****

**BANA 7030: Simulation Modeling & Methods**

**Monte Carlo Simulation of 2019-20 Premier League Season**

**Singh, Ashutosh**

MS Business Analytics

M13433470

**OBJECTIVE**

Monte Carlo Simulation is one of the most popular techniques when it comes to simulating all kinds of outcomes and probabilities. These simulations are widely used in sports forecasting, since the multitude of possibilities makes direct calculation of playoff probabilities infeasible.

In this project, we will try to accomplish the same thing – A model which can predict Ranking and Points of teams playing in a Football (Soccer) league for this season. Due to Covid-19, the current season has been abandoned or suspended everywhere which provides us with the perfect opportunity to observe what our model predicts has this season been completed.

**Keywords**: *Monte Carlo Simulation, Soccer, Predictions*

**Introduction**

Football or “Soccer” as they say here in America, is one of the most popular games across the world. It is played by almost 200+ countries globally with European leagues as the main centre of attraction. Among the European leagues, **“Premier League”** stands out to be the best comprising of famous clubs like Manchester United, Arsenal, Liverpool, Chelsea etc.

Every year **20 teams** play against each other to fight for the coveted title. A total of **380 matches** are played and the team finishing the league with highest tally of points is declared as winner. Manchester City were the winners of **2018-19 season**.

Unfortunately, due to the present Covid-19 situation the **2019-20 season** has been suspended till the situation gets normal. However, that does not stop us from predicting the final table of this season. With the help of simulation modelling, we can simulate what would the table of 2019-20 would have looked like based on that particular team performance from the previous year.

**About the Data**

For the study, the PL Season 2018/19 statistics has been used. The source for the dataset is: <http://www.football-data.co.uk/englandm.php>

It contains information about all the 380 fixtures with following columns as the main features:

* Home Team
* Away Team
* FTHG – Goals Scored by the Home Team at Full-time
* FTAG – Goals Scored by the Away Team at Full-time

**Procedure**

*In a football match, a team is awarded 3 points for a win, 1 for a draw and 0 if it loses.*

In order to simulate results for the each of the fixture in the next league season, we need to calculate metrics for each of the team. To give you a basic overview of how the results were calculated, let us see the first fixture for the PL 2019/20 season which was:

**West Ham United vs Manchester City on Aug 10,2019**

|  |  |  |
| --- | --- | --- |
|  | **Goals For** | **Goals Against** |
| **Average PL Home Team** | **1.568421053** | **1.252631579** |
| **West Ham (Home)** | 1.684210526 | 1.421052632 |
| **Man City (Away)** | 2 | 0.578947368 |

From the above table, we can see the respective statistics against the league average. We can compare how both Home and Away team compares in relation to Goal scoring and conceding goals. Using the above stats, we calculate the respective Attacking and Defending Index for each team.

|  |  |  |
| --- | --- | --- |
|  | **ATT Rating** | **DEF Rating** |
| **West Ham (Home)** | **1.073825503** | **1.134453782** |
| **Man City (Away)** | **1.596638655** | **0.369127517** |

Now since we have the Attacking (Goal Scoring) and Defending (Goal Conceding) metrics for both home and away team, we can calculate what would be the expected goals scored by each team in this match. The Expected goals are calculated quite simply – we multiply the attacking rating of a team with defending rate of opposite team.

|  |  |
| --- | --- |
|  | **ExpG** |
| **West Ham (Home)** | **0.396378542** |
| **Man City (Away)** | **1.81131276** |

**Building Initial Model – (Poisson Distribution)**

Now since we have the expected goals for each team, we can apply Poisson distribution to calculate probabilities of scoring goals from 0 to 5 for each team.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Goals Scored** | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** |
| **West Ham (Home)** | **0.6728** | **0.2667** | **0.0529** | **0.007** | **0.0007** | **0.0001** |
| **Man City (Away)** | **0.1634** | **0.296** | **0.2681** | **0.1619** | **0.0733** | **0.0266** |

