DSA-210 FINAL REPORT

EDES RIZA CAKMAK

30779

FIFA 25 PLAYER RATING ANALYSIS:

TEAM PERFORMANCE & STATISTICAL IMPACT

SUPERVISED BY

ÖZGÜR ASAR

SABANCI UNIVERSITY

FIFA 25 Player Ratings Analysis Report

1. Project Objective

This project aims to evaluate whether the highest-rated players in EA SPORTS FC 25 (top 375 overall ratings) are concentrated in teams that performed well in the 2023/24 season. Furthermore, it investigates which in-game stats contribute most significantly to a player's overall rating. The results are obtained using both statistical correlation and machine learning approaches.

2. Methodology

To conduct this study, we utilized two datasets:
- A dataset listing the top 375 players in FIFA 25, including their club, position, and key in-game attributes (Pace, Shooting, etc.).
- A dataset summarizing club performance statistics from the 2023/24 season, including Wins, Goals Scored, and Total Points.

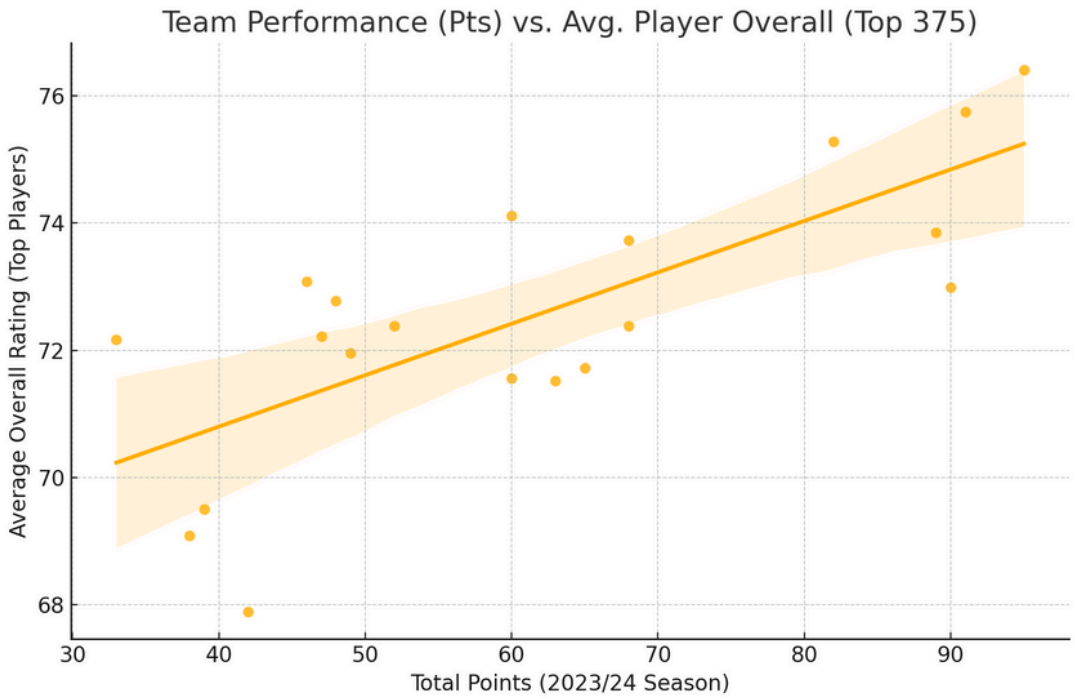
We performed the following steps:
1. Calculated the average overall rating per club based on top-rated players.
2. Merged this with team performance data (from Bundesliga, La Liga, Premier League, Serie A, and Ligue 1).
3. Visualized the relationship between team success and player ratings using regression plots.
4. Trained a Random Forest Regressor to determine which in-game stats most heavily influence the 'Overall' rating.

3. Are Higher-Rated Players on Better Teams?

We hypothesize that teams with better performance metrics (e.g., higher points or goal difference) will also host more highly-rated players in the game. We calculated each teams average player rating and plotted it against their league points.

The scatterplot below illustrates this relationship:

FIFA 25 Player Ratings Analysis Report



As the graph shows, while there is some positive correlation between team performance and average overall rating of top players, the relationship is not perfectly linear. Clubs like Real Madrid and Manchester City score highly both in terms of league performance and player ratings, supporting the hypothesis. However, some clubs like Leverkusen show exceptional team performance with slightly less representation in the top 375 player ratings, possibly due to their players not being as individually famous or well-distributed across FIFA's ranking.

