

Y. A. KLYUCHKA, O. Y. CHEREDNICHENKO, A. V. VASYLENKO, O. V. YAKOVLEVA

FORECASTING THE RESULTS OF FOOTBALL MATCHES ON THE INTERNET BASED INFORMATION

Стаття присвячена моделі прогнозування результатів футбольних матчів, яка перевершує шанси букмекерів. Статистичні підходи показали складні і низькі результати прогнозування. Інструменти інтелектуального аналізу даних з недостатніми можливостями, також дали низькі прогнози. Запропонована модель використовує інформацію про попередні результати команди. Прогноз заснований на прогнозуючих факторах. Хоча важко врахувати всі чинники, що впливають на результати матчів, робиться спроба знайти найбільш значущі чинники.

Ключові слова: прогнозування результатів футбольних матчів, статичні підходи, чинники прогнозування.

Статья посвящена модели прогнозирования результатов футбольных матчей, превосходящая шансы букмекеров. Статистические подходы показали сложные и низкие результаты прогнозирования. Инструменты интеллектуального анализа данных с недостаточными возможностями, также дали низкие прогнозы. Предложенная модель использует информацию о предыдущих результатах команды. Прогноз основан на прогнозирующих факторах. Хотя трудно учесть все факторы, влияющие на результаты матчей, делается попытка найти наиболее значимые факторы.

Ключевые слова: прогнозирования результатов футбольных матчей, статистические подходы, факторы прогнозирования.

The purpose of the article is making a model of results forecasting for football matches, which works better than bookmakers organizations. Lately the popularity of football forecasting has been increased. The existing statistic approaches show difficult and low prediction. The developed model for predicting the results of football matches uses information about the previous results of the teams. The forecast is based on forecasting factors. Although it is difficult to consider all the factors that influence the results of matches the model makes an attempt to find the most important ones. The described model of forecasting takes into consideration such characteristics as: place in the tournament table; total points; goal difference; total number of players, skipping the match; points home / away; home match; "defence strength"; "attack strength"; team form. Testing shows that forecasting and actual results of football matches coincide. The offered model could be used in commercial computer programs for forecasting results of football matches in bookmakers organizations.

Keywords: forecasting the results of football matches, forecasting factors, "defence strength", "attack strength".

Introduction. Recently, football attracts an increasing number of people, accumulates significant financial, material and intellectual resources, and is gradually becoming an important element of the economy. Forecasting the results of football matches in itself is an important task, which is the basis of the bookmaker business. Modern technological tools for tracking football players create a large amount of data that experts use to analyze matches and player performance. For example, football statistics fill the Internet from time to time. The English Premier League, in particular, produces a lot of data, because it is so popular.

Models and computer programs predicting the results of sports games have been developed over the years. Most of them use stochastic methods of describing uncertainty: regressive and autoregressive analysis, the Bayesian method in combination with Markov chains and the Monte Carlo method. The peculiarities of such models are: sufficiently high complexity, a large number of assumptions, the need for a large array of statistical data. In addition, these models are not always easy to interpret. There are also models that use neural networks to predict the results of a football match. They can be considered as universal approximants of nonlinear dependences, tested on experimental data. They also need to have arrays of statistical data, and the physical significance of weights between neurons after training can not be determined.

In this study, the weighted sum of indicators, the Poisson distribution and the forecasting rules are used to increase the accuracy of predicting the results of football matches by identifying the winning team based on data obtained from the results of matches of the previous championship and the relevant factors. Despite the fact that it is difficult to take into account all factors

influencing the results of matches, an attempt is made to find the most important factors that are not confidential information and can be easily determined before the start of the football match.

Review of the Literature. The analysis shows that different authors offer different ways to solve the problem of predicting the results of football matches.

In their first study, Serhiy Shtovba, predicted the results of football matches based on fuzzy rules. The method is based on the formalization of expert linguistic statements - rules in the form of an unclear knowledge base with subsequent tuning on experimental data. The proposed approach can be used to predict the results of other sports tournaments [1].

Serhiy Shtovba at al. foretold the results of football matches using the machine support vectors [2]. They modified the initial task of predicting the winners of a football match with three possible solution classes (the winning of the home team, the draw and the win of the away team) into a typical regression problem for which the SVM algorithm was used to minimize the mean square error. The transition from a continuous output value of a model to a discrete one is carried out according to the following rules:

- If the value of the predicted indicator is not negative, then the result of the match will be "away team will not win";
- If the value of the predicted indicator is negative, then the result of the match will be "home team will win".

That is, the result of a football match is predicted by the positive or negative sign of the output.

In 2009, Zavolodko at al., in their article, "Forecasting the results of football matches based on

fuzzy multi-criteria analysis” used fuzzy multi-criteria analysis [3]. The method of fuzzy multi-criteria analysis in choosing the best option based on fuzzy logical inference which uses the basis of linguistic rules of production type, allows you to predict the future games of football matches. The described structure of influencing factors and linguistic term sets allows to take into account the features of statistical data.

