Machine Learning OTB

Applying Classification Algorithms to Horse Racing

Metis Bootcamp
Project 4 - Classification
Brandon McNeil

Context - Hong Kong Jockey Club:

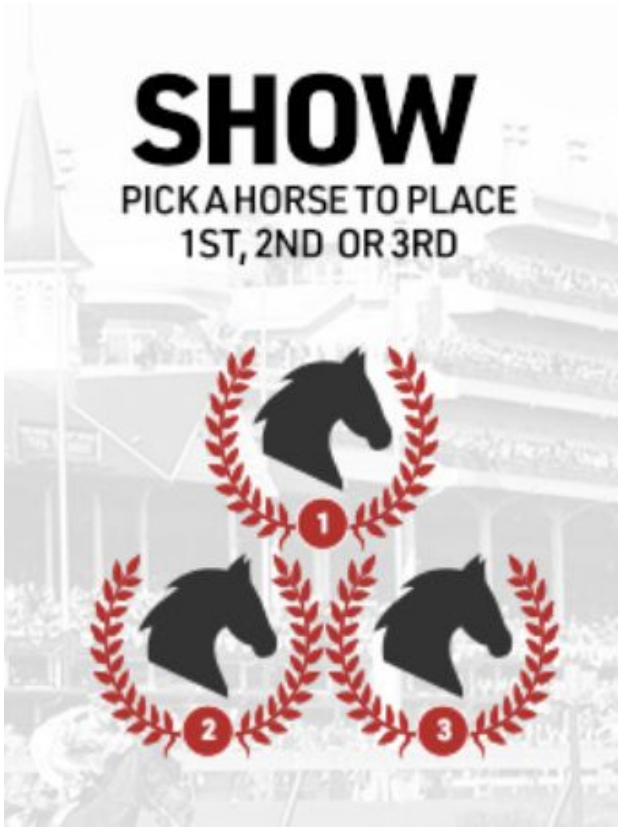
The Hong Kong Jockey Club is one of the most lucrative horse racing institutions in the world, attracting more than **17.86 million dollars** of bets per race.

Bill Benter - a professional gambler - and his team are alleged to have won around **$1 billion dollars over the course of 40 years** with the help of statistical algorithms.

Project Purpose:

- Build a classification model that can calculate the odds of a horse finishing in either **1st, 2nd or 3rd.**

- Gain the statistical edge while betting on a horse to **Show** (finish in 1st, 2nd, or 3rd place).



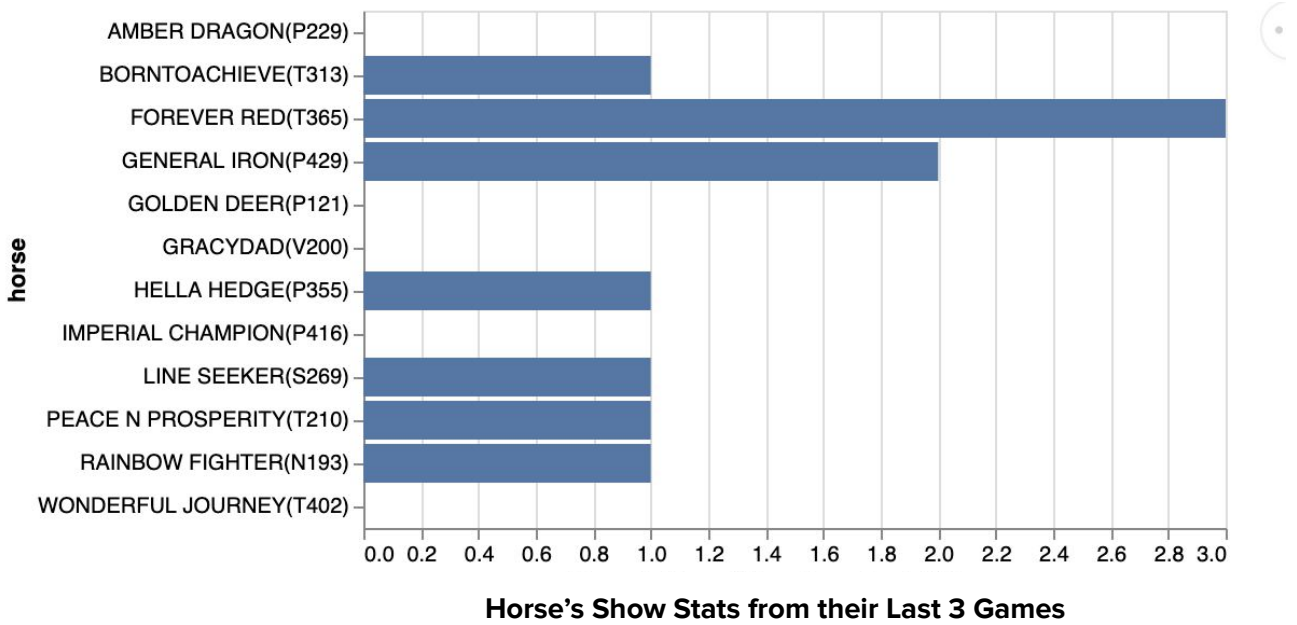
Our Dataset:

