NHL HALL OF FAME PREDICTOR

KRUEGER HOCKEY CONSULTING

OCTOBER 2023

INTRODUCTION

- **Krueger Hockey Consulting** is a boutique consulting firm that assists both professional hockey teams and leagues, such as the National Hockey League (NHL) in evaluating players

- We have developed a number of models to evaluate hockey players and will be discussing the following today:

  Hall of Fame Predictor

  - What is the probability that a given player is a hall of fame caliber player and what factors contribute most to hall of fame potential?

DATA OVERVIEW

- Source: Professional Hockey Database from Kaggle[1]

- Timeframe: 1917 – 2011

- Includes 22 csv files that contain team, player, and league information and statistics, of which, our analysis uses the following:

  1. Master: names and biographical information

  2. Scoring: scoring statistics by player, year, and team

  3. Teams: Team regular season statistics

  4. TeamsPost: Team postseason statistics

  5. SeriesPost: Postseason series results and statistics

  6. HOF: Hall of Fame information

- Resulting final dataset contains 4,722 rows and 41 columns containing biographical player information and career statistics for players who retired from the NHL anytime prior to 2007[1] and excluding goalies

HALL OF FAME STATISTICS
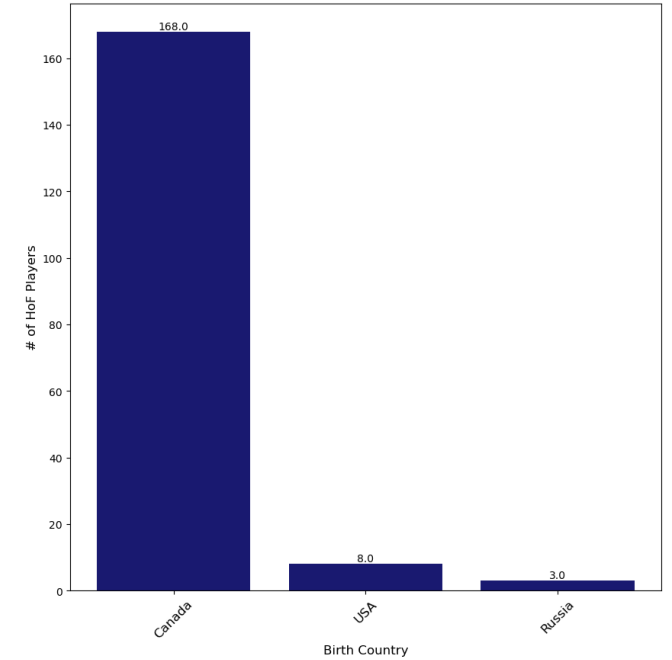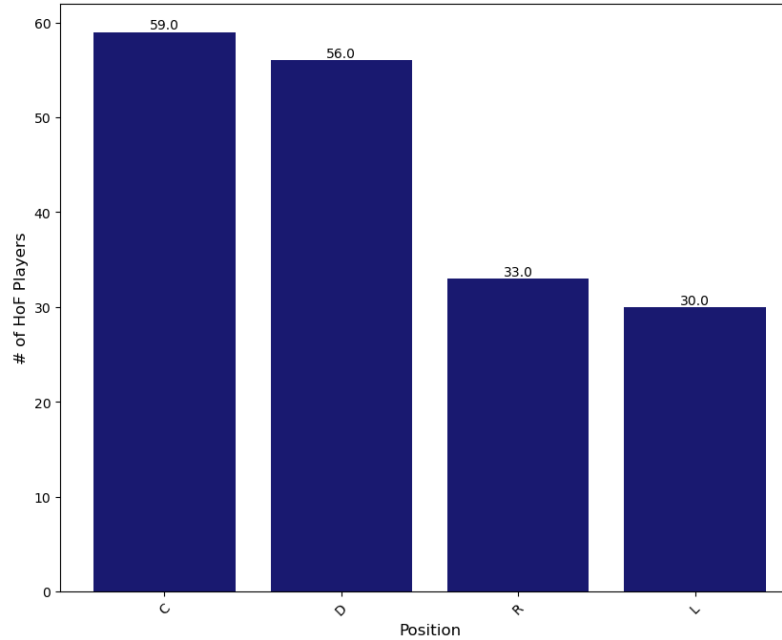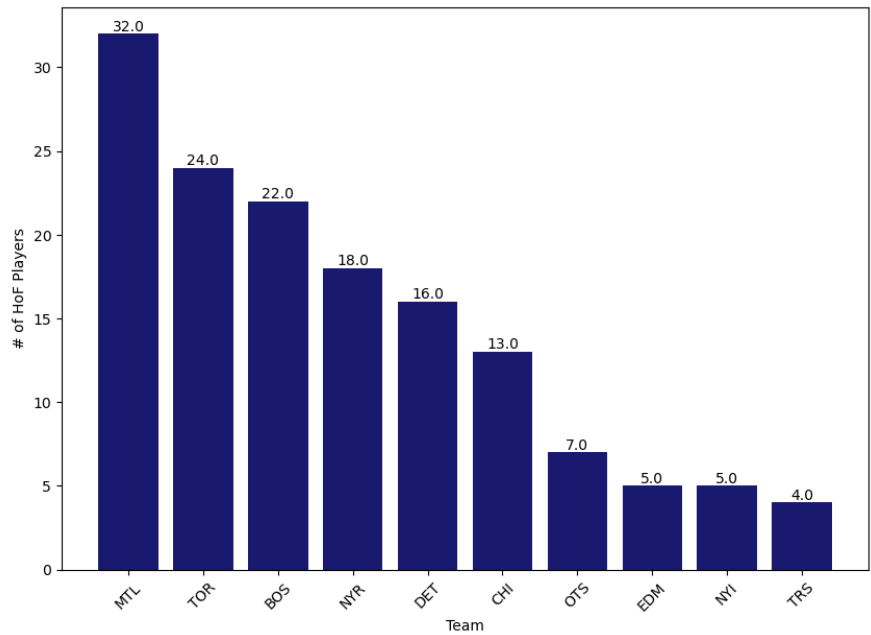
index	Yearsinlg	GP	+/-	Post+/-	SCcount	G_per_szn	PostG_per_szn	A_per_szn	PostA_per_szn	Pts_per_szn	PostPts_per_szn	height	weight
mean	14.0	750.8	71.4	2.3	2.3	15.3	1.4	21.7	2.1	37.0	3.5	71.0	182.3
std	6.0	457.0	146.2	12.6	2.3	10.6	1.4	16.0	2.1	25.0	3.3	2.1	17.9
25%	11.0	449.0	0.0	0.0	1.0	7.0	0.4	8.3	0.5	16.9	0.9	70.0	170.0
50%	14.0	696.0	0.0	0.0	2.0	13.5	1.1	18.5	1.5	31.7	2.4	71.0	182.0
75%	18.0	1125.0	87.0	0.0	3.0	22.1	2.0	32.9	3.2	54.1	5.6	72.0	195.0

