BASEBALL HALL OF FAME VOTING WITH XGBOOST

ERIC ALBERS & AIDAN LORENZ
FINAL REPORT

Contents

1. Introduction

The purpose of this project was to use XGBoost to predict yearly Major League Baseball Hall of Fame voting. It makes sense to first give an overview of how voting for the Baseball Hall of Fame works. The voters consist of members of the so-called Baseball Writers' Association of America (BBWAA), each of whom is a journalist/writer for a qualifying newspaper, magazine, website, etc. On a yearly basis, there are around 550 members of the association that cast a ballot; each voter is allotted ten votes to cast. For a player to be inducted into the Hall, he must receive 75% of the total number of votes. Any player who has played 10 years or more in the MLB is eligible to appear on a ballot, the first of which is 5 years after his retirement. With that being said, the voters predetermine who will even appear on the ballot each year. Currently, a player may appear on up to 10 successive ballots (barring successful election, of course) under the prerequisite that they must receive at least 5% of the votes in order to qualify for the following year's ballot. This mark actually changed in 2015, before which players could appear on up to 15 ballots. Additionally, a player who has been banned from Major League Baseball (e.g. Pete Rose) cannot appear on any ballot.

The problem at hand could be phrased in a couple of different ways. It could be considered as a pure classification problem in that the only predictions being done would be whether a player makes it into the Hall or not. On the other hand, it could be considered as a regression problem in that what is being predicted is the percentage of votes a player will get. Note that the classification approach of the former can easily be deduced from the latter. For this reason (along with interest from the authors), the regression approach was taken. Due to its historically strong performance on regression problems (in Kaggle competitions and so on), XGBoost was used to carry out the regression.

One interesting and noteworthy feature of this problem is that there is a clear distinction between identifying "good" players and identifying players whom the voters *see as* good. There certainly are examples of objectively strong players who have not made the Hall of Fame (e.g. Keith Hernandez and Kenny Lofton), as well as the opposite (Bill Mazeroski). The goal of the project was to predict how voters would actually vote, not to predict what players actually deserved.

2. The Data

Our model ran on statistics gathered from Baseball Reference. Through this link one finds a list of yearly hall of fame voting results, along with various statistics for each player, encapsulating their contributions to their team(s) over their career. These tables were exported to CSV files in the process of creating our dataset. Due to the fundamental difference in important statistics for pitchers and other players, we ran separate models on all pitchers and on other players. Due to the fact that these other players are, in large part, judged on their offensive ability, we will refer to them as batters.

2.1. Sabermetric and Position-Neutral Statistics.

As noted earlier, a player can appear on up to ten consecutive ballots in his efforts to be inducted into the Hall of Fame. This led to an interesting decision as to whether or not individuals should appear only once in our data, with the model's task to attempt to predict the percentage of votes a player receives each of their ten years on the ballot, or if a player should appear in the dataset as many times as they appeared on a ballot, with an input variable being how many ballots they have already been on. The latter was chosen for the purpose of this project, mostly due to the format of the tables on Baseball-Reference.com.

Despite the aforementioned difference between hitting and pitching statistics, some of the more advanced metrics used in this project can be compared across positions. We now describe the position-neutral sabermetric statistics which we employed.

- YoB: *Years on ballot*. Simply the number of times the given player has been on a Hall of Fame ballot.
- HOFm: *Hall of Fame monitor*. Created by renowned baseball mind Bill James, the Hall of Fame monitor attempts to gauge how likely a player is to get into the Hall of Fame. Scores are calculated from "milestone" events which players reach; for example, 40 points are awarded for reaching 3000 career hits and 35 points are awarded for tallying at least 35 career wins. A complete list of all considered achievements and their associated point values can be found here. Per Baseball-Reference, a total HOFm score of 100 yields the player a good chance of making the Hall whereas a score of 130 all but guarantees the player admittance.
- HOFs: *Hall of Fame Career Standards*. Another product of Bill James, the Hall of Fame Career Standards is another benchmark-type statistic similar to HOFm. According to James, this metric is to measure how the overall quality of a player's career whereas the above Hall of Fame monitor quantifies how *likely* a player is to be inducted. One can find the breakdown in how Hall of Fame Standards is scored through this link.

