

КЛІЄНТ-СЕРВЕРНЕ ЗАСТОСУВАННЯ ДЛЯ ПЕРЕДБАЧЕННЯ РЕЗУЛЬТАТІВ СПОРТИВНИХ ЗМАГАНЬ

¹ Вінницький національний технічний університет;

Анотація

Розроблено алгоритм прогнозування результатів спортивних подій, з використанням якого дозволяє оптимізувати процес підбору подій, підвищуючи таким чином прибуток від ставок. Алгоритм реалізований у вигляді відповідного програмного забезпечення. Програмне забезпечення клієнтської та серверної частини було написано на мові програмування JavaScript. Щоб зберігати дані користувачів було обрано базу даних Firestore. Розроблений алгоритм був втілений на серверній частині проекту. На основі тестових даних було проведено експериментальні дослідження. Для створення моделі машинного навчання з метою удосконалення зв'язку між даними використано сервіс AutoML. В результаті, коефіцієнт кореляції було збільшено на 20%, а відсоток повернених інвестицій у середньому збільшився на 15% і став позитивним.

Ключові слова: передбачення результатів спортивних змагань, інформаційні системи.

Abstract

An algorithm for predicting the results of sporting events has been developed, using which it is possible to optimize the process of selecting events, thus increasing the profit from betting. The algorithm is implemented in the form of appropriate software. The client and server software was written in the JavaScript programming language. The Firestore database was selected to store user data. The developed algorithm was implemented on the server part of the project. Based on test data, experimental studies were conducted. The AutoML service was used to create a machine learning model to improve data communication. As a result, the correlation coefficient was increased by 20%, and the percentage of return on investment increased by an average of 15% and became positive.

Keywords: prediction of sports results, information systems.

Вступ

У науці для дослідження процесів різної природи широкого розповсюдження набули методи інтелектуального аналізу даних [1]. З їх допомогою можна вирішувати різноманітні задачі аналізу даних – прогнозування, класифікації, кластеризації тощо. Для створення прогнозу на певну подію потрібно враховувати вплив багатьох факторів в системі та факторів зовні системи. Також певні параметри можуть впливати не тільки на вихідний результат, а і чинити вплив на інші параметри. Через це встановлюється внутрішній зв'язок який важко прогнозувати використовуючи тільки експертні думки. Непередбачуваність та неочевидність зв'язків в системі та між собою є головною проблемою прогнозування. З розвитком нейромереж стало можливим враховувати всі ці параметри та їх вплив на результат. Виходячи з цього і збільшилась точність прогнозування результатів певних подій. На сьогоднішній день машинне навчання та нейромережі почали використовуватись у всіх галузях життя за рахунок їх точності, зручності та гнучкості в налаштуванні автоматичному підстроюванні під конкретні задачі. Основною властивістю нейромереж [2, 3] є паралельна обробка інформації одночасно багатьма нейронами. Завдяки цьому досягається збільшення швидкості обробки інформації. Іншою особливістю нейромереж є здатність до навчання і узагальнення інформації. Завдяки цьому досягається схожість з роботою головного мозку людини.

Метою створення системи є спрощення процесу вибору подій, на які слід зробити ставки, за рахунок передбачення результатів спортивних змагань.

Результати дослідження

Для пошуку оптимальних стратегій передбачення результатів спортивних змагань спочатку був обраний кореляційний аналіз. Однак, він не показав достатньо хороших результатів при великій кількості характеристик. Після цього було прийнято рішення залучити Google AutoML для того, щоб даний сервіс давав свої прогнози для кожен матч по усім маркетам.

Щоб AutoML створив якнайточніше модель необхідно підготувати велику кількість даних. Окрім стандартних коефіцієнтів на кожен матч також було обрахована деяка статистика кожної граючої команди в момент матчу в сезоні. Кожній команді була власна статистика скільки ігор зіграно (gp). Першою було додано clean_sheets та failed_to_score, які відповідають за ні одного пропущеного м'яча та ні одного забитого м'яча відповідно. Псевдокод збору даних:

```

1  if (score.away === 0) {
2  homeStats.clean_sheets_percentage = ((homeStats.clean_sheets_percentage *
3  (homeStats.gp) + 1) / homeStats.gp + 1) || 0;
4  awayStats.failed_to_score_percentage = ((awayStats.failed_to_score_percentage *
5  (awayStats.gp) + 1) / awayStats.gp + 1) || 0; 6  }
7  if (score.home === 0) {
8  awayStats.clean_sheets_percentage = ((awayStats.clean_sheets_percentage *
9  (awayStats.gp) + 1) / awayStats.gp + 1) || 0;
10 homeStats.failed_to_score_percentage = ((homeStats.failed_to_score_percentage *
11 (homeStats.gp) + 1) / homeStats.gp + 1) || 0; 12 }

```

Схожим чином були зібрані дані про відсоток перемог, програвів та нічиїх. Надавши ці утворені дані AutoML створено 3 моделі по трьом маркетам: 1. 3 результати (перемога домашньої команди, нічия чи перемога команди на виїзді); 2. Чи заб'ють команди в сумі більше 2 голів за гру, чи ні; 3. Чи обидві команди заб'ють хоча б по одному голу.

