

---

# Prediction on Soccer Matches using Multi- Layer Perceptron

---

ECE 539 Project Report

---

Author: Yue Weng Mak  
ID: 903-051-7735

---

**Contents**

Introduction .....3

Motivation .....4

Experiment.....4

Results.....5

Improvements.....6

Factors .....7

Conclusion.....8

Reference .....8

## Introduction

Soccer has becoming increasingly popular over the years. During the last decade, soccer's biggest event, the FIFA World Cup, has attracted millions of fans worldwide; in fact, the viewership of the FIFA World Cup match arguably surpasses the Super Bowl event as soccer is the more widely played sport in the world. To showcase its international reach, the World Cup was hosted in Japan and South Korea in 2002. South Africa hosted its first World Cup in 2010, and Brazil will host the next World Cup in 2014. Even in the Middle East, soccer appeals to the audience in that part of the world, with Qatar hosting the World Cup in 2022. The exposure of soccer to different continents highlights the growing popularity for the sport that originated in Europe.

With great viewership comes great opportunity for soccer viewers to bet on scores too. This is especially true in Asia, where soccer betting is common among soccer fans. For instance, in Singapore, the Singapore Pools, which is Singapore's legalized gambling institution, allows Singapore citizens to bet on almost anything, with soccer betting being one of its largest revenue generators. These bets range from predicting the outcome of the score, the exact result of the score, the winning margin and whether a certain player will score first. They are based on analyses including whether a certain player is playing, the coach's record against the opposition and the strategies that the coach has been using. These factors are all based on human analyses with a tinge of biasness in them, which is inevitable for any sport that involves human judgment.

## Motivation

In this project, I will attempt to introduce a scientific discipline in predicting the outcome of matches by proposing a model based on statistics to remove the biasness element in the prediction. With this proposal, I hope to achieve an accurate prediction of the outcome of certain soccer matches by introducing a methodical method of predicting soccer scores. Getting all my predictions correct is what I hope to achieve, but realistically that is hard to achieve. Based on the results of the outcome, I may improve my model so that a higher percentage of accuracy can be achieved with regard to the older model.

## Experiment

For this project, the model that I will be using will be a Multilayer Perceptron (MLP). It will be a 3 layer MLP with 13 inputs, 3 hidden neurons and a single output.

These 13 inputs will be:

- 1) Last 5 matches from Team 1.
- 2) Last 5 matches from Team 2.
- 3) Last 3 encounters between Team 1 and Team 2.

A win for the team will be credited as a 1, a draw 0 and a loss -1. There will also be a multiplying factor, with the most recent game played having the highest weight. This is the case because the outcome of the previous game will affect how the team plays in the next game.

To construct this model based on multiplying factor, the following attributes will be observed:

- The most recent game will have a weight of +2.5
- The next most recent game will have a weight of +2.0, and so on.

In the second layer, there will be 3 hidden neurons. These 3 hidden neurons will be the result of team 1, the result of team 2 and the result between team 1 and team 2. The weights of the 2<sup>nd</sup> layer are determined as follows:

- The first neuron (Team 1) will have a weight of 1;
- The second neuron (Team 2) will have a weight of -1 and
- The third neuron (Team 1 and Team 2) will have a weight of 1.

This is done to determine the outcome of the output, and the output will determine the result of Team 1. If it is highly positive, this means that the result of the first neuron is much greater than the result of the second neuron, resulting in a victory for Team 1. If it is slightly positive and slightly negative, this means that the result of the first neuron is about the same as the result of the second neuron, resulting in a draw for Team 1 and if it is highly negative, this means that the result of the first neuron is much lesser than the result of the second neuron, resulting in a loss for Team 1.

