NFL Wagering – Predictive Analysis

Doug Marcum

January 9, 2021

Abstract

Sports betting is not a new phenomenon in the United States, but due to recent legislation, it is experiencing a surge. With 25 states now allowing their residents to wager on sporting events in varying capacities, consumers are searching for tools to assist in making predictions or rather, to gain an edge in their selection process. Resources are available, but often they are not validated, are high priced, or do not provide any true guidance. With the work being conducted under this proposal, the hope is that through data science, the average gambler can be informed in making their selections.

By analyzing ten, 10, NFL seasons consisting of 256 unique games each season, classification models can be implemented and tested to see if accuracy rates exceeding 52.4% can consistently be processed. If this is successful, future research into other professional arenas (basketball, baseball, soccer, hockey) can be reviewed. Sports betting, once considered to be taboo and sinful, has grown into a true corporate enterprise. Given this overall popularity, the field is just being explored, and the need for data scientists to engage and discover insights will continue to increase. This paper illustrates that machine learning can successfully predict winning and losing teams (approximate accuracy of 65%), as well as provide an edge in predicting if a team will cover a spread or if a game will be under or over an estimated total score.

Background

Humans love games, especially games of chance that provide a reward or prize when the correct outcome is selected. We enjoy the highs and lows that come along from engaging in games of chance, gambling. Gambling triggers the brain's reward system which are linked primarily to the pleasure and motivation centers and releases dopamine into the body. This makes the gambler feel elated while they are putting it on the line and taking risks. Dopamine is the dominant power driver and the chief neurotransmitter in the reward system. Gambling stimulates a “thrill” which triggers the reward system to release up to ten times more than the amount natural rewarding experiences would produce.

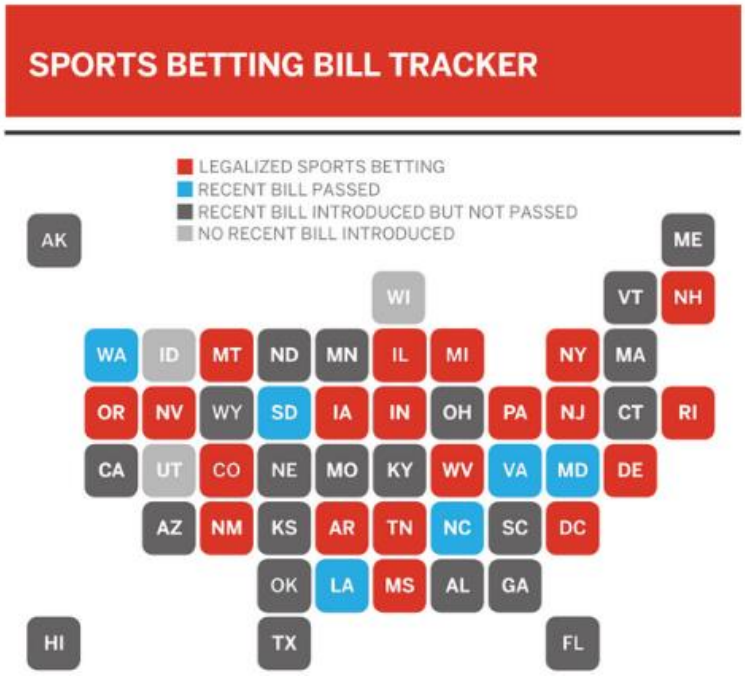
With our chemical drive to gamble, and the U.S. Supreme Court ruling, May 2018, in *Murphy v. National Collegiate Athletic Association* that the Professional and Amateur Sports Protection Act violated the 10th Amendment of the U.S. Constitution, thus leading to each state’s ability to legalize sports betting, the industry has witnessed tremendous growth. Publicly traded companies like DraftKings, FanDuel, and Penn Gaming have been able to operate sportsbooks and mobile sports betting applications in 25 states and the District of Columbia currently. By the year 2025, sports betting is predicted to generate $8 billion in revenue in the U.S, with some estimates as high as $25 billion.[1]

This industry is unlike the casino games of chance, as the gambler is able to use statistical know how to make a prediction, as opposed to a guess in say a game of roulette. As bookmakers gather more and more data, and play out tens of thousands of potential game scenarios, is it possible for the average better, armed with readily available statistics create a simple model that can provide them an edge in making their selection? This proposal is to determine just that.

Through analysis and modeling, it is hoped that the following research questions are answered:

- What features bear the most importance in determining a winning strategy?
- Is it possible to create a model that is simple enough to allow an average user to engage with and interpret the results without being overly complex?
- Can a model be generated that can make selections with an accuracy rate higher that 52.4%?

[1] Associated Press. (2019, November 4). Sports Betting Market Expected to Reach $8 Billion by 2025. Retrieved from https://www.marketwatch.com/story/firms-say-sports-betting-market-to-reach-8-billion-by-2025-2019-11-04.



[2]

Football, specifically the NFL, has been selected for the basis of this proposal as the style of play (multiple players participating separately in offense, defense, and special teams) discourages the effect one player can have on the outcome of each game. Additionally, with a shorter season than Major League Baseball and the National Basketball Association, resting players do not need to be accounted for on a game-to-game basis.

Data

The data, as with most major sports, is vast and comprehensive. What has been selected allowed for a great deal of comparison in determining multiple factors of success. The data set utilized was obtained from Stathead.com and Pro-Football-Reference.com, both professional sporting statistics resources and subsidiaries of Sports Reference. Information pertaining to Sports Reference can be found here: https://www.sports-reference.com/about.html.

With changes to rules of play occurring from decade to decade, the focus was placed on games played from the 2010 season through the completion of the 2019 season of the National Football League (NFL). Only regular season games were reviewed; no preseason or postseason games were included. Posted below is a sample of the initial data set features. All data files can be found at: https://github.com/MarcumDoug/NFL_Wagering-Predictive_Analysis.

Time	Game Time, Eastern	**Spread**	Vegas Line
Day	Day of Week	**vs Line**	Team's W-L-T record against the spread
Week	Week number in season	**Over/Under**	Over/Under
Game	Game number of team	**OU Result**	Was the final score over or under the O/U?

A complete glossary of features obtained and created are listed in Appendix A and Appendix B.

[2] Rodenberg, R. (2020, November 3). United States of Sports Betting: An Updated Map of Where Every State Stands. Retrieved from https://www.espn.com/chalk/story/_/id/19740480/the-united-states-sports-betting-where-all-50-states-stand-legalization.

Data Preparation

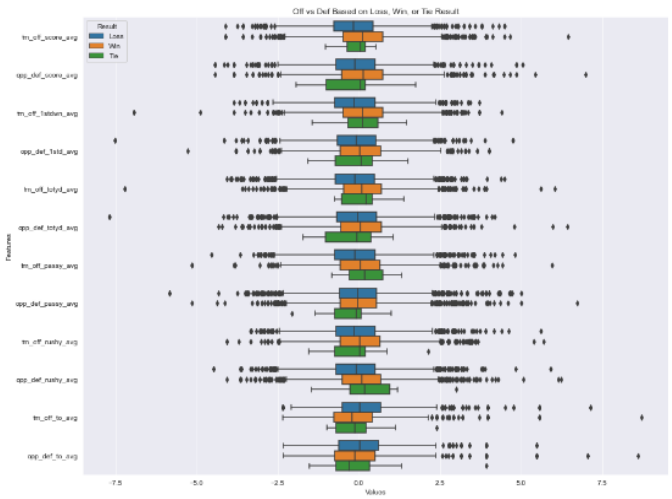
With Sports Reference requesting that their sites not be scraped for data, in addition to the necessary data being located in multiple locations, much of the data was manually collected and stored in CSV files. Extensive transformation and cleaning of the data was required to move forward with analysis. With much of the data needing to be numeric and categorical, conversions for objects were necessary on half of the initial features. Additionally, numerous new features were created to allow for in-depth analysis. Metrics (i.e., Scoring Average, Passing Yards, Turnovers) for teams and opponents needed to be calculated, averaged, and then coordinated from week to week. Features pertaining to winning and losing streaks, as well as streaks related to records against the spread and over / under, were created. All information and actions taken for data preparation is available via Jupyter Notebooks, coded in Python, on the GitHub repository page associated with this project. Those files can be found here: https://github.com/MarcumDoug/NFL_Wagering-Predictive_Analysis/tree/main/EDA.

Exploratory Data Analysis

Exploring the data led to some basic and interesting discoveries. The data consists of 5632 records, 2816 games from Season 2010 Week 1 through Season 2019 Week 17, which is comprised of one row for each team, as well as a row for the opponent, respectively.

When looking at the composition of our three targets (Wins, Spread, Over/Under) the variance is minimal for each. When first reviewing wins, losses, and ties, it was discovered that 49.84% of the games end in either a win or a loss, with ties making up the decision of 0.32% (this equates to 9 games over the course of eleven seasons). In regard to if teams did not cover, covered, or pushed, the percentages were 48.72% either cover or did not cover and 2.56% end in a push. The final category, under, over, push, is where a bit of variance was uncovered. Under results comprised 49.43% of the outcomes, over results 49.25%, and a push equating to 1.31%.

To explore further, the data for each category was separated into offensive compared to defensive averages, normalized [(*Feature – Feature mean) / Feature standard deviation*)], and charted into boxplots. The visuals for each category are posted below. The boxplots illustrate how the cliché that football is a game of inches is true. From the 25th percentile (lower quartile) to the 75th percentile (upper quartile), the statistical difference from a winning and losing team or a team that covers or does not is minimal. The difference appeared in the areas of the outliers. The separation from the top performing teams and teams with substandard results are drastic.



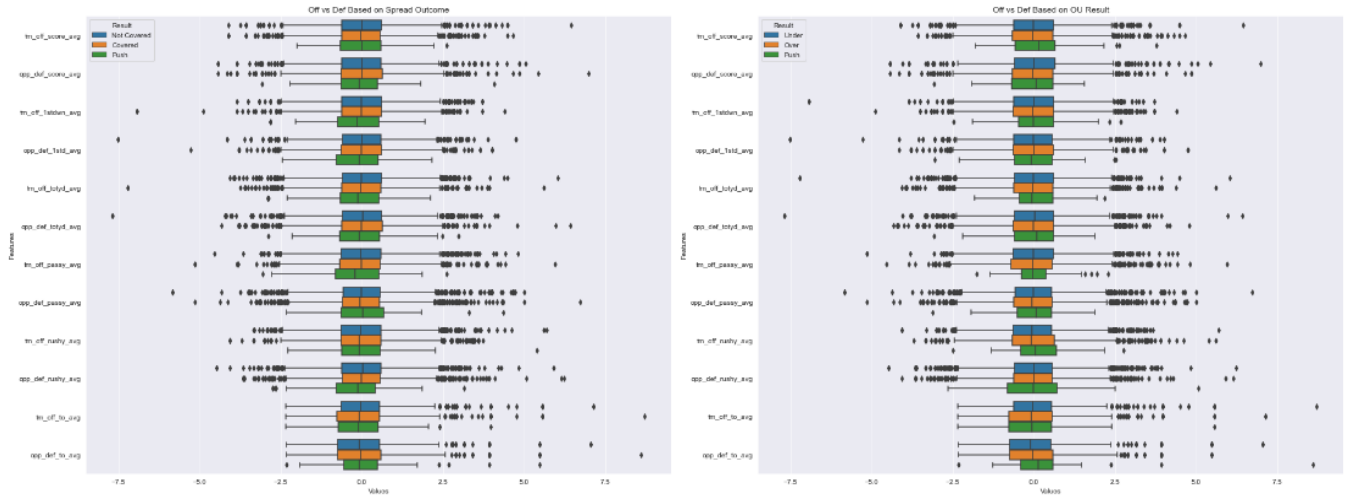


Figure 1, 2, 3: Larger versions of each visual can be found in Appendix C.

The games resulting in a tie or a push were removed for training and testing purposes. With the results being minimal for these outcomes, the purpose of the models is to classify the results of the game into two categories (win or lose, over or under, and cover or no cover). Additionally, week 1 of each season was dropped, as week one of each season establishes the moving averages for each team as the season progresses.

Feature Selection and Modeling

Feature selection was the most important and challenging portion of the analysis. Initially, the number of features were almost limitless, as the world of professional athletics tracks and accounts for every factor that can impact the game. In order to keep features limited, the rule of retaining a model that a layman could utilize was a guiding force. With this, it was vital to identify the most important features, but also not to eliminate those with minor significance. This was needed to preserve the integrity of the data, but also to provide a broader scope in each model's ability to accurately make consistent, valid predictions.

Standard Scaler was deployed for transforming the data. By utilizing Standard Scaler (z-score normalization), features were standardized by removing the mean and scaling to unit variance. The resulting features had a standard deviation of 1 and a mean close to zero. This allowed for the features to form nearly normal distributions. Outliers were a concern, but scaling was able to retain, yet stabilize them.

With three data sets (All Data, Over/Under Data, Spread Data) consisting of 66 - 67 features, a Random Forest Classifier model was utilized to determine feature importance. The Random Forest Classifier model collects the feature importance values so that the same can be accessed via the feature_importances_ attribute after fitting the Random Forest Classifier model. This is illustrated in the chart below, with results of the Over/Under feature importance analysis being displayed.

After features of importance were determined, additional analysis was completed to determine if multicollinearity was present in said features. Correlation heatmaps were constructed to quickly visual features with high correlations. By removing features that had correlations higher than or equal to 0.75, initial features were reduced substantially (Spread Models – 29 features, Over/Under Models – 31 features, Winner Models – 25 features).

Data was randomly split into 80% training and 20% testing data sets. With the purpose of this project to create simple models for the average gambler, classification not regression will be the focus. Determining a win or loss probability, as opposed to predicting specific scoring outcome, can be easier to understand. Since this was a classification prediction problem, the following models were selected for evaluation: Random Forest Classifier, Logistic Regression, K Neighbors Classifier, Gaussian Naïve Bayes, and an Artificial Neural Network. A simple function was created for displaying each model's metrics. Each model was fit and tuned using the training data set, and a prediction was made and evaluated using the testing data. Models were constructed via Python.

