Will Vladimir Guerrero make it in 2017?

Predicting First Ballot Hall of Fame Hitters

Michael Laster

January 2017:

Should mlb.com create Vladimir Guerrero memorabilia To meet Demand if selected?
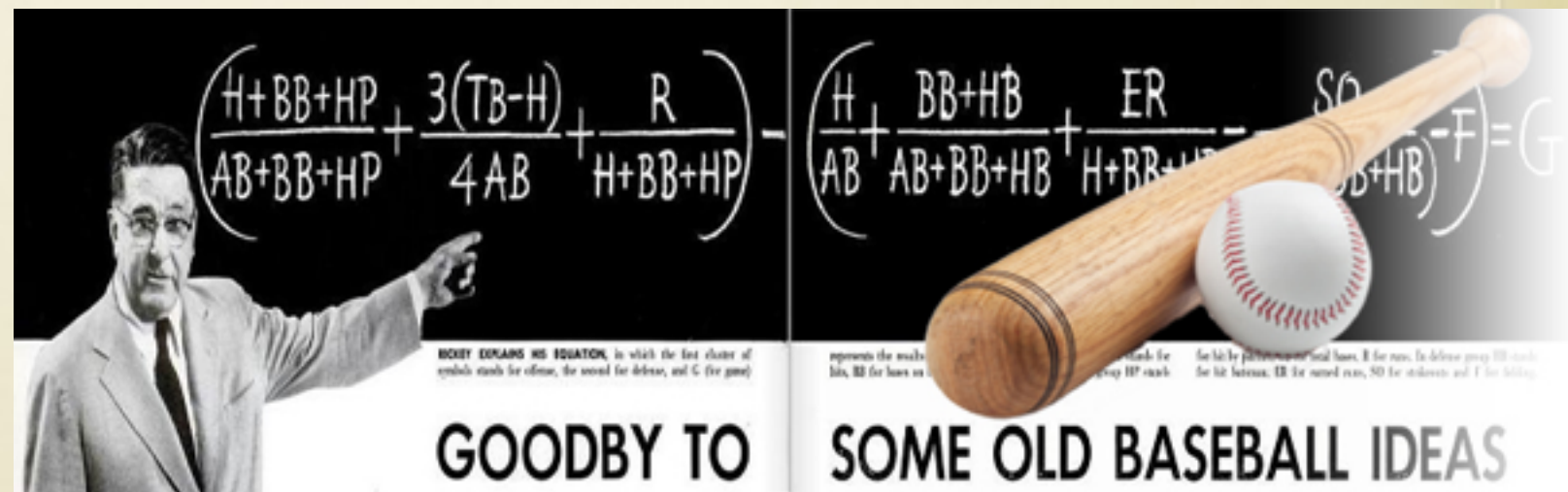
Goal: Predict First Ballot Hall of fame hitters

- Imbalanced class: Only 36/720 are inducted on first ballot

- Metrics to focus on:

  - Precision: If I invest in inventory for a projected induction, I don't want it to go to waste

  - Recall: If a player will be inducted, I don't want to miss the opportunity to sell their HOF-branded jerseys
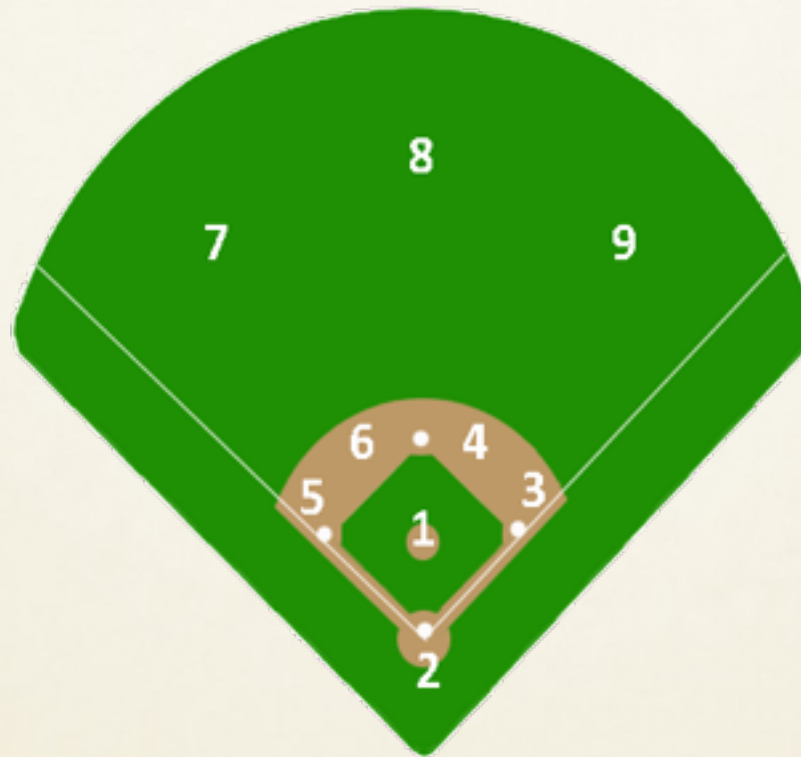
Features to use

