

English Premier League (EPL) Soccer Matches Prediction using An Adaptive Neuro-Fuzzy Inference System (ANFIS) for

Amadin, F. I¹ and Obi, J.C.²

Department of Computer Science, University of Benin, Benin City, Nigeria.
frankamadin@uniben.edu¹; tripplejo2k2@yahoo.com²

ABSTRACT

Prediction of English Premiership League (EPL) matches has been on the heart and minds of researcher over the pass decades, but none has sufficiently introduced and Adaptive Neuro-Fuzzy Inference System (ANFIS) approach for these prediction which has served as the focal aim of this research paper using seven premier league teams and nine linguistic values. Matric Laboratory (MATLAB) 7.0 served as the tool of implementation highlighting various views. The ANFIS training was successful completed at epoch 2 and having an error of 1.41237e-006. The model was further used to predict the outcome of 7 matches with a successful rate of 71%.

Keywords: ANFIS, Premiership, Soccer, Predication, Layers

1 Introduction

It is no doubt that the most viewed sport in the world is soccer having above 10 billion fan all over the world (Tony, 2014). The English league is the most watched soccer league all over the world having more than 2 billion fans (Jonathan, 2014 and Tony, 2014). It comprises of 20 teams and each team plays 38 match in a season which spans ten months. A team plays 19 matches at home (at the city in which the club originates) and 19 match away (Jonathan, 2014). A win for any match regardless of home or away is 3 points a draw is 1 point while a loss earn no point. At the end of each season the 3 teams at the lower end (bottom) of the league are relegated to a lower division (Tony, 2014). The outcome of any match is a win, a draw or a loss (Jonathan, 2014 and Tony, 2014). To calculate the number of games played in a season using the hand shaking lemma we represent each team as a node and a matched played as an edge. There are two parallel edges between each node (one for home match and the other for away match). So, the number of game played = $(38*20) / 2 = 380$ matches. Since the edge contribute twice to the degree of the graph.

2 Review of Related Literature

In predicting the outcome of a soccer match, various approaches have been used some of which include Statistical method, Probabilistic method, Bayesian network, Multilayer perceptrons.

Yue (2003) used multilayer perceptron model to predict the outcome of a soccer match. The neural network had 13 inputs, 3 hidden layers and 1 output layer. The network was feed with the most recent 5 matches of team A and team B. The model was used to predict the outcome of 9 matches and 4 out of the nine matches gave accurate prediction. Based on the efficiency of the result, the model was re-

implemented with two additional hidden layers (a layer for the home team with a weight of 1 and the other for the position of the team on the table of ranking) which was utilized in predicting nine matches, with an 80% accuracy Aditya et al (2013) used a Multi Nominal Logic Regression (MNL) and Support Vector Machines (SVM) to predict the outcome of a match utilizing 3 fundamental approaches. In their first approach Multi nominal Logistic was enacted for the training phase in which performance optimization metrics derived from current matches were considered, rather than taking the average of previous matches and during testing they predicted the outcome between two teams using the number of matches past played. In the second approach, they trained in the same way they later test the data and instead of using the feature vector as the performance metric vector corresponding to the current match, they used Key Performance Parameters (KPP). The third approach tried to find a global set of parameter which was independent of the competing teams.

Ian and Phil (2013), forecasted international; soccer matches result using bivariate discrete distribution. They collected a total of 8,735 international soccer results from two main sources. The data for the period 1993-2001 was obtained from the archive of International Soccer Results (ISR) and the data for the period 2001 to 2004 were obtained from the Record Sport Soccer Statistics Foundation (RSSSF) archive. Data on the Federation of International Football Association (FIFA) world rankings were collected from the FIFA website for each month during the years from 1993 to 2004. The model was based on copula functions. The copula regression model forecasts 41.8% of the results correctly.

From the review of related literature An Adaptive Neuro-Fuzzy Inference System (ANFIS) approach has not been adopted previously for EPL matches prediction which is serving as the focal point of our research.