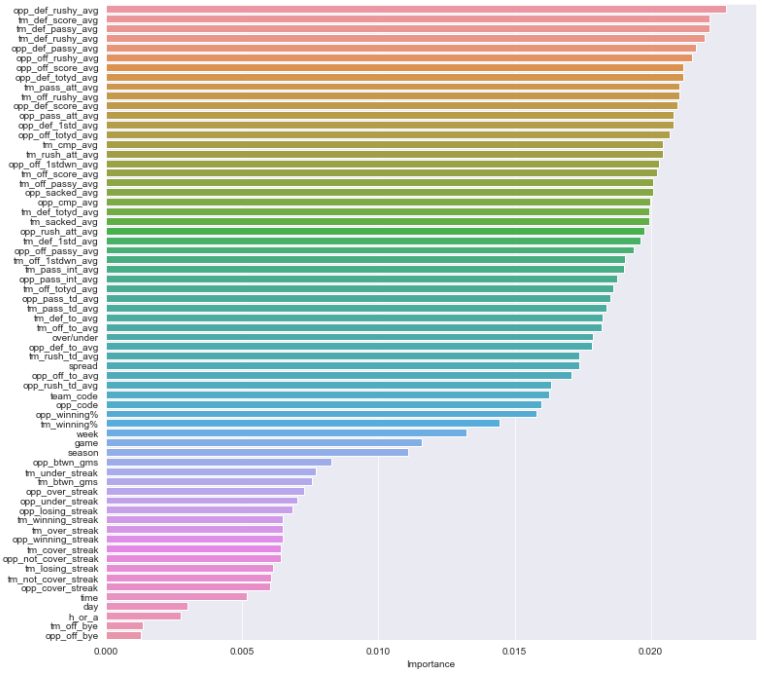


Figure 4: Larger version of visual can be found in Appendix D.

Details pertaining to feature selection and model construction can be found here: https://github.com/MarcumDoug/NFL_Wagering-Predictive_Analysis/tree/main/Models.

Results

Results varied from category to category. Looking first at the results when predicting a winning team, the models performed reasonably well and consistent across the board. Accuracy ranged from 64 – 66%, with Precision, Recall, and F1-Score all scoring relatively the same. This was better than expected, as the predictions were solely based off the previous seasons data and no new data being added during the course of the 2020 season.

The models for predicting a team's ability to cover or not cover a point spread and those to determine if a total score will be over or under have been inconsistent. Each model's accuracy score was based off performing 10 cross validations. The results of the Over / Under predictions are as follows:

Random Forest	Logistic Regression	K Nearest Neighbor	Gaussian NB	Artificial Neural Network
51.68%	52.31%	51.38%	52.73%	52.64%

As can be seen, the results are only slightly better than 50%, but two models, GaussianNB and Artificial Neural Network (ANN), were able to slight come out ahead of the 52.4% threshold. The results of each model's prediction ability relating to the spread are listed below:

Random Forest	Logistic Regression	K Nearest Neighbor	Gaussian NB	Artificial Neural Network
49.86%	49.53%	51.13%	50.28%	53.16

Only the ANN model was able to exceed the 52.4% mark this round. The others hovered around 50%. Even with these results, the models were tested against the 2020 season through week 16, on a weekly basis. The results of each models' weekly results can be viewed in Appendix E.

After each model completed the 2020 weekly season cycle and the results reviewed, an additional 'ensemble model' was created. This model was designed to look at games in which the model had a high level of confidence in its prediction and not to review the prediction of every game. This was done as it is extremely

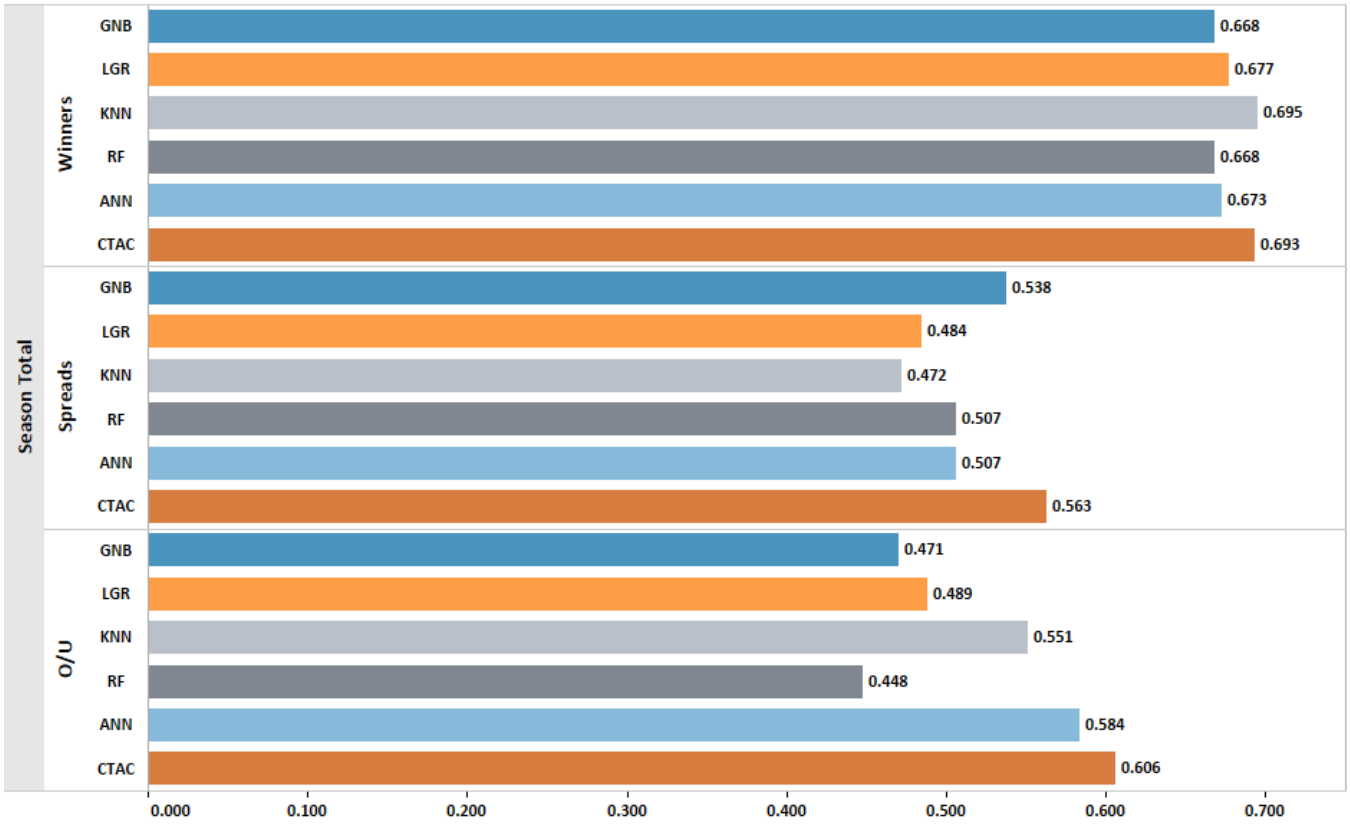
uncommon for a gambler to place a wager on every game, every week. Instead, a gambler is looking for games with a high confidence level or one that the wagering odds are advantageous. While the new 'model' is not a true ensemble model or an actual model, certain metrics were used in evaluating games:

1. The confidence threshold was established after week 2 predictions were finalized. These baseline confidence levels can be seen here: https://github.com/MarcumDoug/NFL_Wagering-Predictive_Analysis/tree/main/Results.