-Our data set contains the race results from the Hong Kong Jockey Club's 2014 - 2017 season.

-Contains over 13,000 rows of data collected from over 1,000 races.

Feature Engineering:

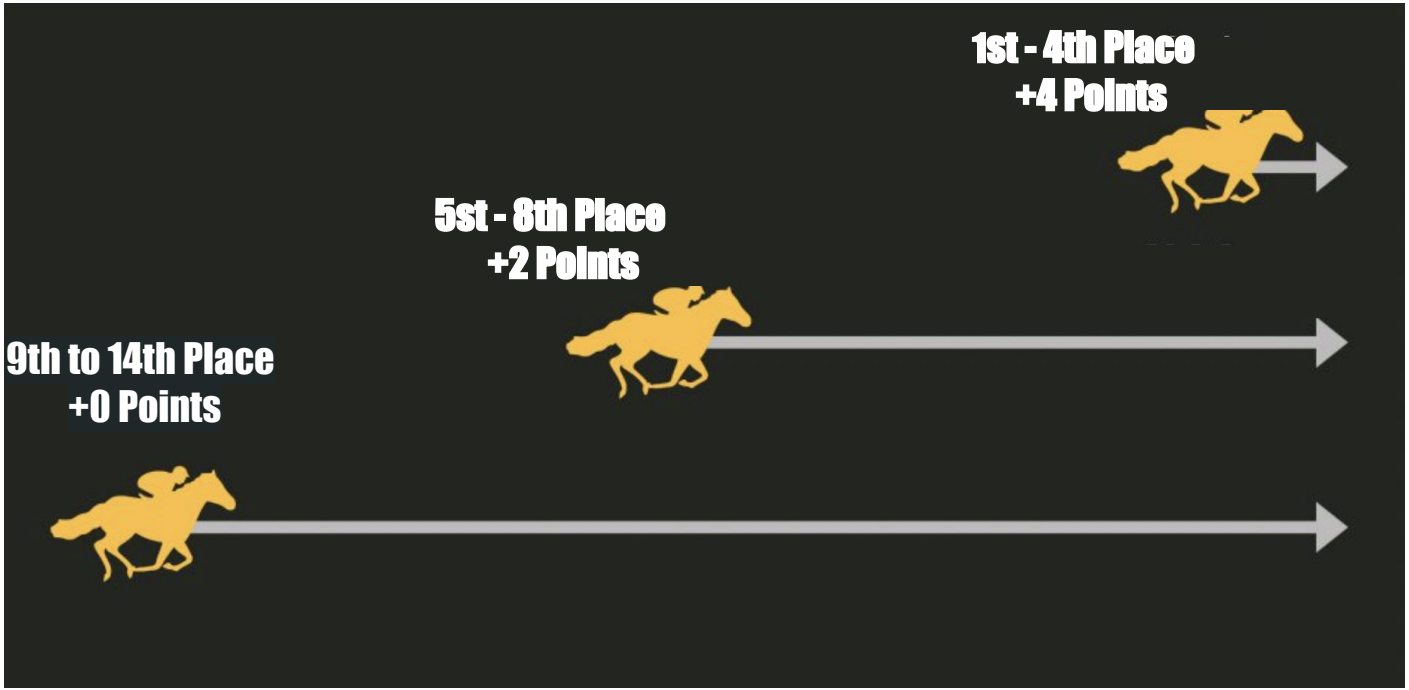
- Aggregated stats on the # of times a horse finished in 1st, 2nd or 3rd in their last 3 races.



