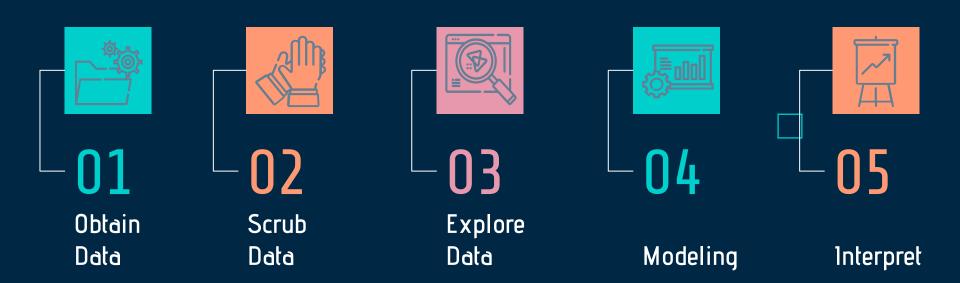


UNDERSTANDING THE PROBLEM



THE PROCESS - OSEMN

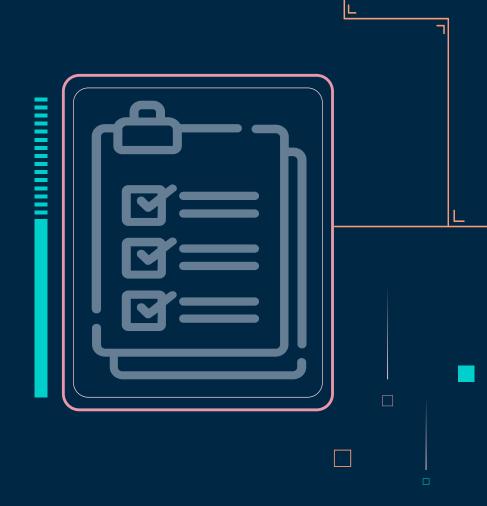


THE DATA

Obtained From Baseball Reference (Stat Head)

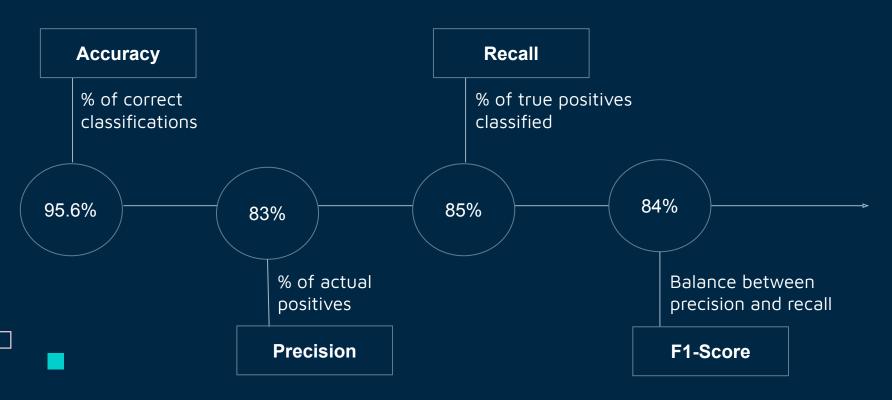
All Hall-of-Famers (197), Retired All-Stars (770), Active All-Stars (136)

1103 Total players



THE MODEL

XGBoost Model:

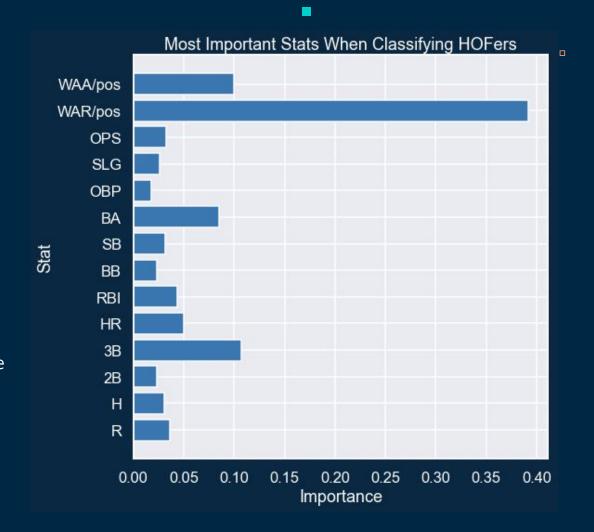


THE RESULTS

Which Stats are Most Important?

WAR/pos (Wins Above Replacement) is overwhelmingly most important

WAA/pos (Wins Above Average), BA (Batting Average), 3B (Triples) are the three secondary stats



OUR RECOMMENDATIONS

How we decided:

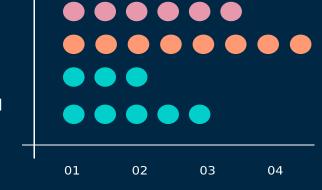
Project career numbers until average retirement age for each player

Predicted "Hall-of-Famers" using model

Identified undervalued players

Compared prices to HOF player with similar statistics

*These investments represent potential ceilings and are not guaranteed



OUR RECOMMENDATIONS





Trevor Story

2016 Topps Chrome
PSA 10
\$80.52



ROI: 331%



Mike Piazza 1992 Fleer PSA 10 *\$346.69*



ALONSO

Pete Alonso 2019 Topps PSA 10 \$47.20



ROI: 185%



Jim Thome
1991 Bowman PSA 10
\$135.16





Alex Bregman
2017 Topps PSA 10
\$39.33



ROI: 236%



Larry Walker 1990 Leaf PSA 10 *\$132.33*



FUTURE WORK

162-GAME DATA MODEL



Mapping projections based on average stats per full season

ERA-BASED MODEL



Accounts for changes to the game of baseball over the years

RISK ASSESSMENT



Finding potential floors if careers do not pan out (injury, decline, etc.)



hoganbyun@gmail.com 410 236 4663

THANK YOU!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik