



**В. М. Кулик, Т. О. Коротєєва**

*Національний університет "Львівська політехніка" м. Львів, Україна*

## АЛГОРИТМ ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ ФУТБОЛЬНИХ МАТЧІВ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Досліджено використання одного із способів машинного навчання, а саме нейронної мережі у прогнозуванні результатів спортивних подій. На цей час сфера машинного навчання перебуває на піку інтересу серед широкого загалу і використовується у вирішенні великого спектра проблем сучасного світу. Серед основних переваг і одночасно причиною, яка стала основною у виборі способу вирішення поставленої проблеми, є можливість застосування нейронних мереж для вирішення задач, в яких не простежується чіткий алгоритм розв'язання. На основі зібраних даних про архітектури та способи застосування різних типів нейронних мереж, алгоритмів навчання, а також і загалом методів, які можуть бути використані для вирішення проблем прогнозування, досліджено та розроблено алгоритм прогнозування. Враховуючи відсутність у відкритому доступі схожих засобів для прогнозування результатів спортивних подій, оцінку алгоритму проведено за допомогою порівняння точності прогнозування з іншими алгоритмами прогнозування, які не передбачають використання нейронних мереж. Таке тестування здійснено на основі загальнодоступного набору даних, отриманого з відкритих джерел.

**Ключові слова:** нейромережа; багатосаровий перцептрон; алгоритм навчання; обернене розповсюдження похибки.

**Вступ.** Із стрімким розвитком технічної частини комп'ютерів відкриваються дедалі нові можливості для розв'язання різноманітних задач, які було неможливо розв'язати раніше, беручи до уваги недостатню обчислювальну потужність. Недарма популярність машинного навчання, і особливо нейронних мереж, досягнула свого піку зовсім недавно, адже алгоритми машинного навчання були описані давніше, ніж їхнє використання стало достатньо ефективним навіть для комп'ютерів, які може дозволити собі практично кожна людина. Машинне навчання є надзвичайно перспективною галуззю, зокрема з використанням нейронних мереж (Gorban & Rossiev, 1996).

Ще однією сферою, яка на відміну від машинного навчання, користується популярністю вже багато років, є сфера беттингу. Незважаючи на своє довге існування і доступність, статистика показує щорічне зростання популярності цієї сфери, що змушує її розвиватись і з кожним роком використовувати нові підходи та методи для заохочення щораз більшої категорії осіб.

Беручи до уваги постійну популярність цієї сфери та існування на ринку багатьох букмекерських контор, можна зробити висновок про відсутність на цей час алгоритму, який зміг би зі стовідсотковою точністю робити прогнози відносно спортивних подій, які відбудуться. Такий алгоритм навряд чи буде створено, адже на результат спортивного матчу впливає надзвичайно велика кількість факторів, багато із яких не піддаються прогнозуванню. Незважаючи на це, чимало провідних

університетів у світі розробляють алгоритм, який мав би високу точність передбачення. Отже, сама спортивна подія, наприклад футбольний матч, а точніше його результат, є результатом збігу багатьох факторів, кожен з яких має різні ступені впливу. Відсутність очевидних зв'язків між факторами впливу робить неможливим стовідсоткове прогнозування результату експертами, які базуються на власних логічних висновках.

У цій роботі досліджено алгоритм прогнозування результатів футбольних матчів на основі нейронних мереж. Для коректної роботи алгоритму спочатку потрібно провести "тренування" нейронної мережі, а пізніше здійснювати його тестування (Galushkin, 2000). Для тренування нейронної мережі було відібрано та використано різноманітні фактори, які формують результат спортивного матчу. За результатами тренування нейронної мережі, можна застосувати алгоритм на відібраній вибірці даних і зробити висновки про його ефективність.

Отже, основним дослідженням цієї роботи є ефективність використання нейронної мережі для прогнозування результатів футбольних матчів, на основі відібраних загальнодоступних факторів. Метою є визначення якості прогнозування на основі нейронної мережі та використання розробленого алгоритму для розроблення програмного засобу.