Feature Engineering Continued!:

Position Score:

Generated a position score using a horse's recorded positions throughout their last 2 games.



Feature Engineering Continued!:

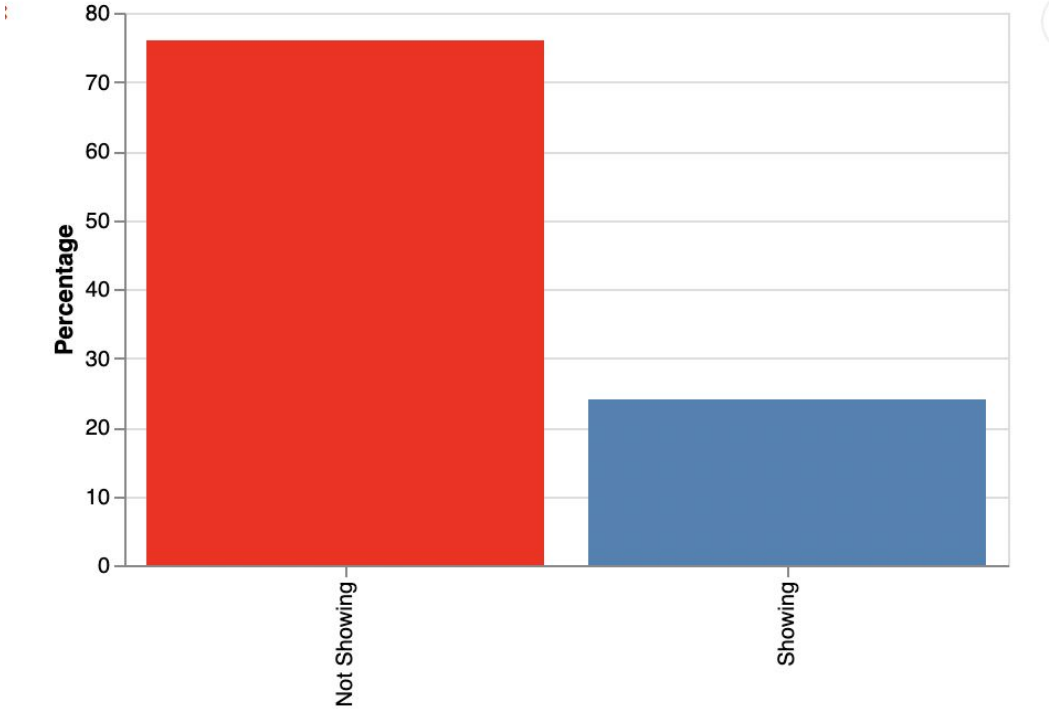
Pla.	Horse No.	Horse	Jockey	Trainer	Act. Wt.	Declar. Horse Wt.	Dr.	LBW	Running Position	Finish Time	Win Odds
10	8	SURE WINNER(D488)	K H Chan	K L Man	114	1021	8	19-3/4	3 5 10	1:12.15	57

4 + 2 + 0 = <u>6 position score</u>

Imbalanced Target Values:

