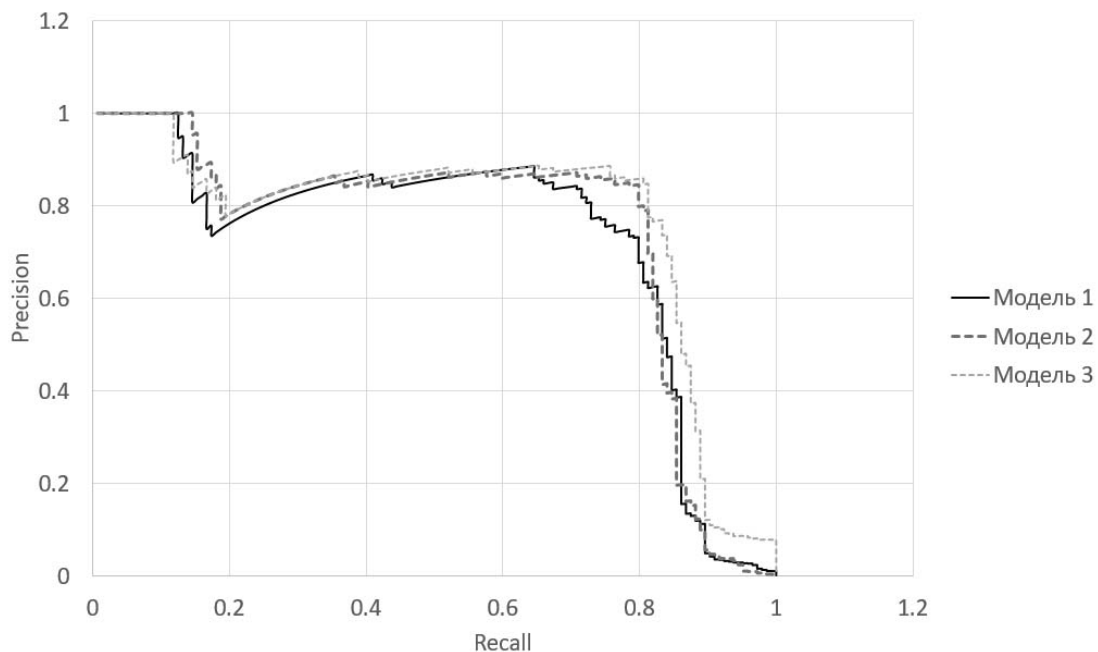


Рис. 6 – Криві precision-recall для побудованих моделей



В табл. 2 порівняно побудовані моделі за значеннями AURPC та F1 score [14]. Значення F1 score визначається за формулою:

$$F1 = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Для всіх моделей вказано максимальне значення F1 score.

Таблиця 2 – Результати експериментів

№	AURPC	F1 score	precision	recall	коефіцієнт узагальнення	час навчання (хв)
1	0,728	0,772	0,8	0,716	7232	18
2	0,744	0,821	0,846	0,799	5647	24
3	0,773	0,831	0,859	0,806	5241	33

За результатами проведених експериментів можна зробити висновок, що запропонована архітектура нейронної мережі, не зважаючи на збільшення часу навчання, дозволяє підвищити якість нейромоделей при приблизно однаковому значенні коефіцієнта узагальнення.

Висновки

У роботі вирішено завдання удосконалення процесу побудови нейромоделей шляхом поєднання переваг багатосарових і згорткових нейронних мереж.

Наукова новизна роботи полягає у тому, що запропоновано архітектуру згорткової нейронної мережі, що може використовуватися для роботи з даними, в яких вхідні параметри не пов'язані між собою. Модифікація метода полягає у додаванні повнозв'язного шару на початку нейромережі. Це дозволяє збільшити якість побудованих моделей.

Практична цінність отриманих результатів полягає у тому, що побудовано нейронні мережі із запропонованою архітектурою, проведено їх навчання та тестування, а також вирішено практичне завдання виявлення шахрайських операцій з кредитними картками.

Перспективи подальших досліджень полягають у визначенні оптимальних параметрів нейронних мереж, а також у застосуванні запропонованого метода на більш широкому класі задач.