## **Results**

Based on the model I proposed, I did a trial run on a number of games played over one weekend. These are the games that I predicted and the results that follow:

Home	Away	Predictions Made by my model	Actual Results
<b>Arsenal</b>	Fulham	Arsenal	Arsenal
<b>Birmingham</b>	Tottenham	Tottenham	Draw
<b>Blackburn</b>	Wolves	Blackburn	Blackburn
<b>Chelsea</b>	Everton	Draw	Draw
<b>Manchester City</b>	Bolton	Draw	Manchester City
<b>Wigan</b>	Stoke	Stoke	Draw
<b>Sunderland</b>	West Ham	West Ham	Sunderland
<b>WBA</b>	Newcastle	Draw	WBA
<b>Liverpool</b>	Aston Villa	Liverpool	Liverpool

As can be seen from the table, 4 out of 9 matches are accurately predicted.

## Improvements

Based on the predictions made by my model, it is not ideal as more than half of the matches are incorrectly predicted. Therefore I decided to improve my model, hopefully increasing the percentage accuracy of my prediction. I decided to add in 2 additional inputs to the model that I proposed. These 2 inputs will be the inputs that will be coming into the hidden layer in layer 2. They are:

- 1) Home Advantage Input (Team 1 will have an additional input of 1 and Team 2 will have an additional input of 0)
- 2) Rankings Input (Team that has a higher ranking than the opposition will have an additional input of 1 while the opposition will have an additional input of 0).

These 2 inputs are chosen because they are factors that can affect the outcome of the game.

For the first input, home support will help spur the home team to play better and may also

intimidate the opposition team which will result in a favorable outcome for the home team. For the second input, the team that is higher up in the table will be more consistent. Hence they are favored to win the game.

With this new model in place, the same games will be used to predict the outcomes.

The result is as follows:

Home	Away	Predictions Made by my old model	Predictions Made by my new model	Actual Results
<b>Arsenal</b>	Fulham	Arsenal	Arsenal	Arsenal
<b>Birmingham</b>	Tottenham	Tottenham	Tottenham	Draw
<b>Blackburn</b>	Wolves	Blackburn	Blackburn	Blackburn
<b>Chelsea</b>	Everton	Draw	Draw	Draw
<b>Manchester City</b>	Bolton	Draw	Manchester City	Manchester City
<b>Wigan</b>	Stoke	Stoke	Stoke	Draw
<b>Sunderland</b>	West Ham	West Ham	West Ham	Sunderland
<b>WBA</b>	Newcastle	Draw	WBA	WBA
<b>Liverpool</b>	Aston Villa	Liverpool	Liverpool	Liverpool

As can be seen from the table, the prediction made by my new model is more accurate than the old model that was being proposed. 2 additional matches are accurately predicted, on top of the 4 matches that are correctly predicted by my old model.

## Intangible Factors

However, there are certain matches that are still inaccurately predicted. These are not due to the statistics but factors that can influence a match during game time. These factors include

referee decisions, where a referee gives a wrong decision which affects the outcome of the game, player and team performance, where the 'star' player and the team as a whole are not performing due to random occurrences, causing the team to play badly. Other factors include the change of strategies used in the game that does not work as planned, such as where the opposition team playing in a defensive manner resulting in a draw instead of a win for the home team. Finally, luck may come into play, where a highly favored team will be shocked by the opposition team.

## **Conclusion**

As can be seen from the model, by increasing the number of inputs, the accuracy rate improves. This highlights the fact that the more the inputs, the more likely the prediction will be correct. However, they are not the only factors that influence the outcome of the game. Many other factors have to be considered too. Therefore there is always an unpredictable factor in the game.

## **Reference**

P. Rotshtein, M. Posner and A. B. Rakityanskaya, "*Football Predictions Based on a Fuzzy Model with Genetic and Neural Tuning*," *Cybernetics and Systems Analysis Journal* (2005)

Alan McCabe, "*An Artificially Intelligent Sports Tipper*," in *Proceedings : 15th Australian Joint Conference on Artificial Intelligence* (2002)

<http://www.soccernet.com>