We are assuming goals scored event to be independent of each other, so for a particular score line scenario, we are multiplying both probabilities. (*E.g if score line is 0-0, we multiply = P(West Ham scores 0) x P(Man City scores 0)*)

|  |  |
| --- | --- |
| **ScoreLine** | **Probability** |
| 1-0 | 0.04357878 |
| 2-0 | 0.00864386 |
| 3-0 | 0.0011438 |
| 2-1 | 0.0156584 |
| 3-1 | 0.002072 |
| 4-2 | 0.00018767 |
| 5-2 | 0.00002681 |
| 4-3 | 0.00011333 |
| 5-3 | 0.00001619 |
| 0-0 | 0.10993552 |
| 1-1 | 0.0789432 |
| 2-2 | 0.01418249 |
| 3-3 | 0.0011333 |
| 0-2 | 0.18037768 |
| 1-2 | 0.07150227 |
| 0-3 | 0.10892632 |
| 0-4 | 0.04931624 |
| 1-4 | 0.01954911 |

**Away Team Wins (P < 0.72)**

**Match draws (0.72 < P < 0.92)**

**Home Team Wins (P > 0.92)**

Now by looking at the above table, we have the relative probabilities of team winning, drawing and losing. All is left to perform **Monte Carlo Simulation:**

* Generate 1000 random numbers between 0 to 1
* Check where the number lies with the mentioned probability range
* Assign points accordingly (3 for Win, 1 for draw, 0 for loss)
* Repeat the process for all the match fixtures

**Results**

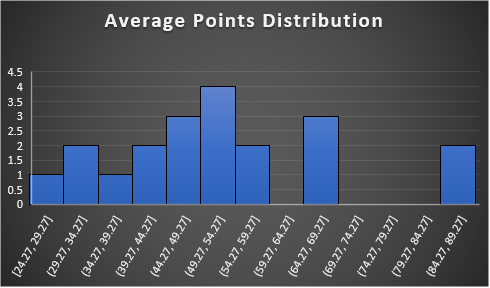
By repeating the process 1000 times, we get the average points accumulated by a team during the course of the season. The following is the final league table which we get after completion of all 380 fixtures:

|  |  |  |
| --- | --- | --- |
| **League Position** | **Teams** | **Points** |
| **1** | **Man City** | **85.87** |
| **2** | **Liverpool** | **84.52** |
| **3** | **Tottenham** | **66.59** |
| **4** | **Chelsea** | **65.25** |
| 5 | Arsenal | 64.06 |
| 6 | Man United | 57.93 |
| 7 | Everton | 55.72 |
| 8 | Leicester | 53.72 |
| 9 | Wolves | 51.61 |
| 10 | Crystal Palace | 50.67 |
| 11 | West Ham | 48.04 |
| 12 | Bournemouth | 47.33 |
| 13 | Watford | 47.32 |
| 14 | Newcastle | 47.08 |
| 15 | Southampton | 41.24 |
| 16 | Burnley | 40.11 |
| 17 | Brighton | 37.52 |
| **18** | **Cardiff** | **32.59** |
| **19** | **Fulham** | **28.84** |
| **20** | **Huddersfield** | **24.57** |

**Predicted Title Winner:** Manchester City

**Teams to Finish in Top Four:** Manchester City, Liverpool, Tottenham, Chelsea

**Teams to be Relegated to Lower Division:** Cardiff, Fulham, Huddersfield



We can observe that points distribution is slightly right-skewed because of high points accumulated by the top teams.

However before comparing the table with the present table (as of March 2020), we can see that there is quite close competition between Man City and Liverpool (Difference of ~ 1 point). So, it would be interesting to look the frequency of title winners for each of the simulation.

We can see that out of 1000 simulations, Man City is predicted to win the title **55%** of times whereas Liverpool **43.6%.**

