Deep Elastic Net for Prediction the Winner of Football Matches

Sergei Anfilets[1[0000-0001-6478-3524]], Sergei Bezobrazov[2[0000-0001-6436-2922]], Vladimir Golovko[3[0000-0003-2615-289X]], Raman Dolny[4[0000-0003-0437-0896]], Valery Kasyanik[5[0000-0002-4369-516X]], Aleksandr Melushev[6[0000-0002-2622-1157]], Egor Mikhno[7[0000-0002-7667-7486]], Dmitry Yakimovich[8[0000-0001-6974-847X]]

[1,2,3,5,7] Brest State Technical University, Moskovskaya 267, 224017, Brest, Belarus
[1] anfilets_sv@mail.ru, [2] bescase@gmail.com, [3] vladimir.golovko@gmail.com, [5] val.tut@gmail.com, [7] dzinushi.kun@gmail.com
[4,6,8] Allbebet OU, Majaka 26, 11411 Tallinn, Estonia
[4] raman@allbebet.io, [6] cto@allbebet.io, [8] dmitry@allbebet.io

Abstract. In this work, we draw attention to prediction of football (soccer) match winner. We develop a system based on Deep Elastic Net that predict the winner of the English Premier League football matches. Our main interest is to forecast whether the match result is a win, a loss or a draw for each team. In our experimental study we show that using open access limited data such as team shots, shots on target, yellow and red cards etc. the system have a good prediction accuracy and profitability. The system was developed for Allbebet OU Company and forming the basis of the engine of oracle for prediction the matches outcomes.

Keywords: soccer analytics, forecasting, neural networks.

1 Introduction

The appearance and development of modern intellectual technologies has allowed many branches of human activity to reach a qualitatively new level and 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 intellectual flows and resources. All this allowed 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 (or, soccer), is a multi-billion international market with a very extensive infrastructure. The cost of transferring of the best players consists of the tens of millions of dollars. For example, the “Paris Saint-Germaine” 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 important and interesting components in sport is the prediction of the outcomes of sports events. This task, which forms the basis of the betting business, and is extremely difficult due to the unpredictable nature of sporting events. 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 predicting the outcomes of sporting events, that can include such problems as big data analysis, data processing, clustering, classification etc. is a good testing playground for testing various methods and approaches.

1.1 Related work

To date, many techniques and methods to predict the results of sporting events was developed. The most popular of them are Bayesian networks [1, 2]; k-nearest neighbor method [3]; support vector machine [4]; stochastic methods for describing uncertainties, such as regression and autoregressive analysis [5-7], Markov chains [8-10], the Monte Carlo method [11] 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 [12-16]. Such systems can be considered as universal approximators of nonlinear dependences. However, for their training and functioning large volumes of all kinds of statistical data are necessary. Today, the largest providers of sports data for football are Wyscout [17] and Opta Sports [18]. 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 our research in the development of the oracle for prediction the results of football matches. We tested the system on English Premier League football matches. The system based on artificial neural network methods and use limited open access information about football team statistics, such as team shots, shots on target, yellow and red cards etc. The remainder of this paper is organized as follows. Section 2 describes the theoretical studies of the applied methods. The results of forecasting and comparative analysis are contained in the Section 3. Conclusion finishing the paper.

2 Methods

2.1 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 choose such parameters that have a high degree of information and importance for the description of a 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 chosen time interval, the number of goals conceded etc. Table 1 shows the parameters that was 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 [20] and include the information about of all Premier League games since 2002. Fig. 1 shows the detailed information that provides this resource.

HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS	AS	HST	AST	HHW	AHW	HC	AC	HF	AF	HO	AO
Blackburn	Sunderland	0	0	D	0	0	D	D Elleray	15	7	5	3	14	11	9	1	1	2	0	0
Charlton	Chelsea	2	3	A	2	1	H	G Barber	5	21	5	12	10	12	3	6	0	3	1	0
Everton	Tottenham	2	2	D	1	0	H	N Barry	13	10	9	5	18	4	10	5	1	1	0	0
Fulham	Bolton	4	1	H	3	1	H	A Wiley	13	3	6	1	16	12	7	4	1	2	0	0
Leeds	Man City	3	0	H	2	0	H	G Poll	13	18	8	10	13	13	2	7	1	1	0	0
Man United	West Brom	1	0	H	0	0	D	S Bennett	20	6	13	5	9	12	9	1	1	1	0	1
Southampton	Middlesboro	0	0	D	0	0	D	B Knight	12	11	5	5	11	14	3	4	0	0	0	0
Arsenal	Birmingham	2	0	H	2	0	H	M Riley	15	7	7	1	6	11	9	2	0	1	0	1
Aston Villa	Liverpool	0	1	A	0	0	D	A D'Urso	11	12	5	6	9	5	6	8	2	2	0	0
Newcastle	West Ham	4	0	H	0	0	D	P Durkin	13	7	10	4	7	9	7	1	1	2	0	0
Chelsea	Man United	2	2	D	2	1	H	G Poll	11	9	5	7	13	9	3	3	3	2	0	0
Birmingham	Blackburn	0	1	A	0	1	A	D Gallagher	15	13	10	9	10	17	6	4	1	2	0	0
Bolton	Charlton	1	2	A	1	1	D	M Messias	15	8	9	4	9	11	4	6	1	1	0	0
Liverpool	Southampton	3	0	H	1	0	H	J Winter	16	7	9	1	11	8	5	8	0	1	0	0
Man City	Newcastle	1	0	H	1	0	H	U Rennie	20	9	12	6	12	14	6	8	1	1	0	0
Middlesboro	Fulham	2	2	D	1	0	H	M Dean	6	10	3	4	18	7	8	11	4	2	0	0
Sunderland	Everton	0	1	A	0	1	A	R Styles	13	11	3	5	10	15	10	4	1	0	0	0
Tottenham	Aston Villa	1	0	H	1	0	H	C Wilkes	8	18	3	8	20	15	2	5	2	0	0	0
West Brom	Leeds	1	3	A	0	1	A	S Dunn	12	11	2	4	9	11	7	3	1	1	0	0

Fig. 1. The example of the statistical data from Football-Data.co.uk

We do not use the information about only the last match, because it is not enough for constructing the full information about conditions of the teams. Instead of this, we observe last 35 matches and calculates the aggregated statistical indicators. In our opinion, such indicators are more informative and can be used for the estimation of current condition of a team.

The aggregation of the input pattern is performed by averaging of the each parameter from Table 1 for 35, 15, 10, 5 and 5-10 matches. Fig. 2 shows the example of the aggregation (averaging) of the input pattern using Team Shots indicator.

In addition, we calculate the standard deviation for 35 matches and add it to the input pattern. The standard deviation give us the information about the team stability and very important for the result forecasting.

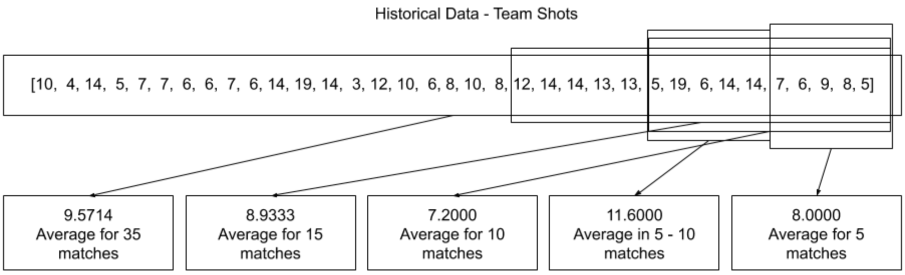


Fig. 2. Example of aggregation of an input pattern

After the aggregation we receive 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.

[9.57, 5.46, 13.23, 5.31, 1.54, 1.17, 1.46, 0.57, 0.63, 4.14, 2.72, 2.59, 2.38, 1.13, 1.03, 1.32, 0.73, 0.90, 8.93, 5.13, 13.13, 6.00, 1.27, 1.27, 1.73, 0.67, 0.87, 7.20, 4.80, 12.80, 6.30, 1.40, 0.60, 2.20, 0.40, 1.20, 8.00, 5.20, 13.60, 5.80, 1.80, 0.80, 2.00, 0.40, 1.00, 3.63, 2.64, 2.06, 2.79, 1.17, 0.75, 1.10, 0.49, 1.10, 9.33, 5.33, 12.67, 4.00, 2.00, 1.00, 2.00, 0.67, 1.33, 11.31, 6.29, 11.69, 6.43, 1.74, 1.31, 1.43, 0.51, 0.69, 3.55, 3.08, 3.03, 3.76, 1.59, 1.17, 1.29, 0.69, 0.85, 11.27, 5.80, 12.33, 6.67, 1.87, 1.00, 1.47, 0.40, 0.87, 11.30, 5.60, 13.50, 6.60, 2.20, 0.80, 1.10, 0.40, 0.50, 13.00, 5.80, 13.60, 8.20, 1.80, 1.40, 0.60, 0.80, 0.20, 4.05, 1.72, 2.42, 2.48, 1.17, 1.20, 0.49, 0.98, 0.40, 13.00, 5.00, 13.67, 8.33, 2.00, 0.67, 1.00, 0.00, 0.33]

Fig. 3. Example of an 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 μ – expected value;
σ – standard deviation.

