Technical Project Proposal

Title: Build your dream team - FIFA 2019

Author: Sibin Sygal

Supervisor: Shagufta Henna

Degree: Masters in Big Data Analytics & Artificial Intelligence

Problem Description

Soccer or Football is not only a game for many its an emotion. Football Association & EA Sports have always been the top selling brand for football simulation video game FIFA. No less than their religion, people follow their favorite clubs.But not everyone has the idea of how much the player's are paid or how much they are valued in market. A model will be created using a data set used to compute attributes of a FIFA 19 player, the main objective is to see the correlation between multiple variables to drive the overall rating of a player. Once we have a high level of confidence with the model, we will use it to predict overall rating for data set available.By comprehensive football experience: comprehension and contextualized knowledge encourage users to play FIFA smartly, choose a better team to the Fantasy Premier League, or improve their betting chances

Data Description

Dataset URL : https://www.kaggle.com/karangadiya/fifa19

The Dataset is designed for people who enjoy data science and have grown up playing soccer and are passionate about FIFA.Dataset contains information from the latest FIFA 19 version of all 18207 players. There are 89 characteristics including personal information such as age, gender, ethnicity, image, team, salary, etc., as well as knowledge about player skills such as ball control, dribbling, passing, scoring, GK skills, etc. It is possible to derive from the dataset many observations and associations between player value, salary, age, and performance.We have an integer of 0 to 100 for each attribute, which measures how good a player is at that attribute. The dataset description from Kaggle is as follows

"Dataset includes latest edition FIFA 2019 players attributes like Age, Nationality, Overall, Potential, Club, Value, Wage, Preferred Foot, International Reputation, Weak Foot, Skill Moves, Work Rate, Position, Jersey Number, Joined, Loaned From, Contract Valid Until, Height, Weight, LS, ST, RS, LW, LF, CF, RF, RW, LAM, CAM, RAM, LM, LCM, CM, RCM, RM, LWB, LDM, CDM, RDM, RWB, LB, LCB, CB, RCB, RB, Crossing, Finishing, Heading, Accuracy, ShortPassing, Volleys, Dribbling, Curve, FKAccuracy, LongPassing, BallControl, Acceleration, SprintSpeed, Agility, Reactions, Balance, ShotPower, Jumping, Stamina, Strength, LongShots, Aggression, Interceptions, Positioning, Vision, Penalties, Composure, Marking, StandingTackle, SlidingTackle, GKDiving, GKHandling, GKKicking, GKPositioning, GKReflexes, and Release Clause".[1]

Methodology

Inorder to have an in-depth analysis on FIFA 19 player statistics , Apache Spark is used as the Big Data Analytics framework. Apache Spark is a cluster-computing lightning-fast Big Data development system. It is designed to rapidly quantify the encoding of large data sets. Python will be used as the programming language for the analysis.Spark operates separately or in the cloud on Hadoop, Apache Mesos, Kubernetes. It can connect to various sources of data. Planned analytical algorithms to be used are as follows

1. Prediction of Overall Rating - Linear Regression
2. Predicting player value - Random forest regression
3. Finding players with similar skills/potential - k-means clustering

Goals

The main goals that has been focussed in this technical research are as follows

1. Check out which teams are most likely to have the highest potential.
2. Find out which teams are youngest / oldest based on players
3. Try to find "negotiations"; that is, if someone has the same skills / potentials, can they be identified for a deal?
4. Could we estimate the ability of a player on the basis of his skills?
5. Build a dream team based on the formations
6. Playing postion prediction using player statistics

References

[1] ‘Apache SparkTM - Unified Analytics Engine for Big Data’. [Online]. Available: https://spark.apache.org/. [Accessed: 22-Dec-2019]. [2] ‘FIFA 19 complete player dataset — Kaggle’. [Online]. Available: https://www.kaggle.com/karangadiya/fifa19. [Accessed: 22-Dec-2019].