Усього нами було виділено 74 характеристики. Коефіцієнти букмекерів на домашню команду, на нічию та на коефіцієнт на домашню команду стали найбільш важливими, що є логічно, бо букмекери дуже ретельно підбирають свої коефіцієнти. Модель, результат тренування якої ми бачимо на рисунку 1, виявилася дуже точною. Це пояснюється тим, що чим менший коефіцієнт букмекерської контори – тим більша реальна імовірність перемоги. Так як середній коефіцієнт на перемогу будь-якої команди, імовірність перемоги якої у нашій моделі є більшою за 0.5 становить приблизно 1.5. Таким чином, нам потрібно не програвати як мінімум кожні 2 гри. З ростом порогу імовірності коефіцієнт зменшується. При подальшому використанні створеного додатку це варто враховувати.

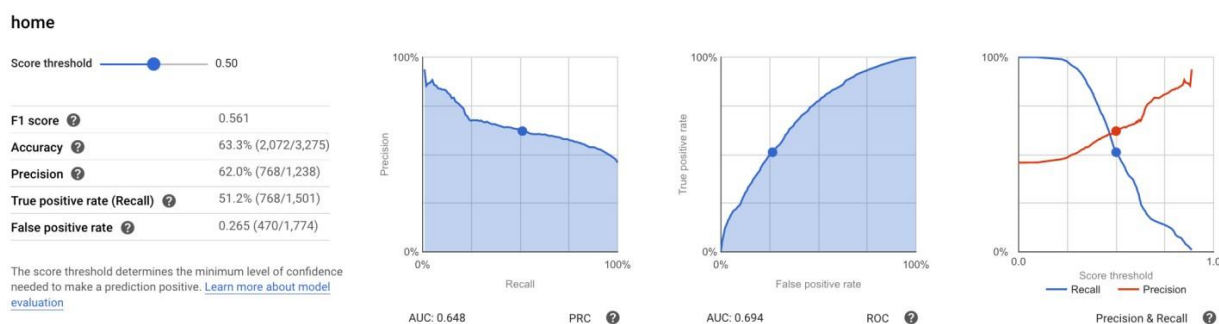


Рис. 1. Результати тренування моделі для трьох результатів

Висновки

Розроблено алгоритм прогнозування результатів спортивних подій, з використанням якого дозволяє оптимізувати процес підбору подій, підвищуючи таким чином прибуток від ставок. Алгоритм реалізований у вигляді відповідного програмного забезпечення. Програмне забезпечення клієнтської та серверної частини було написано на мові програмування JavaScript. Щоб зберігати дані користувачів було обрано базу даних Firestore. Розроблений алгоритм був втілений на серверній частині проекту. На основі тестових даних було проведено експериментальні дослідження. Для створення моделі машинного навчання з метою удосконалення зв'язку між даними використано сервіс AutoML. В результаті, коефіцієнт кореляції було збільшено на 20%, а відсоток повернених інвестицій у середньому збільшився на 15% і став позитивним.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Барский А.Б. Нейронные сети: распознавание, управление, принятие решений. — М.: Финансы и статистика, 2004. — 176 с.

2. Ясницкий Л.Н. Введение в искусственный интеллект. — Учебное пособие для вузов. 2-е издание, испр. — М.: Академия, 2008. — 176 с. — ISBN 978-5-7695-5390-5..

3. Храмов В.В. Интеллектуальные информационные системы. Интеллектуальный анализ данных. Часть 2. — Учебное пособие. — Ростов н/Д.: Ростовский государственный университет путей сообщения, 2012. — 134 с.

Мозгальов Андрій — студент групи ІАКІТ-20м, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: andreymozgaliiov@gmail.com

Науковий керівник: **Кулик Ярослав Анатолійович** — доцент, кандидат технічних наук кафедри комп'ютерних систем і автоматики, Вінницький національний технічний університет, м. Вінниця

Mozgaliiov Andriy – student of group ІАКІТ-20m, faculty of computer systems and automation, Vinnytsia National Technical University, Vinnytsia, e-mail:

Supervisor: **Kulik Yaroslav** . - Associate Professor, Candidate of technical sciences of Computer Control & automatization Systems Department, Vinnytsia National Technical University, Vinnytsia