**Comparing our Results with 2018/19 Season:**

Since our results were simulated on 2018/19 findings, let us see how accurate we were in comparison to original 2018/19 premier league standings.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **Position** | **Teams** | **Points** | | **1** | **Man City** | **85.87** | | **2** | **Liverpool** | **84.52** | | **3** | **Tottenham** | **66.59** | | **4** | **Chelsea** | **65.25** | | 5 | Arsenal | 64.06 | | 6 | Man United | 57.93 | | 7 | Everton | 55.72 | | 8 | Leicester | 53.72 | | 9 | Wolves | 51.61 | | 10 | Crystal Palace | 50.67 | | 11 | West Ham | 48.04 | | 12 | Bournemouth | 47.33 | | 13 | Watford | 47.32 | | 14 | Newcastle | 47.08 | | 15 | Southampton | 41.24 | | 16 | Burnley | 40.11 | | 17 | Brighton | 37.52 | | **18** | **Cardiff** | **32.59** | | **19** | **Fulham** | **28.84** | | **20** | **Huddersfield** | **24.57** | | |  |  |  |  | | --- | --- | --- | --- | | **Pos** | **Team** | **Pld** | **Pts** | | **1** | [**Manchester City (C)**](https://en.wikipedia.org/wiki/Manchester_City_F.C.) | **38** | **98** | | **2** | [**Liverpool**](https://en.wikipedia.org/wiki/Liverpool_F.C.) | **38** | **97** | | **3** | [**Chelsea**](https://en.wikipedia.org/wiki/Chelsea_F.C.) | **38** | **72** | | **4** | [**Tottenham Hotspur**](https://en.wikipedia.org/wiki/Tottenham_Hotspur_F.C.) | **38** | **71** | | 5 | [Arsenal](https://en.wikipedia.org/wiki/Arsenal_F.C.) | 38 | **70** | | 6 | [Manchester United](https://en.wikipedia.org/wiki/Manchester_United_F.C.) | 38 | **66** | | 7 | [Wolves](https://en.wikipedia.org/wiki/Wolverhampton_Wanderers_F.C.) | 38 | **57** | | 8 | [Everton](https://en.wikipedia.org/wiki/Everton_F.C.) | 38 | **54** | | 9 | [Leicester City](https://en.wikipedia.org/wiki/Leicester_City_F.C.) | 38 | **52** | | 10 | [West Ham United](https://en.wikipedia.org/wiki/West_Ham_United_F.C.) | 38 | **52** | | 11 | [Watford](https://en.wikipedia.org/wiki/Watford_F.C.) | 38 | **50** | | 12 | [Crystal Palace](https://en.wikipedia.org/wiki/Crystal_Palace_F.C.) | 38 | **49** | | 13 | [Newcastle United](https://en.wikipedia.org/wiki/Newcastle_United_F.C.) | 38 | **45** | | 14 | [Bournemouth](https://en.wikipedia.org/wiki/A.F.C._Bournemouth) | 38 | **45** | | 15 | [Burnley](https://en.wikipedia.org/wiki/Burnley_F.C.) | 38 | **40** | | 16 | [Southampton](https://en.wikipedia.org/wiki/Southampton_F.C.) | 38 | **39** | | 17 | [Brighton & Hove Albion](https://en.wikipedia.org/wiki/Brighton_%26_Hove_Albion_F.C.) | 38 | **36** | | 18 | [Cardiff City (R)](https://en.wikipedia.org/wiki/Cardiff_City_F.C.) | 38 | **34** | | 19 | [Fulham (R)](https://en.wikipedia.org/wiki/Fulham_F.C.) | 38 | **26** | | 20 | [Huddersfield Town (R)](https://en.wikipedia.org/wiki/Huddersfield_Town_A.F.C.) | 38 | **16** | |

*Simulated Results Actual PL 2018/19 Table*

Our simulation did quite good in predicting the winners, the Top 4 and the teams which were relegated in that season. One major area where there was a shortcoming that it was not able to accurately predict the points of some teams like Man City and Liverpool which had 10 more points than our simulated results.