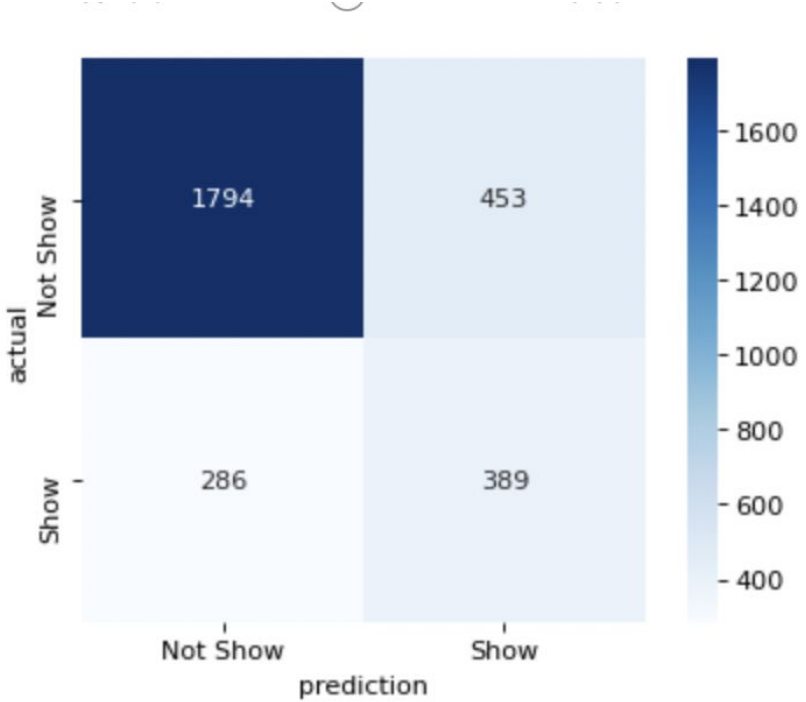
- To be expected!

- Only 3 horses can finish in the top 3 slots out of 14 contending horses.



Model Selection:

- Random Forest Model was the best performing model.
- Class Weights adjusted via hyperparameters to “balanced” to account for imbalanced target.