2. After the new data frames were created, each game was evaluated to determine a consensus prediction from the models. A majority selection (i.e., 2-1) was necessary for a game to be allowed into a selection. If there was no clear consensus (i.e., 2-2), the game was labeled a push and no selection metric was made. The model was titled Confidence Threshold Adjusted Consensus (CTAC).

The results from this model were impressive. While the number of games selected to meet the criteria of selection was less than the other five, the number of games selected was still significant. The final season outcomes for all models are listed below.

Model Prediction Accuracy



ANN, CTAC, GNB, KNN, LGR and RF for each Type broken down by Week. Color shows details about ANN, CTAC, GNB, KNN, LGR and RF. The view is filtered on Week, which keeps Season Total.

Conclusion

Overall, it is clear that predictive analytics are a tool that can be utilized in the sports betting industry. Oddsmakers have been using statistics for decades to set the table, and now, they are using algorithms to become more precise in their estimates. Today, with a little knowhow, the average gambler has access to more information and analysis than any other time in history. The models constructed for this project focused on maintaining a simplistic approach to data gathering, cleaning, and formatting. Limited advanced analytics were put in place, so that a novice to data science could follow due to their understanding of football betting.

The questions this project set out to answer were tackled head on. Important features and variables were easily identified, a number of models were able to consistently exceed the accuracy rate of 52.4% (allowing them to technically be profitable if used in a gaming situation), but it is still questionable if the average gambler could accomplish this task. The data is readily available, but the skills needed to gather, screen, and analyze the information is most likely out of the scope of someone without programming and data analysis skills. That is not to say with basic research the task could not be accomplished, but rather, research and training would be needed. Additionally, more features and parameters would be in order for more sophisticated models, particularly if regression were to be utilized to predict actually scores by both teams.

As gambling continues to grow in popularity and accessibility throughout the United States, it would be in poor judgement not to advise discretion in making decisions pertaining to one's finances and wagering on sporting events. This project uses statistics and advanced analytics to generate predictions and not guarantee any results. Please only wager for fun and please act responsibly.

References

Associated Press. (2019, November 4). Sports Betting Market Expected to Reach $8 Billion by 2025. Retrieved from https://www.marketwatch.com/story/firms-say-sports-betting-market-to-reach-8-billion-by-2025-2019-11-04.

Bartels, R. (2019, March 8). Beating the Bookies with Machine Learning. Retrieved from https://www.kdnuggets.com/2019/03/beating-bookies-machine-learning.html.

Browne-Anderson, H. (2019, February 18). Marco and Hugo Discuss the Role of Data Science in Large-Scale Bets and Bookmaking. Retrieved from https://www.datacamp.com/community/blog/data-science-gambling-bookmaking.

Buchdahl, J. (2018, January 5). The Problem with Data Mining in Sports Betting. Retrieved from https://www.pinnacle.com/en/betting-articles/educational/data-mining-in-sports-betting/4FC29M3VJY59Y4Q3.

Buchdahl, J. (2020, January 10). Using Bayes Factor to Assess Betting Skill. Retrieved from https://www.pinnacle.com/en/betting-articles/educational/bayes-factor-betting-skill-part-one/W9CJUFMSGHQDWKWJ.

Eulig, S. (2019, September 24). Scraping and Exploring Sports Betting Data – Is Arbitrage Possible? Retrieved from https://towardsdatascience.com/scraping-and-exploring-sports-betting-data-is-arbitrage-possible-a-hands-on-analysis-with-code-2ba656d7f5b.

Hubáček, O. & Sourek, G. & Železný, F. (2019, February). Exploiting sports-betting market using machine learning. International Journal of Forecasting. Retrieved from https://www.researchgate.net/publication/331218530_Exploiting_sports-betting_market_using_machine_learning.

Kurcharski, A. (2016, April 30). The Perfect Bet: How the Science of Gambling Influences Everything Around Us. Retrieved from https://www.independent.co.uk/life-style/perfect-bet-how-science-gambling-influences-everything-around-us-a7000741.html.

Lemmon, B.G. (2019, April 28). A Whole New World: Data Science and Sports Gambling. Retrieved from https://blog.usejournal.com/a-whole-new-world-data-science-and-sports-gambling-ad87dc064162.

Rodenberg, R. (2020, November 3). United States of Sports Betting: An Updated Map of Where Every State Stands. Retrieved from https://www.espn.com/chalk/story/_/id/19740480/the-united-states-sports-betting-where-all-50-states-stand-legalization.

Sayre, K (2020, September 14). The NFL is Back and Sports Bettors Are Following. Retrieved from https://www.wsj.com/articles/the-nfl-is-back-and-sports-bettors-are-following-11600112703.

Stapleton, A. (2020, September 4). The NFL is Now Betting on Once Taboo Gambling Industry. Retrieved from https://www.aol.com/article/news/2020/09/04/the-nfl-is-now-betting-big-on-once-taboo-gambling-industry/24611449/.

Appendix A

Initial Data Features with Glossary

Feature	Definition	Feature	Definition
Season	Year Game Occurs	**Team_Off_1stDwn**	Game Number First Downs - Team
Week	Week Game Occurs	**Team_Off_TotYd**	Game Total Offensive Yards - Team
Game	Game Number of Team	**Team_Off_PassY**	Game Passing Yards - Team
Day	Day Game Occurs	**Team_Off_RushY**	Game Rushing Yards - Team
Date	Game Date	**Team_Off_TO**	Game Turnovers - Team
Time	Game Time, Eastern	**Team_Def_1stD**	Game Number First Downs Allowed - Team
Off Bye	Previous Week Team Did Not Play	**Team_Def_TotYd**	Game Total Offensive Yards Allowed - Team
Result	Win, Loss, or Tie	**Team_Def_PassY**	Game Passing Yards Allowed - Team
OT	Did Game Have an Overtime Period	**Team_Def_RushY**	Game Rushing Yards Allowed - Team
Wins	Total Wins of Team	**Team_Def_TO**	Game Turnover Recovered - Team
Losses	Total Losses of Team	**Offense**	Expected Offense Points - Team
H_or_A	Home or Away Game	**Defense**	Expected Defense Points - Team
Team	Team	**Sp. Tms**	Expected Special Teams Points - Team
Opp	Opponent Team is Playing	**Spread**	Vegas Line
Team_Score	Final Team Score of Game	**Spread_Outcome**	Team's W-L-T record against the spread
Opp_Score	Opponent Team Score of Game	**Over/Under**	Over/Under
OU Result		Was the final score over or under the O/U?	

