ABSTRACT

Master's Thesis: 91 pp., 13 figs., 41 tables, 1 supplement, 67 sources.

Topicality. Today, the field of betting and bookmaking is popular with a wide range of sports fans.

Issues of predicting the outcome of future events are and will be relevant for everyday life, sports, politics and more. With the increase in the number and quality of methods of intellectual analysis, the idea of predicting the results of sports events through mathematical algorithms, which can help us to obtain more accurate predictions of results than to listen to the subjective predictions of football experts, has come to fruition.

Probably the biggest problem for people who are professionally engaged in betting is to predict the outcome of events. The sports betting market is growing more and more dynamic every year. Every day, thousands of people place their bets on an event, using different methods to estimate likelihood.

With the increase in market volume and the increasing number of bookmakers, it is becoming increasingly difficult for people to analyze and find successful event and market options.

The paper introduces the concept of betting and describes in general terms the task of bookmaking. It defines the purpose of the research and the tasks that must be completed in order to achieve it. Existing research results of different scientists who have researched this problem are analyzed.

There are four basic principles for predicting the outcome of sports events. Different approaches to the task have been considered and our own method has been proposed. Methods such as Poisson distribution, simulation modeling of the Markov Monte Carlo chain, and many other research methods have been considered. The formulation of the problem is formulated and the properties of the problem are investigated. A backtesting algorithm was developed and described as a mechanism for presenting team statistics at any point in time for a particular season to collect sports event data. Correlation analysis for the selected

parameters was shown to show a moderate correlation of data and the use of Google AutoML to identify patterns between the data was described.

The importance of using machine learning in solving this problem is substantiated. A system has been developed that collects event data and calculates statistics for each team at each point of time using the backtesting algorithm. A service has been developed to create and test the quality of the strategy. The results of experimental studies of task efficiency are presented, where we conducted experimental sets of strategies with and without adding the result of the AutoML service and for each strategy the Pearson correlation coefficient was calculated based on the results of two past seasons. The results obtained are analyzed.

Relationship with working with scientific programs, plans, themes. The work was performed at the Department of Automated Information Processing and Management Systems of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" within the topic "Effective methods of solving the theory of schedules. State Registration Number ”(No. DR 0117U000919).

The purpose of the system is to simplify the process of choosing the events to bet on, to improve the balance statistics of the betting people.

To achieve this goal it is necessary to perform the following tasks:

- review the known results of solving the task;
- develop an event data collection process;
- integrate with the service to calculate predictions for events;
- to develop a backtesting algorithm;
- to develop software implementation of the backtesting algorithm;
- to analyze the results obtained.

The object of study is the process of selecting sports events.

The subject of the study - methods of predicting the results of sports events.

Scientific novelty of the obtained results

A backtesting method was developed to collect data and formulate a strategy that allows predicting the results of sporting events and verifying the reliability of a particular strategy through the use of correlation analysis. It is shown that the use of created models in AutoML system allows to get better correlation indicators between the results of different seasons.

Publications. Theses have been published in the international scientific-practical conference "Mathematical and imitation modeling of the MODS 2019 systems" and in the all-Ukrainian scientific-practical conference of young scientists and students "Information systems and control technologies" (ISTU-2019). The article is published in the scientific journal "Bulletin of modern information technologies".

BETTING, BOOKMAKING, BECTESTING, DATA PROCESSING, FILTRATION, FORECASTING, CORRELATION ANALYSIS, MACHINE TRAINING.