INFO 6205 Ranking System Final Project

——EPL Search and Ranking System
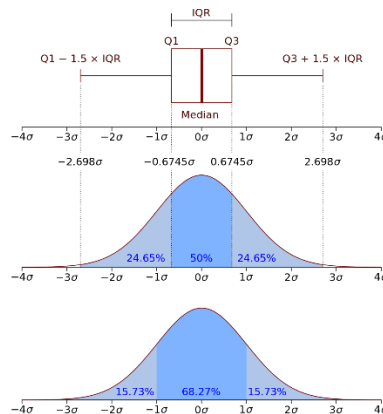
Yujun Zheng

NUID: 001306410

Introduction

Due to the spread of Covid-19, the EPL (English Premier League) has abruptly terminated, this project is to design a system to provide user to look up history match data and use the match data since 2000 to 2020 to predict the result of the terminated 2020 season.

Background knowledge

- What is the PDF(probability density function)?

In probability theory, a probability density function (PDF), or density of a continuous random variable, is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample.

Pic1 a normal PDF function

- EPL basic rule

The English Premier League is the top level of the English football league system. Contested by 20 clubs, it operates on a system of promotion and relegation with the English Football League (EFL). Seasons run from August to May with each team playing 38 matches (playing all 19 other teams both home and away).[1] Most games are played on Saturday and Sunday afternoons.

There are 20 teams in the EPL and each team will play as the home team and away team twice with another team. The winner will get 3 points, the loser will get 0 points. If the two teams draw, both teams will get 1 point. This EPL Ranking System is using this basic rule to rank and sort the team then predict the result of this season.

Class

- EPLRankingSystem.java

  The driver class of this EPL ranking system. Use the Scanner to detect user input.

  ```java
  */
  System.out.println("######Welcome to EPL data search and forecast simulation system!#########");
  Scanner input=new Scanner(System.in);
  while (true){
      System.out.println("Input '1' to search history data.");
      System.out.println("Input '2' to enter 2020EPL Simulation System.");
  ```

- EPLTeam.java

  EPLTeam class is used to store a team's data, including name, wins, loses, draws and points.

  ```java
  public class EPLTeam{
      private String name;
      private int winCount = 0;
      private int loseCount = 0;
      private int drawCount = 0;
      private double teamPoints = 0;
  ```

  Using addWin(), addLose(), addDraw() to update a teams status, automatically update team's point.

  ```java
  public void addWin() {
      this.winCount += 1;
      this.teamPoints += 3;
  }
  public void addLose() {
      this.loseCount += 1;
  }
  public void addDraw() {
      this.drawCount += 1;
      this.teamPoints +=1;
  }
  ```

- EPLMatchResult.java

  EPLMatchResult class is used to store a match result, including below data.

  ```java
  public class EPLMatchResult {
      private String matchDate;
      private String kickOffTime;
      private String homeTeam;
      private String awayTeam;
      private int ftHomeGoals;
      private int ftAwayGoals;
      private String ftResult;
      private int htHomeGoals;
      private int htAwayGoals;
      private String htResult;
  ```

- CSVUtil.java

  ```java
  public static List<EPLMatchResult> historyRead(String filename)
  ```

  Read CSV file through filename and return a list of match history result.

  Using BufferedReader to read data in csv file.

- DataStatistcs.java

  This class has 6 functions, 4 below are using to sort and cauculate the history data.

  They can return the most wins, points team in a history and sort all the team in a history by points or wins.

  ```java
  public static EPLTeam WinMostTeam(List<EPLMatchResult> history)
  ```

```java
public static EPLTeam PointMostTeam(List<EPLMatchResult> history)
public static List<EPLTeam> WinMostSort(List<EPLMatchResult> history)
public static List<EPLTeam> PointMostSort(List<EPLMatchResult> history)
```
below function is used to sort a team list by points:
```java
public static List<EPLTeam> TeamPointSort(List<EPLTeam> teamStatistics)
```
and another private data processing function is to process history data in the class.
```java
private static List<EPLTeam> DataProcessing(List<EPLMatchResult> history)
```
In the function, I use Set to save the name of teams to avoid duplicate!

- DataStatistcsDemo.java
  A demo class to show DataStatistics fuction.
- TwoTeamHistory.java
  historyBetweenTwo function is to search in a history by two team name. return a int array that int[0] means wins of team1, int[1] means wins of team2, int[2] means draws of two team
  ```java
  public static int[] historyBetweenTwo(List<EPLMatchResult> history, String team1, String team2)
  ```
- TwoTeamProbForcast.java
  probBetweenTwo fuction is to cauculate two teams' win or draw probability if they met in a head to head matchup at neutral territory.
  It returns a double array, represents two teams' win rate or draw rate
  ```java
  public static double[] probBetweenTwo(List<EPLMatchResult> history, String team1, String team2)
  ```
- EPLSimulation.java
  ```java
  public static List<EPLTeam> eplSimulation2020(List<EPLMatchResult> history)
  ```
  Using a history to forcast the result of 2019-2020 season.
  Simulation the season for 1000 times and cauculate the average points of each team.

Data

- History Data

This project use the data downloaded from http://www.football-data.co.uk/englandm.php.
Every season has a csv file to store the data.

