Predict Football Match Results from Player Stats

Sylvia, Emanuel, Riah

Introduction

- Overview:
  - Develop models to predict the outcome of soccer matches given the player attributes of teams
- Goal:
  - Provide better models for professional teams and analysts

Introduction - Why's this important?

Improve team salary cap allocation

- If teams know what they need to win, they can more effectively spend their money

Help Sports Analysts be better at predicting games

- Make these models more interpretive so they're easier for non-statisticians to explain to other non-statisticians

Data

- Source: Kaggle

Using three related datasets

- FIFA players stats, 2018 and 2019 (17,000+ observations)
- Match results from major leagues in Europe, 2018/2019 season (~3000 observations)

Data - Player Stats Sets

Each row is a player while each column is a variable

Variables include

- Players position
- Attributes: Attacking, Defending, Pace, Dribbling , ect
- Personal Data: Name, Wage, Worth, Nationality, etc

Data - Match Results Sets

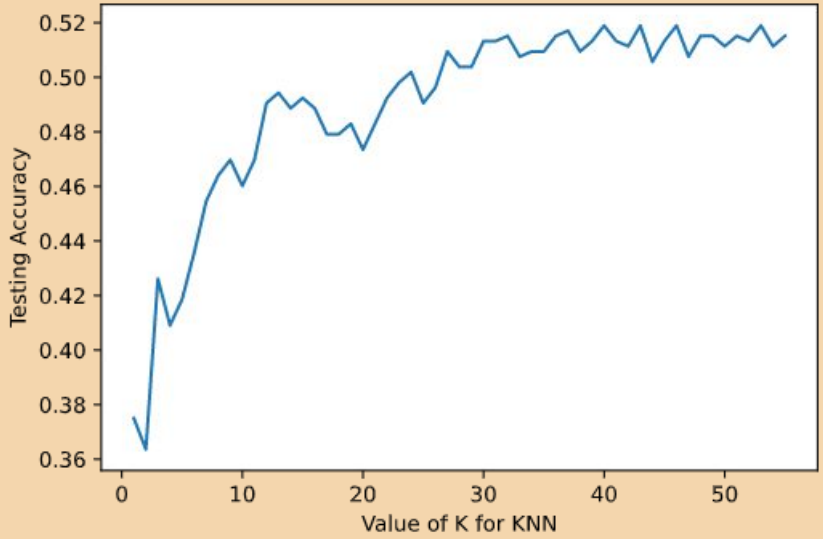
Each row is a match and each column is a variable

Variables include

- Goals scored
- Match result
- Number of cards
- Betting rates

Models 1: K-Nearest Neighbors

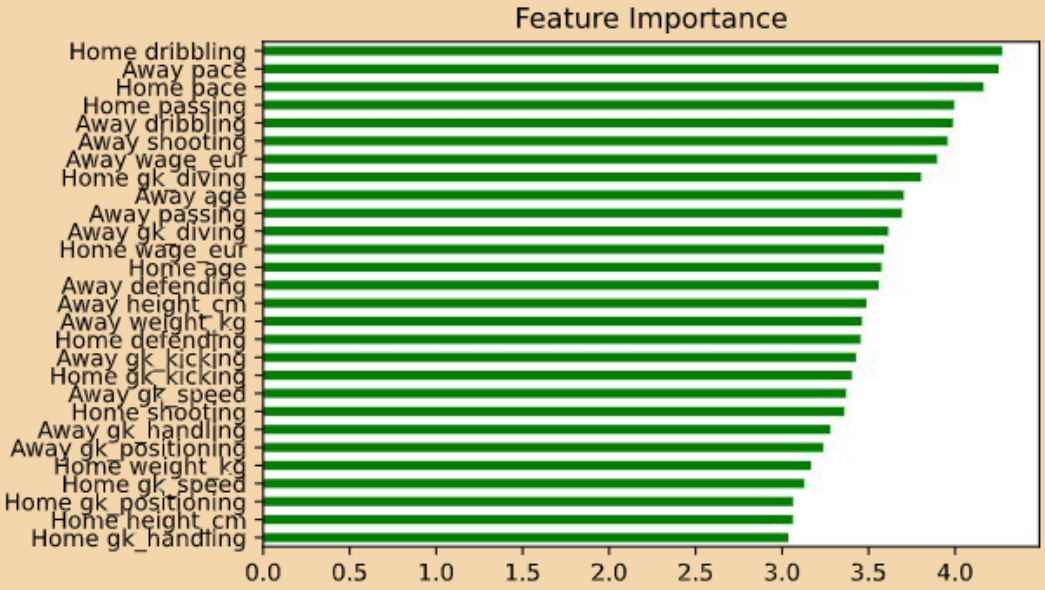
- Validation set approach:
+ Best K: 40
+ Accuracy: 51.89%
- 10-fold cross-validation:
+ Best K: 55
+ Accuracy: 50.18%
- Value of random seed will affect the validation => affect accuracy



Models 2: Decision Tree

- Gradient Boosting:
+ Leaning rate: 0.01
+ Accuracy: 51.70%

- Random forest:
+ Number of tree: 500
+ Accuracy (10-fold) : 46.97%



Conclusion

- Random guess: 33.33%
- Best accuracy: 51.89%
- Generally, goalkeeping stats is less important

Drawbacks

- Observations might not be completely independent
- Might have done better data cleaning
- Skewed data

References

Gambling Sites. “Sports Betting Research & Analysis - How Important Is It?” *GamblingSites.ORG*, 3 Dec. 2019, www.gamblingsites.org/sports-betting/research-and-analysis/.

Palmer, Catherine. *Sports Betting Research: Literature Review*. www.communities.tas.gov.au/__data/assets/pdf_file/0025/43945/Sports_Betting_Literature_Review.pdf.

Sichel, Adam Robert. *The Economic and Social Effects of Legalized Sports Gambling*. 2020, kb.osu.edu/bitstream/handle/1811/91813/1/Adam_Sichel_HonorsThesis.pdf.