**Матеріали та методи дослідження.** Незважаючи на принцип роботи нейронної мережі, існує велика кількість підходів до її реалізації, варіантів архітектури та

### Інформація про авторів:

**Кулик Владимир Михайлович**, магістрант кафедри програмного забезпечення. Email: uavova@ukr.net

**Коротєєва Тетяна Олександрівна**, канд. техн. наук, доцент кафедри програмного забезпечення.

Email: maxim04041997@gmail.com

**Цитування за ДСТУ:** Кулик В. М., Коротєєва Т. О. Алгоритм прогнозування результатів футбольних матчів на основі нейронних мереж. Науковий вісник НЛТУ України. 2017. Вип. 27(9). С. 111–114.

**Citation APA:** Kulyk, V. M., & Koroteyeva, T. O. (2017). Algorithm of the Football Match Result Prediction on the Basis of Neural Networks. *Scientific Bulletin of UNFU*, 27(9), 111–114. <https://doi.org/10.15421/40270924>

алгоритмів навчання (Adams, 2017). Одні нейронні мережі простіші для реалізації, інші є не тільки складними у реалізації, але й передбачають наявність значних ресурсів для навчання. Ці твердження також можна застосувати і до алгоритмів навчання нейронних мереж. Проаналізувавши літературу, можна зробити висновок про високу ефективність нейронної мережі прямого поширення, а також те, що складність її реалізації є задовільною, порівняно з іншими архітектурами нейронних мереж (Kruglov & Bogisov, 2001).

Нейронна мережа прямого поширення є однією з найпопулярніших архітектур нейронних мереж не тільки через складність своєї реалізації, але й продуктивність та ефективність своєї роботи (Badde, Gupta & Patki, 2009). Найбільшою відмінністю цього типу нейронних мереж є те, що її складники не формують циклу, на відміну від рекурентних нейронних мереж, для прикладу.

У нейронній мережі прямого поширення нейрон отримує дані від інших нейронів, які відносно нього знаходяться на рівні вище, і передає результат на рівень нижче. Отже, згідно з назвою нейронної мережі, інформація в ній має тільки один шлях поширення, а саме від входу до виходу, проходячи при цьому через всі приховані шари нейронної мережі.

Формування результату в цій нейронній мережі є стандартним підходом, коли інформація передається через синапси з певними вагами в нейрони, які сумують отримані результати і застосовують до отриманих даних функцію активації (передавальна функція) (Liubun, 2006). Такий процес є незмінним та відбувається крізь усі приховані шари нейронної мережі.

Як зазначено вище, нейронна мережа містить приховані шари. Прихованими вважають шари нейронної мережі, які розміщені між вхідним та вихідним шарами. Приховані шари, а саме їхня кількість, є одним з вирішальних факторів ефективності нейронної мережі, які мають безпосередній вплив на точність результату (Kohonen, 2013). Чим більше прихованих шарів, тим вимогливіша до ресурсів нейронна мережа. Тому ефективною нейронною мережею можна назвати таку, в якій правильно підібрана низка факторів, враховуючи кількість прихованих шарів.

Значний вплив на точність результату нейронної мережі має також і обраний алгоритм її тренування. Тренування є ключовим етапом у створенні нейронної мережі, який відбувається безпосередньо після її реалізації. Зважаючи на кількість типів нейронних мереж, алгоритмів їх навчання є також кілька. Базовий принцип навчання нейронної мережі є по суті корегування ваг на синапсах для мінімізації похибки, якою є різниця між результатом, отриманим на виході нейронної мережі, та еталонними даними.

Після ретельного аналізу літератури, алгоритмом, який використано для тренування нейронної мережі, обрано метод оберненого розповсюдження похибки (Hryhorukiv, Yaroshenko & Filipchuk, 2012). Такий метод є ітеративним градієнтним алгоритмом, принцип роботи якого заснований на передаванні сигналів похибки від виходів нейронної мережі до її входів. Цей метод навчання нейронної мережі є один із найпопулярніших та найпродуктивніших серед описаних.

