

DESIGN OF SUCCESS PREDICTION MODEL USING ARTIFICIAL NEURAL NETWORK METHODS

ALİ KELEŞ

¹Selcuk University, Karapınar Aydoğanlar Vocational School, Konya, Turkey
E-mail: ali.keles@selcuk.edu.tr

Abstract: In this study, it is aimed to model the prediction of sporting success by using artificial neural network models by using some in-match performance data and revenues of the leading sports clubs of the football branch, which has the largest share of the sports industry. In this study, the performance statistics of sports clubs with high revenues in the European continent are analyzed with the use of artificial neural network algorithms in relation to their success in recent years. With the artificial neural network model design developed, predictions of the success of sports clubs in recent years were made with the variables analyzed. With the performance data, the accuracy percentages of the machine learning model trained with the algorithms were found. Success prediction analysis with machine learning methods was performed with MATLAB program.

Keywords - Artificial Neural Networks, Football Industry, Statistics, Machine Learning, Sports Success.

I. INTRODUCTION

With the introduction of the computer into every day and business life, the concept of data has become widespread with the collection of simple data and its use in routine operations[1]. The word data is derived from the Latin word "datum" meaning "to give". Raw data comes in shape, number, and image-like formats. The concept of data mining has become used in the academic field, business field and medical field in the 1990s[2]. In the field of data mining, there are studies in different areas such as statistics, artificial neural networks and data visualization[3].

Deep learning is a popular area of research that is rapidly advancing and delivering impactful results. Unlike neural networks, which in the early days consisted of fewer than five layers, deep learning has many layers[4]. Machine learning strategy: With the rapid spread of information over the internet and the creation of large databases, advances in technology involving the processing, memory, storage, and richness of data have produced new strategies. In recent years, there have been advances in the field of medicine and health sciences[5]. While MATLAB and R coding programs were not used for the use of data mining algorithms in the early years, these algorithms have been given priority in recent years. In the creation of neural networks, artificial neural networks with the MATLAB program provide a flexible and useful environment[6, 7].

With the proliferation of data collected in the field of sports, different methods have been tried to be developed to extract data with high importance value. While more studies are being done abroad, there are not many studies on data mining in the field of sports in our country[8]. The work of researchers working in the field of data mining in the field of sports has reached remarkable points. Data mining techniques and examples of work in the field of sports have been found with more economical intervals in research such

as talent prediction, muscle movement structures, performance measurements and strategy analysis in competition matches[9]. In the field of sports, the formation of teams and the selection of players is a valuable process. When selecting athletes, it is important to transfer their emotional states, behaviors, attitudes, fitness, motoric characteristics, abilities, training, and competition data to the information environment. With the data provided by talent screening tests, people can be directed to sports using data mining rather than random methods[8].

Football is played worldwide by men and women from children to adults at different levels of participation. Professional soccer depends on multiple factors, including physical, biological, and psychological parameters. With the advantages of prediction models, it can be decided which of the variables involved in determining football performance is a better selector with prediction models[10]. It can be said (Wilson, 1995) that neural networks can better predict sporting performance (end-of-season ranking of teams) regardless of the biases of sports experts[11].

The aim of this study is to predict the French Ligue 1 and English Premier League end-of-season team rankings for the 2022-2023 season according to the goals scored and conceded with the artificial neural network model developed according to 2 different input variables.

II. METHOD

2.1. Research Model

In this study, the end-of-season league rankings of the teams that finished the season in the top 5 in three different leagues in the European continent at the end of the 2022-2023 season were predicted by machine learning methods. Matlab software was used for the model.

2.2. Data Set

Within the scope of the study, the variables of the number of goals scored and conceded by the teams that finished the 2022-2023 season in the top 5 places in the English Premier League, Italian Serie A and French Ligue 1 leagues were used.

2.3. Neural Network Variables

A total of 38 matches are played in these leagues during a season. In the artificial neural network model, goals scored, and goals conceded were entered as input variables. End-of-season league standings were used as the output variable. Normalization is done for the input variables and the data is made ready for the model[12].