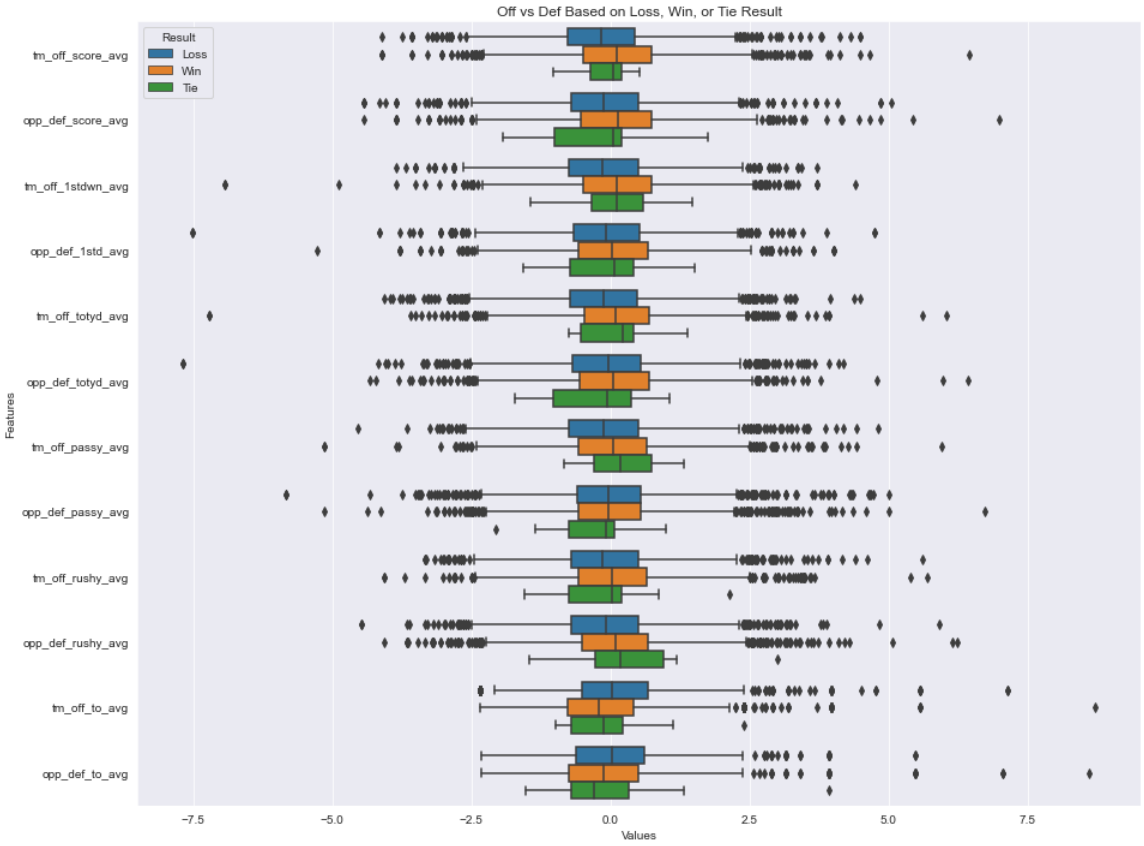
Appendix B

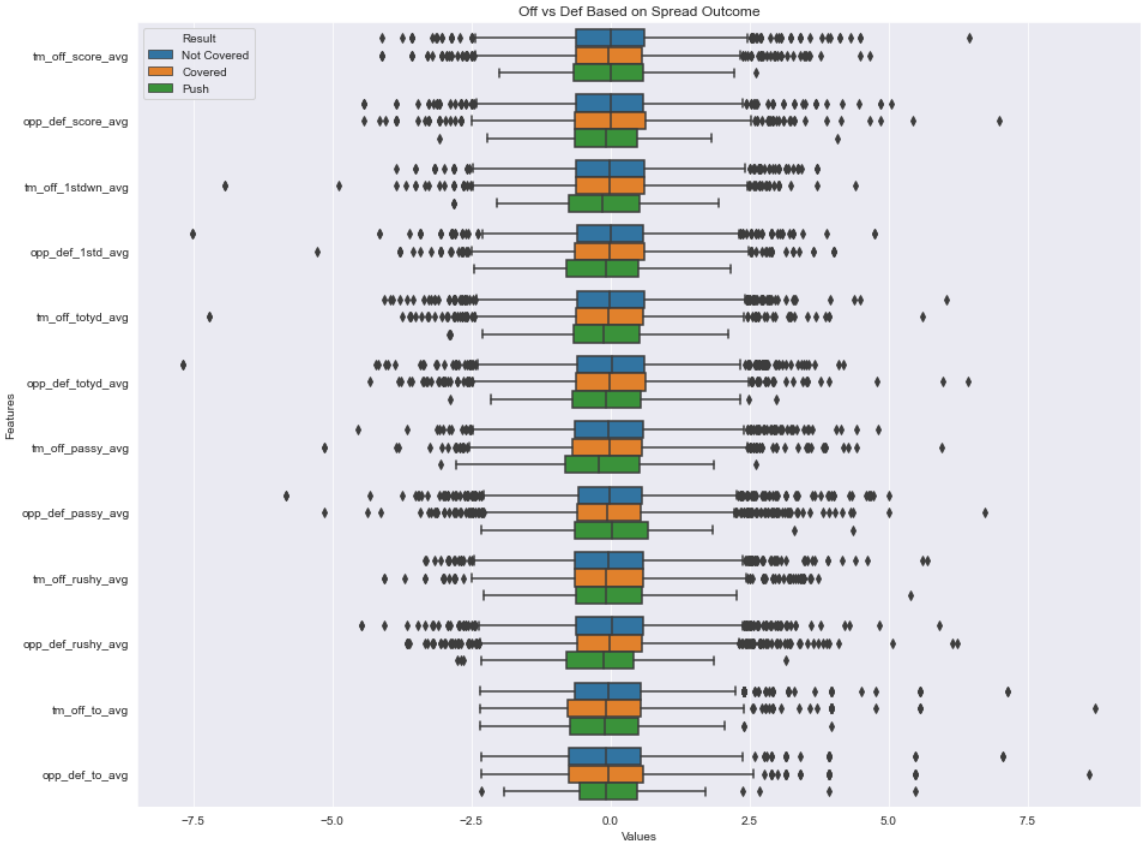
Final Data Features with Glossary

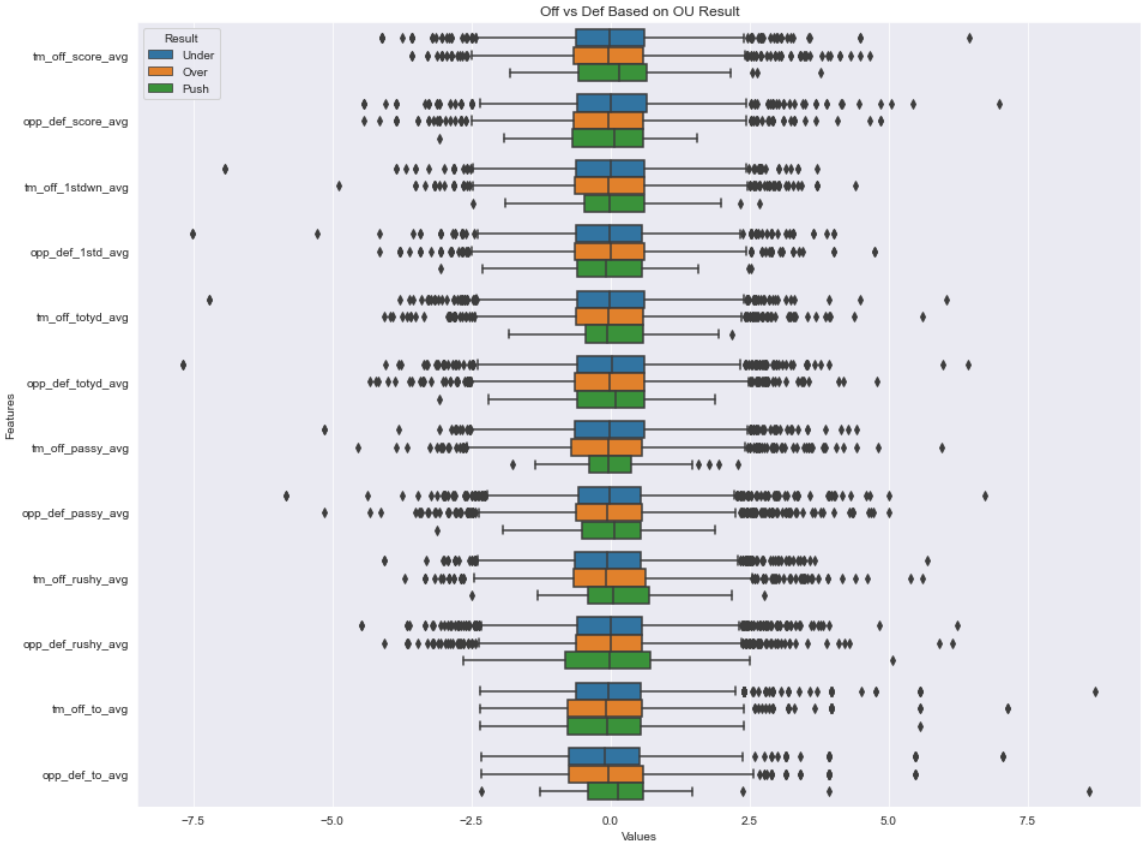
Feature	Definition	Feature	Definition
season	Year Game Occurs	**tm_not_cover_streak**	Team Consecutive Games Not Covering Spread
week	Week Game Occurs	**tm_cover_streak**	Team Consecutive Games Covering Spread
game	Game Number of Team	**tm_under_streak**	Team Consecutive Games Total Score of Under
day	Day Game Occurs	**tm_over_streak**	Team Consecutive Games Total Score of Over
time	Game Time, Eastern	**opp_off_score_avg**	Opponent Scoring Average Per Game
tm_off_bye	Previous Week Team Did Not Play	**opp_def_score_avg**	Opponent Scoring Allowed Average Per Game
opp_off_bye	Previous Week Opponent Did Not Play	**opp_off_1stdwn_avg**	Opponent First Down Average Per Game
tm_btwn_gms	Team Days Off Between Games	**opp_off_totyd_avg**	Opponent Offensive Yards Average Per Game
opp_btwn_gms	Opponent Days Off Between Games	**opp_off_passy_avg**	Opponent Passing Yards Average Per Game
h_or_a	Home or Away Game	**opp_cmp_avg**	Opponent Completion Average Per Game
team	Team	**opp_pass_att_avg**	Opponent Passing Attempts Average Per Game
opp	Opponent Team is Playing	**opp_pass_td_avg**	Opponent Passing Touchdowns Average Per Game

tm_winning%	Team Winning Percentage	**opp_pass_int_avg**	Opponent Passing Interceptions Average Per Game
opp_winning%	Opponent Winning Percentage	**opp_sacked_avg**	Opponent Sacks Against Average Per Game
tm_off_score_avg	Team Scoring Average Per Game	**opp_off_rushy_avg**	Opponent Rushing Yards Average Per Game
tm_def_score_avg	Team Scoring Allowed Average Per Game	**opp_rush_att_avg**	Opponent Rushing Attempts Average Per Game
tm_off_1stdwn_avg	Team First Down Average Per Game	**opp_rush_td_avg**	Opponent Rushing Touchdown Average Per Game
tm_off_totyd_avg	Team Offensive Yards Average Per Game	**opp_off_to_avg**	Opponent Turnovers Average Per Game