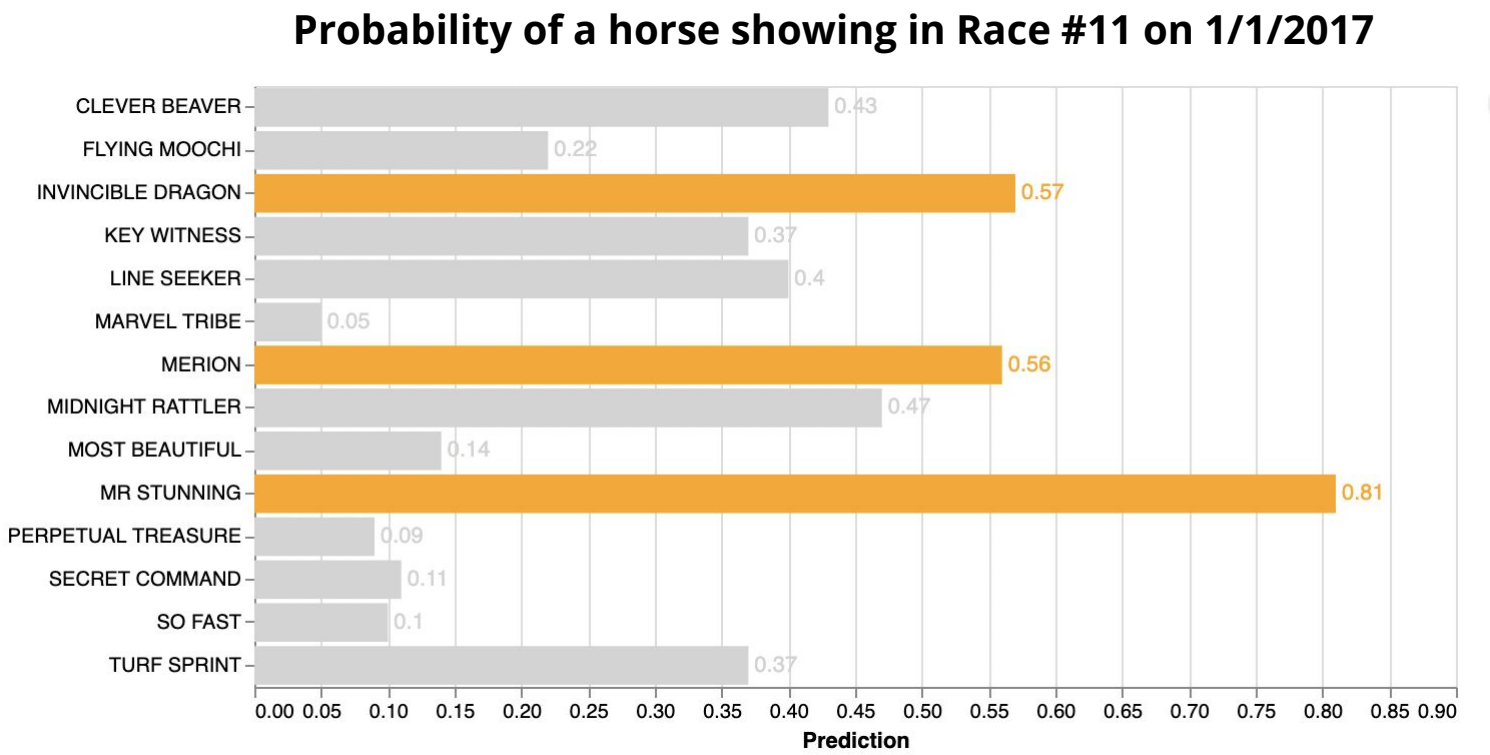


Recall Score: **0.58%**

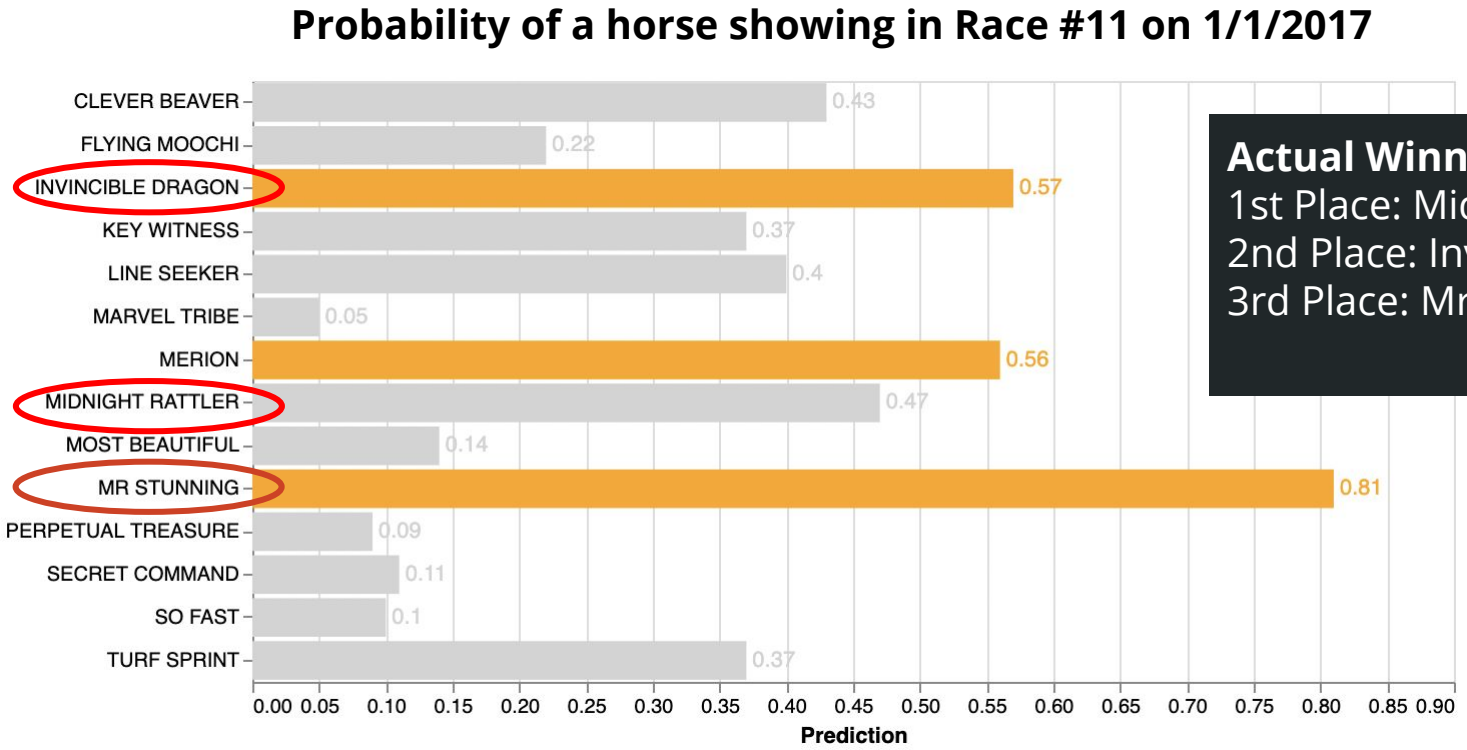
Precision Score: **0.46%**

ROC AUC Score: **0.77%**

Model Results:

