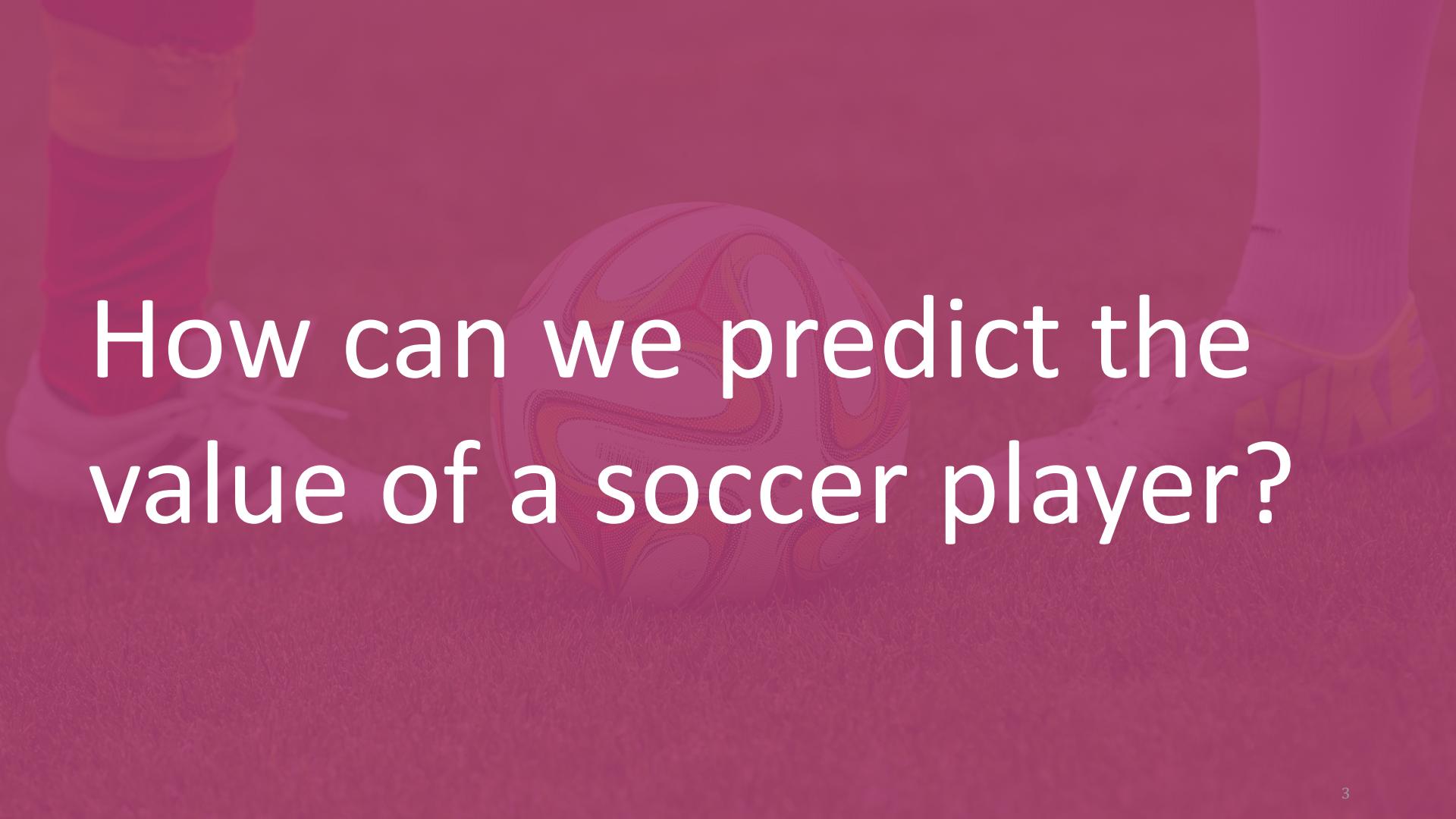


Predicting Soccer Player Values

Elaborated by Elvis Davalillo
July 2019

Summary

- Predict the value of a soccer player depending of his features/performance (+35)
- Analysis included the 3 largest soccer leagues by revenue in the world:
 - English Premier League
 - Spanish Primera Division
 - Italian Serie A

A photograph showing the lower legs and feet of two soccer players in motion. One player is wearing red socks and white shoes, while the other is wearing blue socks and yellow shoes. A red and white soccer ball is positioned between them on a green grass field.