- Yrs: *Years.* The number of years that a given player played in the MLB.
- JAWS: *Jaffe WAR Score System.* JAWS is a metric created by writer Jay Jaffe which attempts to measure a player's total contributions throughout their career, as well as their contributions during their best seasons (in an effort to quantify how good a player was at their best). He does this using perhaps the most important sabermetric invention, a metric known as Wins Above Replacement.

  Wins Above Replacement attempts to be a holistic evaluation of all of a given player's contributions to their team over the course of a season. The basic principle of Wins Above Replacement starts with the following run-expectancy table based on the 24 base-out states that can occur in any inning.

Table 1. Run Expectancy, By The 24 Base/Out States, 1999–2002

1B	2B	3B	0 Outs	1 Out	2 Outs
--	--	--	0.555	0.297	0.117
1B	--	--	0.953	0.573	0.251
--	2B	--	1.189	0.725	0.344
--	--	3B	1.482	0.983	0.387
1B	2B	--	1.573	0.971	0.466
1B	--	3B	1.904	1.243	0.538
--	2B	3B	2.052	1.467	0.634
1B	2B	3B	2.417	1.650	0.815

The above table reports the average runs a team scored in an inning after the point in which a certain base-out state occurred. With this in mind, one can quantify events that happen on the field in terms of runs as follows: suppose it is the beginning of the inning and there is nobody on base with nobody out. The above table gives that on average teams score 0.555 runs from this point forward in the inning. If the current batter hits a single, the state changes to a runner on first with nobody out; a state which has a run expectancy of .953 runs. Thus the single had a net effect of $.953 - .555 = .398$ runs and thus we credit the batter with .398 runs for the single. Of course, events can have a negative impact, for example if there is a runner on first with nobody out and the next batter decides to sacrifice bunt, the situation changes to a runner on second with 1 out. The change in run expectancy from this event if $.725 - .953 = -.228$, and hence the event causes teams to score less runs on average. This is one of the many reasons why modern baseball teams bunt less often in today's game.

One can then quantify the average effect of every event on the baseball field, averaging over all base-out states in order to be context neutral. The reason the statistic is chosen to be context-independent is because batters have little control on the current base-out state for any of their plate appearances. The first step in calculating wins above replacement is to sum the total contribution in terms of runs a player made throughout the season by using these average effects for each event. This gives a context-neutral, flat number of runs a player provided their team over the season.

The next step in calculating wins above replacement is measuring the contributions of the so-called replacement player. The key observation here being that if a player is injured he is not replaced by an average major league player, but instead by someone off of a team's bench, who is typically considerably worse than the average big-leaguer. One can then quantify the contributions of the average bench player in the same manner discussed above, in order to compare any major leaguer to a replacement level player. It should be noted here that the replacement player is done for each position, since it is much easier to replace a first baseman (the position requires less skill on defense) than a shortstop. As such each player is only compared to the average replacement player at their position.

After subtracting off the total run contribution of the league average replacement player for a given player's position, the final calculation is converting this value of runs above replacement, to wins above replacement by the conversion rate 1 win = 10 runs. The reason for this specific conversion is because calculations show if Team A scores 10 more runs than Team B over the course of an entire season, they are expected to win 1 more game. This is the detailed calculation of Wins Above Replacement and the table below demonstrates what a "good" WAR is in a given season.

Scrub	0-1 WAR
Role Player	1-2 WAR
Solid Starter	2-3 WAR
Good Player	3-4 WAR
All-Star	4-5 WAR
Superstar	5-6 WAR
MVP	6+ WAR

WAR is a statistic that simply accumulates over time and as such career WAR totals are often used to evaluate a player's entire career. A typical benchmark is a total of 60 WAR for a player to have a good chance of making the hall of fame (with the exception being Catchers and Relief Pitchers who typically require less). Jay Jaffe's JAWS metric takes the average of a player's total career WAR along with the WAR total over their 7 best seasons in an attempt to quantify both a player's total career success and their "peak" success. Jaffe then averages the JAWS of current hall of famers at each position (assigning a player a position based simply on the one they played the most throughout their career) and says a player is worthy of being inducted into the Hall if their career JAWS is higher than the average Hall of Famer's JAWS at that position.