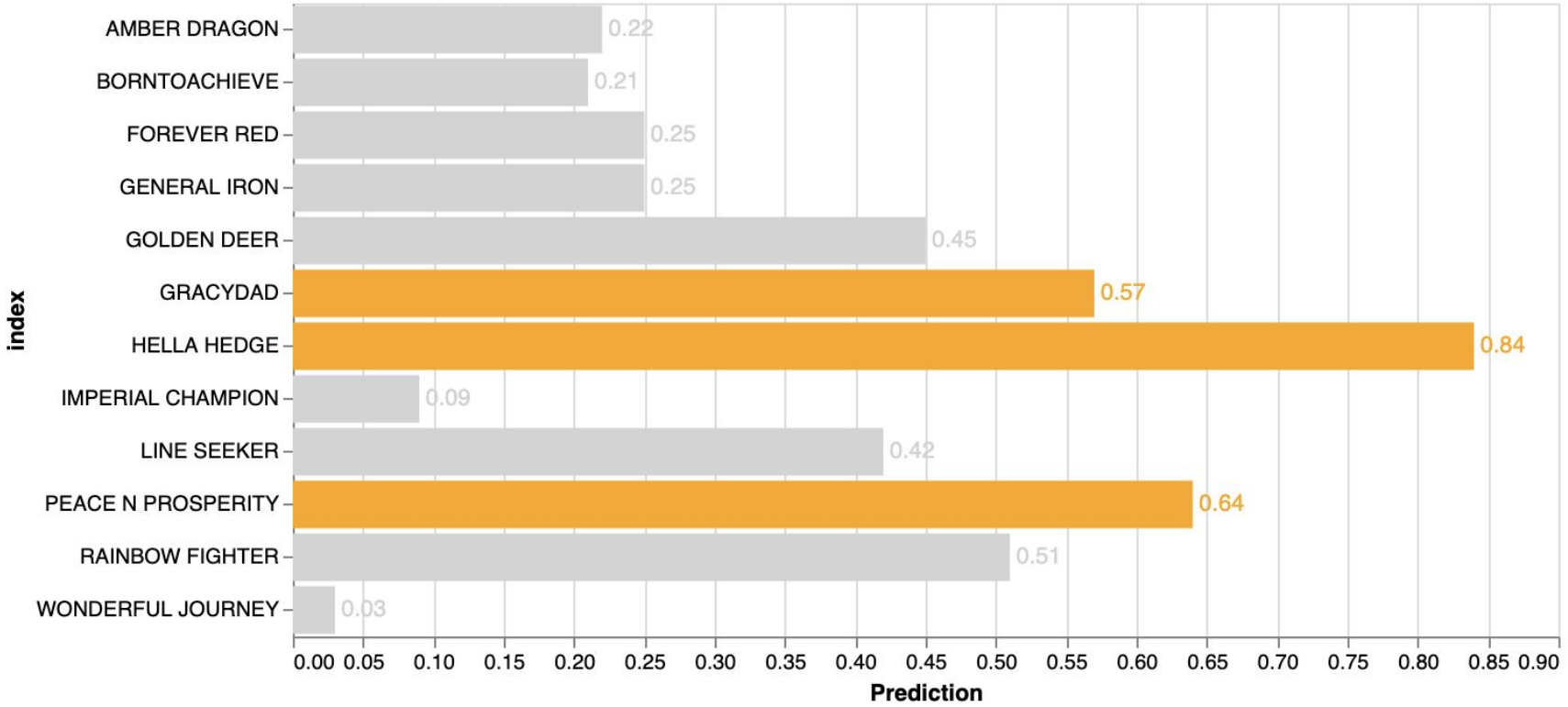


Model Results:



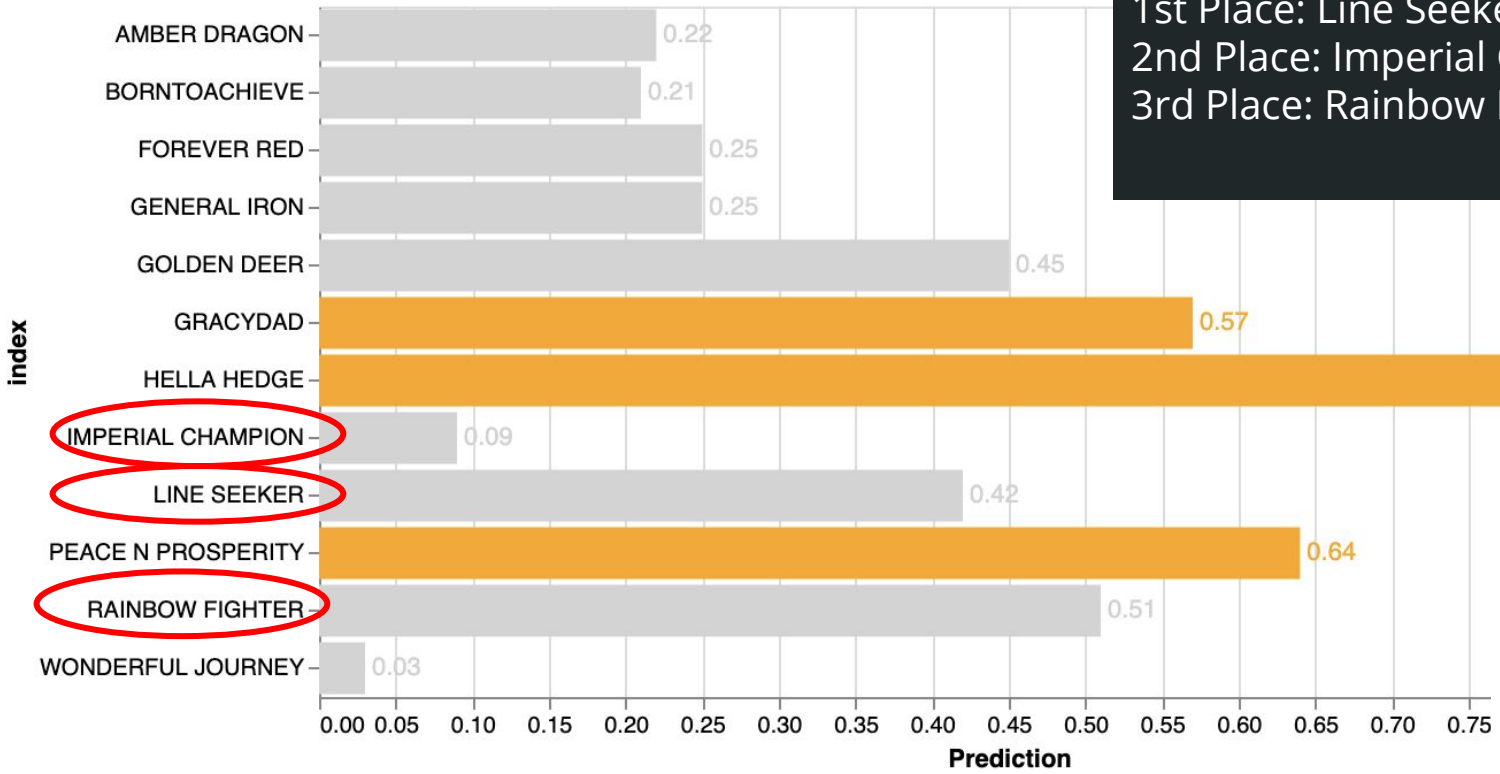
Actual Winners
1st Place: Midnight Rattler
2nd Place: Invincible Dragon
3rd Place: Mr Stunning

