Football ~~Soccer~~ Player Transfer Value Prediction

DurianCandy

Sunny Kwong Charles Siu
Sean Tey Andrew Young

Agenda

- Intro
- Dataset
- Data Wrangling
- Model Selection
- Findings & Conclusion


UEFA
CHAMPIONS
LEAGUE
adidas
UEFA
CHAMPIONS
LEAGUE
OFFICIAL
MATCH BALL


PS4
adidas
EA SPORTS
FIFA 19
FIFA
OFFICIAL LICENSED PRODUCT
EVERYONE
TODOS
E
ESRB

Why are **some soccer players** worth more than others?

Dataset

Title: FIFA 19 Complete Player Dataset

Source: Kaggle, originally from sofifa.com

Link: https://www.kaggle.com/karangadiya/fifa19

Shape:

- 89 Columns x 18,207 Observations
- Each observation represents a soccer player in FIFA 19
- Each column represents an attribute of the player

Can we predict the
value of a soccer player
in the transfer market from their
ratings and characteristics ?

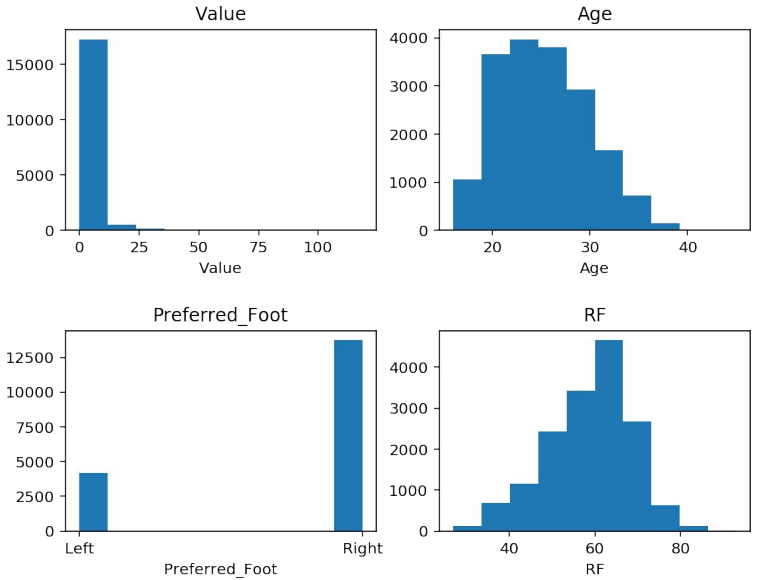
Data Wrangling

Columns

- Added columns related to soccer leagues
  e.g. *Country*, *Pyramid Level*, *Revenue Per Team*
- Dropped extraneous ID and corrupted columns
  e.g. *Player Name*, *Body Type*, *Photo*
- Derived columns based on existing columns
  e.g. *Loan Out* from *Loaned From* and *Club*

Rows

- Dropped rows that have zeros for *Value (target)*



Model Selection

We wanted to be able to interpret our models so we tried:

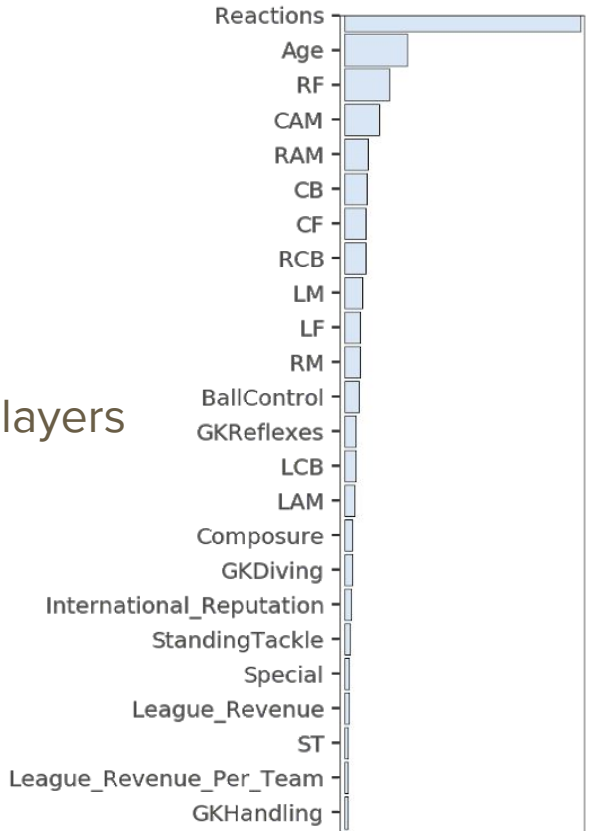
- Basic Linear Regression
- L1 Regularized Regression
- Random Forest

Random Forest seems to perform the best with MedAE of (15, 330 euros).

We are considering using Symmetric mean absolute percentage error (sMAPE)

Findings & Conclusion

- Reaction is the most important
- Age matters
- Attacking players are more predictive than defensive players
- Other important attributes
  - Ball control
  - Goalkeeper reflexes



Thank you