HALL OF FAME PREDICTOR MODELS

Baseline Model[1]
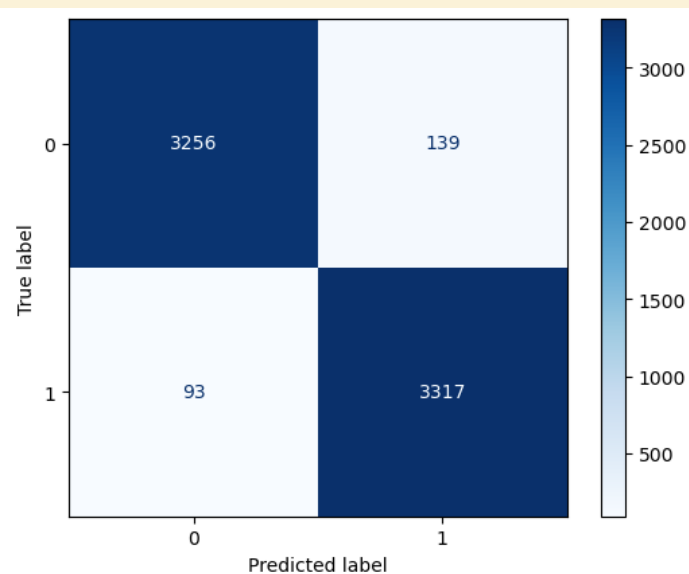
Accuracy: 96.6%   Recall: 97.2%
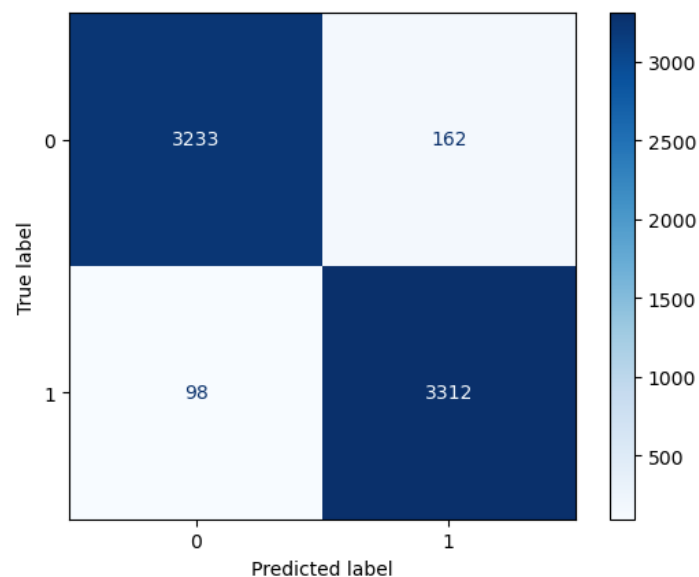Precision: 96.0%   F1: 96.6%