**Comparing our Results with the present Season (2019-20):**

The real test of our simulation is whether we would be able to predict the present season which cannot be completed because of Covid-19 situation.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **Position** | **Teams** | **Points** | | **1** | **Man City** | **85.87** | | **2** | **Liverpool** | **84.52** | | **3** | **Tottenham** | **66.59** | | **4** | **Chelsea** | **65.25** | | 5 | Arsenal | 64.06 | | 6 | Man United | 57.93 | | 7 | Everton | 55.72 | | 8 | Leicester | 53.72 | | 9 | Wolves | 51.61 | | 10 | Crystal Palace | 50.67 | | 11 | West Ham | 48.04 | | 12 | Bournemouth | 47.33 | | 13 | Watford | 47.32 | | 14 | Newcastle | 47.08 | | 15 | Southampton | 41.24 | | 16 | Burnley | 40.11 | | 17 | Brighton | 37.52 | | **18** | **Cardiff** | **32.59** | | **19** | **Fulham** | **28.84** | | **20** | **Huddersfield** | **24.57** | | |  |  |  |  | | --- | --- | --- | --- | | **Pos** | **Team** | **Pld** | **Pts** | | 1 | [Liverpool](https://en.wikipedia.org/wiki/Liverpool_F.C.) | 29 | **82** | | 2 | [Manchester City](https://en.wikipedia.org/wiki/Leicester_City_F.C.) | 28 | **57** | | 3 | [Leicester City](https://en.wikipedia.org/wiki/Leicester_City_F.C.) | 29 | **53** | | 4 | [Chelsea](https://en.wikipedia.org/wiki/Chelsea_F.C.) | 29 | **48** | | 5 | [Manchester United](https://en.wikipedia.org/wiki/Manchester_United_F.C.) | 29 | **45** | | 6 | [Wolves](https://en.wikipedia.org/wiki/Wolverhampton_Wanderers_F.C.) | 29 | **43** | | 7 | [Sheffield United](https://en.wikipedia.org/wiki/Sheffield_United_F.C.) | 28 | **43** | | 8 | [Tottenham Hotspur](https://en.wikipedia.org/wiki/Tottenham_Hotspur_F.C.) | 29 | **41** | | 9 | [Arsenal](https://en.wikipedia.org/wiki/Arsenal_F.C.) | 28 | **40** | | 10 | [Burnley](https://en.wikipedia.org/wiki/Burnley_F.C.) | 29 | **39** | | 11 | [Crystal Palace](https://en.wikipedia.org/wiki/Crystal_Palace_F.C.) | 29 | **39** | | 12 | [Everton](https://en.wikipedia.org/wiki/Everton_F.C.) | 29 | **37** | | 13 | [Newcastle United](https://en.wikipedia.org/wiki/Newcastle_United_F.C.) | 29 | **35** | | 14 | [Southampton](https://en.wikipedia.org/wiki/Southampton_F.C.) | 29 | **34** | | 15 | [Brighton & Hove Albion](https://en.wikipedia.org/wiki/Brighton_%26_Hove_Albion_F.C.) | 29 | **29** | | 16 | [West Ham United](https://en.wikipedia.org/wiki/West_Ham_United_F.C.) | 29 | **27** | | 17 | [Watford](https://en.wikipedia.org/wiki/Watford_F.C.) | 29 | **27** | | 18 | [Bournemouth](https://en.wikipedia.org/wiki/A.F.C._Bournemouth) | 29 | **27** | | 19 | [Aston Villa](https://en.wikipedia.org/wiki/Aston_Villa_F.C.) | 28 | **25** | | 20 | [Norwich City](https://en.wikipedia.org/wiki/Norwich_City_F.C.) | 29 | **21** | |

*Simulated Results PL 2019/20 Season (as of Mar, 2020)*

**Key Insights:**

* Liverpool was set to win the league in the present season. Our table however predicted it to be second but still it was close second to Man City in terms of number of points (~ 1 point) and there were about 43.6% cases in our simulations where Liverpool won. So, our model performed quite well in this regard.
* There were some other accurate predictions in top teams like Chelsea, Man United but model failed in case of teams like Tottenham, Arsenal which performed considerably worse this season.
* Some Mid-table and Lower-Table teams like Leicester City, Burnley performed considerably good in comparison to our model results

**Conclusion**

In conclusion, our model was able to perform reasonably well for some teams for this season. However, there are some obvious shortcomings, some other factors which need to be included for better results.

**Future Scope**

The model performance can be further improved:

* By incorporating the injury probabilities happening for a team. This will affect a team’s performance for the season. We can do a **Sensitivity Analysis** to incorporate all possible scenarios for a team (if team is at its full strength or not)
* Using distributions other than **Poisson**. We can compute the **odds ratio** of scoring vs conceding and use **binomial distribution** to predict whether a team has scored goal or not.