NBA Hall of Fame Probability Predictor

Quantifying your favorite player's chances at landing in the Hall

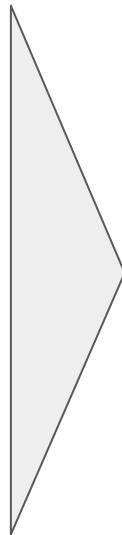MSiA 423 Final Project
Max Holiber

Motivation

NBA Popularity

- Social Media Presence:
  - **38M**
  - **28M**
  - **12M** YouTube
- Over **20M more** Instagram followers than next highest major sports league
- **22/30** teams more than **1M** Instagram followers

Outcome

- Engage NBA fans who want insights on favorite players
- Increase social media following of NBA players
- Marketing tool for NBA, teams, and endorsement companies to promote future HOF players

LIVE DEMO

Data Description

Scope

- 1950 - 2017
- NBA only

Specs

- 53 columns (e.g. points, rebounds, etc)
- 24.7K rows (per player per season)
- 3.9K players (3% HOF)

Fields of Interest

Player Name	HOF *
# Years played	Points
Rebounds	Assists
Steals	Blocks
Win Shares (WS)	3-point FG made

* binary target variable

	Player	G	Year_x	WS	3P	TRB	AST	STL	BLK	PTS	HOF
0	A.C. Green	1278.0	16	9.4	0.097027	7.412363	1.095462	0.808294	0.427230	9.648670	0
3	Aaron McKie	793.0	13	5.9	0.489281	3.262295	2.680958	1.150063	0.206810	7.403531	0
4	Aaron Williams	715.0	14	5.2	0.001399	3.902098	0.678322	0.394406	0.753846	5.762238	0
5	Adam Keefe	617.0	9	6.5	0.001621	4.142626	0.692058	0.497569	0.280389	5.035656	0
6	Adonal Foyle	733.0	12	4.0	0.000000	4.721692	0.469304	0.360164	1.627558	4.077763	0

Modeling

① HOF Probability Prediction

- Gradient Boosted Tree
  - 10 reps, 10-Fold CV parameter tuning
- Optimal Model
  - Learning Rate: **0.2**
  - # Trees = **85**
  - Max_Depth = **6**
- Results:
  - CV F-Score on HOF class: **69%** ✅
  - Original success criteria: **60%** ✅
- Other models considered:
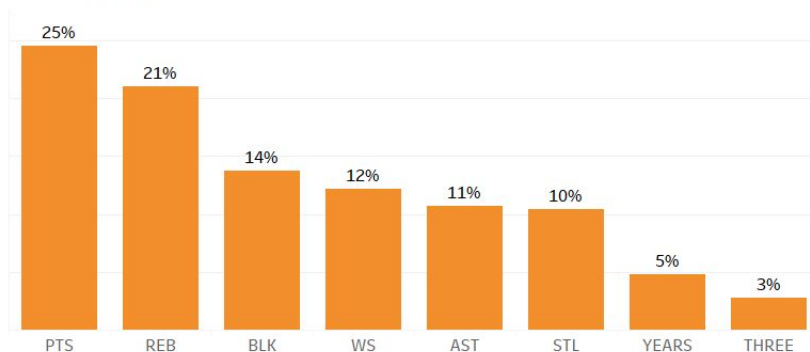 - Logistic Regression, Random Forest

② Player Similarity

- K Nearest Neighbors
  - # Neighbors: **10**
  - Distance Function: **Euclidean**

Insights

- **HOF Predicted Class:**
  - **27 players (8%)**
- **Feature Importance:**
  - Top 2: **Points & Rebounds**
  - Surprises: **Blocks (3)**, **WS (4)**
- **HOF Averages:**
  - **20 PTS/G**
  - **7 REB/G**
  - **4 AST/G**

Feature Importance

PTS	25%
REB	21%
BLK	14%
WS	12%
AST	11%
STL	10%
YEARS	5%
THREE	3%

Avg Value for Predicted HOF Class

AST per Gm	4.40
BLK per Gm	0.80
PTS per Gm	20.20
REB per Gm	7.00
STL per Gm	1.30
THREE per Gm	1.10
WS (max. over career)	14.60
YEARS (# seasons)	13.20

Contact

Thanks for listening!

Name: Max Holiber

Email: maxholiber2019@u.northwestern.edu

Github: https://github.com/maxholi/MSIA-423-Project.git

LinkedIn: https://www.linkedin.com/in/max-holiber-7497a162