



# Predicting the Potential of players in FIFA 2019

Chandan Arora

May 18, 2019

## 1. Introduction

Football has a huge fanbase all over the world. Every year, after the end of season, the transfer window allows different clubs to buy, sell or loan players to/from other clubs. Also, the betting agencies start placing their bets on whether a player will complete his rumoured transfer to the club in picture. Before spending millions on a player, the clubs analyse the player's past performance and his future potential to make sure if he will be worth the hassle or not. EA Sports' FIFA 19 is the latest version of their football simulation game. FIFA provides ratings of the players based on the performance in the past season, and his potential based on attributes like passing accuracy, dribbling, crossing, finishing, height, weight, etc. Different clubs would certainly want to predict the potential of the player before finalizing a deal to get him to their club.

### • Problem

The aim of this project is to predict the potential of a player based on the data present in the dataset. The dataset contains the details of players, their nationality, and other attributes such as dribbling, acceleration, stamina, shot-accuracy, etc.

### • Interest

Football clubs around the world want in-depth analysis before putting in a bid for the player in question. The scouting teams from different clubs scout players extensively before recommending a player to the club. The clubs would, therefore, be very interested in predicting the potential of a player before buying.

## 2. Data Source

The players' data for FIFA 19 can be found on [kaggle.com](https://www.kaggle.com). The complete dataset was downloaded from this [link](#). This dataset contains the players' details with attributes that would be useful in predicting the potential of the player.

### • Data Cleaning

The dataset contains complete details of the players attributes such as age, preferred foot, weak foot, wages, skill moves, crossing, finishing, stamina, header accuracy, shot accuracy, etc. Some of the attributes such as stamina, strength, acceleration have a few null entries. These null entries have been replaced by the mean of the attribute to remove any discrepancy.

### • Feature Selection

A few attributes such as body type, face, flag, etc. have been removed as they will not be used to predict the potential of the player. The features that have been chosen to predict the accuracy potential are age, overall, international reputation, stamina, strength, aggression, composure, ball control, dribbling, acceleration.

```
[427]: df_new=df[['ID','Name','Age','Overall','Potential','International Reputation', 'Stamina','Strength','Aggression','Composure','BallControl', 'Dribbling','Acceleration']]
[428]: df_new.head()
```

	ID	Name	Age	Overall	Potential	International Reputation	Stamina	Strength	Aggression	Composure	BallControl	Dribbling	Acceleration
0	158023	L. Messi	31	94	94	5.0	72.0	59.0	48.0	96.0	96.0	97.0	91.0
1	20801	Cristiano Ronaldo	33	94	94	5.0	88.0	79.0	63.0	95.0	94.0	88.0	89.0
2	190871	Neymar Jr	26	92	93	5.0	81.0	49.0	56.0	94.0	95.0	96.0	94.0
3	193080	De Gea	27	91	93	4.0	43.0	64.0	38.0	68.0	42.0	18.0	57.0
4	192985	K. De Bruyne	27	91	92	4.0	90.0	75.0	76.0	88.0	91.0	86.0	78.0