How can we predict the
value of a soccer player?

Web-scrappling

Scraped data from 2,000 soccer players from sofifa.com.

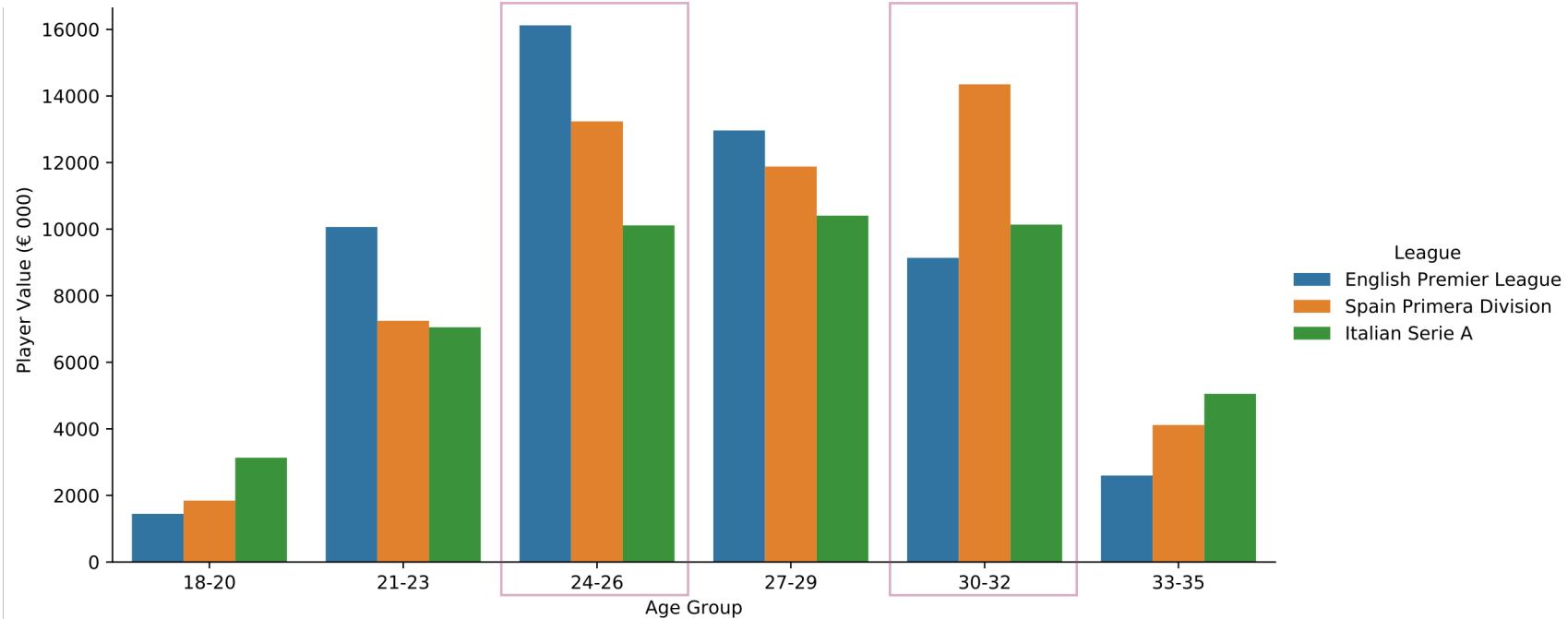
EDA

Analyzed +35 features that included age, position, etc.

