



## DEEP MULTILAYER NEURAL NETWORK FOR PREDICTING THE WINNER OF FOOTBALL MATCHES

Sergei Anfilets <sup>1)</sup>, Sergei Bezobrazov <sup>1)</sup>, Vladimir Golovko <sup>1)</sup>, Anatoliy Sachenko<sup>2)</sup>, Myroslav Komar <sup>2)</sup>, Raman Dolny <sup>3)</sup>, Valery Kasyanik <sup>1)</sup>, Pavlo Bykovyy <sup>2)</sup>, Egor Mikhno <sup>1)</sup>, Oleksandr Osolinskyi <sup>2)</sup>

1) Department of Intelligent Information Technology, Brest State Technical University, 267Moskovskaya st., Brest, 224017, Belarus,  
anfilets\_sv@mail.ru, bescase@gmail.com, vladimir.golovko@gmail.com, val.tut@gmail.com, dzinushi.kun@gmail.com

2) Department for Information Computer Systems and Control, Ternopil National Economic University, 11 Lvivska st., Ternopil, 46020, Ukraine,  
as@tneu.edu.ua, mko@tneu.edu.ua, pb@tneu.edu.ua, oso@tneu.edu.ua

3) Allbebet OU,  
26 Majakast., Tallinn, 11411, Estonia  
raman@allbebet.io

### Paper history:

Received 18 August 2019  
Received in revised form 14 February 2020  
Accepted 17 February 2020  
Available online 31 March 2020

### Keywords:

soccer analytics;  
forecasting;  
neural networks;  
Deep Elastic Net;  
open access dataset;  
data preprocessing;  
artificial intelligent.

**Abstract:** In this work, we draw attention to prediction of football (soccer) match winner. We propose the deep multilayer neural network based on elastic net regularization that predicts the winner of the English Premier League football matches. Our main interest is to predict the match result (win, loss or draw). In our experimental study, we prove that using open access limited data such as team shots, shots on target, yellow and red cards, etc. the system has a good prediction accuracy and profitability. The proposed approach should be considered as a basis of Oracle engine for predicting the match outcomes.

*Copyright © Research Institute for Intelligent Computer Systems, 2020. All rights reserved.*

## 1. INTRODUCTION

The appearance and development of modern intelligent technologies promoted many branches of human activity to reach a qualitatively new level as well as achieve previously unthinkable results. A vivid example of the active evolution and integration of up-to-date technologies is modern sport. Today, sport involves more and more people, increases financial, material and intelligent flows and resources. All this enabled the sport to become an important political and economic component of the modern world.

It is no secret that one of the most popular kind of sport – football (soccer), is a multi-billion international market with a very extensive infrastructure. Transfer value of the best players is estimated of tens of millions of dollars. For example, the “Paris Saint-Germain” football club bought a

player Neymar from “Barcelona” for \$260.9 million. According to Forbes rating, the total cost of the ten most expensive football clubs in 2017 is \$23.29 billion.

One of the most attractive sports activities is predicting the outcomes of sports games. This task forms the basis of betting business, and it is extremely difficult due to the unpredictable nature of sporting events, especially football games (matches). There are many determining factors of scoring a goal including the strength of attack and defense, the home ground advantage and others taking place during the match. Along with this, such unpredictable factors as the removal of a player (red card, a penalty, the judge’s rigor and many others can affect the final score. As a result, the task of forecasting the outcomes of sporting events, that can include such problems as big data analysis, data

processing, clustering, classification and etc. [1, 2, 26] is a good playground for testing the various methods and approaches.

## 2. RELATED WORK

Nowadays, many techniques and methods to predict the results of sporting events have been developed. The most popular of them are Bayesian networks [3, 4]; k-nearest neighbor method [5]; support vector machine [6]; stochastic methods for describing uncertainties, such as regression and autoregressive analysis [7-9], Markov chains [10-12], the Monte Carlo method [13], and others.