Model Results Part 2:



Model Results Part 2:

Actual Winners
1st Place: Line Seeker
2nd Place: Imperial Champion
3rd Place: Rainbow Fighter



Future Work:

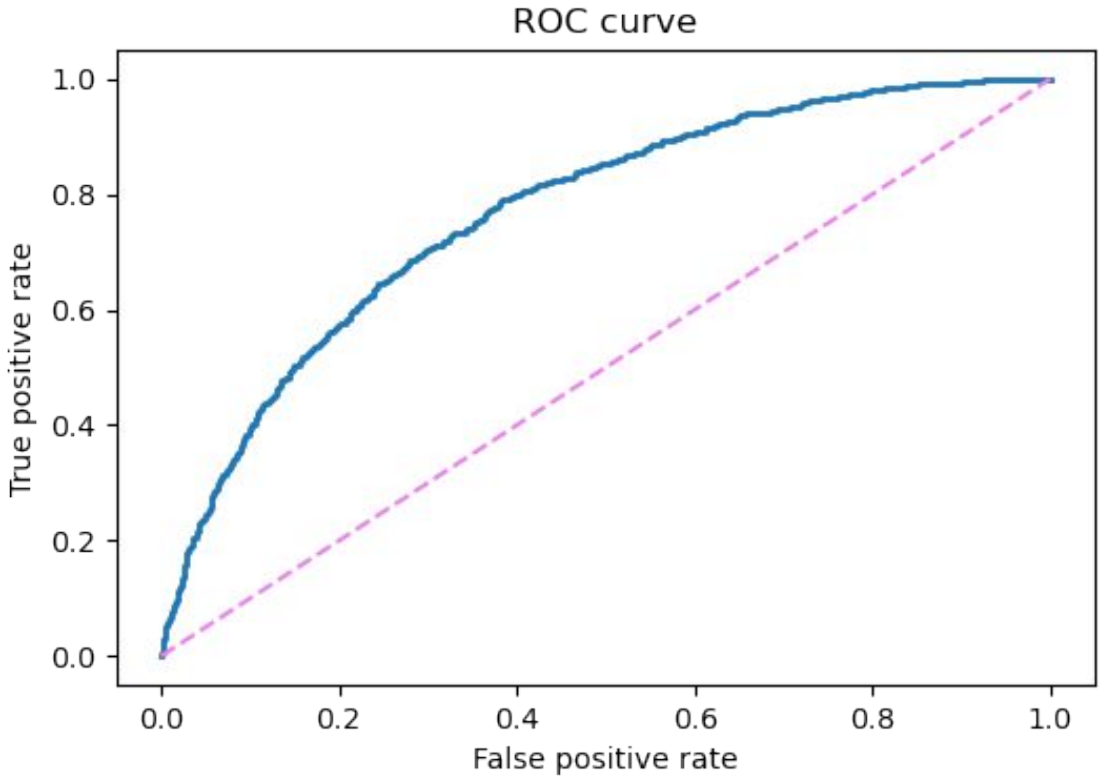
- Explore additional feature engineering to try to raise the ROC AUC score.

- Adjust Target Values to be a horse finishing a race in 1st or 2nd place and eventually only 1st place.

Appendix:

ROC / AUC Score:

- ROC AUC score of **.77%** on unseen data.



Feature Importance: