

# EA Sports' FIFA

## Problem statement

You are given the following public dataset from <https://sofifa.com>. The data lists soccer players from the 2015-2020 editions of EA Sports FIFA video game. You are tasked with 4 different objectives for this project using these

## Content

- Every player available in FIFA 15, 16, 17, 18, 19, and also FIFA 20
- 100+ skill attributes
- Player positions, with the role in the club and in the national team
- Player attributes like: Attacking, Skills, Defense, Mentality, GK Skills, etc.
- Player personal data
  - Nationality
  - Club
  - Date Of Birth
  - Wage
  - Salary
  - Height
  - Weight
  - Preferred foot
  - Real Face
  - Url

Note: club changes by these players can be tracked using the unique Sofifa id across all years. The problem statements are as follows:

## Part A - Rank the Players

- Use data from FIFA 19. Predict the **Overall** (OVR ) skill Statistic for players in the FIFA 20 Edition. Train on all players from FIFA 19
- Graphically represent the **Overall** data for the players in Test Set

### **Instructions :**

- Sofifa Id, URL, Name, Potential, Real face, URL variables cannot be used during the prediction or learning.
- In case of a change in playing position, alter FIFA 20 position to match FIFA 19 Position
- The final score will be assigned using MSE.
- Test Set: Players from the following **Clubs**
  - FC Barcelona
  - FC Bayern München
  - Real Madrid
  - Paris Saint-Germain
  - Juventus
  - Manchester City
  - Liverpool FC

## Part B - Classify player position

- Train on data from FIFA 19. Predict the **player\_position** variable using other skill statistics for players in the FIFA 20 Edition. Train on all players from FIFA 19
- Graphically represent the predicted **player\_position** for the players in Test Set

### **Details :**

- Sofifa Id, URL, Wage, Salary, Name, Real face, URL variables cannot be used during the prediction or learning. \*
- The final score will be assigned using **Classification accuracy**
- Test Set: Players from the following **Clubs**
  - FC Barcelona
  - FC Bayern München
  - Real Madrid
  - Paris Saint-Germain
  - Juventus
  - Atlético Madrid
  - Manchester City
  - Liverpool
  - Manchester United

## Part C - Which Club has the Best Staff

- Study player data from *Division 1 European League\** players from the last 5 Years. Analyze changes in player stats and value. Rank the clubs according to best increase in statistics of a player
- Graphically represent the scores for the test set.

### **Details :**

- Sofifa Id, URL, Wage, Salary, Name, Real face, URL variables cannot be used during the prediction or learning.
- Assign a score to all clubs out of 100, and represent it in an appropriate visualization. With 100 being the highest
- For the test set, your model will be scored using MSE against the average of all models of the class. (Score =  $MSE(\text{Your Values}, \text{Average of all models values})$ )
- The test set is as follows for this problem statement (Division 1 European League):

#### **League =**

- Premier League - English Premier League
- Bundesliga German 1. Bundesliga
- Ligue 1 French Ligue 1
- La Liga Spain Primera Division
- Serie A Italian Serie A

## Part D - How will things be in 2021

- Going through data for player skill changes, predict the new **skill stats** for the test set players in the 2020-21 season

### **Restrictions :**

- Sofifa Id, URL, Wage, Salary, Name, Real face, URL variables cannot be used during the prediction or learning.
- For the test set, your model will be scored using MSE against the average of all models of the class. (Score =  $MSE(\text{Your Values}, \text{Average of all models values})$ )
- The test set is as follows for this problem statement (Division 1 European League\*):

#### **League =**

- Premier League - English Premier League
- Bundesliga German 1. Bundesliga
- Ligue 1 French Ligue 1
- La Liga Spain Primera Division
- Serie A Italian Serie A