Cross Validation Avg. Precision Scores:
Train: 96.0%
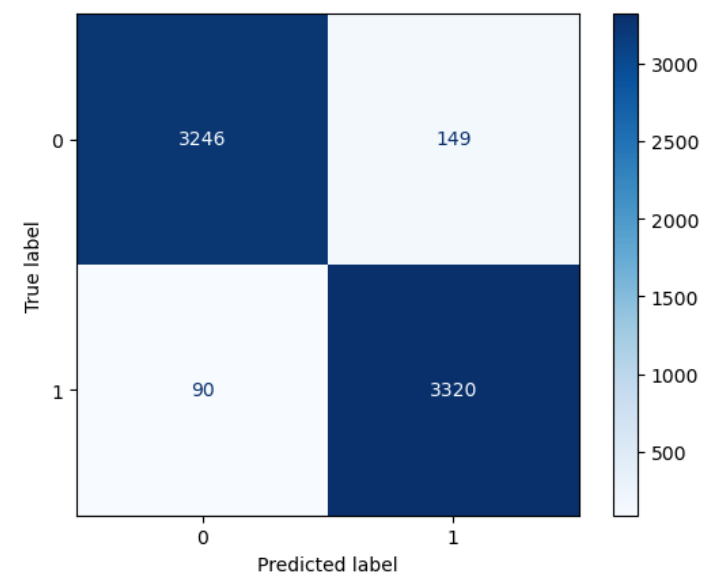Test: 95.1%

Ridge Regularization Model[1]

Accuracy: 96.2%   Recall: 97.1%
Precision: 95.3%   F1: 96.2%

Cross Validation Avg. Precision Scores:
Train: 95.4%
Test: 94.8%

Lasso Regularization Model[2]