3 The ANFIS Model for Predicting the Outcome of English Premiership League (EPL) Soccer Match

The Adaptive Neuro Fuzzy Inference System (ANFIS) is one of the many hybrids of neural and fuzzy system. ANFIS combines the learning capability of the neural network and the explanatory power of the fuzzy system. The ANFIS architecture uses the sugeno type inference system Jang (1991). An example of the sugeno inference is given in equation 1:

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B, \text{ then } f_1 = p_1x + q_1y + r_1 \tag{1}$$

Where x and y is the input variable and A and B is the fuzzy set of linguistic variables and q, p and r , are consequent parameters.

ANFIS ARCHITECTURE

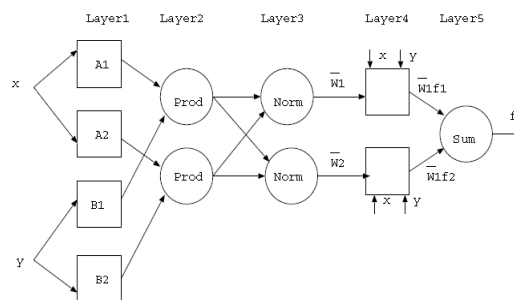


Figure 1: ANFIS ARCHITECTURE

The ANFIS architecture in Figure 1 comprises of varied layers which includes:

- a. **Layer 1 (Input Layer):** this is the first layer it is called the input layer. Each node generates the membership grades of a linguistic label (Jang, 1993). An example of a membership function is the generalized bell function is given in equation 2:

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\frac{(x - c_i)^2}{a_i^2} \right]^{b_i}} \quad (2)$$

Where {a, b, c} is the parameter set and x stands for the individual value. As the values of the parameters change, the shape of the bell-shaped function varies. Parameters in that layer are called premise parameters.

- b. **Layer 2 (Fuzzification layer):** The second layer of the Adaptive Neuro Fuzzy Inference System (ANFIS) model is fuzzification layer it is also called the membership function layer. In this layer the output of the first layer is mapped to a fuzzy set using bell-shaped membership function (Jang, 1997 and Nauck et al., 1997). The Bell membership function provides smooth and non-linear functions that can be used by the learning systems. Each node calculates the firing strength of each rule using the min or prod operator show in equation 3. In general, any other fuzzy AND operation can be used.

$$O_{2,i} = w_i = \mu_{A_i}(x) \times \mu_{B_i}(x) \quad i = 1, 2 \quad (3)$$

- c. **Layer 3 (Rule Layer):** The third layer is the rule layer. Each neuron in this layer corresponds to a single first-order Sugeno fuzzy rule receiving signal from the membership function layer and computes the truth value of the rule (Jang, 1993 and 1997). The nodes calculate the ratios of the rule's firing strength to the sum of all the rules firing strength using the equation 4. The result is a normalised firing strength.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \quad (4)$$

- d. **Layer 4 (Normalization layer):** The fourth layer is the normalization layer. Each neuron in this layer receives signals from all rule neurons in the third layer, and calculates the normalized firing strength of a given rule. The nodes compute a parameter function on the layer 3 output using equation 5. Parameters in this layer are called consequent parameters.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (5)$$

- e. **Layer 5 (Defuzzification layer):** The fifth layer is the defuzzification layer. Each neuron in this layer is connected to the respective normalization neuron in the fourth layer (Jang, 1997 and Nagnevitsky, 2002). It is a single node that aggregates the overall output as the summation of all incoming signals using equation 6

$$O_{5,1} = \text{overall output} = \sum_i \bar{w}_i f_i \quad (6)$$

3.1 Generating Dataset

In generating the dataset we relied on 5 factors which then constituted the ANFIS parameters which was then used to predict the outcome of the match, they are:

- 2 of the last most recent match played by team A
- 2 of the last most recent match played by team B
- The point of both teams in the table of ranking
- Their popularity
- Home advantage.