- G: *Games.* The number of games in which a given player appeared.

2.2. **Batters.** We now discuss the statistics we used to assess only the batters (not the pitchers).

- R: *Runs.* The number of runs a given player has scored.

- H: *Hits*. The number of hits a given player got.
- HR: *Home Runs*. The number of home runs a given player hit.
- RBI: *Runs Batted In*. The number of runs a given player batted in.
- SB: *Stolen Bases*. The number of stolen bases a given player had.
- BB: *Bases on Balls*. The number of times a given player got walked.
- BA: *Batting Average*. Batting average is calculated by the following formula:

$$\frac{\text{H}}{\text{PA - BB - HBP}}$$

  where PA is number of plate appearances and HBP is the number of hit by pitches.
- OPS+: *Normalized On Base Percentage Plus Slugging Percentage*. On base percentage is the following:

$$\frac{\text{H} + \text{BB} + \text{HBP}}{\text{AB} + \text{BB} + \text{HBP} + \text{SF}}$$

  where all abbreviations are as above and SF is the number of sacrifice flies. Slugging percentage is the following:

$$\frac{\text{1B} + 2 \times \text{2B} + 3 \times \text{3B} + 4 \times \text{HR}}{\text{AB}}$$

  where 1B is the number of singles, and so on. OPS is simply the sum of the on base and the slugging percentages. Now, OPS+ is a given players OPS divided by the league average OPS for that year, adjusted for ballpark factors (and multiplied by 100). Above 100 signifies an above average hitter and vice versa.

2.3. Pitchers.

- W: *Wins*. Though it is falling out of fashion among analysts these days, wins is a historically important statistic for starting pitchers. For a starting pitcher to qualify for a win, he must pitch at least five innings, leave the game with his team in the lead, and his team must win, never having relinquished the lead after his departure from the game. If the starting pitcher does not qualify for a win, it is up to the official scorer of the game which relief pitcher will get the win (usually the last pitcher of record on the winning team when his team takes the lead for good).
- ERA+: *Normalized Earned Run Average*. Earned run average (ERA) is the number of earned runs a pitcher gives up (earned as in not resulting from an error) per nine innings he pitches. ERA+ is calculated according to the following formula:

$$\text{ERA+} = \frac{\widetilde{\text{ERA}}}{\text{ERA}} \times 100$$

  where $\widetilde{\text{ERA}}$ is the league average ERA for that season. Ballpark factors are also taken into account. Above 100 signifies an above average pitcher, and vice versa.
- SV: *Saves*. Saves are a relevant statistic specifically for closing pitchers. A pitcher gets a save if he is the finishing pitcher in a game in which his team wins, he pitches for at least 1/3 of an inning, he comes in with a three run lead or smaller.
- IP: *Innings Pitched*. The number of innings the given pitcher pitches.
- BB: *Walks*. The number of walks the given pitcher gives up (fewer is better).
- SO: *Strike Outs*. The number of batter the given pitcher strikes out.

- WHIP: *Walks and Hits per Innings Pitched.* The number of walks plus the number of hits the given pitcher gives up divided by the total number of innings they pitch (considered, in our statistics, over the course of a season).

2.4. **Steroids.** Before discussing the results of this project, a short digression on steroid use in baseball is required. While there is historical precedent for players taking substances to obtain a competitive advantage from the very beginning of baseball, it is the outbreak of steroid use in the 1990s and early 2000s that has left a lasting effect on Hall of Fame voting. This is likely due to the media coverage on steroid use in baseball during this time, along with players like Barry Bonds and Roger Clemens testifying in court on their use of steroids. Although the league seemed to turn a blind eye to steroid use throughout the 90s as Mark McGwire and Sammy Sosa put up astronomical home runs numbers which helped MLB's ratings immensely, the members of the BBWAA have been less forgiving over the last few years. Many writers view the steroid epidemic during the turn of the century as a lowpoint for baseball and, although many players never failed a drug test (mostly due to Major League Baseball waiting until 2004 to institute more frequent drug testing along with more substantial punishments for steroid use) another group of writers refuse to vote for those they believed used steroids.