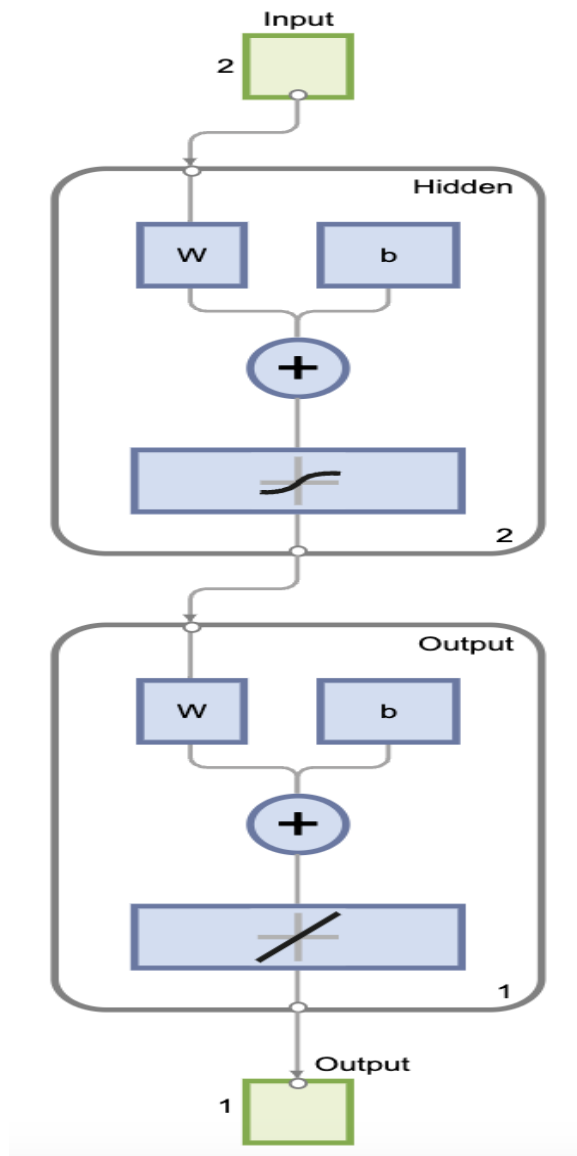


Figure 1. Artificial neural network model.

The sequence of the neural network model is modeled as two input variables and one output variable in two layers from left to right.

III. RESULTS

This section includes the prediction values obtained from the models, performance indicator values of the models and normalization values of the data.

	Club	G	CG	GN	CGN
1	Manchester City	94	33	1,00	0,00
2	Arsenal	88	43	0,90	0,22
3	Manchester United	58	43	0,43	0,22
4	Newcastle	68	33	0,59	0,00
5	Liverpool	75	47	0,70	0,31
6	Brighton	72	53	0,65	0,44
7	Aston Villa	51	46	0,32	0,29
8	Tottenham	70	63	0,62	0,67
9	Brentford	58	46	0,43	0,29
10	Fulham	55	53	0,38	0,44
11	Crystal Palace	40	49	0,14	0,36
12	Chelsea	38	47	0,11	0,31
13	Wolverhampton	31	58	0,00	0,56
14	West Ham	42	55	0,17	0,49
15	Bournemouth	37	71	0,10	0,84
16	Not. Forest	38	68	0,11	0,78
17	Everton	34	57	0,05	0,53
18	Leicester	51	68	0,32	0,78
19	Leeds United	48	78	0,27	1,00
20	Southampton	36	73	0,08	0,89

Table 1. England Premier League 2022-2023 end of season standings, goals scored, goals conceded and normalization values.

G: Goal, CG:Conceded Goals, GN:Goals ScoredNormalization, CGN:Conceded Goal Normalization.

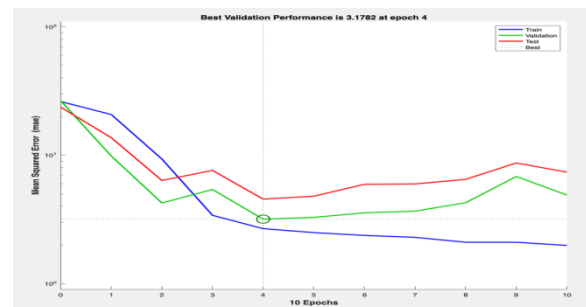


Figure 2. Performance Indicators Values of the Model developed for the Premier League.

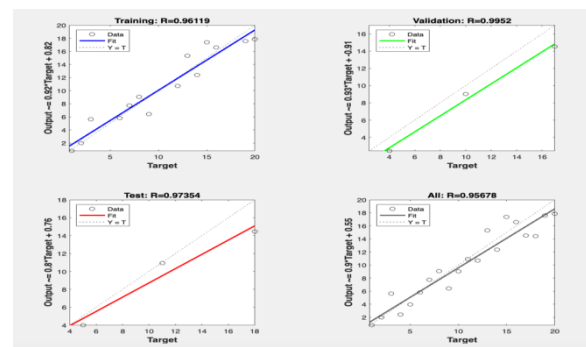


Figure 3. Training, Validation, Testing and All Values of the Model Developed for the Premier League

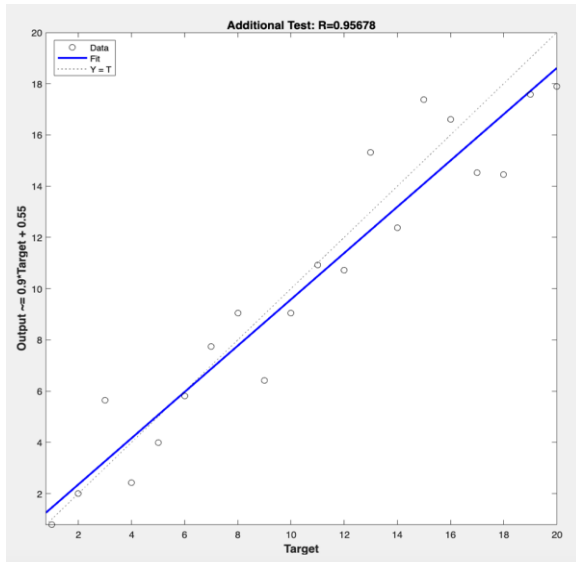


Figure 4. Additional Test Values of the Model Developed for the Premier League.

	Club	G	CG	GN	CGN
1	Paris S.G.	89	40	1,00	0,21
2	Lens	68	29	0,68	0,00
3	Marseille	67	40	0,67	0,21
4	Rennes	69	39	0,70	0,19
5	Lille	65	44	0,64	0,29
6	Monaco	70	58	0,71	0,56
7	Lyon	65	47	0,64	0,35
8	Clermont	45	49	0,33	0,38
9	Nice	48	37	0,38	0,15
10	Lorient	52	53	0,44	0,46
11	Reims	45	45	0,33	0,31
12	Montpellier	65	62	0,64	0,63
13	Toulouse	51	57	0,42	0,54
14	Brest	54	54	0,47	0,48
15	Rc Strasbourg	59	59	0,55	0,58
16	Nantes	55	55	0,48	0,50
17	Auxerre	35	63	0,18	0,65
18	Ajaccio	23	74	0,00	0,87
19	Troyes	45	81	0,33	1,00
20	Angers	33	81	0,15	1,00

Table 2. France Ligue 1 2022-2023 end of season standings, goals scored, goals conceded and normalization values.

G: Goal, CG: Conceded Goals, GN: Goals Scored Normalization, CGN: Conceded Goal Normalization.

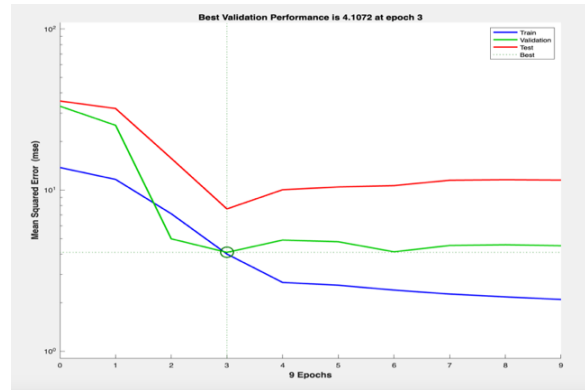


Figure 5. Performance Indicators Values of the Model developed for the French Ligue 1.

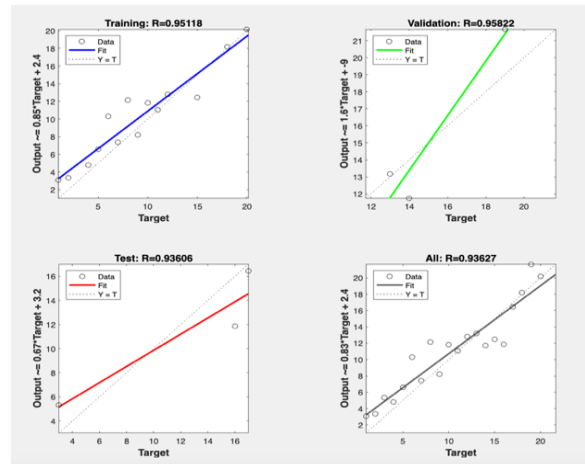


Figure 6. Training, Validation, Test and All Values of the Model Developed for the French Ligue 1

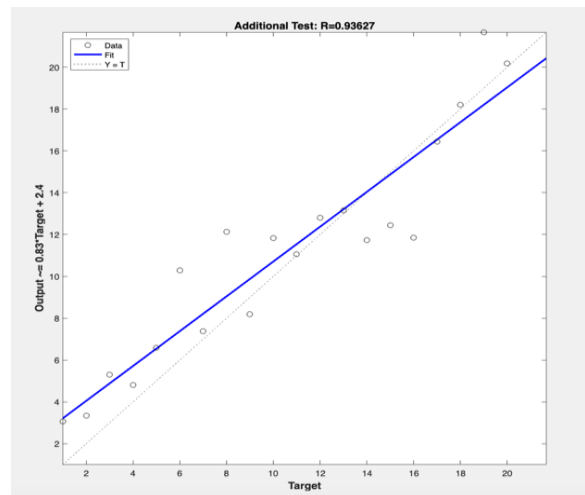


Figure 7. Additional Test Values of the Model Developed for the French Ligue 1.