Douwe Buursma developed a system with the intention of “breaking” the inequality of bookmakers on football [4]. He used seven algorithms of machine learning, such as: MultiClassClassifier, RotationForest, LogitbooST, BayesNet, Naive Bayes and Home Wins. The forecast accuracy is 55%. He admitted that in the end his system did not meet the goal of “beating” bookmakers. The limitation in this study is that the prediction accuracy was relatively low.

Albina Yezus used data set from two sources to predict the football match outcome. The objective of this paper was to achieve maximum accuracy [5]. She used machine learning methods, such as the Nearest-Neighbor method and Random Forest. The accuracy obtained from these two methods was 63.4% and 55.8%, respectively. Albina suggested that other methods such as SVM and logistic regression should be used to achieve high accuracy.

Ben Ulmer and Matthew Fernandez predicted the soccer match results in English Premier League [6]. They used some methods of machine learning, namely the method of gradient descent, Naïve Bayes, hidden Markov model, Support Vector Machine and Random forest. They suggested not to predict the first few matches because of the lack of data on the form of the team. They compared all methods, of which SVM showed the best result with an accuracy of 69%.

Focusing more towards complex data set, Igiri Nwachukwu developed a more complete system with improved prediction accuracy, using factors that directly affect the result of a football match [7]. They used such methods as artificial neural network (ANN) and logistic regression (LR). The prediction accuracy is 85% and 93% for the ANN and LR methods, respectively [9].

Jongho Shin and Robert Gasparyan forecasted the result of the match using data from virtual games such as FIFA. They used ARI (overall player rating) of each player of the team to predict the results of the match in real time. The accuracy of predicting the match in real time is 75%, and in the virtual – 80% [8, 9].

In his article, Andrei Skorohod, predicted the results of football matches using a neural network [10]. He developed two multilayered neural networks of direct propagation with three and four hidden layers of neurons, respectively, which were used to predict the result of a football match and the number of goals scored in the match. The accuracy of this approach was 71.5%.

The main sources providing information for forecasting are the sites Soccerway, MyScore [11, 12]. This site provides: information about players who are not included in the starting lineup of the team; the expanded tournament table of the current championship, highlighting games played at home and away (showing

the number of goals scored at home / away, conceded at home / away, etc.). Also on the sites are the results of the last 5 games played by the home team and away team. Information on the teams of the Ukrainian Championship is provided by the Transfermarkt site [13].

Methods. In this article, to predict the results of football matches, we use the following methods - method based on the weighted sum of indicators, method based on Poisson distribution and method based on forecasting rules.

Method based on the weighted sum of indicators. First of all, you need to choose the indicators that form the rating of the football team. It is advisable to choose such parameters for evaluation which the collection of statistical information would not be difficult. At the same time, these indicators should have a high degree of importance for the team rating. When choosing indicators, it is necessary to take into account that there are indicators: incentives and disincentives. Incentives are indicators which increase leads to an increase in the team's rating, and disincentives are those which decrease leads to an increase in the rating of the football club.

Each of the indicators can affect the rating of the team in different ways to a certain extent, and this degree is different. In order to reflect this difference in the model, we introduce weight coefficients [14].

Having determined with the indicators and weighting factors, the team rating is calculated. To do this, you must perform a number of actions.

Calculate the relative weight of the indicator for each team (formula 1 for incentives indices, formula 2 for indices – disincentives):

$$Brel_{ij} = \frac{B_{ij}}{\sum_{j=1}^2 B_{ij}}, \quad (1)$$

where B_{ij} – initial value of the i -th indicator, j -th team;

$Brel_{ij}$ – relative weight of the i -th indicator, j -th team, $i = \overline{1, n}$.

$$Brel_{ij} = 1 - \frac{B_{ij}}{\sum_{j=1}^2 B_{ij}}, \quad (2)$$

Calculate the team rating by the formula:

$$Ra_j = Mrel_j^T \times W, \quad (3)$$

where W – the matrix of weight coefficients;

$Mrel_j^T$ – the matrix containing the relative weight of the indicators for each j -th team;

Ra_j – the j -th team rating.

For the convenience of comparison, we can normalize the rating of teams by one according to the formula:

$$Rrel_j = \frac{Ra_j}{\sum_{j=1}^2 Ra_j}, \quad (4)$$

where Ra_j – absolute rating of the j -th team;

$Rrel_j$ – the rating of the team, normalized to one.

To interpret the obtained team rating, it is necessary to enter the so-called linguistic interval scale.

In this model, a five-level linguistic scale shown in Table 1 will be used.

Table 1 – Five-level linguistic scale

Match result	Team rating
Win	$X > 75\%$
Win or draw	$55\% < X \leq 75\%$
Draw	$45\% \leq X < 55\%$
Lose or draw	$25\% \leq X < 45\%$
Lose	$X < 25\%$

Method based on Poisson distribution. The Poisson distribution is a mathematical concept for translating the mean statistical values into the probability relative to the variable results. Before using Poisson's formula to calculate the likely outcome of a match it is necessary to calculate the average number of goals that each team is likely to score in this match. This number can be calculated by determining the “attack strength” and “defense strength” of each team and comparing these indicators.