Artificial neural networks have proven themselves in such tasks as prediction, pattern recognition, classification, control, robotics, etc.

In our opinion, the sport event predictive systems based on artificial neural networks are the most promising. The advantage of such systems is their flexibility, versatility and accuracy of prediction [14-18]. Such systems can be considered as universal approximators of nonlinear dependences. However, for their training and functioning the large volumes of all kinds of statistical data are needed. Nowadays, the largest providers of sports data for football are Wyscout [19] and Opta Sports [20]. While Opta Sports collects and distributes full, time-stamped, contextual data live, featuring complete x/y coordinates (as well as z coordinates where applicable, such as shots in football), and a granularity of event type unique amongst data providers, Wyscout is focused on collecting sports video data.

In this paper, we present the research outcomes in developing the oracle for forecasting the results of football matches. We tested the system on football matches of English Premier League. A proposed approach is based on artificial neural network methods and it uses the limited open access information about football team statistics, such as team shots, shots on target, a number of yellow and

red cards, etc. The remainder of this paper is organized as follows. Section 3 describes the theoretical studies of applied methods. The results of forecasting as well as the comparative analysis are contained in Section 4. Conclusion is given in Section 5.

## 3. METHODS

### 3.1 DATA PREPROCESSING

Many different indicators for description the strengths of the football teams can be used. The selection of indicators that form a football team rating is an important task. It is necessary to select such parameters that have a high degree of information and importance for the description of the team. The most significant parameters are a standing place, the number of points at a chosen time interval, the number of goals scored for a selected time interval, the number of conceded goals, etc. Table 1 shows the parameters that were selected for our system.

**Table 1.** The list of the important parameters.

Abbreviation	Parameter
S	Team Shots
ST	Team Shots on Target
C	Team Corners
F	Team Fouls Committed
CS	Team Yellow Cards and Red Cards
GS	Team Goals Scored
HTGS	Team Goals Scored in first part
GC	Team Goals Conceded
HTGC	Team Goals Conceded in first part

The described parameters can be downloaded from Football-Data.co.uk [22] and include the information about all Premier League games since 2002 (Fig. 1).

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	Home Team	Away Team	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS	AS	HST	AST	HHW	AHW	HC	AC	HF	AF	HO	AO
2	Blackburn	Sunderland	0	0	D	0	0	D	D Ellerav	15	7	5	3	14	11	9	1	1	2	0	0
3	Charlton	Chelsea	2	3	A	2	1	H	G Barber	5	21	5	12	10	12	3	6	0	3	1	0
4	Everton	Tottenham	2	2	D	1	0	H	N Barry	13	10	9	5	18	4	10	5	1	1	0	0
5	Fulham	Bolton	4	1	H	3	1	H	A Wiley	13	3	6	1	16	12	7	4	1	2	0	0
6	Leeds	Man City	3	0	H	2	0	H	G Poll	13	18	8	10	13	13	2	7	1	1	0	0
7	Man United	West Brom	1	0	H	0	0	D	S Bennett	20	6	13	5	9	12	9	1	1	1	0	1
8	Southampton	Middlesboro	0	0	D	0	0	D	B Knight	12	11	5	5	11	14	3	4	0	0	0	0
9	Arsenal	Birmingham	2	0	H	2	0	H	M Riley	15	7	7	1	6	11	9	2	0	1	0	1
10	Aston Villa	Liverpool	0	1	A	0	0	D	A D'Urso	11	12	5	6	9	5	6	8	2	2	0	0
11	Newcastle	West Ham	4	0	H	0	0	D	P Durkin	13	7	10	4	7	9	7	1	1	2	0	0
12	Chelsea	Man United	2	2	D	2	1	H	G Poll	11	9	5	7	13	9	3	3	3	2	0	0
13	Birmingham	Blackburn	0	1	A	0	1	A	D Gallagher	15	13	10	9	10	17	6	4	1	2	0	0
14	Bolton	Charlton	1	2	A	1	1	D	M Messias	15	8	9	4	9	11	4	6	1	1	0	0
15	Liverpool	Southampton	3	0	H	1	0	H	J Winter	16	7	9	1	11	8	5	8	0	1	0	0
16	Man City	Newcastle	1	0	H	1	0	H	U Rennie	20	9	12	6	12	14	6	8	1	1	0	0
17	Middlesboro	Fulham	2	2	D	1	0	H	M Dean	6	10	3	4	18	7	8	11	4	2	0	0
18	Sunderland	Everton	0	1	A	0	1	A	R Styles	13	11	3	5	10	15	10	4	1	0	0	0
19	Tottenham	Aston Villa	1	0	H	1	0	H	C Wilkes	8	18	3	8	20	15	2	5	2	0	0	0
20	West Brom	Leeds	1	3	A	0	1	A	S Dunn	12	11	2	4	9	11	7	3	1	1	0	0