Тренування відбувається на еталонній вибірці даних, які в цьому разі представлені факторами, що без-

посередньо впливають на формування результату футбольного матчу. Для забезпечення необхідної продуктивності та ефективності тренування нейронної мережі проведено детальний аналіз факторів та обрано 12 з них, які мають максимальний вплив та які доцільно включити у вибірку.

Оскільки в цьому разі невідомою є оптимальна кількість прихованих шарів, було прийнято рішення реалізувати нейронні мережі з двома та чотирма прихованими шарами відповідно, для експериментального порівняння ефективності роботи таких нейронних мереж.

**Результати дослідження та їх обговорення.** Нейронна мережа була реалізована за допомогою мови програмування Python, зважаючи на її популярність, достатню продуктивність та вбудований функціонал. У реалізації було застосовано бібліотеку NumPy, яка дала змогу використати необхідні математичні функції та забезпечила підтримку великих багатовимірних масивів.

Для оцінки роботи нейронних мереж обрано матчі, результати яких були заздалегідь відомими, тобто які на час тестування вже відбулись. Однаковий набір тестових даних застосовували до обох нейронних мереж для порівняння результатів їхньої роботи.

Серед результатів матчів, які обрали для тренування нейронної мережі, як вже описано вище, були лише ті, які мають значний вплив на формування результату футбольного матчу, отже, серед них можна виділити такі:

- статистика останніх п'яти матчів кожної з команд;
- статистика зустрічей поточних команд;
- статистика ігор вдома команди, яка приймає зустріч;
- статистика ігор в гостях команди, яка грає на виїзді.

Ці фактори увійшли до групи тих, які мають найбільший вплив. Для прикладу, одним з основних факторів, від якого можна відштовхуватись під час складання прогнозу, є саме форма команди, а вона визначається, зазвичай, за статистикою останніх п'яти ігор команди, тому цей фактор і потрапив до групи з найбільшим впливом, інші фактори по аналогії. Отже, максимальний ефект очікується саме використовуючи ці фактори. Але є також і такі, які мають значно менший вплив на формування результату, серед них:

- погода, під час якої відбувається матч;
- період дня, коли саме відбувається матч;
- турнірне положення;
- статистика стосовно фолів, штрафних, кутових.

Щоб оцінити доцільність та правильність підбору параметрів для тренування нейронної мережі, а також перевірити їх вплив на формування результату на практиці, було вирішено застосувати обидві вибірки факторів і порівняти їх.

Отже, після формування двох тестових груп, які складаються із всіх факторів впливу та відповідно з тих, які було віднесено до групи факторів з найбільшим впливом, наступним етапом став етап тренування нейронної мережі, результатом якого стала натренована нейронна мережа, готова для здійснення прогнозів відносно представлених вхідних даних.

На рис. зображено процес тренування нейронної мережі. Змінні "Epoch" та "Error" відповідають за число пройдених ітерацій та похибку відповідно. З кожною ітерацією тренування, похибка зменшується, що свідчить про успішність процесу тренування.

```

Windows PowerShell
Epoch: 1143; Error: 0.49685512378;
Epoch: 1144; Error: 0.496608986177;
Epoch: 1145; Error: 0.496397610571;
Epoch: 1146; Error: 0.49620163895;
Epoch: 1147; Error: 0.496188344164;
Epoch: 1148; Error: 0.496166811316;
Epoch: 1149; Error: 0.496135998734;
Epoch: 1150; Error: 0.49609992031;
Epoch: 1151; Error: 0.49605612916;
Epoch: 1152; Error: 0.495984064596;
Epoch: 1153; Error: 0.495940883582;
Epoch: 1154; Error: 0.49587009406;
Epoch: 1155; Error: 0.49577167639;
Epoch: 1156; Error: 0.495713701557;
Epoch: 1157; Error: 0.495665027042;
Epoch: 1158; Error: 0.495623260488;
Epoch: 1159; Error: 0.495580964156;
Epoch: 1160; Error: 0.495467634378;
Epoch: 1161; Error: 0.495370938381;
Epoch: 1162; Error: 0.495290438732;
Epoch: 1163; Error: 0.495187384955;
Epoch: 1164; Error: 0.495107194614;
Epoch: 1165; Error: 0.495020945441;
Epoch: 1166; Error: 0.49499918828;
Epoch: 1167; Error: 0.494967732024;
Epoch: 1168; Error: 0.49494378072;
Epoch: 1169; Error: 0.49492958351;
Epoch: 1170; Error: 0.494909048254;