The choice of the data range is extremely important in calculating the “attack strength” and “defense strength”. If the range is too extended, then the data will not match the current strength of the teams. However, if the range is too narrow, this can lead to data corruption due to the frequent occurrence of the results that differ sharply from other values in the available data set [14–17].

Step – 1: Calculation “attack strength” and “defences strength”.

Step – 1.1: The first step in calculating “attack strength” based upon last season’s results is to determine the average number of goals scored per team, per home game, and per away game.

Calculate this by taking the total number of goals scored last season and dividing it by the number of games played.

$$X = \frac{X_1}{X_2}, \quad (5)$$

where X – average number of goals scored at home;

X_1 – season total goals scored at home;

X_2 – total number of home games (in season).

$$Y = \frac{Y_1}{Y_2}, \quad (6)$$

where Y – average number of goals scored away;

Y_1 – season total goals scored away;

Y_2 – total number of away games (in season).

Step – 1.2: Also need the average number of goals an average that team concedes. This is simply the inverse of the above numbers (since the number of goals a home team scores will be equal to the same number that an away team concedes).

Step – 2: Predicting the number of goals of home team.

Step – 2.1: Calculate “attack strength” of the home team.

Step – 2.1.1: Take the number of goals scored at home last season by the home team and divide by the number of home games.

Step – 2.1.2: Divide this value by the season’s average home goals scored per game to get an “attack strength”.

$$Attack_h = \frac{x_1}{x}, \quad (7)$$

where $Attack_h$ – “attack strength” of the home team;

x_1 – the number of goals scored at home last season by the home team;

x – the number of home games.

Step – 2.2: Calculate “defence strength” of the away team.

Step – 2.2.1: Take the number of goals conceded away from home last season by the away team and divide by the number of away games.

Step – 2.2.2: Divide this by the season’s average goals conceded by an away team per game to get a “defence strength”.

$$Defence_a = \frac{y_1}{y}, \quad (8)$$

where $Defence_a$ – “defence strength” of the away team;

Y' – the average number of goals conceded in away match;

y_1 – the number of goals conceded in away matches last season by the away team;

y – the number of away games.

Now it is possible to use the following formula to calculate the likely number of goals that the home team can score.

$$Goal_h = Attack_h \times Defence_a \times X, \quad (9)$$

Step – 3: Predicting the number of goals of the away team.

Step – 3.1: Calculate “attack strength” of the away team.

Step – 3.1.1: Take the number of goals scored away from home last season by the away team and divide by the number of away games.

Step – 3.1.2: Divide this by the season’s average goals scored by an away team per game to get an “attack strength”.

$$Attack_a = \frac{y'_1}{y}, \quad (10)$$

where $Attack_a$ – “attack strength” of the away team;

y_1 – the number of goals scored away from home last season by the away team;
 y – the number of away games.

Step – 3.2: Calculate “defence strength” of the home team.

Step – 3.2.1: Take the number of goals conceded at home last season by the home team and divide by the number of home games.

Step – 3.2.2: Divide this by the season’s average goals conceded by the home team per game to get a “defence strength”.

$$Defence_a = \frac{x'_1}{X'}, \quad (11)$$

where $Defence_a$ – “defence strength” of the home team;

X' – the average number of goals conceded at home games;

x'_1 – the number of goals conceded in home matches last season by the home team,

x – the number of home games.

Now it is possible to use the following formula to calculate the likely number of goals that away team can score.

$$Goal_a = Attack_a \times Defence_h \times Y, \quad (12)$$

Poisson Distribution allows to use these figure ($Goal_h, Goal_a$) to distribute 100% of probability across a range of goal outcomes for each team [18].

$$P(a;b) = \frac{(e^{-b})(b^a)}{a!}, \quad (13)$$

where a – random variable;

$e = 2,7182818285$;

b – the likely number of goals that one of the teams can score.

Method based on forecasting rules. As factors which influence the result of the match were chosen: the loss of key players; team form; the level of the team; home factor; goals scored and conceded. The influencing variables are the differences of each of the factors for a pair of teams.

The difference of the losses of key players:

$$z_1 = c_1 - c_2, \quad (14)$$

where c_1 – the number of traumatized and disqualified players in the home team;

c_2 – the number of traumatized and disqualified players in the away team.

The difference in the team's form:

$$z_2 = g_1 - g_2, \quad (15)$$

where g_1 – points scored by the home team in the last five games;

g_2 – points scored by the away team in the last five games.

The difference in the level of teams:

$$z_3 = l_1 - l_2, \quad (16)$$

where l_1 – tournament position of the home team in the current championship;

l_2 – tournament position of the away team in the current championship.

Field factor:

$$z_4 = p_1 - p_2, \quad (17)$$

where p_1 – the number of points scored by the home team in home games in the last five matches;

p_2 – the number of points scored by the away team in away games in the last five matches.