**Figure 1 – The example of statistical data from Football-Data.co.uk**

We are not limited to information only for the last match. We observe last 35 matches and calculate the aggregated statistical indicators. In our opinion, such indicators are more informative and they can be used for estimating the current state of a team.

The aggregation of the input pattern is performed by averaging the each parameter in Table 1 for 35, 15, 10, 5 and 5-10 matches. Fig. 2 shows the

example of aggregating (averaging) the input pattern using the Team Shots indicator.

In addition, we calculated the standard deviation for 35 matches and added it to input pattern. The standard deviation provides information about the team stability and it is very important for the result forecasting.

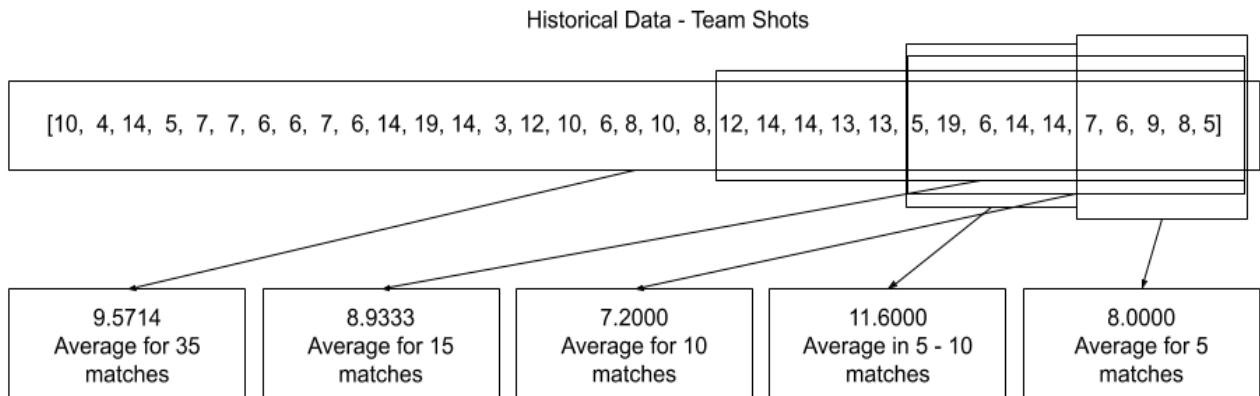


Figure 2—Example of aggregating the input pattern

After the aggregation we get the input vector consisted of 108 values – 54 values for both teams.

Fig. 3 demonstrates the example of an input pattern for one football match.