Accuracy: 96.5%   Recall: 97.4%
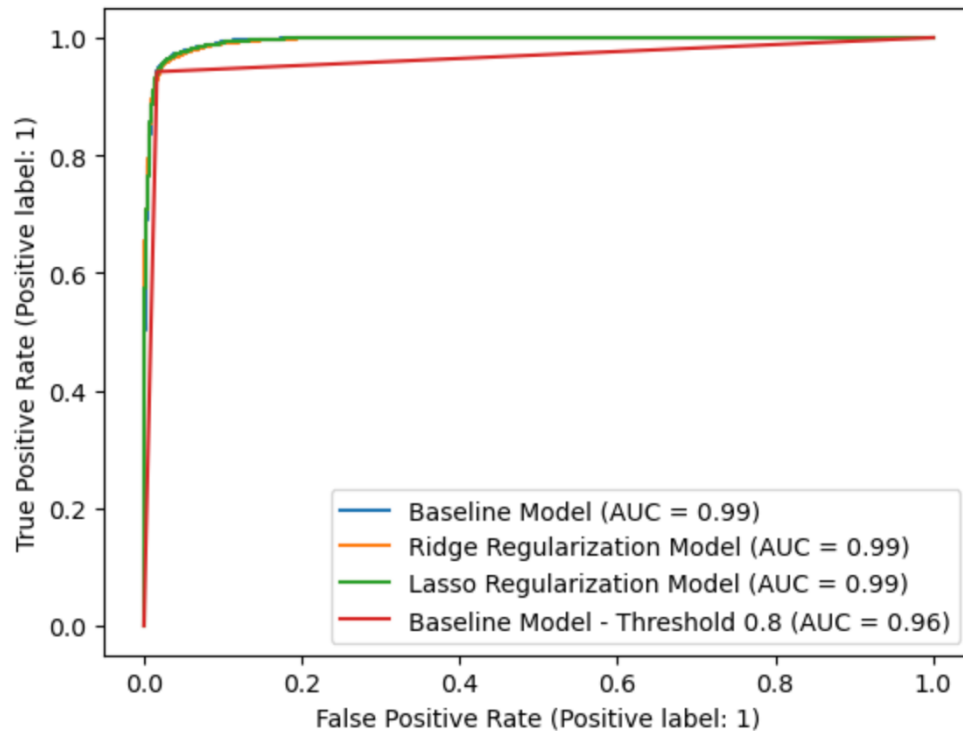Precision: 95.7%   F1: 96.5%
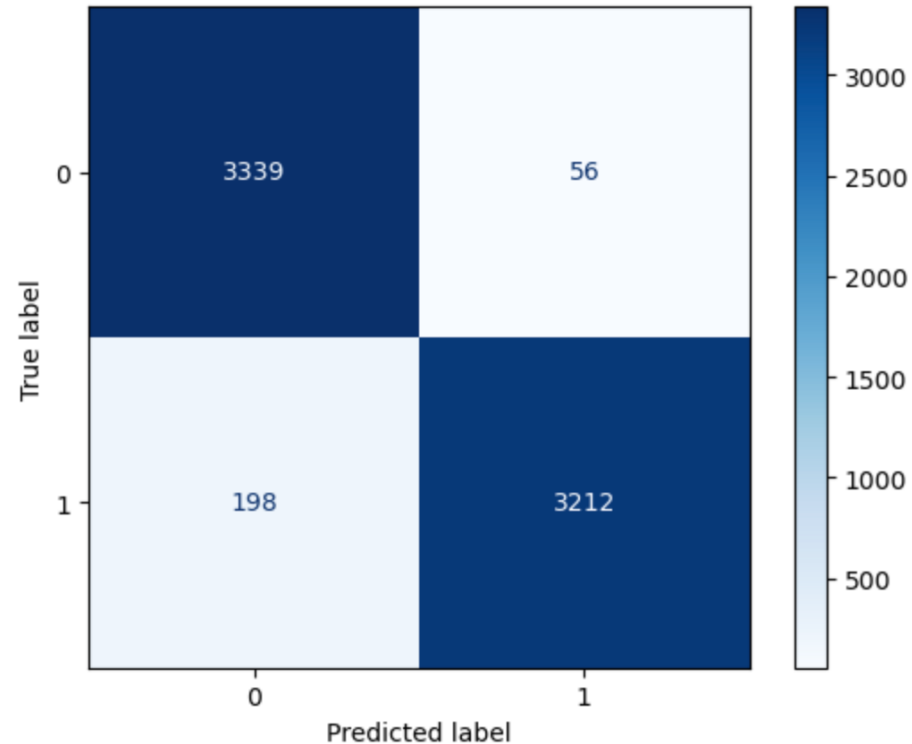
Cross Validation Avg. Scores:
Train: 95.8%
Test: 95.0%

1) Max iterations: 5,000, random_state: 42
2) Max iterations: 5,000, random_state: 42, solver: liblinear

BASELINE MODEL WITH CUSTOM THRESHOLD OF 0.8

Accuracy: 96.2%

Precision: 98.3%

Recall: 94.2%

F1_score: 96.2%

Cross Validation Avg. Precision Scores:
Train: 98.3%
Test: 97.5%

*Sacrificing accuracy to improve precision and reduce false positives*

Increased false negatives, reduced false positives

1) Max iterations: 5,000, random_state: 42
2) Max iterations: 5,000, random_state: 42, solver: liblinear

