Player Valuation and Segmentation in EA FC 25

A Proposal report for the BDM capstone Project

Submitted by

Name: OAS PODDAR

Roll number:23F3003433

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project Title "**Player Valuation and Segmentation in EA FC 25**". I extend my appreciation to **Electronic Arts**, for providing the necessary resources that enabled me to conduct my project.

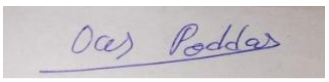
I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate: 

Name: OAS PODDAR

Date: 11/10/2025

1 Executive Summary

Electronic Arts is an online gaming company which operates globally around the world that develops and provides sports-based video games such as EA SPORTS FC. It focuses on creating realistic and data-driven gaming experiences for millions of players worldwide through advanced technologies and player analytics.

The organization faces challenges in valuing players accurately and segmenting them based on their attributes and performance. If the ratings of the players do not reflect their performance as in the real world, it can lead to unfair player comparisons and unbalanced gameplay. In order to analyze the player statistics, we need to identify key factors which affect the overall ratings and also to group similar players effectively.

This project aims to apply data analytics techniques such as correlation analysis, regression modeling, and clustering on the EA FC 25 player dataset. Such techniques will help in improving the relationships between the players attributes and their ratings and can group the similar players in an effective manner. The outcome helps in valuing the players more accurately, balanced gameplay and improved decision making in future EA SPORTS releases.

2 Organization Background

Electronic Arts (EA SPORTS) is one of the world's leading gaming companies, which is known for creating realistic sports video games like EA SPORTS FC, Madden NFL, and F1. It is founded in 1982 and headquartered in Redwood City, California, EA SPORTS focuses on building an engaging digital platform for users who enjoy football, basketball, and car racing games. The company uses real player data, modern graphics, and innovative technologies to make the gameplay feel as close to real-life sports as possible. It also updates player ratings, skills, and team dynamics regularly based on real-world performances, which keeps the games exciting and authentic. EA SPORTS has built a strong global community of players and fans by constantly improving its games and offering an interactive experience.

3 Problem Statement

3.1 **Player Valuation Analysis**: Many players in EA SPORTS FC 25 appear to be undervalued or overvalued. Due to this inconsistency, the attributes of the players do not reflect the true rating of each player which makes it difficult for the users to make informed decisions in team building and transfers.

3.2 **Player Segmentation for Performance Insights**: EA SPORTS FC 25 lacks a structured approach to group players based on their attributes and ratings which helps in facilitating better talent identification, tactical role assignment, and balanced team formation strategies.

4 Background of the Problem (200 Words)

The two problems stated in the Problem statement arises due to the improper and ineffective utilization of the large and detailed dataset available in the database of EA SPORTS FC 25 database. Despite the information is given in detailed form such as player ratings, attributes, and performance metrics, the dataset in not being analyzed in a systematic manner in order to derive meaningful insights about player potential, team efficiency or overall game realism. Due to the absence of structured analytical methods, it limits the ability of identifying trends, compare players, or optimize gameplay balance based on accurate data.

Major cause: The main cause is the lack of integrated data management and analytical frameworks which helps in transforming raw data into meaningful insights that reflect real world player performance.

Internal Problems

1. Inconsistent player updates: The player ratings and statistics are not being updated on a regular basis which causes data inaccuracy across different seasons.
2. Lack of data integration: The connection between player form, match outcomes and attributes is missing due to which it reduces the depth of analysis.
3. Outdated analytical tools: The absence of automated tracking systems leads to outdated and incomplete insights.

External Problems

1. Real-world performance variations: There are many situations due to which there is variation in the players' performance which causes data inaccuracy.

2. Fan and media influence: Many a times due to the views of public over a particular player there is some biasness on the players ratings and updates.
3. Dynamic football environment: The evolution of football tactics and incoming of new players in the football world requires adjustments in the database to maintain realistic data.

5 Problem Solving Approach

Under this section we will be discussing about how we will be addressing the key challenges of player valuation and segmentation in EA SPORTS FC 25 database where a structured and data driven problem solving approach is being followed. The approach is divided into multiple stages such as data collection, preprocessing, exploratory data analysis, model development, and interpretation of results. Following are the sequential steps for the approach to be used in solving the problem statements discussed before:

1. Data collection and understanding: The first step to be under taken is to collect the dataset of EA SPORTS FC 25 which includes player ratings, attributes, position, clubs and performance statistics. The dataset will be examined carefully to understand the relationships between the features of the dataset such as passing, dribbling, shooting, pace and defense which describes the overall ratings of the player. If any missing, duplicate or inconsistent data is found, that will be identified for correction.

2. Data preprocessing and cleaning: After the dataset is collected, the next step is to clean the dataset by handling missing values, removing duplicates, normalizing numeric attributes, and encoding categorical variables like nationality, position, and club. The cleaned and preprocessed data will be prepared for further analytical modelling.

3. Exploratory Data Analysis: With the completion of task in the cleaning and preprocessing of data, our next step is to conduct data analysis on this dataset. For the analysis we will using an analysis approach which is Exploratory Data Analysis which helps in identifying correlations between player attributes and their overall ratings. Various kinds of visualization tools such as histograms, scatter plots, and correlation heatmaps will be used to detect patterns and outliers.