This has led to a backlog of qualified candidates on recent Hall of Fame Ballots. For example, on a purely statistical basis, Barry Bonds is clearly worthy of being inducted into the Hall. Nevertheless, he has appeared on the ballot for the previous seven years and has never received more than 60% of the vote. As a result, over the past few years, one could make a reasonable hall of fame case for far more than 10 players on recent ballots. As such, some writers are unable to vote for players who they believe worthy of induction into the Hall of Fame. Recently the Hall of Fame seems to be rebounding from this effect however, having inducted the most players over a three year span in the Hall's history from $2017-2019$, but for many years in the early 2010s few candidates were inducted.

To account for the depressed voting totals for players like Bonds and Clemens, an extra data point was added to all players in this project, a simple binary entry where 1 indicates that a player has had connections to Steroid use throughout his career with a 0 if their record is relatively clean. This was only done for players who played in the 1980s or later, since they are the one's whose steroid use has effected their Hall of Fame voting results. While we could speculate on the use performance enhancing substances from players in earlier decades, it seemed a rather fruitless endeavor to the authors, since it seemed to have little impact on voting during earlier time periods.

3. Results

3.1. **First Attempts.** Our first attempt consisted of splitting the batters and pitchers, training an XGBoost regressor on a portion of the data, doing a gridsearch to find the near-optimal hyperparameters and testing on a portion of the data. An example of the code is shown below:

```
[ ] pitcher_data = pd.read_csv('PITCHERS.csv')
    pitcher_data = pitcher_data.drop(['Name','Votes','%vote','WAR','WAR7','G','AB','R','H','HR','RBI','SB'
                                     ,'BB','BA','OBP','SLG','OPS','OPS+','L','ERA','G.1','GS','H.1','HR.1'], axis=1)

    pitcherX = pitcher_data.iloc[:, 0:14]
    pitcherY = pitcher_data.iloc[:, 14]
```

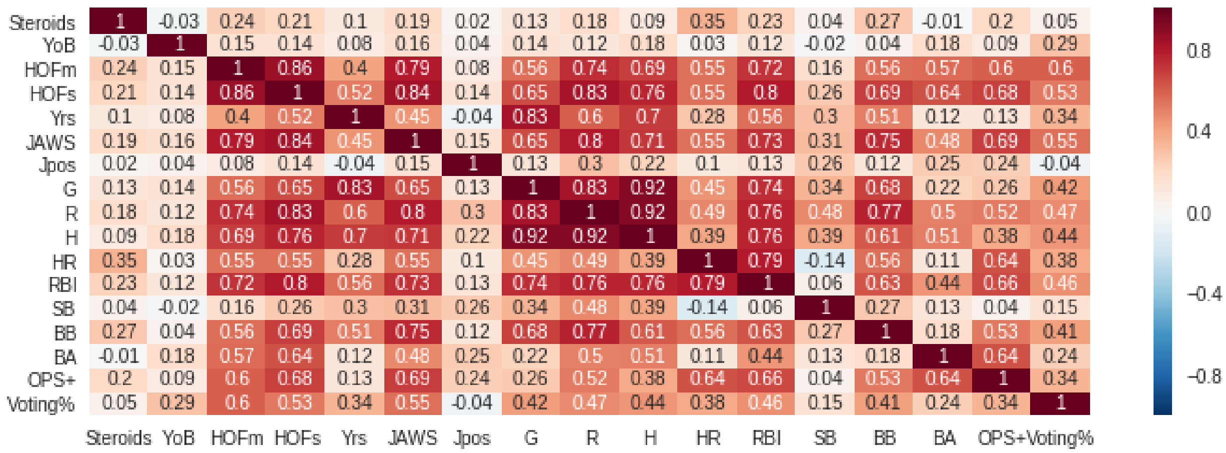
```
[ ] pitcherX_train, pitcherX_test, pitcherY_train, pitcherY_test = train_test_split(pitcherX, pitcherY, test_size = 0.3,
                                                                                    random_state=20)
```

```
[ ] params = [{'learning_rate': [0.4],
    'max_depth':[4], 'min_child_weight':[1], 'gamma':[0.3], 'subsample':[0.6,0.7,0.8,0.9,1],
                 'colsample_bytree':[0.6,0.7,0.8,0.9,1]}]
```

```
[ ] rgr = GridSearchCV(xgb.XGBRegressor(), params, cv=5)
```

```
[ ] rgr.fit(pitcherX_train, pitcherY_train)
```

One interesting feature is that we can see correlation matrices. Below, one can see the correlation matrix for the batter data.