tm_off_passy_avg	Team Passing Yards Average Per Game	**opp_def_1std_avg**	Opponent First Downs Allowed Average Per Game
tm_cmp_avg	Team Completion Average Per Game	**opp_def_totyd_avg**	Opponent Total Offensive Yards Allowed Average Per Game
tm_pass_att_avg	Team Passing Attempts Average Per Game	**opp_def_passy_avg**	Opponent Total Passing Yards Allowed Average Per Game
tm_pass_td_avg	Team Passing Touchdowns Average Per Game	**opp_def_rushy_avg**	Opponent Rushing Yards Allowed Average Per Game
tm_pass_int_avg	Team Passing Interceptions Average Per Game	**opp_def_to_avg**	Opponent Turnovers Caused Average Per Game
tm_sacked_avg	Team Sacks Against Average Per Game	**opp_losing_streak**	Opponent Consecutive Loses
tm_off_rushy_avg	Team Rushing Yards Average Per Game	**opp_winning_streak**	Opponent Consecutive Wins
tm_rush_att_avg	Team Rushing Attempts Average Per Game	**opp_not_cover_streak**	Opponent Consecutive Games Not Covering Spread
tm_rush_td_avg	Team Rushing Touchdowns Average Per Game	**opp_cover_streak**	Opponent Consecutive Games Covering Spread
tm_off_to_avg	Team Turnovers Average Per Game	**opp_under_streak**	Opponent Consecutive Games Total Score of Under
tm_def_1std_avg	Team First Downs Allowed Average Per Game	**opp_over_streak**	Opponent Consecutive Games Total Score of Over
tm_def_totyd_avg	Team Total Offensive Yards Allowed Average Per Game	**result**	Win, Loss, or Tie
tm_def_passy_avg	Team Total Passing Yards Allowed Average Per Game	**team_score**	Team Final Score of Game
tm_def_rushy_avg	Team Rushing Yards Allowed Average Per Game	**opp_score**	Opponent Final Score of Game
tm_def_to_avg	Team Turnovers Caused Average Per Game	**spread**	Vegas Line
tm_losing_streak	Team Consecutive Loses	**spread_outcome**	Team's W-L-T record against the spread
tm_winning_streak	Team Consecutive wins	**over/under**	Over/Under
OU Result		Was the final score over or under the O/U?	