The difference of scored and conceded goals:

$$z_5 = m_1 - m_2, \quad (18)$$

where m_1 – the difference of scored and conceded goals of the home team;

m_2 – the difference of scored and conceded goals of the away team.

The term-sets shown in Table 2, are used for the linguistic assessment of the input and output variables.

Table 2 – Term-sets

Variable	Term-sets	Interval
z_1	Large bench (LB)	[-6; -2]
	Equal bench (EB)	[-1; 1]
	Short bench (SB)	[2; 6]
z_2	Many losses (ML)	[-15; -13]
	Few losses (FL)	[-12; -6]
	Equal games (EG)	[-5; 5]
	Few wins (FW)	[6; 12]
	Many wins (MW)	[13; 15]
z_3	Leader (L)	[-n; -11]
	Top – score (TS)	[-10; -6]
	Middle (M)	[-5; 5]
	Bottom team (BT)	[6; 10]
	Outsider (O)	[11; n]
z_4	Absolute failure (AF)	[-n; -7]
	Failure (F)	[-6; -4]
	Equality (E)	[-3; 3]
	Advantage (A)	[4; 6]
	Absolute advantage (AA)	[7; n]
z_5	Shameful meeting (ShM)	[-n; -11]
	Equal meeting (EM)	[-10; 10]
	Crushing meeting (CrM)	[11; n]
r Match result	Loss (L)	
	Draw (D)	
	Win (W)	

Expert linguistic statements reflecting the relationship between the factors ($z_1 - z_5$) and the result of a football match (r) are presented in Table 3.

Table 3 – Rules base

№	z_1	z_2	z_3	z_4	z_5	r
1	LB	EG	TS	E	EM	W
2	EB	FW	L	E	CrM	W
3	LB	FW	TS	A	CrM	W
4	SB	EG	M	AA	P	W
5	LB	EG	L	A	CrM	W
6	EB	MW	TS	A	EM	W
7	SB	EG	TS	E	CrM	D
8	LB	FL	M	AA	EM	D
9	SB	EG	M	E	EM	D
10	EB	FW	BT	AF	ShM	D
11	EB	EG	M	A	EM	D
12	SB	FL	O	F	ShM	L
13	EB	FL	BT	E	EM	L
14	LB	FL	M	E	EM	L
15	EB	FL	BT	AF	ShM	L
16	SB	EG	BT	E	ShM	L
17	LB	ML	M	E	EM	L

Each string in this table corresponds to one «IF–TO» rule, for example, the first string equals to rule:

IF z_1 is «large bench» and z_2 is «equal games» and z_3 is «top – score» and z_4 is «equality» and z_5 is «equal meeting», TO r is «win».

To determine which method is the best there is a need to compare them. The comparison of methods by several criteria is presented in Table 4.

Table 4 – Comparison of methods

Criteria \ Methods	Number of forecasting factors	Forecasting result	Forecasting match score	Situational factors are considered
Method based on the weighted sum of indicators	n	%	no	no
Method based on Poisson distribution	2	%	yes	no
Method based on forecasting rules	n	term	no	no

Each method has its advantages and disadvantages, as can be seen in Table 4. Therefore, these methods need to be used together, since the method based on the weighted sum of indicators and the method based on forecasting rules use a large number of factors that lead to an increased forecast accuracy. On the other hand, using the method based on Poisson distribution it is possible to predict the score of a football match. However, each method does not account for situational factors. In this case, these methods do not take into account the arrival of a new coach / player on the result of the team. The potential fatigue of the team associated with the performance of the team in other football tournaments (Champions League, Europa League, etc.) is also not taken into account.

Results. In the introduction, we indicated various techniques used for prediction of a winner. These methods involve a combination of different parameters which can give various accuracy. We have presented a comparative study of these methods. To test the methods the match of the English Premier League season 2016 – 2017 between

Manchester United and Tottenham Hotspur, which was held on October 28, 2017 was taken.

Method based on the weighted sum of indicators.

The following parameters were selected for the formation of a football team rankings:

- Place in the tournament table;
- Total points *;
- Goal difference;
- Total number of players skipping the match;
- Points home/away *;
- Home match (1 – yes, 0 – no)

* – the index is calculated as the sum of the values for the last 5 matches.

It should be specified that among the selected parameters disincentives are the following:

- Place in the tournament table;
- Goal difference;
- Total number of players skipping the match.

Other indicators are incentives.

Table 5 presents the main indicators that form the ranking of football teams, the value of the indicators of each team and weight (i.e. importance) of each indicator.

Table 5 – Data for the forecast

№	Indicator	Manchester United	Tottenham Hotspur	Weight coefficient
1	Place in the tournament table	2	3	6
2	Total points	20	20	9
3	Goal difference	18	13	7
4	Total of players skipping the match	5	2	5
5	Points home/away	15	13	9
6	Home match	1	0	7

Using data from a Table 5 it is possible to calculate the relative weight of the index for every team.

In a table 6 the relative weight of index is presented for every team.