The first 2 parameters (2 of the last most recent match played by team have 9 linguistic variable (WW, WD, WL, DW, DD, DL, LW, LD, LL) while the point of both teams and the popularity of the teams have 3 parameters (highA, same, lowA) and the home advantage have 2 linguistic variable (homeA and awayA).

3.2 Variable Used In the Model

Table 1, shows, six linguistic variables and six linguistic values associated with the Anfis proposed architecture for predicting the EPL results and matches

Table 1: Fuzzy Linguistic Variables and Values

Position	Club	Last Two Match Played	Point
1	Chelsea	WD	33
2	Manchester City	WW	25
3	Southampton	WW	22
4	Manchester United	WW	21
5	West ham United	DW	20
6	Arsenal	LW	20

4 Stimulations and Experiment

The stimulation was carried out using MATLAB Fuzzy logic toolbox 2007 and the model was used to predict the outcome of 7 matches. The various simulation modules are specified from Figure 2, 3, 4 and 5 respectively

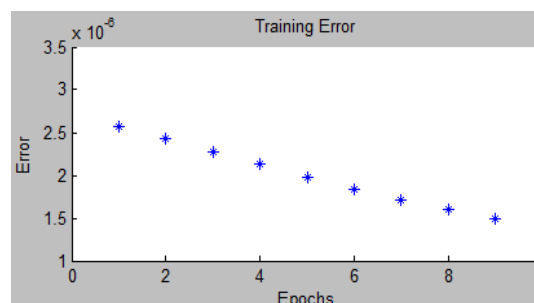


Figure 2: Training Error

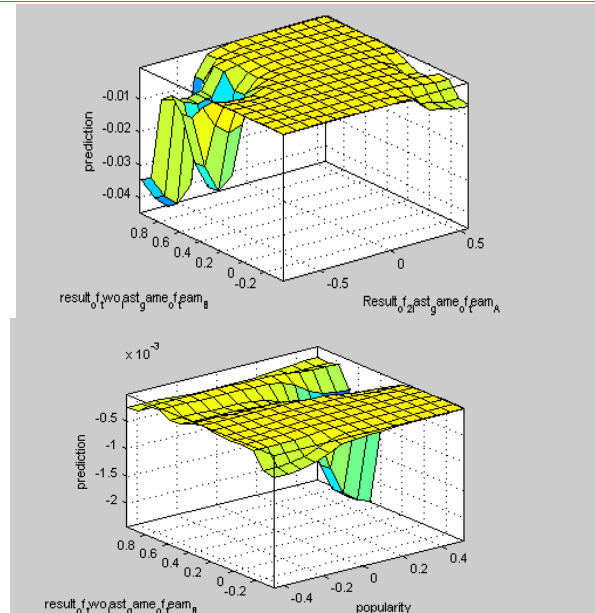


Figure 3: Surface View Prediction 1

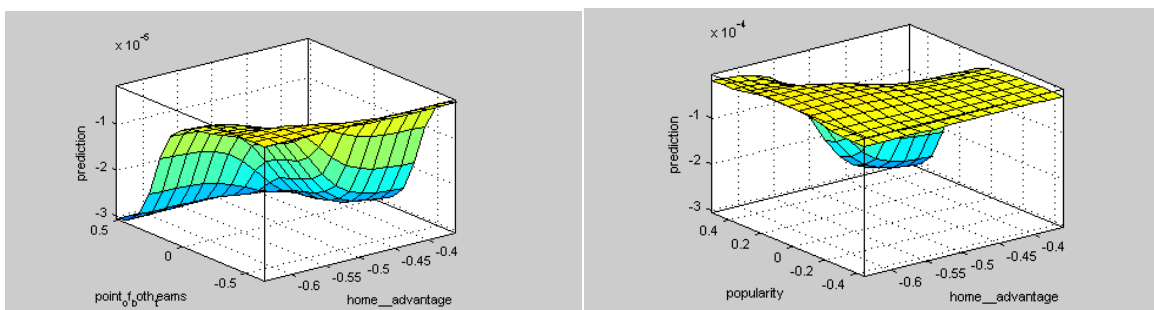


Figure 4: Surface View Prediction 2

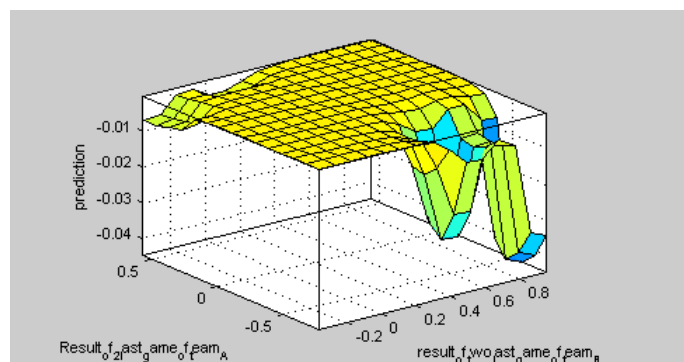


Figure 5: Surface View Prediction

The ANFIS training was completed at epoch 2 and having an error of 1.41237e-006. The model was further used to predict the outcome of 7 matches and 5 outcomes out of the matches were accurately predicted.

4.1 Prediction Results

The EPL prediction matches are specified clearly on Table 2, highlighting both actual result and prediction results

Table 2: English Premier League (EPL) Prediction Matches

Home Team	Away Team	Actual Result	Prediction
Manchester	Stoke	Manchester	Manchester
Chelsea	Tottenham	Draw	Chelsea
Sunderland	Man City	Man City	Man City
Everton	Hall City	Draw	Draw
Arsenal	Southampton	Arsenal	Southampton
Leister City	Liverpool	Liverpool	Liverpool
Swansea	QPR	Swansea	Swansea

