Research Article **GJAI (2019) 1:1**

Global Journal of Artificial Intelligence
(DOI:10.28933/GJAI)

Artificial Neural Network Approach to Football Score Prediction

Balogun O. [1], Ogunseye, A. A[2]

[1]College of Information and Communication Technology, Bells University of Technology, Otta, Nigeria; [2]Department of Electronic & Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

ABSTRACT

Sport betting companies and participants can maximize their profit in the sports betting business if they are able to accurately predict the outcome of football matches. This work seeks to develop such a football match prediction system with Manchester United football club as a case study. The developed system is based on an Artificial Neural Network (ANN) model. Scores from previous matches played by Manchester United were used to train and validate the network. The system has prediction accuracies of 73.72% and 113.5% for goals scored by, and against Manchester United respectively. The performance of the model is reasonably good but it can be improved by training the model with more football scores.

Keywords: ANN, prediction, Football, Betting, Manchester United

***Correspondence to Author:**
Ogunseye, A. A
Department of Electronic & Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

How to cite this article:
Balogun O., Ogunseye, A. A.Artificial Neural Network Approach to Football Score Prediction.Global Journal of Artificial Intelligence, 2019; 1:1.

eSciPub LLC, Houston, TX USA.
Website: http://escipub.com/

1.0 Introduction

Sport is a great source of entertainment, data, and revenue globally. It is such an emotive phenomenon that fans of sports teams have been known to do the most unthinkable to the most unreasonable things to show their love and support for their teams.

A recent dimension to sports in Nigeria comes in the form of sports betting. It is estimated that Nigerians spend up to $25 million per day on betting which double**s** what is used to be in 2014 (AnswersAfrica, 2018) This multi-billion Naira evolving industry is, in this case, fast becoming the most profitable venture for capitalists as evidenced by the number of companies that are springing up. The big players include Nairabet®, Merrybet®, Bet9ja®, Naijabet®, Lovingbet® and Surebet247® (Adebayo, 2018).

The pervasive nature of betting in recent times has made it intuitively important to be able to predict, with reasonable accuracy, the possible scoreline of a match before placing a bet on it. The method of predicting different outcomes in sports have been employed by sporting organizations to make appropriate plans to ensure success during and at the end of a season.

For this work, we make use of an Artificial Neural Network (ANN) model to predict the scoreline of Manchester United matches against opposing teams for matches played in the English Premier League. The ANN model accepts normalized data and is trained with a suitable algorithm.

2.0 Related Works

Haghighat et al. (2013) did an overview of all the data mining techniques used in sports result prediction. They reviewed techniques such as ANN, Decision trees, Support Vector Machines (SVM), Fuzzy Systems, Bayesian methods etc. They looked at the methods employed in the researches they studied and came to the conclusion that the low prediction accuracy encountered showed that there is a need for further research into sports result prediction and that differences in the data sets used prevents researchers from comparing their results with previous studies. They suggest that prediction algorithms that have been used successfully in other fields should be employed to predict sports result coupled with the hybridization of different algorithms. They also add data sets with more features that are comprehensive and collected from reliable sources should be employed to give a platform of result comparisons.

Igiri & Okechukwu (2014) made use of Knowledge Discovery in Databases (KDD) to develop a model for football match prediction making use of 9 features, some of which are a home advantage, the effect of injury, the effect of the external cup on league matches etc. They analyzed statistical and machine learning approaches for prediction and after concluding on these methods, they made use of ANN and Logistics Regression (LR) techniques. Their models, in the end, gave them 85% and 93% accuracy for ANN and LR techniques respectively. The amount of data used in the analysis is limited and the features of the data set do not have an equal impact across the teams analyzed. Research shows that making future predictions becomes more accurate with more data as it affords the model to learn.

Blaikie et al (2011) used ANN to predict NFL and NCAA football scores. They analyzed data to identify the most predictive statistics which were used in the model. Using derivative analysis, results showed that the NFL model outperformed the NCAA model. The NFL model was consistently in the top half compared to other prediction experts.

Zaveri et al (2018) used ANN and other prediction methods to predict the score outcome for matches played in the Spanish La Liga over the course of five seasons. Using the FIFA 18 game database i.e. match history database and Team vs Team database coupled with player career statistics, they predicted the outcome of matches between home and away

teams. They were able to achieve a 71.63% accuracy with LR. Their ANN model achieved an accuracy of 63.1% from the match history database and 69.2% from the combination of the match history database and the Team vs Team database.

(Tümer & Koçer, 2017) made use of Artificial Neural Networks to predict team rankings in volleyball for the male professional league. The data used in the analysis spanned a period of 2 years i.e. 2013-2015. The accuracy obtained was 98% using a layer 4-neuron model which had “logsig” transfer function, “trainlm” training function, and “learned” adaptive learning function.

3.0 The ANN model

The ANN model used is a multilayered feedforward network. It has six (6) input layers, five (5) hidden layers, and two (2) output layers. The input layer consists of Pre-season (game win ratio to give insights about the current form and depth of the team), week number (popularly called match day in the football community), the manager (this determines if the manager is new or an incumbent- this relates also to team performance as there is a likely hood of a new manager if the team is performing below expectation), team played (the rating of team being played against), attendance, and match outcome i.e. if it was a win for Manchester United or not.

The input data was transformed and normalized to a range between 0 and 1. The sigmoid transfer function was also used, coupled with randomized weights (otherwise known as bias) to feed the network forward.