Modelling

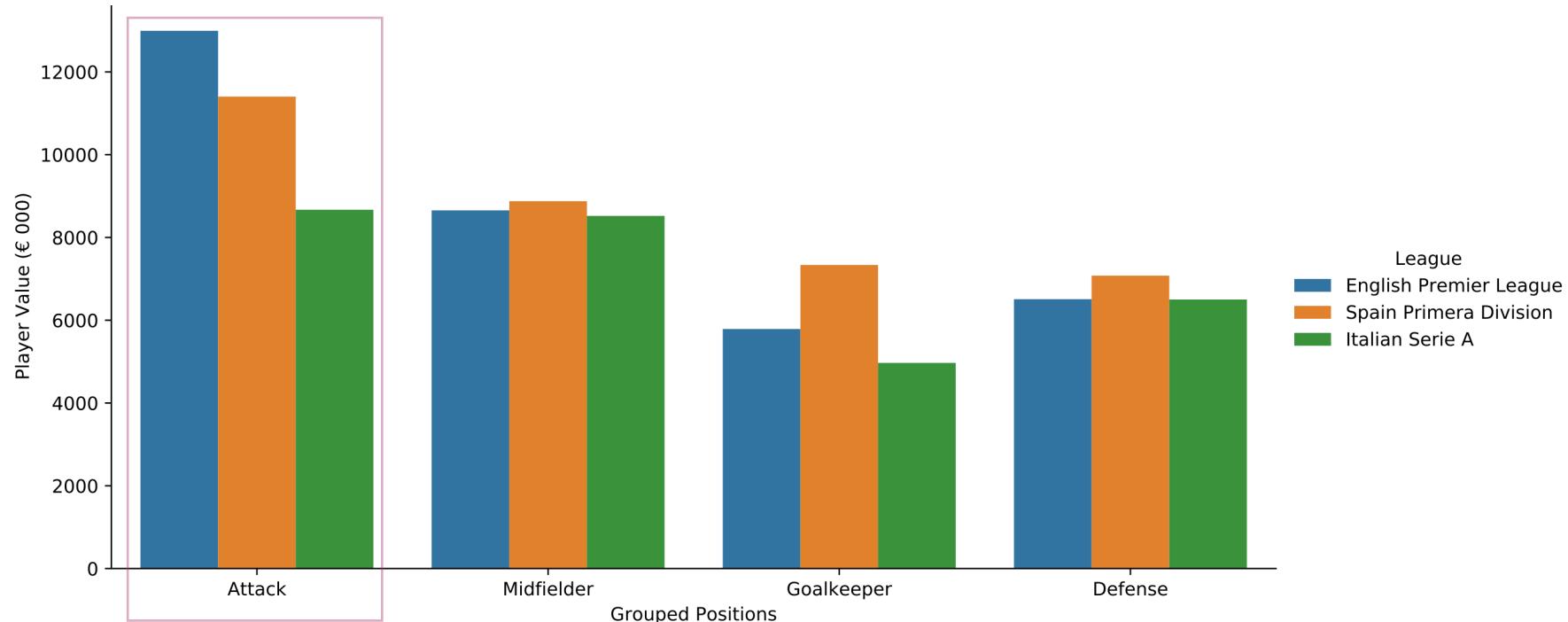
Followed an iterative process to select, fit, validate, and score prediction models

Let's understand the variable we want to predict: Player values by age



Data Source: sofifa.com

Player values by position: Attackers are more expensive than the rest



Data Source: sofifa.com

Because of differences
between leagues,
different models for each
league were created

Premier League

Linear regression with 40
coefficients

$R^2: 0.87$

Position was more predictive
than in other leagues

Spanish Primera

Linear regression with 40
coefficients

$R^2: 0.89$

Features related to
movement, such as
acceleration were more
predictive than in other
leagues