5 Conclusion

The ANFIS model was able to make an accurate prediction of 5 out of 7 matches which is promising but this implies that there might be more factors which were not included in our model that determine the outcome of a soccer match. So, further study is encouraged to increase the accuracy of prediction of a soccer match using more applicable variables.

REFERENCES

- [1] Aditya S. T., Aditya P. and Vikesh K. (2013) "Game ON! Predicting English Premier League Match Outcomes" retrieved from gameON.com
- [2] Ian M. and Phil S (2013), "Forecasting international soccer match results using bivariate discrete distributions" Centre for Operational Research and Applied Statistics, Salford Business School, University of Salford, Salford, Manchester UK.
- [3] Jang, R (1991). "Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm". Proceedings of the 9th National Conference on Artificial Intelligence, Anaheim, CA, USA, pp. 762–767. Retrieved on October 15th 2014 from www.irrolecom/article/00098.pdf
- [4] Jang, J. (1993). "ANFIS: adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man and Cybernetics Volume 23 issue 3. doi:10.1109/21.256541.
- [5] Jang, S (1997) "Neuro-Fuzzy and Soft Computing" Prentice Hall, Pp 335–368, ISBN 0-13-261066-3
- [6] Jonathan T. (2014), "2014-15 English Premier League and MLS TV" retrieved online from <http://www.philly.com/philly/blogs/thegoalkeeper/NBC-Sports-2014-Premier-League-TV-schedule-for-August-through-November.html#sd6Wr3imqE4ehOdH.99>
- [7] Nauck, D., Klawon F. and R. Kruse, (1997), "Foundations of Neuro-Fuzzy Systems", J. Wiley & Sons pages 312

- [8] Negnevitsky, M., (2002) "Artificial Intelligence: A Guide to Intelligent Systems", 2d ed. Harlow, England: Addison Wesley, pp. 90-343.
- [9] Tony M. (2014), "20 English Premiership Teams" <http://www.businessinsider.com/new-premier-league-kits-2014-8?op=1>
- [10] Yue W. M. (2003)" Prediction on Soccer Matches using Multi-Layer Perceptron" ID: 903-051-7735