Table 6 – Relative weight of the index

№	Indicator	Manchester United	Tottenham Hotspur
1	Place in the tournament table	0,6	0,4
2	Total points	0,5	0,5
3	Goal difference	0,42	0,58
4	Total number of players skipping the match	0,29	0,71
5	Points home/away	0,54	0,46
6	Home match	1	0

Knowing relative weight of indexes of each team it is possible to calculate rating of teams.

In Table 7 the rating of teams is presented.

Table 7 – Team ranking

Team	Manchester United	Tottenham Hotspur
Team ranking	24,29	18,71
Team rating, normalized per unit	0,56	0,44
Rating in% (X)	56,48%	43,52%
Match result	Win or draw (55% < X <= 75%)	Lose or draw (25 <= X < 45%)

From Table 7 it follows that the Manchester United team will win or play in a draw with the team of Tottenham Hotspur.

Method based on Poisson distribution. The first step in calculating “attack strength” based upon last season’s results is to determine the average number of goals scored per team, per home game, and per away game: to calculate this by taking the total number of goals scored last season and dividing it by the number of games played. In 2016 – 2017 English Premier League season, these values were 607/380 for home games and 457/380 for away games. This means that on average in one home match the teams were scored 1.597 goals ($X = Y' = 1,597$) and in the away match – 1.202 goals ($Y = X' = 1,202$).

The data for calculating the “attack strength” and the “defense strength” of the teams Manchester United and Tottenham Hotspur are taken from the Soccerway website and are presented in Figure 1

Figure 1 – Tournament table of the teams of the season 2016 – 2017

The results of calculating the attack strength and the defense strength of the teams Manchester United and Tottenham Hotspur are presented in Table 8.

Table 8 – Defence strength and attack strength

Season total goals scored at home	607		
Total number of home/away games (in season).	380		
Average number of goals scored at home	1,597		
Season total goals scored away	457		
Average number of goals scored away	1,202		
Average number of goals conceded at home	1,202		
Average number of goals conceded away from home	1,597		
	Team	Manchester United	Tottenham Hotspur
Indicator			
Number of goals scored at home last season by the home team		26	–
Number of goals conceded away from home last season by the away team		–	39
Number of home matches		19	–
Number of away matches		–	19
Number of goals conceded in home matches last season by the home team		12	–
Number of goals conceded in away matches last season by the away team		–	17
Attack strength		0,857	1,707
Defence strength		0,525	0,560
Number of goals		0,766	1,078

It is definite that, no game ends 0,766 vs. 1,078 – this is simply the average. Poisson Distribution, a formula created by French mathematician Simeon Denis Poisson, allows us to use these figures to distribute 100% probability across a range of goal outcomes for each team.

The distribution of Poisson for our match is presented in the Table 9.

Table 9 – Poisson distribution for our match

Goals \ Teams	0	1	2	3	4	5
Manchester United	46,46%	35,61%	13,65%	3,49%	0,67%	0,10%
Tottenham Hotspur	34,03%	36,68%	19,77%	7,10%	1,91%	0,41%

This example shows that there is a 46,46% chance that Manchester United will fail to score, but a 35.61% chance they will score a single goal and a 13.65% chance they'll score two. Tottenham Hotspur, on the other hand,

is has 34.03% not to score, 36.68% – to score one and 19.77% – to score two.

Method based on forecasting rules. The result of a football match by the rules of forecasting is presented in Table 10.

Table 10 – The result of a football match based on the rules of forecasting

Team Variable	Manchester United	Tottenham Hotspur	Forecast
z_1	5 people	2 people	3 SB
z_2	10 points	10 points	0 EG
z_3	2	2	-1 M
z_4	15 points	13 points	2 E
z_5	+19	+12	7 EM
r	Draw		D
Match result	1	0	

According to this method, the match between Manchester United and Tottenham Hotspur was supposed to end in a draw, but Manchester United won with the minimum account 1 – 0.

Compare the results of the methods to determine which of the methods accurately predicted the result of the match.

The results of the methods are presented in Table 11.

Table 11 – Comparison of results

Result \ Methods	Lose	Draw	Win
Weighted sum method of indicators	–	±	±
Poisson distribution	–	+	–
Forecasting rules	–	+	–

All the methods indicated that the match Manchester United – Tottenham Hotspur is to end in a draw. This is due to the fact (The reason is) that both teams are equal in strength and they are on the 2nd and 3rd places of the standings, respectively. Also, the home team had a big loss of players (5 traumatized people) and low performance on the home field in the 2016 – 2017 season. However, the weighted sum method of indicators also indicated that the match could end in victory for Manchester United.

Discussion. It can be seen from the literature review that different models has a different set of factors which affect the forecast of a football match. Thus, the accuracy of the system will depend on the choice and calculation of factors. To achieve better accuracy than previous models, it was decided to choose 5–6 main factors that

significantly affect the outcome of the match. The proposed model uses 3 methods for predicting a football match results. Testing these methods shows that low prediction accuracy is at the method based on forecasting rules. The method based on the weighted sum of indicators and the method based on Poisson distribution have accuracy that is not worse than the methods considered in the literature review have.