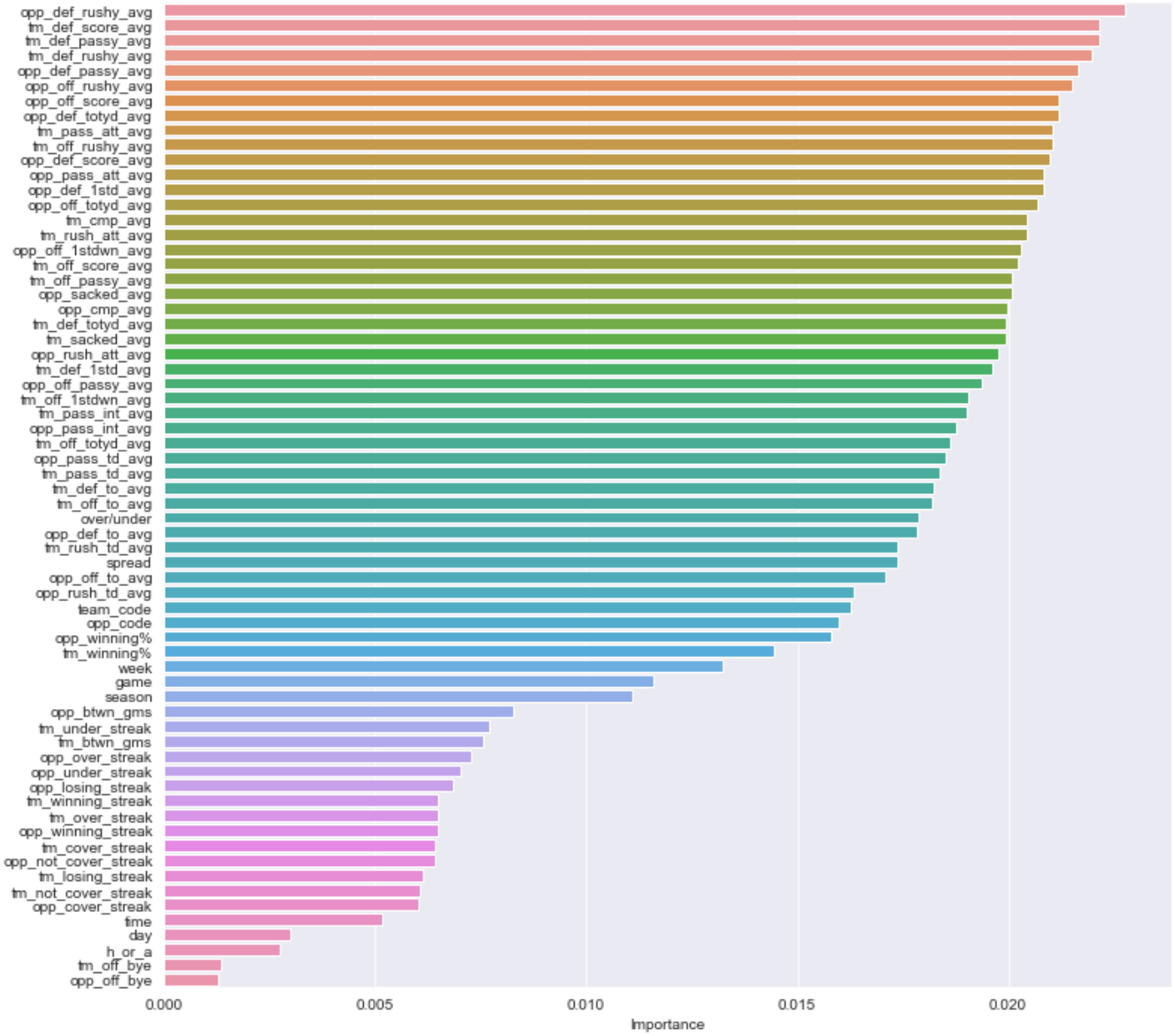
Appendix C


Off vs Def Based on Loss, Win, or Tie Result
Result
Loss
Win
Tie
tm_off_score_avg
opp_def_score_avg
tm_off_1stdwn_avg
opp_def_1std_avg
tm_off_totyd_avg
opp_def_totyd_avg
tm_off_passy_avg
opp_def_passy_avg
tm_off_rushy_avg
opp_def_rushy_avg
tm_off_to_avg
opp_def_to_avg
Features
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
Values


Off vs Def Based on Spread Outcome
Result
Not Covered
Covered
Push
tm_off_score_avg
opp_def_score_avg
tm_off_1stdwn_avg
opp_def_1std_avg
tm_off_totyd_avg
opp_def_totyd_avg
tm_off_passy_avg
opp_def_passy_avg
tm_off_rushy_avg
opp_def_rushy_avg
tm_off_to_avg
opp_def_to_avg
Features
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
Values

