A.M.Sachin Randalle - st20098214

Cardiff MEtropoltan university  BSC IN SOFTWARE ENGINEERING

FINAL Project Proposal

# Acknowledgement

For the Final Project of the BSc in Software Engineering Cardiff Metropolitan University I propose to do a Football Results Predicting System for a real life team. I name this system as Arsenal Fc Premier League Predictor.

I have briefly explained the methods and the techniques which I am going to use to implement this system. I wish I have selected a project which is eligible for a final project submission.

Thank You.

Contents

[Acknowledgement 1](#_Toc457102822)

[**Introduction** 3](#_Toc457102823)

[**Current situation leading to problem identification** 4](#_Toc457102824)

[**Proposed technique to solve the problem** 5](#_Toc457102825)

[**Feasibility** 6](#_Toc457102826)

[**Deliverables** 7](#_Toc457102827)

[**Required Resources** 8](#_Toc457102828)

[**Limitation** 8](#_Toc457102829)

[**Time plan for implementation** 9](#_Toc457102830)

[References 10](#_Toc457102831)

# **Introduction**

Arsenal Football Club is an [English](https://en.wikipedia.org/wiki/England) professional [football](https://en.wikipedia.org/wiki/Association_football) club based in [Holloway](https://en.wikipedia.org/wiki/Holloway,_London), [London](https://en.wikipedia.org/wiki/London), that plays in the [Premier League](https://en.wikipedia.org/wiki/Premier_League), the top flight of English football. The club [has won](https://en.wikipedia.org/wiki/Football_records_in_England#Total_titles_won_.281871.E2.80.93present.29) 12 [FA Cups](https://en.wikipedia.org/wiki/FA_Cup), a joint-record, 13 [League titles](https://en.wikipedia.org/wiki/List_of_English_football_champions#Total_titles_won), two [League Cups](https://en.wikipedia.org/wiki/Football_League_Cup), 14 [FA Community Shields](https://en.wikipedia.org/wiki/FA_Community_Shield), and one [UEFA Cup Winners' Cup](https://en.wikipedia.org/wiki/UEFA_Cup_Winners%27_Cup) and [Inter-Cities Fairs Cup](https://en.wikipedia.org/wiki/Inter-Cities_Fairs_Cup). (Arsenal, 2013)

**What is Premier League?**

The Premier League is an English professional league for men's [association football](https://en.wikipedia.org/wiki/Association_football) clubs. At the top of the [English football league system](https://en.wikipedia.org/wiki/English_football_league_system), it is the country's primary football competition. Contested by 20 clubs, it operates on a system of [promotion and relegation](https://en.wikipedia.org/wiki/Promotion_and_relegation) with the [English Football League](https://en.wikipedia.org/wiki/English_Football_League). Welsh clubs that compete in the English football league system can also qualify.

The Premier League is a corporation in which the 20 member clubs act as shareholders. Seasons run from August to May. Teams play 38 matches each (playing each team in the league twice, home and away), totaling 380 matches in the season.

# **Problem Statement**

The major problem here is predicting a hypothesis which is very hard to pick. Football is a game where 11 players contained team play against another 11 players contained team. At the end of each game even weaker team could win, but the intelligence should be able to predict the most accurate winner considering the past results of each teams.

The team Arsenal FC will face 19 different teams twice a season which means 38 matches altogether. The prediction should have the influence of previous season results too.

# **Proposed technique to solve the problem**

The technique that I propose to solve this problem is Naïve Bayes Algorithm.

It is a classification technique based on [Bayes’ Theorem](https://en.wikipedia.org/wiki/Bayes%27_theorem) with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as ‘Naive’ (Ray, 2015).

Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

Bayes theorem provides a way of calculating posterior probability P(c|x) from P(c), P(x) and P(x|c). Look at the equation below:

[](http://www.analyticsvidhya.com/wp-content/uploads/2015/09/Bayes_rule-300x172.png)

Figure 1

By using this technique can analyze the previous scores and results and then predict an accurate results.

# **Feasibility**

**Scope**

Scope of this project would be predicting the results of the team Arsenal FC throughout the season. And to view the predicted results in a graphical way where a user can read and understand it properly. And the system will have different predicting methods to predict the results.

**Strengths**

Regardless of the team strength (Star Players in the team) system will predict only considering the previous results.

**Threat**

Predicted results will not be accurate all the time.

**Technical**

To build the project I will be using my personal computer which has Matlab installed in it. I have a stable internet connection at home thus any other technical requirement will not be needed.

# **Deliverables**

As I have mentioned earlier the technique that I will be using to predict the results would be Naïve Bayes System. By using this technique there will be couple of manual methods that I’m expecting to include

Such as:

* Without considering home/away games
* Without considering cup games

There will be three outputs when considering a game. Those are Win, Draw and Lose. While predicting for the whole season I’m planning to make a point table which would elaborate the final points in the season for the team.

The results will look like this. (Since I didn’t start to build the project yet, I edited an image like how I’m thinking to give the output with my GUI)



Figure 2

# **Required Resources**

I will be using Matlab to build this project.

# **Limitation**

The system has a limitation which is it can only predict the probabilities either win, draw or lose. It will not predict the scores of the each game.

# **Time plan for implementation**

|  |  |
| --- | --- |
| To get all the required data from Internet | 1 Day |
| To make Flow Chart | 4 Days |
| To do the Mathematical and Algorithm Parts | 15 Days |
| To Finalize the AI part | 10 Days |
| Make the GUI | 7 Days |

# References

Arsenal. (2013, January 2). *Arsenal Football Club PLC*. Retrieved from arsenal.com: http://www.arsenal.com/the-club/the-arsenal-way

League, P. (2016). Retrieved from Premier League: http://www.premierleague.com/about

Ray, S. (2015, September 13). Retrieved from Analytics Vidhya: http://www.analyticsvidhya.com/blog/2015/09/naive-bayes-explained/