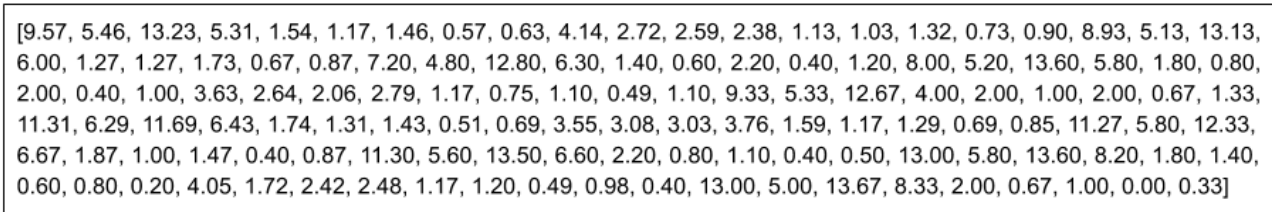


Figure 3 – Example of input pattern

In addition, the input data is normalized according to the following statistical indicators: average value and standard deviation. That allows obtaining a more stable predictive model. The formula for normalization is:

$$x_n = (x - \mu) / \sigma, \tag{1}$$

where  $\mu$  – expected value;  
 $\sigma$  – standard deviation.

### 3.2 DEEP MULTILAYER NETWORK

For forecasting the sport outcomes, a deep multilayer neural network based on elastic net

regularization was proposed. This NN structure is selected because it is capable to produce good predictions using the limited data.

Zou and Hastie [23] considered the Elastic Neural Network as a new method for regularization and variable selection. It produces a sparse model with good prediction accuracy, while encouraging a grouping effect. This method demonstrates good results, especially when the number  $p$  of predictors is much larger than the number of observations  $n$ .

In our study, we used the multilayer neural network (Fig. 4).

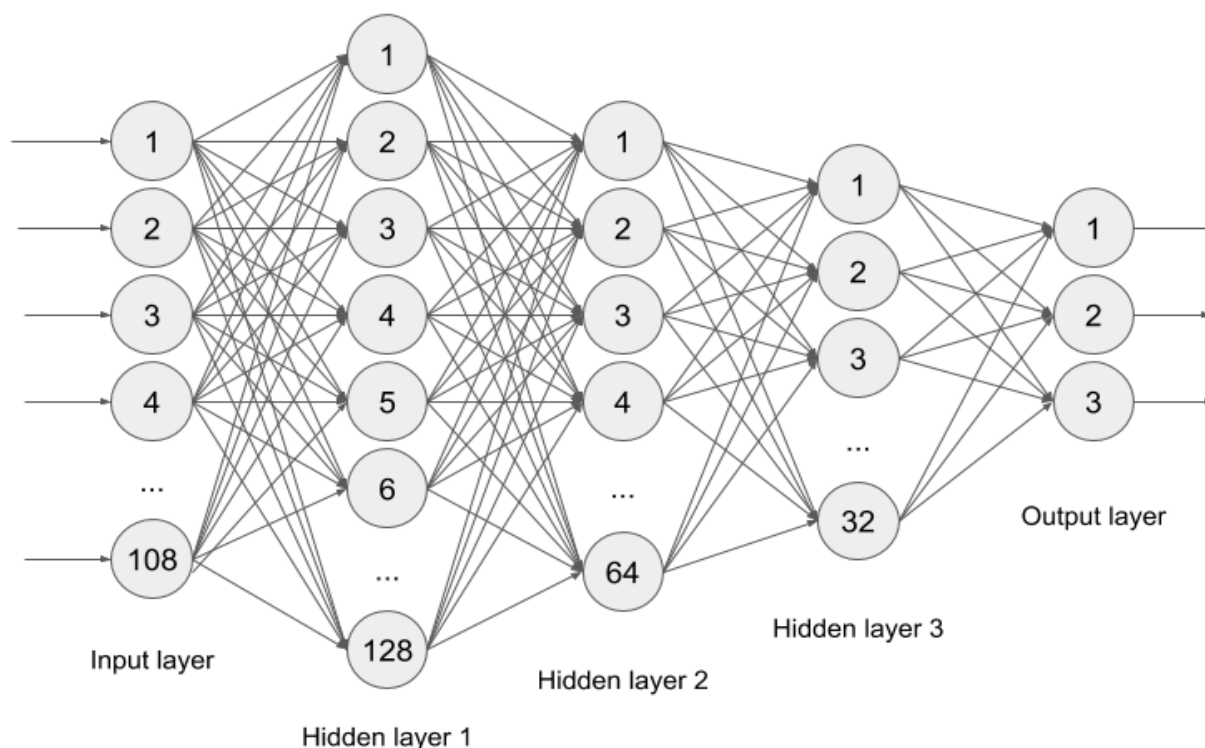


Figure 4 – The structure of developed Deep Elastic Net for forecasting match results