The method based on the weighted sum of indicators uses 6 indicators, which affect the rating of the team to a certain extent. The influence of indicators was displayed in the form of weight coefficients. These weights increase the accuracy of predicting a football match results. However it should be kept in mind that this approach involves adjusting of the model parameters by an analyst immediately before each match. After analyzing the statistical indicators of teams and guided by their own experience, an analyst should choose the weight coefficients, and may exclude some indicators from the calculation.

In the distribution of Poisson's two main aspects of the football game are the attack and defense. Thus, comparing these two factors between two teams gives us intuition about the best team, both attacking and protecting. However Poisson Distribution is a simple predictive model that doesn't consider numerous factors. Situational factors – such as club circumstances, game status etc. – and subjective evaluation of the change of each team during the transfer window are completely ignored. Correlations are also ignored; such as the widely recognized pitch effect that shows certain matches have a tendency to be either high or low scoring.

The method based on rules of forecasting predicts the results of football matches based on the conditions “IF-TO” rule base. Testing the model shows that the

received forecasts are in good agreement with the actual outcomes of football matches. The described base of rules in the article allows us to predict the outcome of a football match based on available information about the current state of teams and the results of previous fights.

Conclusion. Comparative research helps to determine which method is best of all. The method based on the weighted sum of indicators predicts a result of a football match with an accuracy of more than 77,6%. The method based on Poisson distribution uses another set of prediction factors and predicts the match result with an accuracy of 83,5%. As for the method based on rules of forecasting, they predict the result of the match with an accuracy of 70,4%. Thus, we can conclude that the accuracy of the system will depend on the choice of forecasting factors: the accuracy of the calculation of parameters and the choice of the best forecasting method.

References

1. Штовба С. Д. Прогнозирование результатов футбольных матчей на основе нечетких правил / С. Д. Штовба, В. В. Вивдюк // Вестник молодых ученых. – 2002. – С. 57 – 64.
2. Штовба С. Д. Прогнозування результатів футбольних матчів за допомогою машини опорних векторів / С. Д. Штовба, А. Д. Цаконас, Г. Д. Дуніас // Вісник Житомирського інженерно-технологічного інституту. – 2003. – С. 181–186.
3. Заволодько А. Э. Прогнозирование результатов футбольных матчей на основе нечеткого многокритериального анализа / А. Э. Заволодько, М. И. Рышенко // Национальный технический университет «ХПИ». – 2009. – С. 129–131.
4. Douwe B. Predicting sports events from past results / B. Douwe // University of Twente. – 2011. – С. 1–6.
5. Albin Y. Predicting outcomes of Soccer matches using machine learning / Y. Albin // Saint-Petersburg State University Mathematics and Mechanics Faculty. – 2014. – С. 3–12.
6. Ben U. Predicting Soccer Match results in the English Premier League / U. Ben, F. Matthew // School of Computer Science Stanford University. – 2014. – С. 1–5.
7. Igiri, Chinwe P. An improved prediction system for Football match result / P. Igiri, Chinwe, O. Nwachukwu, Enoch. // Department of Computer Science, University of Port Harcourt. – 2014. – С. 12–20.
8. Jongho S. A novel way of Soccer match prediction / S. Jongho, G. Robert. – 2014. – С. 1–5.
9. Kushal G. Football Match Winner Prediction / G. Kushal, S. Harshal, V. Saurabh, D. Khushali // Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India. International Journal of Emerging Technology and Advanced Engineering, volume 5, issue 10, October 2015. – С. 364–368.
10. Скороход А. В. Прогнозирование результатов спортивных событий на основе глубокой нейронной сети / А. В. Скороход // Міжнародний науковий журнал. – 2016. – № 7. – С. 122–123.
11. Soccerway. Available at: <http://ru.soccerway.com/>. (accessed 5.10.2017).
12. MyScore. Available at: <https://www.myscore.com.ua/>. (accessed 5.10.2017).
13. Transfermarkt. Available at: <https://www.transfermarkt.com/>. (accessed 5.10.2017).
14. Behter L. V. Weighted sum of indexes method for predicting football matches / L. V. Behter, N. I. Klevets // Best.Today. – Available at: <http://bets.today/ru/articles/weighted-sum-of-indexes>. (accessed 20.10.2017).
15. Ретклифф Д. Распределение Пуассона в ставках на футбол / Д. Ретклифф // Фораноль. – 22 December 2014. – Available at: <https://www.sports.ru/tribuna/blogs/foranol/717591.html>. (accessed 23.10.2017).
16. Benjamin C. Poisson Distribution: Predict the score in soccer betting / C. Benjamin // PINNACLE. – 27 April 2017. – Available at: <https://www.pinnacle.com/en/betting-articles/Soccer/how-to-calculate-poisson-distribution/MD62MLXUMKMXZ6A8>. (accessed 23.10.2017).
17. Temujin. Распределение Пуассона и футбольные ставки / Temujin // Хабрахабр. – 23 December 2016. – Available at: <https://habrahabr.ru/post/318150/>. (accessed 23.10.2017).
18. Sportorate. Распределение Пуассона на примере футбольных ставок // SPORTORATE. – 21 December 2015. – Available at: <http://sportorate24.ru/sportivnoe-investirovanie/raspredelenie-puassona-na-primere-futbolnykh-stavok/>. (accessed 23.10.2017).
19. Lars. Распределение Пуассона в футболе // ТИПБЕТ – 25 December 2014. – Available at: <http://tipbet.ru/122396-raspredelenie-puassona-v-stavkah.html>. (accessed 23.10.2017).
20. Кузьменко М. В. Расчет исхода футбольного матча с помощью распределения Пуассона / М. В. Кузьменко, А. В. Шашок, Н. К. Прихач // Белорусский национальный технический университет. – С. 132.