2000-2001.csv 2010-2011.csv
2001-2002.csv 2011-2012.csv
2002-2003.csv 2012-2013.csv
2003-2004.csv 2013-2014.csv
2004-2005.csv 2014-2015.csv
2005-2006.csv 2015-2016.csv
2006-2007.csv 2016-2017.csv
2007-2008.csv 2017-2018.csv
2008-2009.csv 2018-2019.csv
2009-2010.csv 2019-2020.csv

Program Result

Welcome page:

```
#######Welcome to EPL data search and forecast simulation system!#########
Input '1' to search history data.
Input '2' to enter 2020EPL Simulation System.
Input '3' to quit.
```

Input 1 to search history data:

Eg1. We are search the team with most wins in 2018:

```
#######Welcome to EPL data search and forecast simulation system!#########
Input '1' to search history data.
Input '2' to enter 2020EPL Simulation System.
Input '3' to quit.
1
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
2018
####Show the team with most wins, input 1:
####Show the team with most points, input 2:
####Show the team sorted with wins, input 3
####Show the team sorted with points, input 4
1
Team: Man City
        Points: 98.0    WinMatches: 32
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
```

Eg2. We input a wrong year or wrong number:

```
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
2025
Illegal Input! Please input again
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
```

Eg3. We are searching the team rank by points in 2005:

```
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
2005
####Show the team with most wins, input 1:
####Show the team with most points, input 2:
####Show the team sorted with wins, input 3
####Show the team sorted with points, input 4
4
Team: Chelsea
        Points: 91.0    WinMatches: 29          Team: Fulham
Team: Man United                                        Points: 48.0    WinMatches: 14
        Points: 83.0    WinMatches: 25          Team: Charlton
Team: Liverpool                                         Points: 47.0    WinMatches: 13
        Points: 82.0    WinMatches: 25          Team: Middlesbrough
Team: Arsenal                                           Points: 45.0    WinMatches: 12
        Points: 67.0    WinMatches: 20          Team: Man City
Team: Tottenham                                         Points: 43.0    WinMatches: 13
        Points: 65.0    WinMatches: 18          Team: Aston Villa
Team: Blackburn                                         Points: 42.0    WinMatches: 10
        Points: 63.0    WinMatches: 19          Team: Portsmouth
Team: Newcastle                                         Points: 38.0    WinMatches: 10
        Points: 58.0    WinMatches: 17          Team: Birmingham
Team: Bolton                                            Points: 34.0    WinMatches: 8
        Points: 56.0    WinMatches: 15          Team: West Brom
Team: West Ham                                          Points: 30.0    WinMatches: 7
        Points: 55.0    WinMatches: 16          Team: Sunderland
Team: Wigan                                             Points: 15.0    WinMatches: 3
        Points: 51.0    WinMatches: 15
Team: Everton
        Points: 50.0    WinMatches: 14
```

Input 0 to back to main menu:

```
###EPL history data System! Input '0' to back to main menu
Input years to search the data of that year(2000-2019)
0
Input '1' to search history data.
Input '2' to enter 2020EPL Simulation System.
Input '3' to quit.
```

Input 2 to simulation: (Need to wait some time, depending on the cauculation ability of the computer):

We can see that there are 20 teams in the list.

The final rank forcast is as below: (Because we simulation 1000 tims and cauculate the average points, so the team points are double!)

```
#########2020EPLSimulation#########
All team:
[Liverpool, Brighton, Norwich, Sheffield Unite
##Result##
Team: Liverpool
        Points: 83.65   WinMatches: 25
Team: Man United
        Points: 81.725  WinMatches: 24
Team: Chelsea
        Points: 80.364  WinMatches: 23
Team: Man City
        Points: 79.11   WinMatches: 23
Team: Arsenal
        Points: 75.105  WinMatches: 21
Team: Tottenham
        Points: 67.755  WinMatches: 19
Team: Leicester
        Points: 58.506  WinMatches: 16
Team: Everton
        Points: 57.396  WinMatches: 14
Team: Sheffield United
        Points: 56.666  WinMatches: 15
Team: Southampton
        Points: 56.501  WinMatches: 14
Team: Wolves
        Points: 51.646  WinMatches: 12
Team: Newcastle
        Points: 51.309  WinMatches: 12
Team: Aston Villa
        Points: 48.626  WinMatches: 12
Team: Burnley
        Points: 48.521  WinMatches: 13
Team: Crystal Palace
        Points: 46.454  WinMatches: 12
Team: West Ham
        Points: 45.499  WinMatches: 11
Team: Watford
        Points: 44.023  WinMatches: 11
Team: Brighton
        Points: 43.258  WinMatches: 9
Team: Bournemouth
        Points: 42.849  WinMatches: 10
Team: Norwich
        Points: 33.159  WinMatches: 7
Input '1' to search history data.
Input '2' to enter 2020EPL Simulation System.
Input '3' to quit.
```

Conclusion

By simulate the 2020 EPL season, we predict that Liverpool will be the Champion.
The rank is below:

Rank	Team	Points	Wins
1	Liverpool	83.65	25
2	Man United	81.725	24
3	Chelsea	80.364	23
4	Man City	79.11	23
5	Arsenal	75.105	21

6	Tottenham	67.755	19
7	Leicester	58.506	16
8	Everton	57.396	14
9	Sheffield United	56.666	15
10	Southampton	56.501	14
11	Wolves	51.646	12
12	Newcastle	51.309	12
13	Aston Villa	48.626	12
14	Burnley	48.521	13
15	Crystal Palace	46.454	12
16	West Ham	45.499	11
17	Watford	44.023	11
18	Brighton	43.258	9
19	Bournemouth	42.849	10
20	Norwich	33.159	7

Reference

1. https://en.wikipedia.org/wiki/Probability_density_function#/media/File:Boxplot_vs_PDF.svg
2. https://en.wikipedia.org/wiki/Premier_League
3. http://www.football-data.co.uk/englandm.php.