FINAL MODEL EVALUATION ON UNSEEN TEST DATA

Baseline Model vs. Baseline w/ Threshold = 0.8
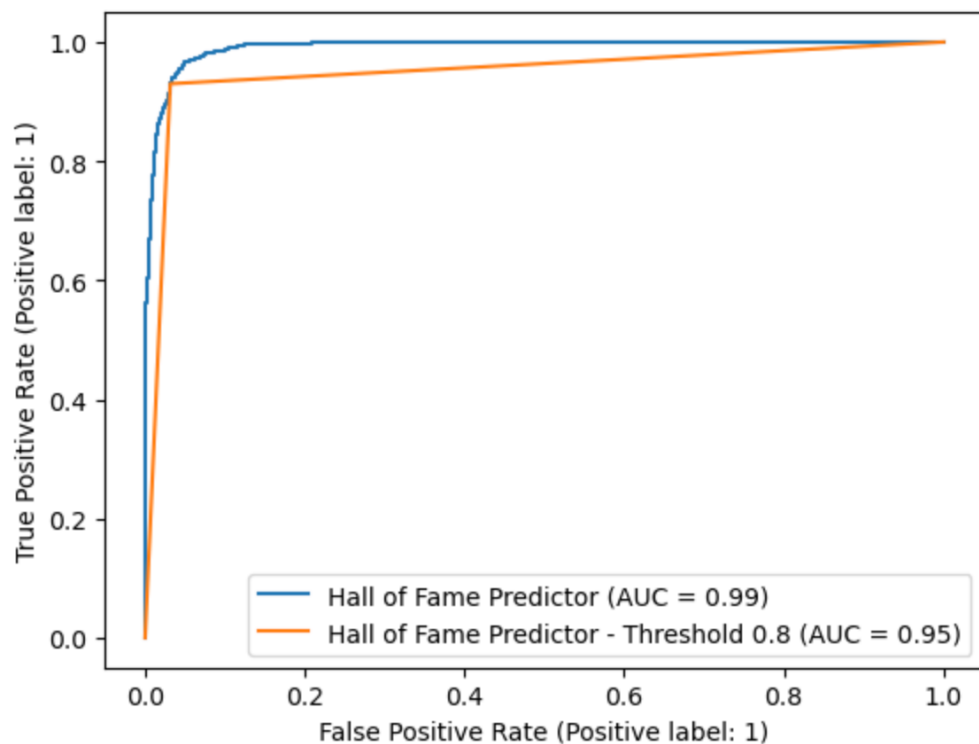
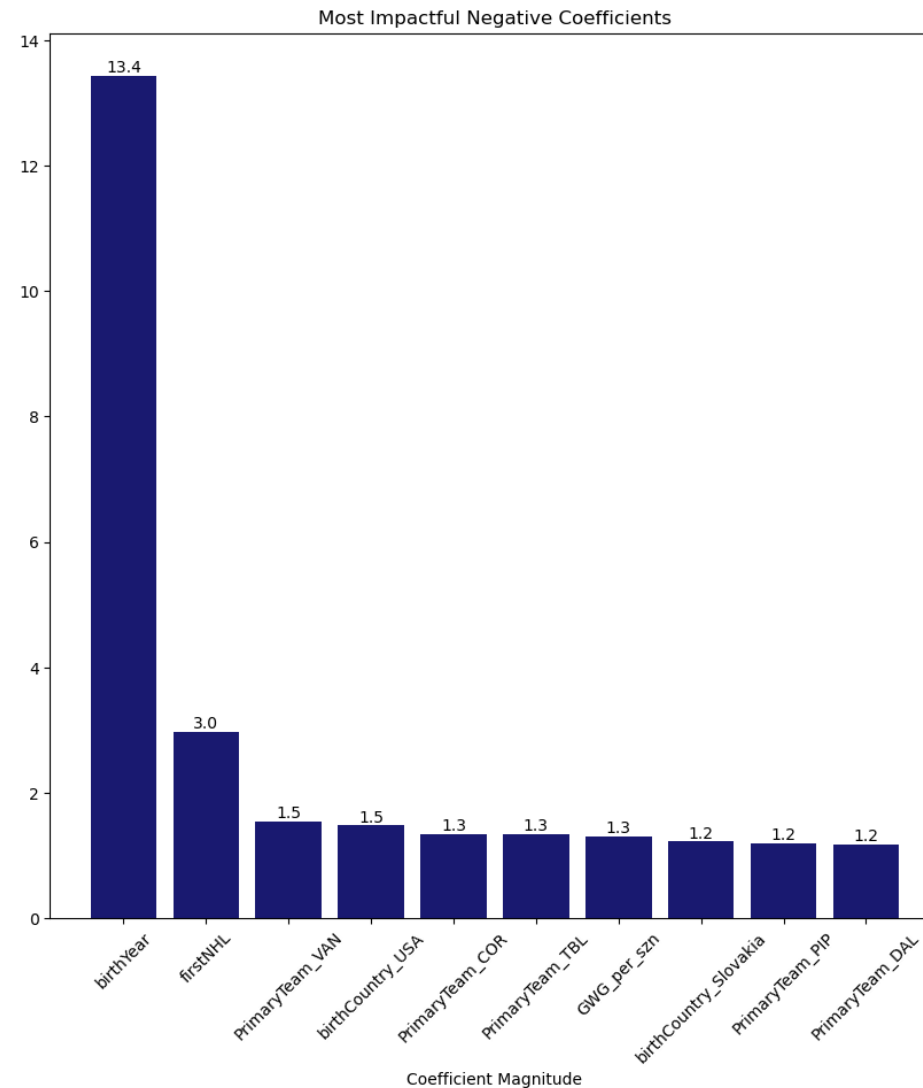