The input layer contains 108 neurons. The dimension of the input pattern determines the number of input neurons. The net has three hidden layers with 128, 64 and 32 neurons in each layer respectively and 3 neurons in the output layer. Such network contains more than two hidden layers and therefore it is called deep multilayer neural network.

A prepared pattern (see Fig. 3) enters to the network input. Further, three hidden layers with *LeakyReLU* activation functions perform calculations on the input pattern. Three output neurons reflect the results of the calculations, interpreted in the next form: win – draw – loss. The first output neuron is responsible for the victory of the home team, the second neuron – for a draw in the match, and the third neuron – for the victory of the guest team, respectively. The *softmax* activation function for the neurons of the output layer is used.

A distinctive feature of Elastic Net is that it uses L1, L2 regularization. While the L1 regularization (also known as Lasso Regression) employs to select parameters, the L2 regularization (also known as Ridge Regression) performs a network overfitting control (overfitting means the growth of model coefficients) in the learning process.

#### 4. CASE STUDY







To test the system that was developed for the forecasting the results of English Premier League football matches, a data set consisting of 5018 patterns was used (i.e., the set contains the information about all games since 2002). To make the early termination of learning process we worked with the validation set with the size of 15% of training set. For the learning the Deep multilayer neural network, the following parameters were selected:

- Coefficient for L1 regularization = 0.002;
- Coefficient for L2 regularization = 0.0005;
- Algorithm of learning – SGD (Stochastic Gradient Descent) with a step equal to 0.01;
- Minibatch size equals 64.

The training process takes approximately 5 minutes on the following PC configuration: GPU NVidia 1070TI, CPU Xeon e5-2680 v2, RAM 32 GB.

The trained system was tested on the last 350 Premier League matches that were not included in training and validation sets. Table 2 demonstrates an example of the predicted outputs of the trained system.

**Table 2.** The example of forecasting results.

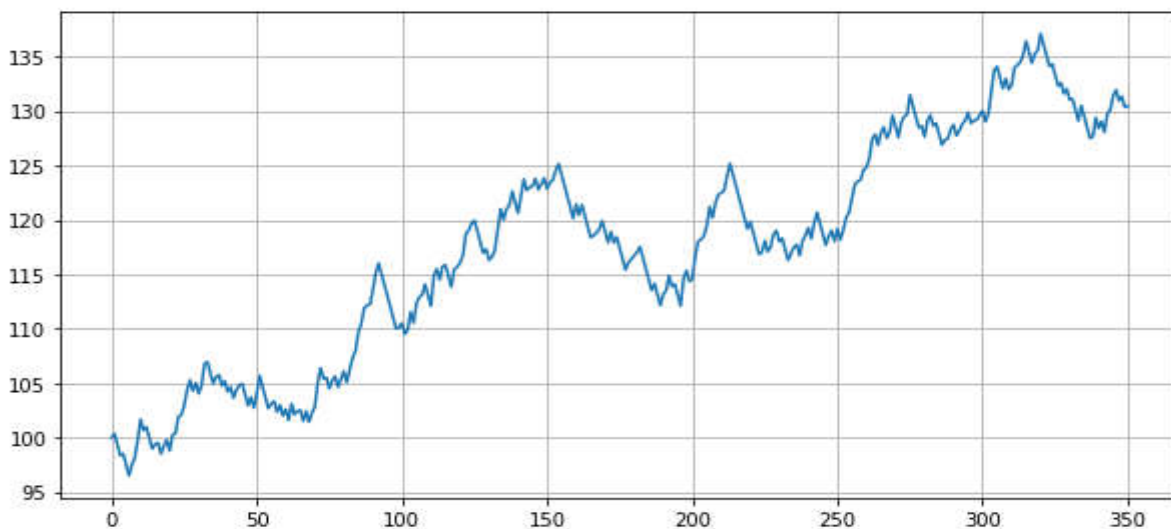
Teams and final score	Parameter
West Ham vs Southampton (score 3:0)	
Huddersfield vs Man United (score 1:1)	
Leicester vs Arsenal (score 3:0)	
Burnley vs Man City (score 0:1)	
Fulham vs Cardiff (score 1:0)	
Wolves vs Fulham (score 1:0)	

The proposed system confirmed the prediction accuracy of 61.14% on the test dataset. This accuracy is calculated by dividing the number of correct predictions by the total matches.

We used as test data set last 250 matches of English Premier League. As it can be seen from the above, the proposed approach is characterized by the higher accuracy and low computational costs in comparison with existing approaches. So, for instance in [24] the deep sparse auto encoder for football match prediction has the accuracy of 51.4%. In [25] the comparisons of various extensions for Bradley–Terry models were made and a hierarchical log-linear Poisson model for predicting the outcomes of soccer matches was built. The best performing hierarchical log-linear Poisson model

showed the 54% accuracy. We employed as test data set the last 250 matches of English Premier League. Using the prediction results of the developed system, users can bet on this or that team in the upcoming match.

Fig. 5 represents the profitability of the sports bets using the outputs of the developed system. The value, which is equal to 100 on a y-axis, means the starting amount in soccer fan account (for instance, 100 EUR). As it can be seen in Fig. 5, the profit reaches 30.47 points (by bet365 ratio) that is a good result. In spite of the fact that the developed system is capable of making a profit, some sections with relative large drawdowns can be seen in the chart. Thus, a trend from 152-th till 190-th matches is downward.



**Figure 5** –Profitability of the sports bets (x-axis – matches; y-axis – profit)

## 5. CONCLUSION

The approach to predict the results of sports competitions, which is based on a Deep Multilayer Neural Network with elastic net regularization and a possibility to be trained on a limited open access dataset, is proposed. Instead of using raw data, authors selected the most valuable parameters and designed the data preprocessing procedure that allows building a representative team conditions and forming the input patterns. The developed system proved good prediction results and it is able to make a profit. So, it can be used as a basis for Oracle to predict the results of sporting events. The main advantage of proposed approach is a simplicity as well as low computer complexity.

The system can be improved by using the specified dataset and providing the paid resources and developing a more complex neural network architecture. For instance, further we are going to employ the deep autoencoder for preliminary data processing and the deep multilayer perceptron for forecasting.

This research starts our ambitious project to develop and implement the advanced artificial intelligent approaches for professional sport.

## 5. ACKNOWLEDGMENTS

We thank the Allbebet OU staff who provided insight and expertise that greatly assisted this research.