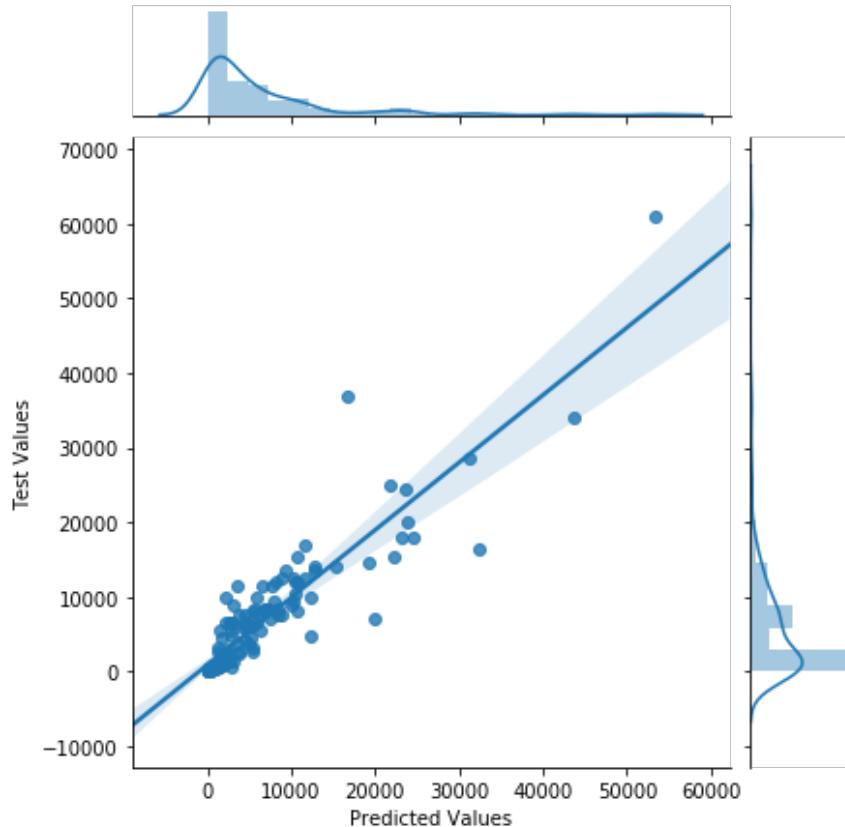
Italian Serie A

Linear regression with Lasso
regularization, 24 coefficients

$R^2: 0.79$

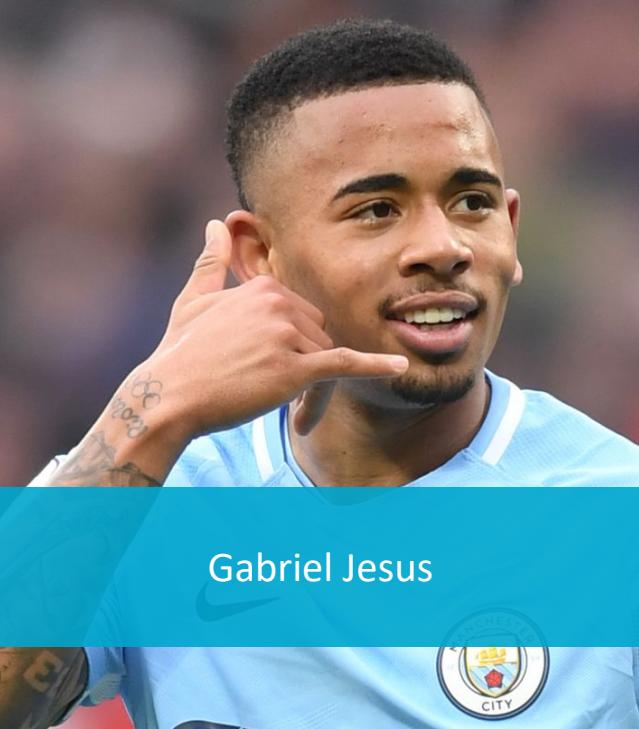
Age, and ball control were
more predictive than in other
leagues

Bonus: predicted vs real values for the English Premier League



Raw Data Source: sofifa.com

Final Model : Linear Regression with 40 coefficients



Let's predict his
market value

40,8

Predicted Value (€ mn)

37,5

Real Value (€ mn)

8,8%

Difference

Gabriel Jesus' value is way above the average, the model predictability tend to slightly decrease as we move away from the value mean



Key Takeaways

- It might be a little late for me to join a professional soccer team
- Additional data that can help to increase predictability:
 - Release contract clauses
 - Physical status (injuries record)
 - Team's budget
 - Excluding outliers (players with very high or very low values)

A soccer ball is positioned in the center of the frame, resting on a patch of green grass. In the background, the lower legs and feet of two players are visible. One player is wearing red socks and white shoes, while the other is wearing white socks and yellow Nike shoes. The scene is set outdoors on a sunny day.

Questions?