4. Predictive modeling for player valuation: After the analysis of the data is being done, our next step is to use regression techniques such as linear regression or random forest regression to predict player valuation or overall rating based on key performance indicators (KPI). The model will be trained and tested using appropriate data splits to ensure reliability. In order to determine which factors will contribute most to player valuation accuracy, feature importance analysis approach will be used.

5. Clustering for Player Segmentation: After the subsequent completion of training and testing the data on regression techniques in the previous step, the next step is to group the similar players using unsupervised learning methods like K-means or Hierarchical clustering is used. The players will be segmented based on their attributes such as pace, physicality and skill level. Due to this segmentation, it will help in identifying distinct profiles of each player such as attacking, defensive and balanced which will be helpful in building a balanced team and player comparison.

6. Evaluation and Insights: The final step in discussing about approach to solve the problem statements is the evaluation of the models based on performance metrics such as R^2 score, Mean Squared Error (for regression), and Silhouette Score (for clustering). The insights derived will enable EA SPORTS to make data-backed adjustments to player ratings and improve fairness in gameplay.

6 Expected Timeline

The dataset is chosen from the official website of Kaggle which is the database of EA SPORTS FC 25. After the dataset is acquired from the Kaggle the dataset is being cleaned for further analysis. Subsequently regression and clustering will be done for player valuation and segmentation. In addition to it data visualization would have a major role in analyzing the dataset. In the final stage the model built will be evaluated to generate meaningful insights valuable to the EA SPORTS in improving its gameplay. Below is the Gantt chart as well as flow chart prepared to show timelines and workflow of our project.

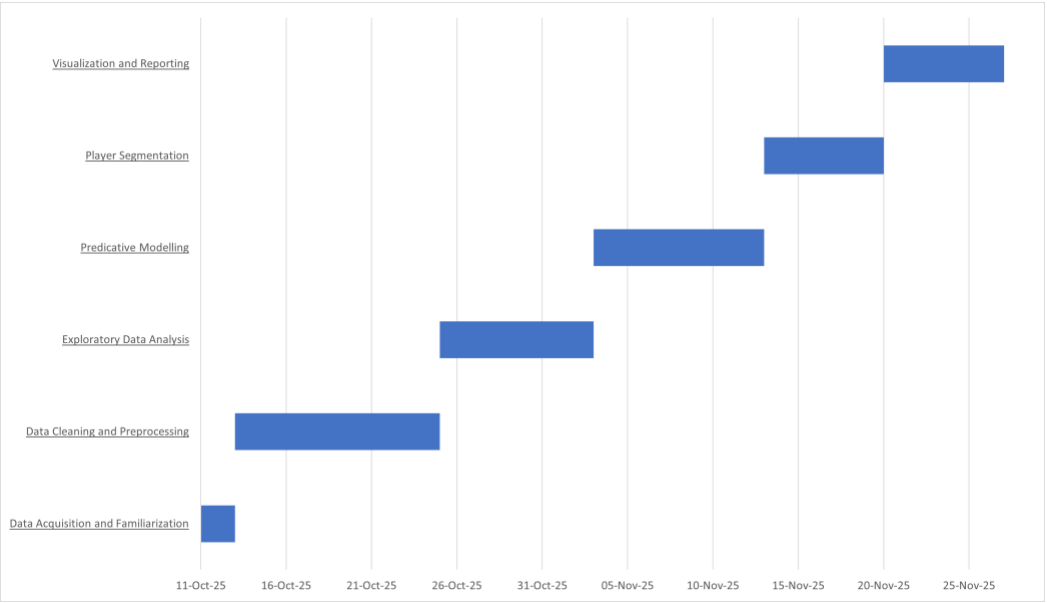


Figure 1 Gantt chart showing the start and duration of each project phase.

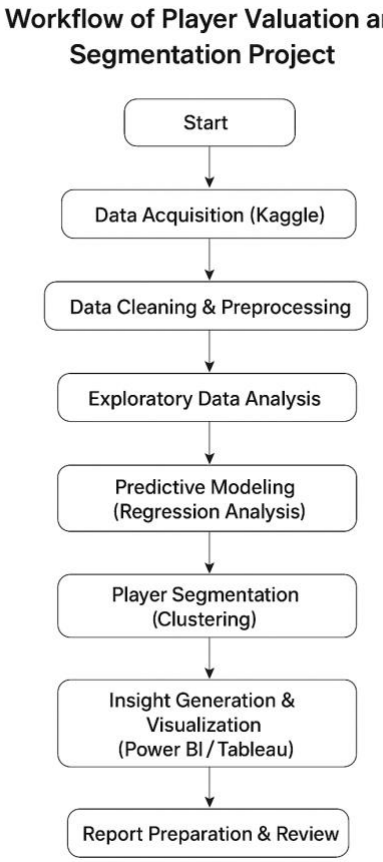


Figure 2 Workflow of the project

7 Expected Outcome

1. Accurate player valuation: This project will help us in building a model which will be able to estimate the players' ratings and market value accurately based on Key Performance Attributes (KPAs), reducing overvaluation and undervaluation in EA SPORTS FC 25.

2. Effective player segmentation: With the help of clustering, players will be grouped into distinct categories such as attacking, defensive, and balanced types, which helps in better comparison, team formation and tactical decisions.

3. Actionable insights and visualization: The final outcome will include visual dashboards and analytical insights that will help EA developers and analysts to enhance the gameplay balance and improve future player ratings systems.