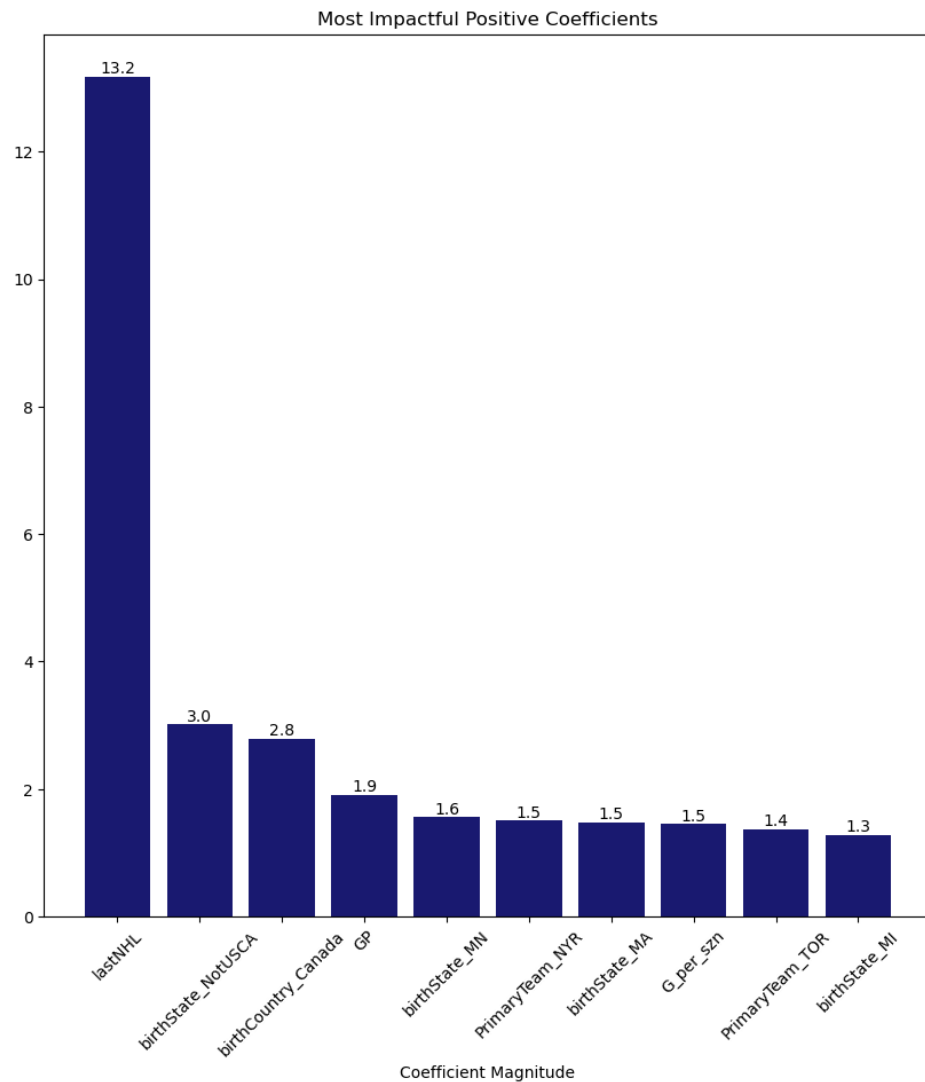**On unseen data, model with threshold = 0.8 performs with 96.7% precision**
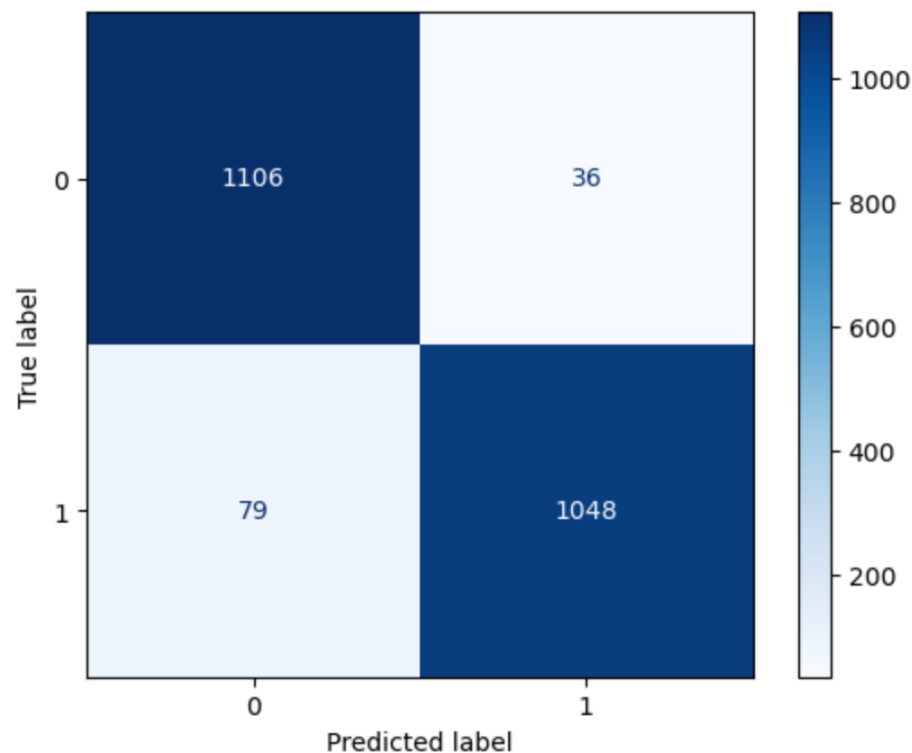
Precision: 96.7%