As expected, the most heavily correlated features are the sabermetric statistics such as HOFm, HOFs, and JAWS.

These first attempts provide us a baseline to which we compare all other methods. This dataset was curated by the authors and thus the optimal performance on this set is unkown. As such, a baseline from running XGBoost on all of our data will have to suffice. For the batters, the mean-squared-error was 78.0 (a number which is hard to interpret by itself, it is only meaningful to compare to other MSEs). Our accuracy within 5% of the ground truth was 86.4% and the accuracy within 10% was 93.0%. As for the pitchers, we saw an MSE of 74.7, an accuracy within 5% of 87.7% and an accuracy within 10% of 94.1%.

3.2. **Splitting By Position.** Our first inclination was to make finer splits than just between batters and pitchers. The idea is that it is better to compare, for instance, all second basemen to one another than it is to compare all positions, as players at different positions have different expectations for their offensive production. The issue, however, is that there is simply not enough data to run models on each and every position separately. To remedy this, we used the common trend in baseball that the players up the middle of the field (catcher, second base, short stop, and center field) are the best fielders and consequently worse hitters, whereas the corner positions (third base, first base, left field, and right field) are the best hitters. We therefore split between the commonly "good hitters" and the commonly "good fielders." We also ran the good fielders once without catchers and once with catchers because catcher statistics are particularly weighted towards the defense, even when considered alongside the other strong fielding positions. As for pitchers, we simply split between starting pitchers and relief pitchers.

We found that even doing such broad splits led to, in some cases, only slight increase in performance and, in others, decreases in performance. The MSE, accuracy with 5%, and accuracy within 10% for the good batters were 90.0, 87.2%, and 93.4% respectively. For good fielders, the numbers were 86.8, 85.8%, and 92.8%. For starting pitchers, they were 145.8, 81.5%, and 89.8%. Since there were so few relief pitcher data points and it is so rare for them to get into the hall of fame, we did not even consider them.

3.3. **Over and under-sampling.** After splitting up our batters by the positions they played proved unsuccessful, we attempted to examine what data points our model was predicting the most inaccurately. We noted that a majority of players in our dataset appeared on only one ballot due to receiving less than 5% of the votes on their first appearance. This is a result of the fact that the standards are quite low to simply appear on a ballot and much higher to actually receive a substantial number of votes from the writers. As such we examined how our model was projecting players who receive more than 50% of the votes from the writers, and it was clear that this was a source of inaccuracy from our model.

To remedy this we turned to a process known as over and under-sampling. Typically used in classification problems, over-sampling is a process by which data points which make up a minority class in a dataset are duplicated in order to better punish a model for classifying members of this class incorrectly. Similarly, under-sampling is a process by which only a percentage of data points which make up a majority class in a dataset are used in order to better balance the data. There are, of course, pros and cons to using these methods. Certainly over-sampling will better punish our model for performing poorly on players who receive a large number of votes, while on the other hand, it has the potential to over fit to the specific data points which we duplicated.

We now detail exactly how our data was over and under-sampled, as well as the results of using these methods. We first split the data into a training and test set, since it is imperative one only over and under-samples data in the training set. On the training set, the data was broken up into the following four classes: those who received less than 5% of the votes, those who received between 5 and 30 percent of the votes, those who received between 30 and 60 percent of the votes and those who received greater than 60 percent of the votes. The class breakdown in our training data for the batters was as follows:

Class	Number of Datapoints in Class
Less than 5%	939
Between 5% and 30%	536
Between 30% and 60%	222
Greater than 60%	112

Table 1. Batter Class Breakdown

To balance the data, only half of the data points from the "less than 5%" class were used, whereas datapoints in the the between 30% and 60% class were duplicated once, and datapoints in the greater than 60% class we duplicated twice. Thus the class breakdown for our training set after these manipulations is as follows:

Class	Number of Datapoints in Class
Less than 5%	470
Between 5% and 30%	536
Between 30% and 60%	444
Greater than 60%	336

Table 2. Batter Class Breakdown after Over/under-sampling

A similar process was done for our pitchers data. Before any changes to the data the class breakdown for the pitchers was as follows:

Class	Number of Datapoints in Class
Less than 5%	510
Between 5% and 30%	243
Between 30% and 60%	79
Greater than 60%	69

Table 3. Pitcher Class Breakdown

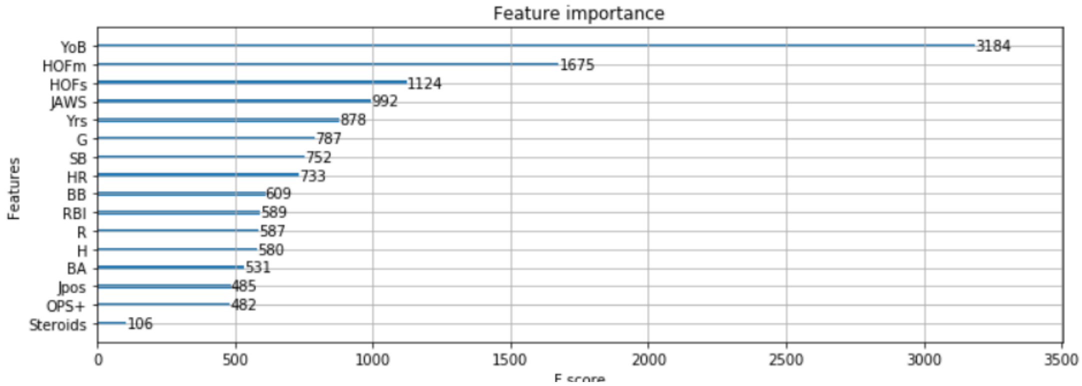
Again, only half the datapoints from the class "less than 5% were used while datapoints in the between 30% and 60% class were duplicated twice and datapoints in the greater than 60% class were duplicated three times. This led to a class breakdown after over/under-sampling as follows:

Class	Number of Datapoints in Class
Less than 5%	255
Between 5% and 30%	243
Between 30% and 60%	237
Greater than 60%	276

Table 4. Pitcher Class Breakdown after Over/Under-sampling

All in all, this led to slightly improved results. For the batters, the mean-squared error between the ground truth and the predictions on the test set was roughly 80.5 (no change from running on the unmodified data), but the accuracy within 5% improved to 88.7% (up from 86.4%) and the accuracy within 10% improved to 94.3% (up from 93.0%). For the pitchers, the mean-squared error was 79.8, while the accuracy within 5% was 88.5% (up from 87.7%), and the accuracy within 10% improved to 94.5% (up from 94.1%).

For a bit of insight into how our model works, we now provide the feature importance tables for the model on the batters. These tables simply tally how often a given statistic is occurs in each of one of the trees made by XGBoost.



We notice that years on ballot is by far the most important feature to our model, with many of the sabermetric statistic also being important features in this dataset.

3.4. **All Players.** A principal issue with this problem is the limited number of datapoints in our datasets. One attempt to remedy this issue is to attempt to fit a model to the data of all players, combining both the pitcher and batter data. In order to do this however, only statistics which are comparable across batters and pitchers can be used. The Sabermetric statistics (JAWS, HOFm, HOFs) are capable of being compared between pitchers and batters and as such were used in this attempt. On top of this, since OPS+ and ERA+ both attempt to measure how much better a player is than the average pitcher or hitter, we included these metrics in our all players dataset, using OPS+ for the hitters and ERA+ for the pitchers. Altogether, the data used was as follows:

- Steroids
- Years on Ballot
- Hall of Fame Monitor
- Hall of Fame Standards
- JAWS
- Years Played
- Position Played
- OPS+/ERA+

We also over/under-sampled the training data for this attempt, as it already demonstrated its usefulness when running on the batters and pitchers separately. The specific method of over/under-sampling this dataset is precisely identical to the method outlined when over-under-sampling the batters dataset.

The results were slightly worse than over/under-sampling the batters and pitchers separately, yet still an improvement on running models on the unmodified data. The mean-squared error for this attempt was roughly 91.0 while we obtained an accuracy within 5% of 87.4% and an accuracy within 10% of 93.5%. The following table summarizes all results.

Model	Mean-Squared Error	Acc. within 5%	Acc. within 10%
Batters No Modifications	**78.0**	86.4%	93.0%
Pitchers No Modifications	**74.7**	87.7%	94.1%
Good Batters	90.0	87.2%	93.4%
Good Fielders	86.8	85.8%	92.8%
Starting Pitchers	145.8	81.5%	89.8%
Batters Over/Under-sampling	80.5	**88.7**%	**94.3**%
Pitchers Over/Under-sampling	79.8	**88.5**%	**94.5**%
All Players	91.0	87.4%	93.5%

Table 5. All Results

4. Future Predictions

One of the main reasons for working on this project is using our model to predict which players will be inducted into the Hall over the next few years. Fortunately Baseball-Reference also has future years ballots which we were able to use to complete this task. Rather than provide an entire ballot which consists mostly of players who have no case for the Hall of Fame, we simply provide projected voting totals for a few select players which the authors believe to have a compelling case for Hall of Fame induction.

We refer the reader to the appendix, where projections for selected players are listed. Altogether, over the next few years our model projects the following players to be inducted: Derek Jeter, Alex Rodriguez, David Ortiz. On top of this, the following players are projected exceed 65% of the vote in one year of voting: Todd Helton, Carlos Beltran, and Curt Schilling. Of note, the model projects that Barry Bonds and Roger Clemens, the two most egregious steroid users in recent memory, will ultimately fall short of induction. This is likely due to the fact that our model is using the Steroids indicator to punish these players, along with the fact that previous years voting for Bonds and Clemens was a part of our data set and so our model is better suited to understand their depressed voting totals.

5. Further Directions

Needless to say, there are countless variations of this project which could be carried out. The most obvious variations come in using different statistics. For example, based on our feature importance charts, we could reasonably run models considering only, for instance, YOB and HOFm. Such a method would drastically reduce the dimension of the problem and thereby simplify it greatly. It should be noted however, that a few attempts were made by the authors to limit the number of input variables which led to worse results than the one's presented in this paper.

In large part, many of the difficulties in this project can be attributed to the small amount of data. This is likely why splitting by position did not yield good results; we simply did not have enough relevant data when it is split among too many data sets. We will, however, be able to collect more data each January when Hall of Fame voting takes place. In theory, with the collection of more data we will be able to get more accurate results and perhaps eventually splitting by position will prove very effective. This would presumably take a significant increase in the amount of data available.

Finally, one might try dropping old data, say, all data before 1950. One would do this for the simple reason that voters in the past cared about different things than modern voters do and in particular likely used less statistics in evaluating baseball player's career. The issue with this approach is that, as mentioned earlier, we are already suffering from lack of data. Cutting out more would work against this issue.

6. Appendix

Year on Ballot	Projected % Votes
1	58.6%
2	64.2%
4	69.0%
7	73.8%
9	**78.9**%

Table 6. Derek Jeter Voting Projections

Year on Ballot	Projected % Votes
1	68.4%
2	**76.3**%

Table 7. Alex Rodriguez Voting Projections

Year on Ballot	Projected % Votes
1	45.2%
2	38.8%
6	55.2%
7	61.1%
9	**80.5**%

Table 8. David Ortiz Voting Projections

Year on Ballot	Projected % Votes
2	29.6%
3	35.6%
6	54.0%
9	65.6%
10	65.9%

Table 9. Todd Helton Voting Projections

Year on Ballot	Projected % Votes
1	46.8%
4	54.0%
5	61.2%
9	70.4%
10	71.6%

Table 10. Carlos Beltran Voting Projections

Year on Ballot	Projected % Votes
8	60.3%
9	60.9%
10	70.1%

Table 11. Curt Schilling Voting Projections