```

Рис. Процес тренування нейронної мережі

Табл. Точність прогнозування нейронних мереж

Кількість прихованих шарів	Кількість факторів	Відсоток точності передбачення
2	12	52,45 %
2	20	55,06 %
4	12	63,41 %
4	20	67,25 %

Згідно з даними табл., можна зробити висновок, що найкраще себе показала нейронна мережа з чотирма прихованими шарами та 20 факторами, яка отримала результат точності передбачення 67,25 %. Нейронна мережа, яка складається із двох прихованих шарів, показала відчутно гірший результат, який становив 55,06 % у разі застосування 20 факторів.

Незважаючи на те, що різниця в кількості факторів є великою, результат, отриманий з використанням різної кількості факторів, показав, що відібрані фактори мають значно більший вплив на формування результату, тому рештою факторів можна нехтувати у разі, якщо фактор продуктивності є критичним.

**Обговорення отриманих результатів.** Дослідивши алгоритми, які також використовують для прогнозування результатів спортивних матчів і отримавши точність розробленого алгоритму від 52 до 67 %, можна зробити висновок, що алгоритм проявив себе позитивно і може бути використаний в подальшому для створення інших програмних засобів, задачею яких є прогнозування.

Покращити точність прогнозу алгоритму можна застосувавши більшу кількість прихованих шарів нейронної мережі, про що свідчить стрімке зростання точності результату прогнозу зі збільшенням числа прихованих

шарів, але тоді і час тренування нейронної мережі зростає відповідно.

**Висновок.** Зважаючи на щорічне збільшення інтересу до сфери беттингу, а також надзвичайно малий асортимент програмного забезпечення який не тільки допомагав би зацікавленим людям у здійсненні їхніх прогнозів відносно конкретних матчів, але й спеціалістам, думці яких довіряють багато інших людей і які займаються цим професійно, можна зробити висновок, що цей алгоритм може бути використано для покращення цієї ситуації шляхом створення на його основі різноманітних програмних засобів прогнозування результатів матчів.

Беручи до уваги інші алгоритми, задачею яких є прогнозування результатів матчів, цей алгоритм не відрізняється своєю точністю, натомість він має великий потенціал розвитку, беручи до уваги можливість його модифікації, наприклад, використавши більшу кількість прихованих шарів для підвищення точності результату.

### Перелік використаних джерел

Adams, T. (2017). *Training an artificial neural network*. Retrieved from: <https://www.solver.com/training-artificial-neural-network-intro>.

Badde, S., Gupta, A., & Patki, K. (2009). Cascade and Feed Forward Back propagation Artificial Neural Network Models for Prediction of Compressive Strength of Ready Mix Concrete. Second International Conference on Emerging Trends in Engineering (SICETE). Retrieved from: [http://www.iosrjournals.org/iosr-jmce/papers/sicete\(civil\)-volume3/26.pdf](http://www.iosrjournals.org/iosr-jmce/papers/sicete(civil)-volume3/26.pdf)

Galushkin, A. I. (2000). *Teoriia neironnykh setei*. Moscow: IPRZhR. 415 p. [in Russian].