IV. CONCLUSION

The aim of this study is to predict the end-of-season team rankings of the French Ligue 1 and English Premier League for the 2022-2023 season according to the goals scored and conceded with the artificial neural network model developed according to 2 different input variables. When similar studies in the literature are examined, based on the data of 2010-

2016, the accuracy rate of 68% was reached with the logistic regression model[13]. The logistic regression model obtained from the data set of the college basketball league for the years 2009-2012 using the number of free throws, defensive rebounds, assists, and turnovers yielded a rate of 68%[14]. In the same way, they reached 62% accuracy with the regression model created with the data set of shooting percentage, 3-point shooting percentage, free throw shooting percentage, offensive rebounding, assists, turnovers, number of successful free throws from the seasons belonging to the American basketball league in the years 2008-2012[15]. When the values in the findings section of the research are examined, it is seen that the performance indicators value of the model developed for the Premier League in Figure 2 is found to be 3.178, while the performance indicators value of the model developed for the French Ligue 1 in Figure 5 is found to be 4.107.

In the results of the research, the training of the model developed for the English Premier League 2022-2023 season in Figure 2 is 0.961, network validation is 0.995, test regression is 0.973 and regression for all inputs is 0.956. The training of the model developed for the French Ligue 1 2022-2023 season was 0.951, network validation was 0.958, test regression was 0.936 and regression for all inputs was 0.936. Additional test values based on the model of the developed model showed high accuracy in the English Premier League (R: 0.956). The additional test values for the French Ligue 1 data reached high values as in the English league (R: 0.936). According to the results determined with the artificial neural network model, the team ranking of the English Premier League 2022-2023 season was predicted with an accuracy of over 99% according to the goals scored and conceded, and the French Ligue 1 season was predicted with an accuracy of over 95%.

When these accuracy rates are considered, it is understood that a higher level of prediction rate can be achieved with the contribution of other variables other than these parameters. In addition to these two

variables, forward-looking prediction values can be presented to clubs, coaches and athletes with models that can be created with the data obtained from competition analysis and training tracking systems.

REFERENCES

- [1] Suat A. Veri, büyük veri ve işletmecilik. Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 2016;19(35):137-154.
- [2] Gorunescu F. Data Mining: Concepts, models and techniques: Springer Science & Business Media; 2011.
- [3] Komar E. Türkiye ve İtalya Voleybol Süper Ligleri 2013-2020 İstatistik Verilerinin Veri Madenciliği Yöntemleriyle Analizi [Yüksek Lisans Tezi]. Giresun: Giresun Üniversitesi Sağlık Bilimleri Enstitüsü; 2023.
- [4] Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics. 2017;37(2):505-515.
- [5] Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920-1930.
- [6] Fernandes ES, do V Filho A, Scremin MAA, Bastos LC, Pezzi S, Bastos RC, et al., editors. Use of artificial neural network to help the valuation of football referees' physical fitness. IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339); 1999: IEEE.
- [7] Mikut R, Reischl M. Data mining tools. Wiley interdisciplinary reviews: data mining and knowledge discovery. 2011;1(5):431-443.
- [8] Murathan T, Devicioğlu S. Veri madenciliği ve spor alanındaki uygulamalar. Spor Bilimleri Dergisi. 2018;29(3):147-156.
- [9] Ersoy İ. Sporda veri analizi [Yüksek Lisans Tezi]. İzmir: Ege Üniversitesi, Fen Bilimleri Enstitüsü; 2021.
- [10] Abdullah MR, Musa RM, Maliki ABHMB, Kosni NA, Suppiah PK. Role of psychological factors on the performance of elite soccer players. Journal of Physical Education and Sport. 2016;16(1):170.
- [11] Wilson RL. Ranking college football teams: A neural network approach. Interfaces. 1995;25(4):44-59.
- [12] Sözen A, Arcaklioğlu E, Özkaymak M. Turkey's net energy consumption. Applied Energy. 2005;81(2):209-221.
- [13] Prasetyo D, editor Predicting football match results with logistic regression. 2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA); 2016: IEEE.
- [14] Magel R, Unruh S. Determining factors influencing the outcome of college basketball games. 2013.
- [15] Jones ES. Predicting outcomes of NBA basketball games: North Dakota State University; 2016.

★★★