2.2 Deep Elastic Net

For the forecasting the sport outcomes a Deep Elastic Net was used. The choice of this structure of the neural network is because it able to produce good predictions using the limited data.

The Elastic Neural Network was proposed by Zou and Hastie [21] 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 demonstrated good results, especially when the number of predictors p is much larger than the number of observations n.

In our study, we use an Elastic Net with the architecture shown on Fig.4.

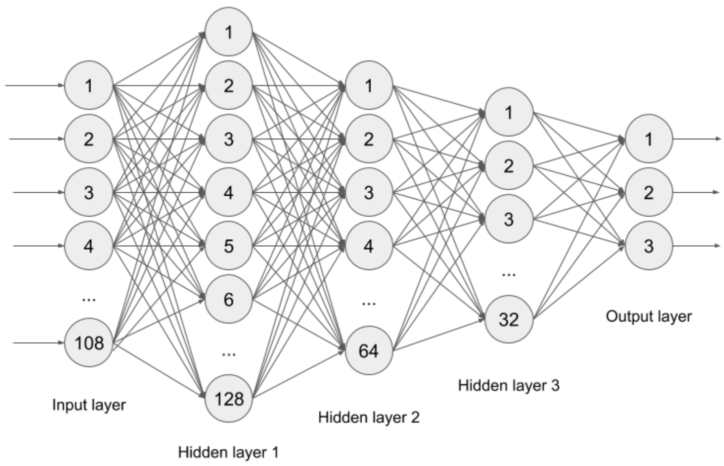


Fig. 4. The structure of the developed Deep Elastic Net for matches results forecasting

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 that contains several hidden layers is called as Deep Elastic Net.

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 L1 regularization (also known as Lasso Regression) is used to select parameters, L2 regularization (also known as Ridge Regression) performs a network overfitting control (overfitting means the growth of model coefficients) at the learning process.

3 Results

To test the system that was developed for the forecasting of 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). For early termination of the learning process, the validation set with the size of 15% of the training set was used. For the learning of Deep Elastic Net, the following parameters were selected:

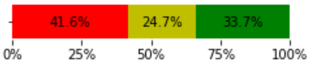
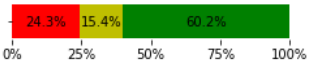
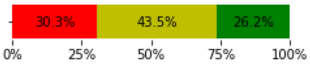
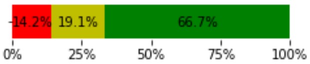
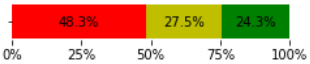
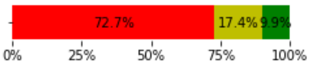
- Coefficient for L1 regularization = 0.002;
- Coefficient for L2 regularization = 0.0005;
- Algorithm of learning - SGD (Stochastic Gradient Descent) with a step equal to 0.01;

- Minibatch size equals to 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 first number that is placed on the red part represents the probability of a winning a home team. The olive part displays the probability of a draw in the match. The green part shows the probability of a winning an away team. The system showed 61.14% of prediction accuracy on the test dataset. 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 equals to 100 on a y-axis means the start sum on your account (for instance, 100 EUR). As can be seen the profit reach 30.47 points (by bet365 ratio) what is a good result.

In spite of the fact that the developed system is capable of making a profit, you can see on the chart some sections with relative large drawdowns. Thus trend from 152-th till 190-th matches is downward.

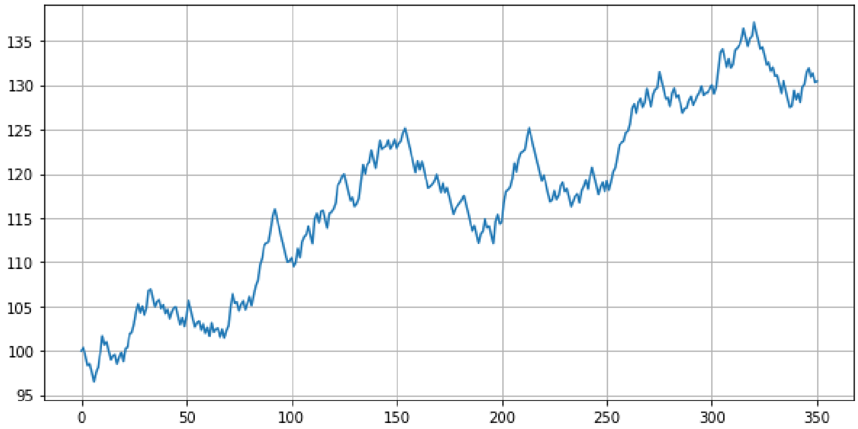


Fig. 5. Profitability of the sports bets (x-axis – matches; y-axis - profit)

4 Conclusion

The paper proposed an approach to predict the results of sports competitions. The proposed system is based on a Deep Elastic Net and can be trained on a limited open access dataset. Instead of using raw data, we chose the most valuable parameters and proposed the data preprocessing procedure that allows building a representative team conditions and forming the input patterns. The developed system was used as a basis for Oracle to predict the results of sporting events showed good prediction results and able to make a profit. The system can be improved by using a more detailed and complete dataset that can be provided by paid resources and by developing a more complex neural network structure.

This research starts our ambitious project to create and implement artificial intelligent approaches for professional sports.

5 Acknowledgments

This research was supported by Allbebet OU. We thank our colleagues from Allbebet OU who provided insight and expertise that greatly assisted the research.

References

1. Constatinou, A., Fenton, N., Neil, M.: Profiting from an inefficient Association Football gambling market: Prediction, Risk and Uncertainty using Bayesian networks. Knowledge-Based Systems (50), 60–86 (2013).

2. Zadrozny, B., Elkan, C.: Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In: Eighteenth International Conference on Machine Learning, pp. 609–616. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA (2001).
3. Hoekstra, V., Bison, P., Eiben, G.: Predicting football results with an evolutionary ensemble classifier (2012).
4. Platt, J.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers 10(3), 61-74 (1999).
5. Clarke, S., Bailey, M., Yelas, S.: Successful applications of statistical modeling to betting markets. Mathematics Today, 44(1), 38-44 (2008).
6. Creal, D., Koopman, S., Lucas, A.: Generalized autoregressive score models with applications. Journal of Applied Econometrics 28(5), 777–795 (2013).
7. Goddard, J.: Regression models for forecasting goals and match results in association football. International Journal of Forecasting 21: 331–340 (2005).
8. Modeling NFL overtime as a Markov chain, http://harvardsportsanalysis.org/2014/01/modeling-nfl-overtime-as-a-markov-chain/, last accessed 2019/06/16.
9. Hirotsu, N., Wright, M.: 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 53(1), 88–96 (2002).
10. Pfeiffer, M., Zhang, H., Hohmann, A.: A Markov chain model of elite table tennis competition. International Journal of Sports Science and Coaching 5(2), 205–222 (2010).
11. Rue, H., Salvesen, O.: Prediction and retrospective analysis of soccer matches in a league. Journal of the Royal Statistical Society. Series D (The Statistician) 49(3), 399–418 (2000).
12. Arabzad, A., Araghi, M., Soheil, S.: Football match results prediction using artificial neural networks: the case of Iran pro league. Int. J. Appl. Res. Ind. Eng., 1(3), 159–179 (2014).
13. Davoodi, E., Khanteymoori, A.: Horse racing prediction using artificial neural networks. Recent Adv. Neural Networks, Fuzzy Syst. Evol. Comput., 155-160 (2010).
14. Fernandez, M., Ulmer, B.: Predicting Soccer Match Results in the English Premier League, (2014).
15. McCabe, A., Trevathan, J.: Artificial intelligence in sports prediction. In: Information Technology: New Generations, ITNG 2008. Fifth International Conference on, IEEE, 1194–1197 (2008).
16. Tax, N., Joustra, Y.: Predicting the Dutch football competition using public data: A machine learning approach. Trans. Knowl. Data Eng., 10(10), 1–13 (2015).
17. Wyscout, https://wyscout.com/, last accessed 2019/06/16.
18. Opta Sports, https://www.optasports.com, last accessed 2019/06/16.
19. Allbebet OU, https://allbebet.com, last accessed 2019/06/16.
20. Football-Data, http://www.football-data.co.uk, last accessed 2019/06/16.
21. Zou, H., Hastie, T.: Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society B. 67(2), 301–320 (2005).