Список використаних джерел

- [1] Russell S. J. Artificial intelligence: a modern approach / S. J. Russell – Pearson, 2010. – 1152 p.
- [2] Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016. 781 p.
- [3] Хайкин С. Нейронные сети: полный курс. Москва: Вильямс, 2006. 1104 с.
- [4] Kumar K. V. Neural networks and fuzzy logic. S. K. Kataria & Sons, 2016. 300 p.
- [5] CS231n: Convolutional Neural Networks for Visual Recognition. [Electronic resource]. – Access mode: <http://cs231n.github.io>
- [6] Bertsekas D. P. Convex Optimization Theory. Athena Scientific, 2009. 256 p.
- [7] Nielsen M. Neural Networks and Deep Learning. Determination Press, 2015. 216 p.
- [8] Louizos C., Welling M., Kingma D. P. Learning sparse neural networks through L_0 regularization. Internation. Conf. on Learning Representations, 2018.
- [9] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014. 15. P. 1929–1958.
- [10] Caruana R., Lawrence S., Giles L. Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping. Neural information processing systems conference, 2000.
- [11] Credit Card Fraud Detection. [Electronic resource]. – Access mode: <https://www.kaggle.com/mlg-ulb/creditcardfraud>
- [12] Abdi. H. & Williams, L.J. Principal component analysis. [Electronic resource]. – Access mode: <https://arxiv.org/pdf/1108.4372.pdf>
- [13] Keilwagen, J., Grosse, I., Grau, J. Area under precision-recall curves for weighted and unweighted data. PLoS One, 2014.
- [14] Ting, Kai Ming. Encyclopedia of machine learning. Springer, 2010. 45p.



References

- [1] Russell S. J. Artificial intelligence: a modern approach / S. J. Russell – Pearson, 2010. – 1152 p.
- [2] Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016. 781 p.
- [3] Haykin S. Neyronnyie seti: polnyi kurs. Moskva: Vilyams, 2006. 1104 p.
- [4] Kumar K. V. Neural networks and fuzzy logic. S. K. Kataria & Sons, 2016. 300 p.
- [5] CS231n: Convolutional Neural Networks for Visual Recognition. [Electronic resource]. – Access mode: <http://cs231n.github.io>
- [6] Bertsekas D. P. Convex Optimization Theory. Athena Scientific, 2009. 256 p.
- [7] Nielsen M. Neural Networks and Deep Learning. Determination Press, 2015. 216 p.
- [8] Louizos C., Welling M., Kingma D. P. Learning sparse neural networks through L_0 regularization. Internation. Conf. on Learning Representations, 2018.
- [9] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014. 15. P. 1929–1958.
- [10] Caruana R., Lawrence S., Giles L. Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping. Neural information processing systems conference, 2000.
- [11] Credit Card Fraud Detection. [Electronic resource]. – Access mode: <https://www.kaggle.com/mlg-ulb/creditcardfraud>
- [12] Abdi. H. & Williams, L.J. Principal component analysis. [Electronic resource]. – Access mode: <https://arxiv.org/pdf/1108.4372.pdf>
- [13] Keilwagen, J., Grosse, I., Grau, J. Area under precision-recall curves for weighted and unweighted data. PLoS One, 2014.
- [14] Ting, Kai Ming. Encyclopedia of machine learning. Springer, 2010. 45p.

УДК 681.787.6

ВОЛОКОННО-ОПТИЧЕСКАЯ СИСТЕМА КОНТРОЛЯ ТЯГИ ГАЗОТУРБИННЫХ ДВИГАТЕЛЕЙ

Сандлер А.К., Карпилов О.Ю.

Национальный университет "Одесская морская академия", г. Одесса, Украина

E-mail: ¹albertsand4@gmail.com, ²kau.onma@gmail.com

Copyright © 2018 by author and the journal "Automation technologies and business - processes.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0>



DOI:

***Аннотация.** На современном этапе эволюции судовой энергетики отмечается определенный рост применения газотурбинных установок (ГТУ). Анализ эксплуатации судов и кораблей с ГТУ показал, что эти установки более эффективны по сравнению с дизельными на быстроходных (со скоростью хода более 20 ... 25 узлов) судах водоизмещающего типа, а также на судах с динамическими принципами поддержания с большой энерговооруженностью - до 20 ... 40 кВт на 1 тонну водоизмещения против 1 ... 2 кВт на 1 тонну у водоизмещающих судов со скоростями движения 12 ... 15 узлов.*

Для судов и кораблей с динамическими принципами поддержания актуальным является точное определение векторов тяги двигателя. Традиционно измерение усилий, передаваемых на корпус, производится в узлах крепления установки либо на специальных динамометрических стендах.

В обоих вариантах контроля сложным является выбор тягоизмерительных устройств (ТИУ), обеспечивающих достаточно высокую информативность и чувствительность измерений.