- Making a "2nd-order prediction"

- "Old-School" stats- hits, hrs, rbi, ab

- Sabrmetrics-"WAR"

$$\left(\frac{H+BB+HP}{AB+BB+HP}+\frac{3(TB-H)}{4AB}+\frac{R}{H+BB+HP}\right)-\left(\frac{H}{AB}+\frac{BB+HB}{AB+BB+HB}+\frac{ER}{H+BB+HP}-\frac{SO}{BB+HB}-F\right)$$

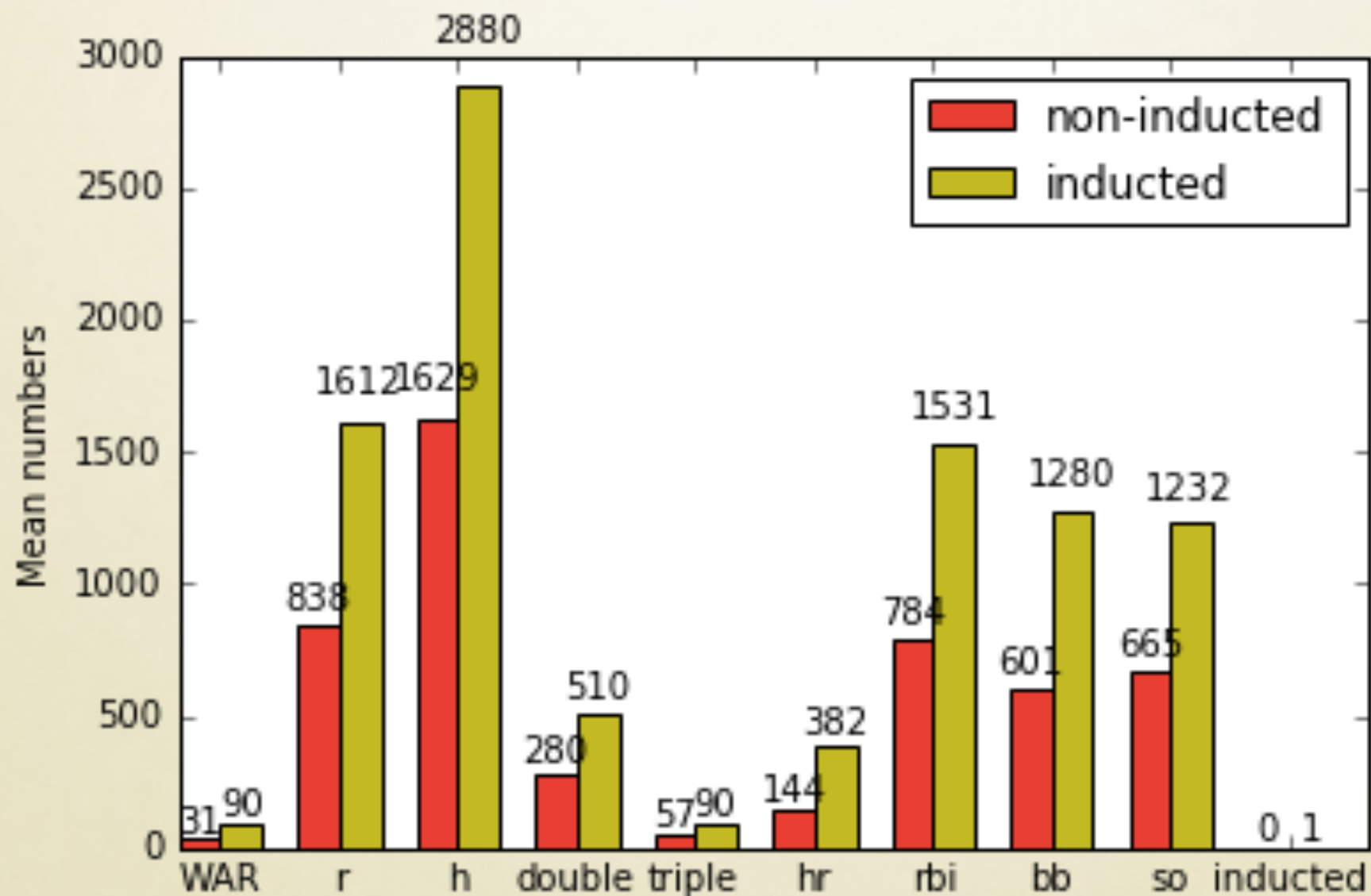GOODBY TO | SOME OLD BASEBALL IDEAS

Feature selection cont'

- Position played

- Era of voting
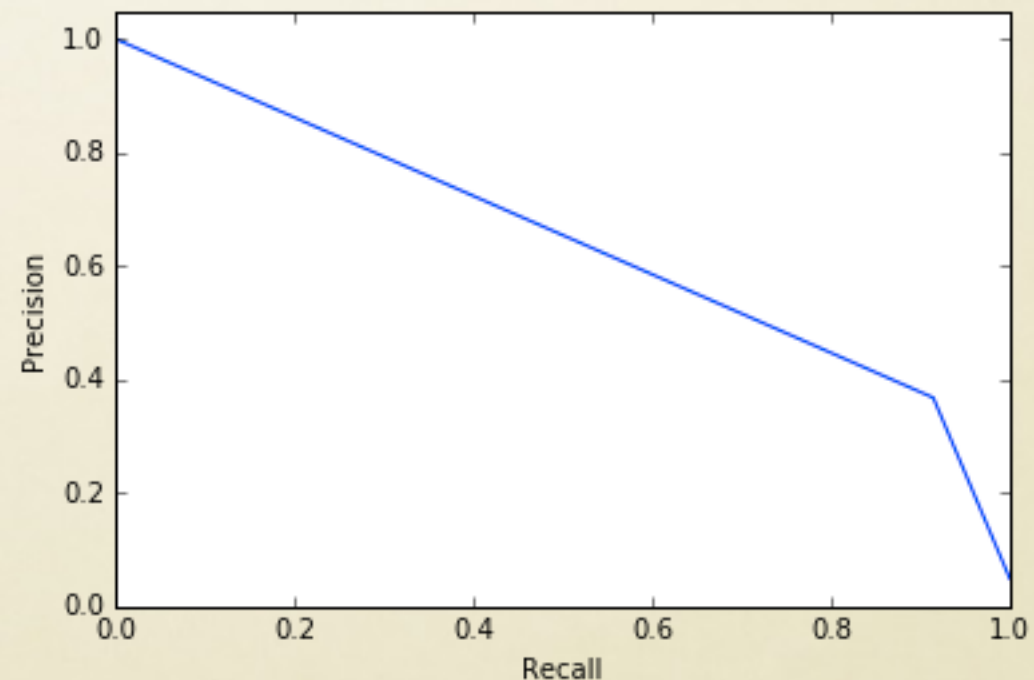
- Steroids/illegal activity

Data Exploration

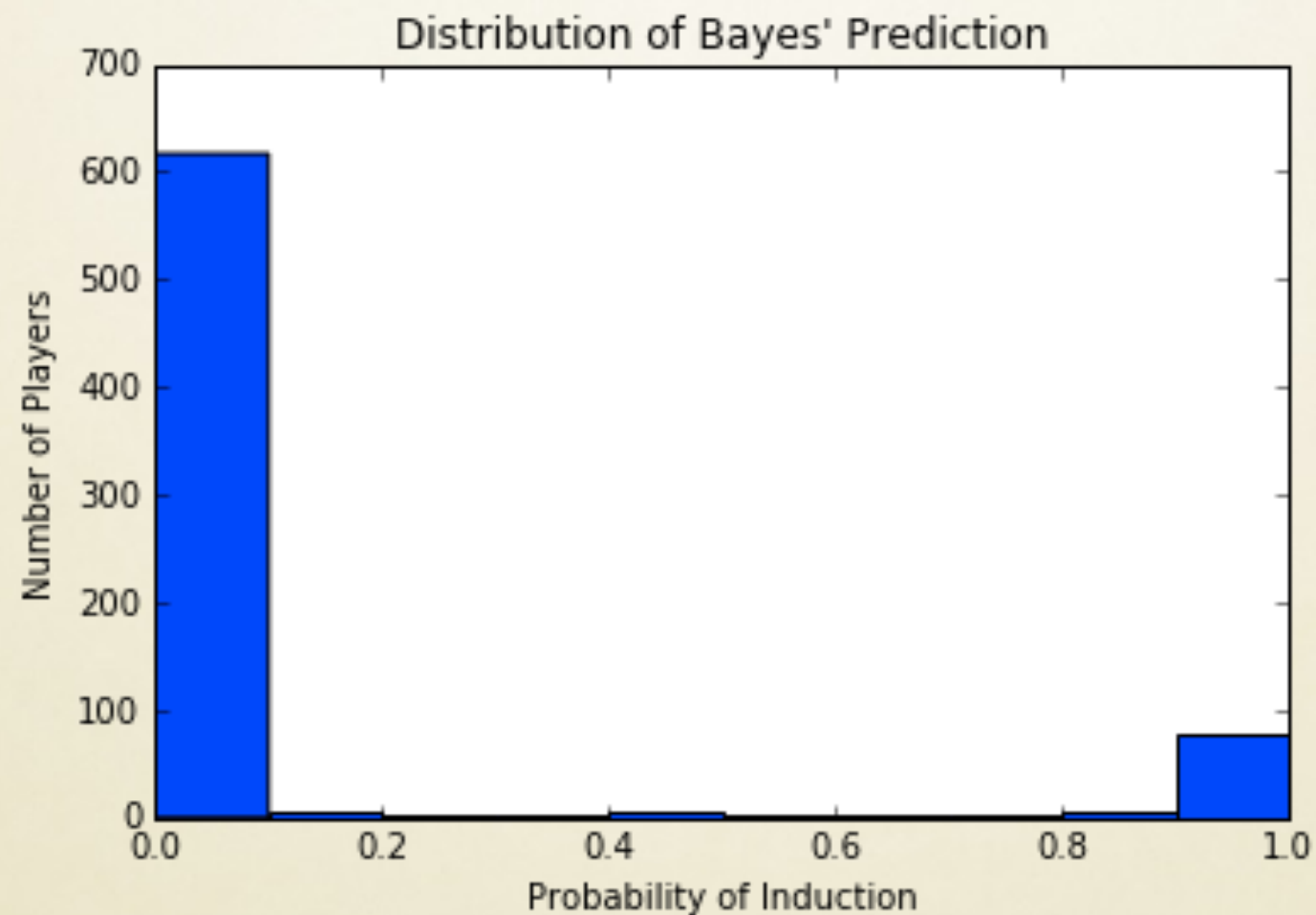Comparison of Mean Stats for Inducted vs. Not

The Best performing algorithm:

Naive Bayes

- Average Precision (cross-val): 0.7155
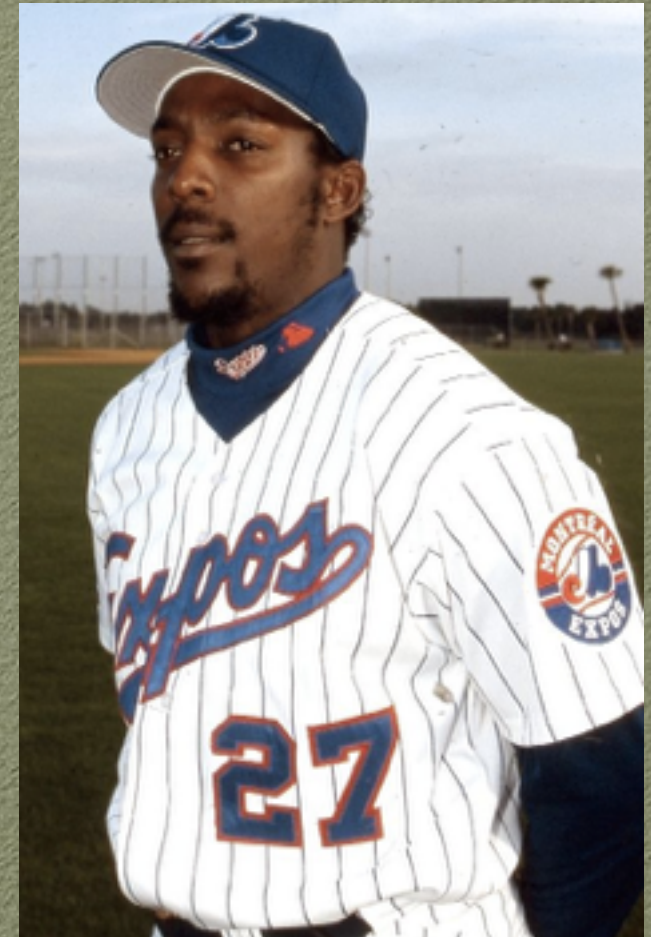
- Precision: .3711

- Recall: .9116

HE MAY BE NAIVE, BUT HE'S VERY OPINIONATED

Further Adjustments

- Adjust sensitivity to different variables, especially era

- Change distributions, not so Gaussian

- More thorough grid search of other algorithms

Guerrero's Chances: 0.9946!

Thank you!