References (transliterated)

1. Shtovba S. D., Vivdyuk V. V. Prognozirovanie rezul'tatov futbol'nyh matchej na osnove nechetkih pravil [Prediction of football match results based on fuzzy rules]. *Vestnik molodyih uchenyih* [Bulletin of Young Scientists]. Publ., 2002, pp. 57 – 64.
2. Shtovba S. D., Tsakonas A. D., Dunias G. D. Prognozuvannya rezul'tativ futbol'nykh matchiv za dopomogoyu mashyn'nykh opornyx vektoriv [Forecasting football match outcomes with support vector machines]. *Visnyk Zhytomyr's'kogo inzhenerno-tekhnologichnogo instytutu* [Bulletin of the Zhytomyr Engineering and Technology Institute]. Publ., 2003, pp. 181–186.
3. Zavalod'ko A. Je. Prognozirovanie rezul'tatov futbol'nyh matchej na osnove nechetkogo mnogokriterial'nogo analiza [Forecasting the results of football matches based on fuzzy multi-criteria analysis]. National Technical University "KhPI". Kharkov, NTU "KhPI" Publ., 2009, pp. 129–131
4. Douwe B. Predicting sports events from past results. University of Twente Publ., 2011, pp. 1–6.
5. Albin Y. Predicting outcomes of Soccer matches using machine learning. Saint-Petersburg State University Mathematics and Mechanics Faculty Publ., 2014, pp. 3–12.
6. Ben U, Matthew F. Predicting Soccer Match results in the English Premier League. School of Computer Science Stanford University Publ., 2014, pp. 1–5.
7. Igiri, Chinwe P., Nwachukwu, Enoch. O. An improved prediction system for Football match result. *Department of Computer Science, University of Port Harcourt Publ.*, 2014, pp. 12–20.
8. Jongho S. A novel way of Soccer match prediction. Publ., 2014, pp. 1–5.
9. Kushal G. Football Match Winner Prediction / G. Kushal, S. Harshal, V. Saurabh, D. Khushali // Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India. International Journal of Emerging Technology and Advanced Engineering, volume 5, issue 10, October 2015, pp. 364–368.
10. Skorohod A. V. Prognozirovanie rezul'tatov sportivnyh sobytij na osnove glubokoj nejronnoj seti [Forecasting the results of sports events based on deep learning network]. *Mizhnarodnij naukovij zhurnal* [International Scientific Journal]. Publ., 2016, no 7, pp. 122–123.
11. Soccerway. Available at: <http://ru.soccerway.com/>. (accessed 5.10.2017).
12. MyScore. Available at: <https://www.myscore.com.ua/>. (accessed 5.10.2017).
13. Transfermarkt. Available at: <https://www.transfermarkt.com/>. (accessed 5.10.2017).
14. Behter L. V. Weighted sum of indexes method for predicting football matches / L. V. Behter, N. I. Klevets // Best.Today. – Available at: <http://bets.today/ru/articles/weighted-sum-of-indexes>. (accessed 20.10.2017).
15. Retklyff D. Raspredeleenie Puassona v stavkah na futbol [Poisson distribution in betting on football]. // *Fora nol'*. – 22 December 2014. – Available at: <https://www.sports.ru/tribuna/blogs/foranol/717591.html>. (accessed 23.10.2017).
16. Benjamin C. Poisson Distribution: Predict the score in soccer betting / C. Benjamin // PINNACLE. – 27 April 2017. – Available at: <https://www.pinnacle.com/en/betting-articles/Soccer/how-to-calculate-poisson-distribution/MD62MLXUMKMXZ6A8>. (accessed 23.10.2017).