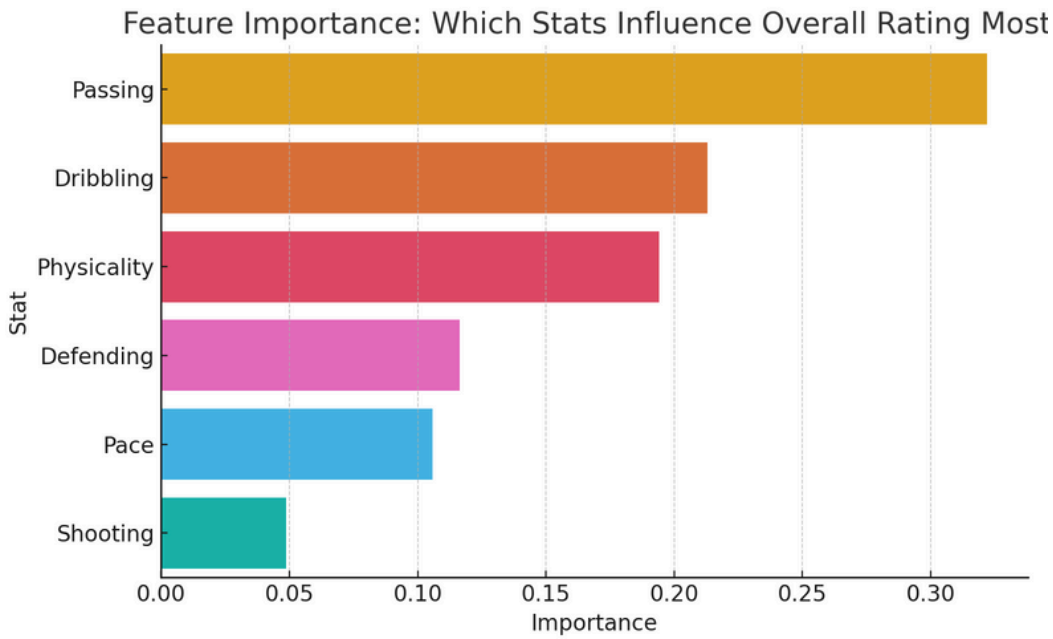
4. Which In-Game Stats Affect Overall Ratings?

We applied a Random Forest Regressor to identify which features (Pace, Shooting, Passing, Dribbling, Defending, Physicality) are most predictive of a players Overall rating. Feature importance scores were calculated using sklearns ensemble.RandomForestRegressor.

We used 80% of the data for training and 20% for testing. The importance score represents the relative contribution of each feature to predicting the target variable.

Below is the feature importance chart based on the model:

FIFA 25 Player Ratings Analysis Report



The model indicates that 'Passing' and 'Dribbling' are the most significant contributors to the overall rating among top players. 'Shooting', surprisingly, had the least impact in this elite group, potentially because creative, agile players are more favored in modern football tactics.

5. Case Comparisons: Mbappe, Haaland, Bellingham

- Kylian Mbappe, with the highest overall rating, plays for Real Madrida team that topped La Liga and performed well in the Champions League.
- Erling Haaland also features highly, and Manchester City were among the best-performing teams in England.
- Jude Bellingham transferred to Real Madrid and had a breakout season, which matches his elevated FIFA rating.

These examples suggest a strong alignment between individual performance, club success, and in-game ratingat least for elite-level players.

6. Limitations and Future Work

This study focuses exclusively on the top 375 players in FIFA 25, which may introduce a selection bias favoring elite clubs and world-renowned players. As a result, the findings might not generalize to lower-ranked teams or emerging talents.

Additionally, team performance data was analyzed at the aggregate level. Individual performance data (e.g., goals, assists, minutes played) for each player was not directly integrated. Including such data could lead to more nuanced conclusions regarding rating accuracy.

Future work may involve:

- Expanding the dataset to include all FIFA-licensed players
- Integrating per-player real-world statistics
- Applying more advanced models like XGBoost or SHAP for better explainability
- Comparing with historical FIFA versions to observe rating evolution over time

7. Ethical Considerations and AI Assistance

All data used in this project was publicly available and cited accordingly. No private or sensitive information was used.

AI assistance via OpenAIs ChatGPT was used only for improving the structure of the code and refining the language used in this report. All data preprocessing, statistical analysis, and modeling steps were conducted by the student. The AI was not used to make analytical decisions, select methods, or evaluate the findings.

Appendix: How the Graphs Were Generated

Graph 1: Team Performance vs. Average Player Overall

This graph visualizes the relationship between a team's total points in the 2023/24 season and the average FIFA 25 overall rating of its top-rated players.

Steps:

1. Calculated each player's 'Overall' score by averaging six core in-game attributes: Pace, Shooting, Passing, Dribbling, Defending, and Physicality.
2. Grouped players by their team and calculated the average overall rating for each club.
3. Merged this data with the team's real-life season performance (Points).
4. Plotted a regression line using seaborn.regplot to observe correlation.

Code Used:

```
players_df['Overall'] = players_df[['Pace', 'Shooting', 'Passing', 'Dribbling', 'Defending', 'Physicality']].mean(axis=1)
team_avg = players_df.groupby('Team')['Overall'].mean().reset_index()
merged = pd.merge(team_data, team_avg, left_on='Squad', right_on='Team')
sns.regplot(data=merged, x='Pts', y='Overall')
```

Graph 2: Feature Importance of In-Game Stats

This graph ranks the in-game attributes based on their importance in predicting a player's overall rating using a machine learning model.

Steps:

1. Selected six features: Pace, Shooting, Passing, Dribbling, Defending, Physicality.
2. Used RandomForestRegressor from sklearn to fit a model on those features against the 'Overall' score.
3. Retrieved feature importances from the model to determine which stats contribute most to the rating.