Accuracy: 94.9%

Recall: 93.0%

F1: 94.8%

MEANINGFUL COEFFICIENTS

Most Impactful Positive Coefficients

Most Impactful Negative Coefficients

- Games played, average goals per season most impactful statistics when assessing hall of fame eligibility

- Those who played mostly for NYR or TOR more likely to be of hall of fame caliber

- Non-US born players more likely to be inducted

- Players with more recent birth years (i.e. players who retired at a younger age) less likely to be of hall of fame caliber

CONCLUSION

- **Recommendations:**

  - Utilize our baseline model with a threshold set to 0.8 as a tool to determine who should be inducted into the hall of fame each year

  - Given model, put greater focus on candidates who are not US born or are from MN, MA, or MI

  - In addition, put focus on recently eligible players who have played many NHL games and have a high average goals per season in comparison to peers

Additional Considerations/Next Steps:

  - Opportunity to retrain model with more recent data given dataset ends in 2011

  - Utilize more advanced models to improve precision

  - Build an additional model to evaluate goaltenders

  - Develop a modified version of this model to assess current players' potential when contemplating trades or evaluating players in free agency

BIOGRAPHY

Emily Krueger

Email: ekrueger1217@gmail.com

M: 732-403-4566

Github: https://github.com/ekrueger1217

LinkedIn: https://www.linkedin.com/in/emily-krueger-058513103/