Sports Analytics Competition- SME, IITJ

Team Name: Prometheus

Members:
1) Samanvay Lakhotia (Team Leader)
2) Nishant Jain
3) Akshat Srivastava

Problem Statement 1

Problem 1:

L. Messi's contract with the club is ending in 2021 and he is assumed to leave the club once this season ends. The coach is having a great pain in deciding whom to select for the position as there are some important matches ahead and as a coach of a leading club, he is responsible to win the La Liga title for the club and the fans.

The coach wants to select some young players (upto 26 years of age) suitable for the position. Based on your analysis of the data, suggest 2-3 players for the position and give the reasoning for the same.

Important consideration points:

1) Lionel Messi is a regular and a very important player **on the field** for FC Barcelona. We need someone to replace him in the team immediately, someone who can take to the pitch immediately.
2) FC Barcelona being a pretty big club, wouldn't mind paying for the transfer clause of excellent players with capabilities like Messi. Hence, their current contract validity won't really matter.
3) Young players required (Age ≤ 26)

Our recommendations:

1. Paulo Dybala

2. Marco Asensio

3. James Rodríguez

Methodology of generation of our recommendations:

Messi is a Right Wing player by position, but plays mostly in attacking zones, scoring goals. For his replacement, we'd desire the same characteristics: Attacking and Midfield presence.

We **generated a new metric** for **"Difference from Messi"** * based on the ratings for RF, RW, RS, RAM, and the various ratings for a player's skill (only related to Attacking and midfielding).

We found that these 3 players had the least difference from Messi in comparison to all the other players. Hence, we select these players for our final recommendation from the entire dataset.

* : For more details, you may see our .ipynb file.

Problem Statement 2

Problem 2:

The team management is also interested in having an idea of how much the wage the club should fix for the incoming players (that you will be suggesting). The club is having a tradition to pay players based on their Overall Score and at par with other players playing in La Liga with similar ability. You have to help the management in deciding the wage for the players.

Important Points to note:

1) Our analysis is only for Wages of La Liga players.

1) We found out that Wage is biquadratically dependant on Overall Score, and linearly dependant on Value of a player, their release clause and their age (years of experience would be a better term).

1) Also, Wage of a player increases with their International Reputation, which is as expected, since clubs would be willing to pay higher amounts for better players of International repute.

Methodology for recommendation of Wages:

We make 2 ML models:

1) Polynomial of degree 4: B/w Wage and Overall Score.
2) Multiple Linear Regression: B/w Wage and Value, Release Clause, Age.

Next, we **ensemble** the learnings of these 2 models, to get a final model with **better predictive power** (without overfitting), by taking the best of both models!

Predicted Wages for our players:

The predicted values of Wages for the selected players are:

1) **Paulo Dybala : €419K**
2) **Marco Asensio : €208K**
3) **James Rodríguez : €236K**

Problem Statement 3

Problem 3:

The club management will also be interested in fixing a transfer clause for the players (that you will be suggesting). Generally clubs fix transfer clause amount based on the player's Potential. You have to assist them in taking the decision.

Important Points to note:

1) We found that Transfer Clause is biquadratically dependant on the Potential of the player. As the problem statement also requires using the Potential, we pay special emphasis on it here.

1) Also, Transfer Clause is linearly dependant on Value of a player, and their Wage.

1) Here, we'll use the complete dataset, not limited to any particular League, as we did in case of Problem Statement 2.

Methodology for recommendation of Transfer Clause:

We again make 2 ML models:

1) Polynomial of degree 4: B/w Transfer Clause and Potential.
2) Multiple Linear Regression: B/w Transfer Clause and Value, Wage of a player.

Next, we **ensemble** the learnings of these 2 models, to get a final model with **better predictive power** (without overfitting), by taking the best of both models!

Predicted Transfer Clauses for our players:

The predicted values of Transfer Clauses for the selected players are:

1) **Paulo Dybala : €161M**
2) **Marco Asensio : €112M**
3) **James Rodríguez : €107M**

Thank you!