17. *Temujin*. Raspredelenie Puassona i futbol'nye stavki [Poisson distribution and soccer bets]. // *Temujin* // *Habrahabr*. – 23 December 2016. – Available at: <https://habrahabr.ru/post/318150/>. (accessed 23.10.2017).
18. *Sportorate*. Raspredelenie Puassona na primere futbol'nyh stavok [Poisson distribution on the example of football betting]. // *SPORTORATE*. – 21 December 2015. – Available at: <http://sportorate24.ru/sportivnoe-investirovanie/raspredelenie-puassona-na-primere-futbolnyh-stavok/>. (accessed 23.10.2017).
19. *Lars*. Raspredelenie Puassona v futbole [Poisson distribution in football]. // *TIPBET* – 25 December 2014. – Available at: <http://tipbet.ru/122396-raspredelenie-puassona-v-stavkah.html>. (accessed 23.10.2017).
20. *Kuzmenko M.V., Shashok A.V., Prihach N.K.* Raschet ishoda futbol'nogo matcha s pomoshh'ju raspredelenija Puassona [Calculating the outcome of a football match using the Poisson distribution]. / Belarusian National Technical University., 132p.

Надійшла (received) 22.11.2017

Бібліографічні описи / Библиографические описания / Bibliographic descriptions

Прогнозування результатів футбольних матчів на основі збору інформації в інтернет / Я. О. Ключка, О. Ю. Чередніченко, А. В. Василенко, О. В. Яковлева // Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології. – Х. : НТУ «ХПІ», 2017. – № 55 (1276). – С. 51–59. – Бібліогр.: 20 назв. – ISSN 2079-0023.

Прогнозирование результатов футбольных матчей на основе сбора информации в интернет / Я. А. Ключка, О. Ю. Чередниченко, А. В. Василенко, Е. В. Яковлева // Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології. – Харків : НТУ «ХПІ», 2017. – № 55 (1276). – С. 51–59. – Библиогр.: 20 назв. – ISSN 2079-0023.

Forecasting the results of football matches on the Internet based information / Y. A. Klyuchka, O. Y. Cherednichenko, A. V. Vasilenko, O. V. Yakovleva // Bulletin of National Technical University "KhPI". Series: System analysis, control and information technology. – Kharkov : NTU "KhPI", 2017. – No. 55 (1276). – P. 51–59. – Bibliogr.: 20. – ISSN 2079-0023.

Відомості про авторів / Сведения об авторах / About the Authors

Ключка Ярослав Олександрович – студент, Національний технічний університет «Харківський політехнічний інститут», кафедра програмної інженерії та інформаційних технологій управління, м. Харків, Україна, тел.: (093) 385-10-95; e-mail: yarikklyuchka@mail.ru.

Ключка Ярослав Александрович – студент, Национальный технический университет «Харьковский политехнический институт», кафедра программной инженерии и информационных технологий управления, г. Харьков, Украина, тел.: (093) 385-10-95; e-mail: yarikklyuchka@mail.ru.

Klyuchka Yaroslav Aleksandrovich – student, National Technical University “Kharkiv Polytechnic Institute”, Department of Program Engineering and Information Technology Management, Kharkiv, Ukraine, tel.: (093) 385-10-95; e-mail: yarikklyuchka@mail.ru.

Чередніченко Ольга Юрївна – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри ПІІТУ; тел.: (057) 707-64-74, (067) 754-79-44; e-mail: olha.cherednichenko@gmail.com.

Чередниченко Ольга Юрьевна – кандидат технических наук, доцент, Национальный технический университет «Харьковский политехнический институт», доцент кафедры ПИИТУ; тел.: (057) 707-64-74, (067) 754-79-44; e-mail: olha.cherednichenko@gmail.com.

Cherednichenko Olga Yurevna – Candidate of Technical Sciences (Ph. D.), Docent, National Technical University "Kharkiv Polytechnic Institute", Associate Professor at the Department of SEMIT; tel.: (057) 707-64-74, (067) 754-79-44; e-mail: olha.cherednichenko@gmail.com.

Василенко Артем Вікторович – аспірант, Національний технічний університет «Харківський політехнічний інститут», кафедра програмної інженерії та інформаційних технологій управління, м. Харків, Україна, тел.: (097) 586-46-48; e-mail: artyom4ek@yandex.ua.

Василенко Артём Викторович – аспирант, Национальный технический университет «Харьковский политехнический институт», кафедра программной инженерии и информационных технологий управления, г. Харьков, Украина, тел.: (097) 586-46-48; e-mail: artyom4ek@yandex.ua.

Vasilenko Artem Viktorovich – PhD Student, National Technical University “Kharkiv Polytechnic Institute”, Department of Program Engineering and Information Technology Management, Kharkiv, Ukraine, tel.: (097) 586-46-48; e-mail: artyom4ek@yandex.ua.

Яковлева Олена Володимирівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інформатики; тел.: (057) 702-14-94; e-mail: olena.yakovleva@nure.ua.

Яковлева Елена Владимировна – кандидат технических наук, доцент, Харьковский национальный университет радиоэлектроники, доцент кафедры информатики; тел.: (057) 702-14-94; e-mail: olena.yakovleva@nure.ua.

Yakovleva Olena Vladimirovna – Candidate of Technical Sciences (Ph. D.), Docent, Kharkiv National University of Radioelectronics, Associate Professor at the Department of Informatics; tel.: (057) 702-14-94; e-mail: olena.yakovleva@nure.ua.