## 6. REFERENCES

- [1] V. Golovko, S. Bezobrazov, A. Kroshchanka, A. Sachenko, M. Komar, A. Karachka, "Convolutional neural network based solar photovoltaic panel detection in satellite photos," *Proceedings of the 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications IDAACS'2017*, Bucharest, Romania, September 21-23, 2017, pp. 14-19.
- [2] M. Komar, V. Golovko, A. Sachenko, S. Bezobrazov, "Intelligent system for detection of networking intrusion," *Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS'2011*, Prague, Czech Republic, 2011, pp. 374-377.
- [3] A. Constatinou, N. Fenton, M. Neil, "Profiting from an inefficient Association Football gambling market: Prediction, risk and uncertainty using Bayesian networks," *Knowledge-Based Systems*, no. 50, pp. 60–86, 2013.
- [4] B. Zadrozny, C. Elkan, "Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers," *Proceedings of the Eighteenth International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc. San Francisco, 2001, pp. 609–616.
- [5] V. Hoekstra, *Predicting Football Results with an Evolutionary Ensemble Classifier*, Master Thesis, VU University Amsterdam, 2012.
- [6] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods." *Advances in Large Margin Classifiers*, vol. 10, issue 3, pp. 61-74, 1999.
- [7] S. Clarke, M. Bailey, S. Yelas, "Successful applications of statistical modeling to betting markets," *Mathematics Today*, vol. 44, issue 1, pp. 38-44, 2008.
- [8] D. Creal, S. Koopman, A. Lucas, "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, vol. 28, issue 5, pp. 777–795, 2013.
- [9] J. Goddard, "Regression models for forecasting goals and match results in association football," *International Journal of Forecasting*, vol. 21, pp. 331–340, 2005.
- [10] Modeling NFL overtime as a Markov chain. [Online]. Available at: <http://harvardsportsanalysis.org/2014/01/modeling-nfl-overtime-as-a-markov-chain/>, last accessed 2019/06/16.
- [11] N. Hirotsu, M. Wright, "Using a Markov process model of an association football match to determine the optimal timing of substitution and tactical decisions," *Journal of the Operational Research Society*, vol. 53, issue 1, pp. 88–96, 2002.
- [12] M. Pfeiffer, H. Zhang, A. Hohmann, "A Markov chain model of elite table tennis competition," *International Journal of Sports Science and Coaching*, vol. 5, issue 2, pp. 205–222, 2010.
- [13] H. Rue, O. Salvesen, "Prediction and retrospective analysis of soccer matches in a league," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 49, issue 3, pp. 399–418, 2000.
- [14] A. Arabzad, M. Araghi, S. Soheil, "Football match results prediction using artificial neural networks: the case of Iran pro league," *Int. J. Appl. Res. Ind. Eng.*, vol. 1, issue 3, pp. 159–179, 2014.
- [15] E. Davoodi, A. Khanteymoori, "Horse racing prediction using artificial neural networks," *Proceedings of the 11th WSEAS Int. Conf. on*

Neural Networks, NN'10, Proceedings of the 11th WSEAS Int. Conf. on Evolutionary Computing, EC'10, Proc. of the 11th WSEAS Int. Conf. on Fuzzy Systems, FS'10, Iasi, Romania; 13-15 June 2010, pp. 155-160.

- [16] B. Ulmer, M. Fernandez, *Predicting Soccer Match Results in the English Premier League*, 2014. [Online]. Available at: [https://pdfs.semanticscholar.org/d4df/bb3d1454416799b72004d3d88bae6bbd660a.pdf?\\_ga=2.134687474.2044286076.1584428703-696523042.1563168651](https://pdfs.semanticscholar.org/d4df/bb3d1454416799b72004d3d88bae6bbd660a.pdf?_ga=2.134687474.2044286076.1584428703-696523042.1563168651)
- [17] A. McCabe, J. Trevathan, "Artificial intelligence in sports prediction," *Proceedings of the IEEE Fifth International Conference on Information Technology: New Generations, ITNG'2008*, 2008, pp. 1194–1197.
- [18] N. Tax, Y. Joustra, "Predicting the Dutch football competition using public data: A machine learning approach," *Trans. Knowl. Data Eng.*, vol. 10, issue 10, pp. 1–13, 2015.
- [19] Wyscout. [Online]. Available at: <https://wyscout.com/>, last accessed 2019/06/16.
- [20] Opta Sports. [Online]. Available at: <https://www.optasports.com>, last accessed 2019/06/16.
- [21] Allbebet OU. [Online]. Available at: <https://allbebet.com>, last accessed 2019/06/16.
- [22] H. Zou, T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society B.*, vol. 67, issue 2, pp. 301–320, 2005.
- [23] Football-Data. [Online]. Available at: <http://www.football-data.co.uk>, last accessed 2019/06/16.
- [24] K. Tsampazis, A. Tefas, "Deep sparse autoencoders for football match and bet prediction," *Proceedings of the 10th Hellenic Conference on Artificial Intelligence SETN'18*, July 2018, article no.: 4, pp. 1–8.
- [25] A. Tsokos, S. Narayanan, I. Kosmidis, G. Baio, M. Cucuringu, G. Whitaker, F. Király, "Modeling outcomes of soccer matches," *Machine Learning*, no. 108, pp. 77–95, 2019.
- [26] T. Maksymyuk, J. Gazda, O. Yaremko and D. Nevinskiy, "Deep learning based massive MIMO beamforming for 5G mobile network," *Proceedings of the 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS-2018)*, Lviv, 2018, pp. 241-244.