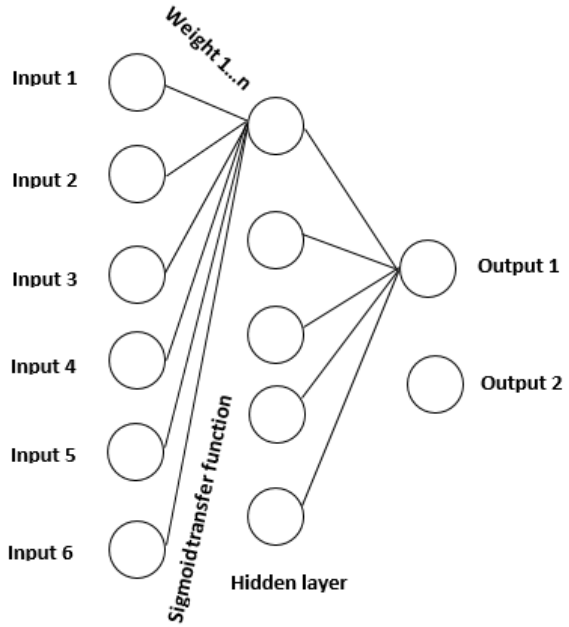

Figure 1: The ANN model showing partial connections in the network

The sigmoid function is given as

$$\text{Sigmoid function} = \frac{1}{(1+e^{-x})} \quad (1)$$

The back propagation algorithm was then used to train the network for it to properly learn.

The back propagation carries out its operation by two processes. First, it does a forward pass where the outputs are calculated and the error at the output units are also calculated. It does this by using the input data and the hidden nodes to calculate the output. It also does a back pass where the output error is used to alter the weights of the output units. The error is adjusted by back-propagating the error at the output units through the weight. It does this by calculating the differences between output results and finding the derivative of the hidden nodes.

3.1 Data

The data used for this research was got from the Wikipedia® page of Manchester United which is frequently updated and shows detailed information of matches played. Manchester United is one of the big teams in the English Premier League but it has been having dwindling fortunes for some years because of the departure of their long-term team manager. In an attempt to improve its performance, the team has changed the team manager four times over the years. This makes Manchester United a good candidate for this research. The data used spans a period of nine years i.e. 2009/2010-2017/2018 seasons.

331 of the data set were used as the training data set while 12 was used for validation. Using the developed team rating table, the opposing teams were given scores in accordance with the developed ratings (squashing) technique.

4.0 Result and Discussion

The data collected and model was developed with Microsoft Excel coupled with a neural network plug-in called 4CastXL®. Excel solver® was also used during the network training phase. Table 1 shows the result at the end of the simulation.

The model gave an accuracy of 73.72% for goals scored by Manchester United but gave a 113.35% for goals scored against Manchester United. This is indicative that although the model is learning, it needs to be adjusted needs for better performance.

Table 1: Results vs. Model output

Result		**Predicted Result**	
Goals For	Goals against	Goals For	Goals against
0	0	0.390894291	0.145250487
1	0.5	0.72208709	0.49961076
0	0.75	4.44063E-12	0.61615116
1	0.5	0.752334138	0.495373241
1	0.5	1	0.634654422
0	0.75	0.780015451	0.491083552
1	0	0.387590157	0.147460142
0	0.5	4.45487E-12	0.618883575
1	0.5	0.816792823	0.484552919
0	0	1.42714E-12	0.025882282
0	0	0.450739859	0.246524794
1	0.25	0.744833648	0.53147979

5.0 Conclusion

This work made use of an artificial neural network to predict score line for Manchester United against other English premier league teams. In the end, the network showed that it was learning but not yet perfect.

In line with the recommendations of Haghighat et al(2013), we intend to continue to improve the models based on our data and new data which will be gotten at the end of the 2018/2019 season.

6.0 References

1. Adebayo, R. (2018, July 6). Economics of $2 billion Sports Betting Industry in Nigeria. Retrieved from nairametrics: https://nairametrics.com/2018/07/06/economics-of-sports-betting/
2. AnswersAfrica. (2018). Betting in Nigeria-By the Numbers. Retrieved from AnswersAfrica: https://answersafrica.com/betting-nigeria-numbers.html
3. Blaikie, A. D., David, J. A., Abud, G. J., & Pasteur, R. D. (2011, August 11). NFL & NCAA Football Prediction using Artificial Neural. Retrieved January 31, 2019, from https://www.semanticscholar.org/paper/NFL-%26-NCAA-Football-Prediction-using-Artificial-Blaikie-Abud/5207ead80e566abd29bf2c171143fa6473c28b6a
4. Haghighat, M., Rastegari, H., & Nourafza, N. (2013, November). A Review of Data Mining Techniques for Result Prediction in. ACSIJ Advances in Computer Science: an International Journal, 2(5), 7-12. Retrieved January 31, 2019
5. Igiri, C. P., & Okechukwu, E. (2014, December). An Improved Prediction System for Football a Match Result. IOSR Journal of Engineering, 4(12), 12-20. Retrieved January 31, 2019
6. Tümer, A. E., & Koçer, S. (2017, May 30). Prediction of team league's rankings in volleyball by artificial neural network method. International Journal of Performance Analysis in Sport, 17(3), 202-211. Retrieved January 31, 2019, from https://www.tandfonline.com/doi/full/10.1080/24748668.2017.1331570
7. Zaveri, N., Zaveri, N., Tiwari, S., Shinde, P., & Teli, L. K. (2018, Febuary). Prediction of Football Match Score and Decision Making Process. International Journal on Recent and Innovation Trends in Computing and Communication, 6(2), 162-165. Retrieved January 31, 2019, from http://www.ijritcc.org/download/browse/Volume_6_Issues/February_18_Volume_6_Issue_2/1520497925_08-03-2018.pdf