

INFO 6205 Ranking System Project

Yuanxin Zhang 001405356

Yongjun Ou 001023025



# Introduction

## What is Ranking System?

xxxxxxxxxx

## About the Premier League

xxxxxxxxxx

### Relevant Terms

Xxxxxxxxxx

### 2018–19 Premier League

xxxxxxxxxx

# Aim of the Project

Designed a Ranking System for the EPL

* Flexible, Dynamic Model from input data source
* Simulate reliable data for predicting the probability to each team pairs fo the result.
* Generating final ranking table with meaningful and logical ranking method.
* Generating table for each team showing their winning probability to other teams.
* EPL season data from EPL official website.

# Project Description

* Possion distribution:

If the value of discrete variable X can be 0,1,2…..and the probability of each value is

A picture containing object, clock

Description automatically generated(  ，we call  X  obey the Possion Distribution with variable  .  is the expectation.

* We have known that the goal of football game match Possion Distribution. Then we using the Poisson distribution model and the average number of goals per game, calculate the theoretical calculation distribution of the number of goals.
* As we have calculated the number of goals through the previous process. And we know the exact distribution of the event that we are going to simulate. We can then do the Monte Carlo method. Randomly generate gaming goals as well as result following Possion Distribution of every , and calculating the probability of win, tie and lose.

# Datasets used in the Project

Json Data Set: season-1920.json (from www.football-data.co.uk/englandm.php)

A group of people

Description automatically generated

# Implementation

Extract every data model & convert into Team Model

Modeling the data to original model data

Read the original data from json

Loop to do 100K simulation through the Possion simulation

Insert data to the getPossionVariable(), get the simulation data

Iterate the data extract from team model

Ranking teams by sorting the gaming result table by (TranScore, Wins, GoalDiff)

Count all team pairs gaming result from their probability

Calculate the result probability for each team pairs

Implementation of get Possion Variable from Lamda

A screenshot of a cell phone

Description automatically generated

Implementation of Monte Carlo method

A screenshot of a social media post

Description automatically generated

Implementation of Sorting

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

Ranking By TranScore

A screenshot of a cell phone

Description automatically generated

Ranking By Game wins

A screenshot of a cell phone

Description automatically generated

Ranking By Game Goal difference

A screenshot of a cell phone

Description automatically generated

Team table (Liverpool)

A screenshot of a cell phone

Description automatically generated

# Result & Analysis

Output csv file:

1. Ranking tables sorting by Number of game wins, number of Goal difference, Transfer Score. (Transfer Score = 1\*wins + 0.5\*tie + 0\* loss)
2. Tables showing for each team. The probability result of gaming with other teams are shown.
3. By analyzing the data table below, EPL season 2019-2020, the team Liverpool ranks #1 on the 3 kinds of final ranking table. But it will lost some games with high probability. The team Norwich lies on the bottom of the ranking tables.

# Reference

[1] <https://en.wikipedia.org/wiki/Ranking>

[2] <https://www.premierleague.com/>

[3] <https://en.wikipedia.org/wiki/2018%E2%80%9319_Premier_League>

[4] <https://en.wikipedia.org/wiki/Poisson_distribution>

[5] <https://en.wikipedia.org/wiki/Monte_Carlo_method>