Gorban, A. N., & Rossiev, D. A. (1996). *Neironnye seti na personalnom kompiutere*. Novosibirsk: Nauka. 276 p. [in Russian].

Hryhorkiv, V. S., Yaroshenko, O. I., & Filipchuk, N. V. (2012). Neironni merezhi ta yikh vykorystannia dlia prohozuvannia tendentsii rynku nerukhomosti [Neural networks and their application for real estate market trends forecasting]. *Scientific Bulletin of UN-FU*, 22(5), 324–330.

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37, 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>

Kruglov, V. V., & Borisov, V. V. (2001). *Iskustvennye neironnye seti: teoriia i praktika*. Moscow: Goriachaia liniia-Telekom. 382 p. [in Russian].

Liubun, Z. M. (2006). *Osnovy teorii neironmerzheh*. Lviv: Vyd. tsentr LNU im. Ivana Franka. 140 p. [in Ukrainian].

**В. М. Кулик Т. А. Коротеєва**

Національний університет "Львівська політехніка", г. Львів, Україна

## АЛГОРИТМ ПРОГНОЗИРОВАНИЯ РЕЗУЛЬТАТОВ ФУТБОЛЬНЫХ МАТЧЕЙ НА ОСНОВИИ НЕЙРОННЫХ СЕТЕЙ

Исследовано использование одного из способов машинного обучения, а именно нейронной сети в прогнозировании результатов спортивных событий. В данный момент сфера машинного обучения находится на пике интереса среди широкой общественности и используется в решении большого спектра проблем современного мира. Среди основных преимуществ и одновременно причиной, которая стала основной в выборе способа решения поставленной проблемы, является возможность применения нейронных сетей для решения задач, в которых не видно четкого алгоритма решения проблемы. На основе собранных данных об архитектуре и способах применения различных типов нейронных сетей, алгоритмов обучения, а также методов, которые могут быть использованы для решения проблем прогнозирования, проведено исследование и разработка алгоритма прогнозирования. Учитывая отсутствие в открытом доступе похожих средств для прогнозирования результатов спортивных событий, оценка алгоритма проведена с помощью сравнения точности с другими алгоритмами прогнозирования, которые не предусматривают использования нейронных сетей. Такое тестирование осуществлено на основе общедоступного набора данных, полученного из открытых источников.

**Ключевые слова:** нейросети; многослойный перцептрон; алгоритм обучения; обратное распространение ошибки.

## **ALGORITHM OF THE FOOTBALL MATCH RESULT PREDICTION ON THE BASIS OF NEURAL NETWORKS**

The authors present the research of using one of the approaches of machine learning called neural network in predicting sporting events results. At the moment, the field of machine learning is at the peak of interest among the general public and is used in solving a large aspect of the problems of the modern world. One of the main reasons for neural networks and machine learning to be very popular today, is the availability of computer resources for everyone, which gave the opportunity to use them for almost everybody who needs to solve their problems, not spending a huge amount of money on the resources for neural networks specifically. Among the main advantages and at the same time the reason, that has become the main one in choosing the method of solving the posed problem is the possibility of using neural networks to solve problems in which there is no clear algorithm provided. Another area that has been in the lead for several years now is the bookmaking industry, which popularity is only growing. This is especially evident concerning the most popular kind of sport in the world – football. This situation was partly caused because of computerization which today allows almost every computer user make a bet remotely any time they wish, even during the sport match. Based on the collected data about architecture and methods of using various types of neural networks, as well as in general methods that can be used to solve prediction problems, research was conducted and prediction algorithm was developed. The authors have explained the lack of similar tools in the open access to predict the results of sporting events; the evaluation of the algorithm was carried out by comparing accuracy with other forecasting algorithms that do not involve the use of neural networks. Such testing was carried out on the basis of a publicly available data set obtained from open sources.

**Keywords:** neural networks; multilayer perceptron; learning algorithm; backpropagation.