**Prof. Vladimir Golovko** received M.E. degree in Computer Engineering in 1984 from the Moscow Bauman State Technical University. In 1990 he received PhD degree from the Belarus State Technical University and in 2003 he received doctor science degree in Computer Science from the United Institute of Informatics problems national Academy of Sciences (Belarus). At present he works as a head of Intelligence Information Technologies Department and Laboratory of Artificial Neural Networks of the Brest State Technical University.

His research interests include Artificial Intelligence, neural networks, autonomous learning robot, signal processing, chaotic processes, intrusion and epilepsy detection. He has published more than 300 scientific papers.



**Sergei Bezobrazov, PhD.**, an Associate Professor at the Intelligent Information Technologies Department, Brest State Technical University. Scientific interests include artificial intelligence, machine learning, information security. He earned his B.Eng. Degree in Computer Science at Brest State Technical University in 2001 and his PhD Degree in Information Security at Belarus State University in 2006.

His main areas of research interest include artificial intelligence, machine learning, information security. He has published over 40 scientific papers.



**Anatoliy Sachenko** is Professor and Head of Dept for Information Computer Systems and Control, and Principal Investigator of the Research Institute for Intelligent Computer Systems, Ternopil National Economic University. He earned his PhD Degree in Electrical Engineering at L'viv Physics and Mechanics Institute, Ukrainian National Academy of Science, in 1978 and his Doctor of Technical Sciences Degree in Electrical and Computer Engineering at Leningrad Electrotechnic Institute in 1988.

His main Areas of Research Interest are Computational Intelligence in Applications, Distributed Measuring Systems and Networks, Intelligent Cyber Security, Wireless Sensor Networks, IT Project Management.



**Dr. Myroslav Komar** graduated from Ternopil National Economic University in 2001 with speciality "Information Systems in Management", 2002 – Master in Economic Cybernetics, 2013 – PhD degree in Information Technology.

He is an Assistant Professor of Dept for Information Computer Systems and Control, and Scientific Secretary of Research Institute for Intelligent Computer Systems.

Areas of scientific interests include Artificial Intelligence, Neural Networks and Immune System, Evolution of Systems, Cybersecurity.



**Dr. Oleksandr Osolinskyi**, Bachelor (2004), Computer Engineering, Ternopil Academy of National Economy, Specialist (2005), Computer Systems and Networks, Ternopil Academy of National Economy, Junior Scientist (2005), Research Institute for

Intelligent Computer Systems, Ph. D. in Computer Systems and Components (2016). A Senior lecturer at Ternopil National Economic University.

Research interests: software development, web-design, distributed systems, computer systems architectures.



**Dr. Pavlo Bykovyy**, Bachelor (2004), Computer Engineering, Ternopil Academy of National Economy, Specialist (2005), Computer Systems and Networks, IEEE Member (2004), IEEE Student Branch Chairman in Ternopil State Economical University (2005),

Ph. D. in Computer Systems and Components (2011). Research interests: security systems, databases, software development.