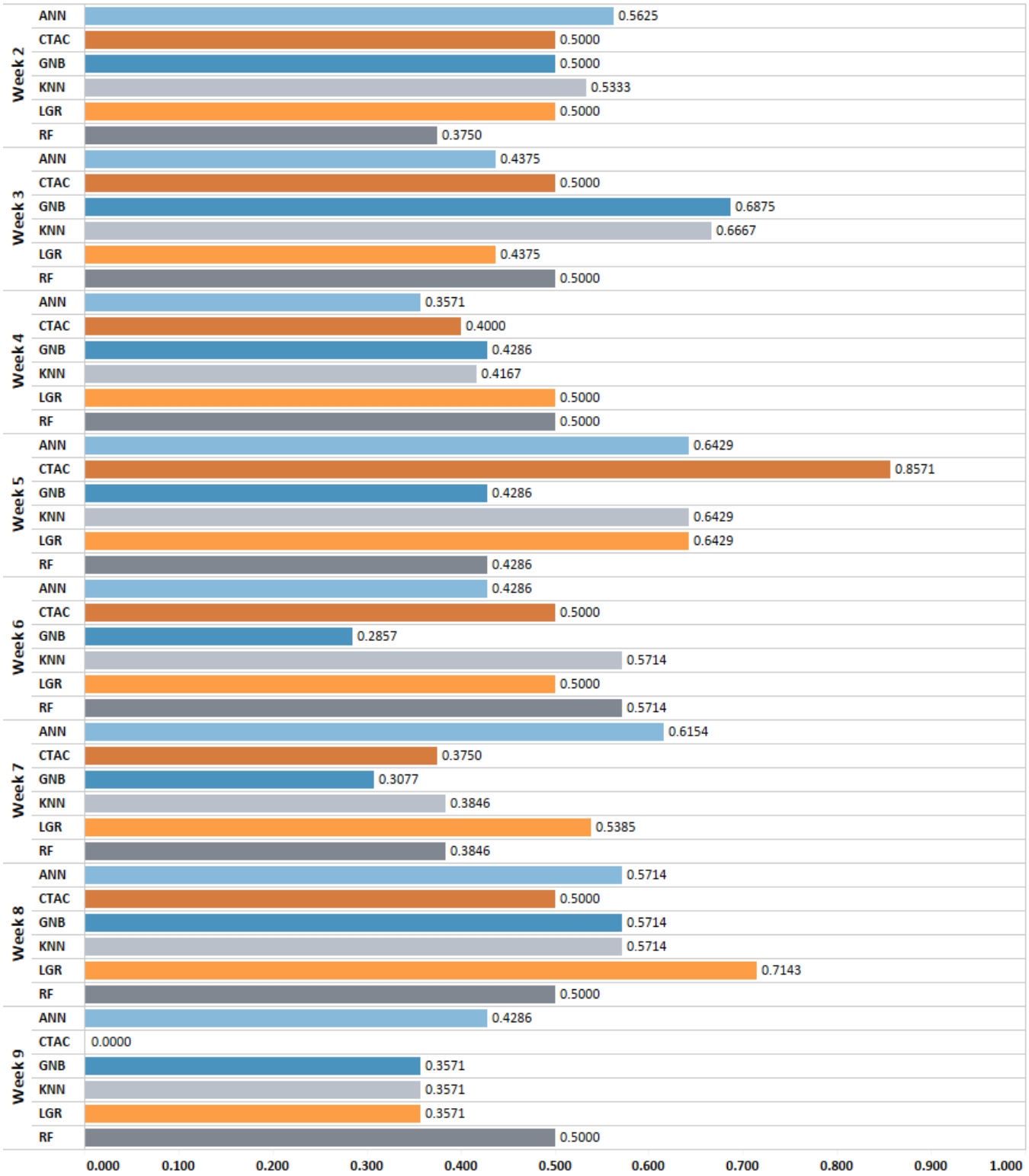

Off vs Def Based on OU Result
Result
Under
Over
Push
tm_off_score_avg
opp_def_score_avg
tm_off_1stdwn_avg
opp_def_1std_avg
tm_off_totyd_avg
opp_def_totyd_avg
tm_off_passy_avg
opp_def_passy_avg
tm_off_rushy_avg
opp_def_rushy_avg
tm_off_to_avg
opp_def_to_avg
Features
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
Values

Appendix D

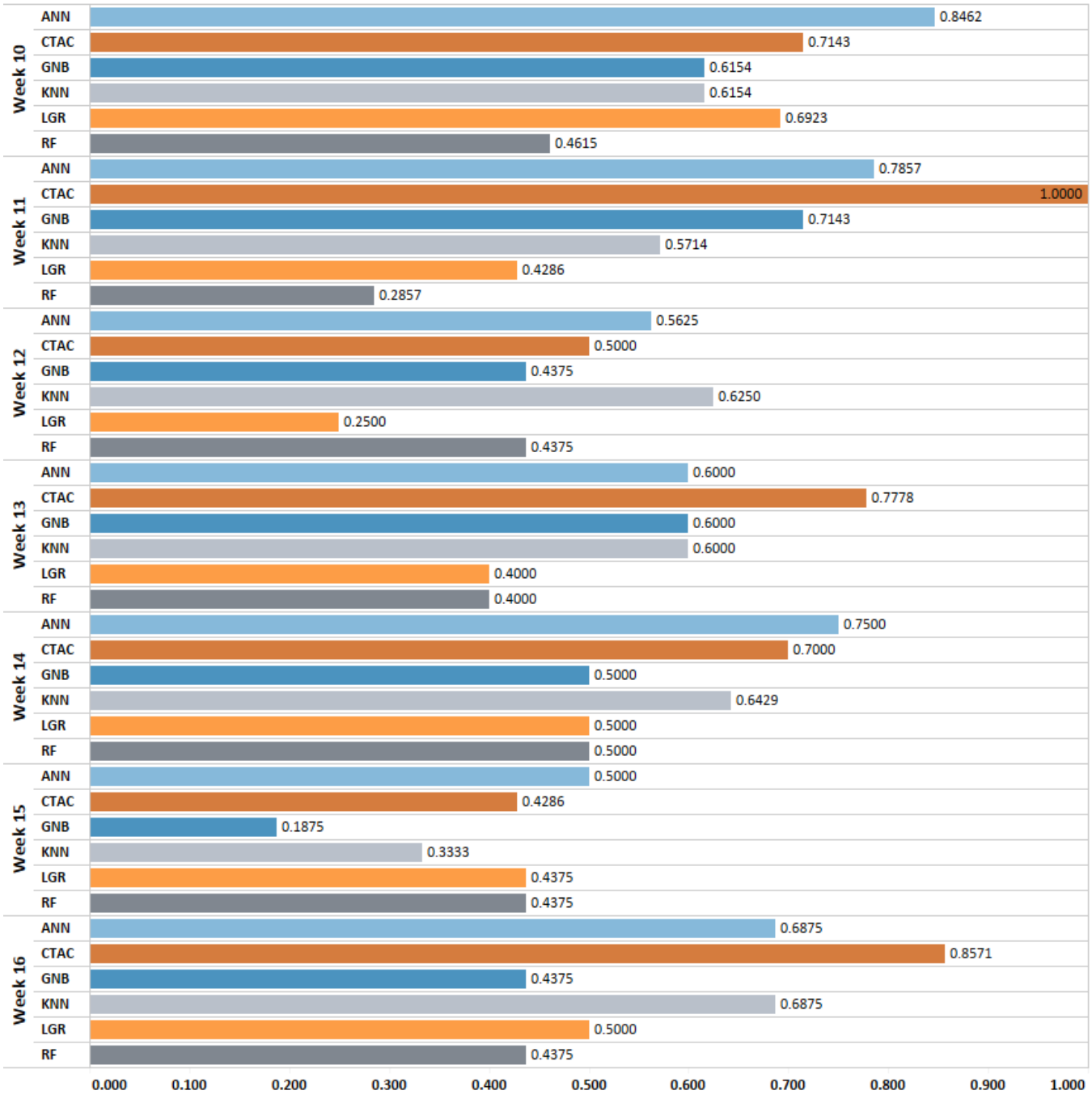

opp_def_rushy_avg
tm_def_score_avg
tm_def_passy_avg
tm_def_rushy_avg
opp_def_passy_avg
opp_off_rushy_avg
opp_off_score_avg
opp_def_totyd_avg
tm_pass_att_avg
tm_off_rushy_avg
opp_def_score_avg
opp_pass_att_avg
opp_def_1std_avg
opp_off_totyd_avg
tm_cmp_avg
tm_rush_att_avg
opp_off_1stdwn_avg
tm_off_score_avg
tm_off_passy_avg
opp_sacked_avg
opp_cmp_avg
tm_def_totyd_avg
tm_sacked_avg
opp_rush_att_avg
tm_def_1std_avg
opp_off_passy_avg
tm_off_1stdwn_avg
tm_pass_int_avg
opp_pass_int_avg
tm_off_totyd_avg
opp_pass_td_avg
tm_pass_td_avg
tm_def_to_avg
tm_off_to_avg
over/under
opp_def_to_avg
tm_rush_td_avg
spread
opp_off_to_avg
opp_rush_td_avg
team_code
opp_code
opp_winning%
tm_winning%
week
game
season
opp_btwn_gms
tm_under_streak
tm_btwn_gms
opp_over_streak
opp_under_streak
opp_losing_streak
tm_winning_streak
tm_over_streak
opp_winning_streak
tm_cover_streak
opp_not_cover_streak
tm_losing_streak
tm_not_cover_streak
opp_cover_streak
time
day
h_or_a
tm_off_bye
opp_off_bye
0.000
0.005
0.010
0.015
0.020
Importance

Appendix E

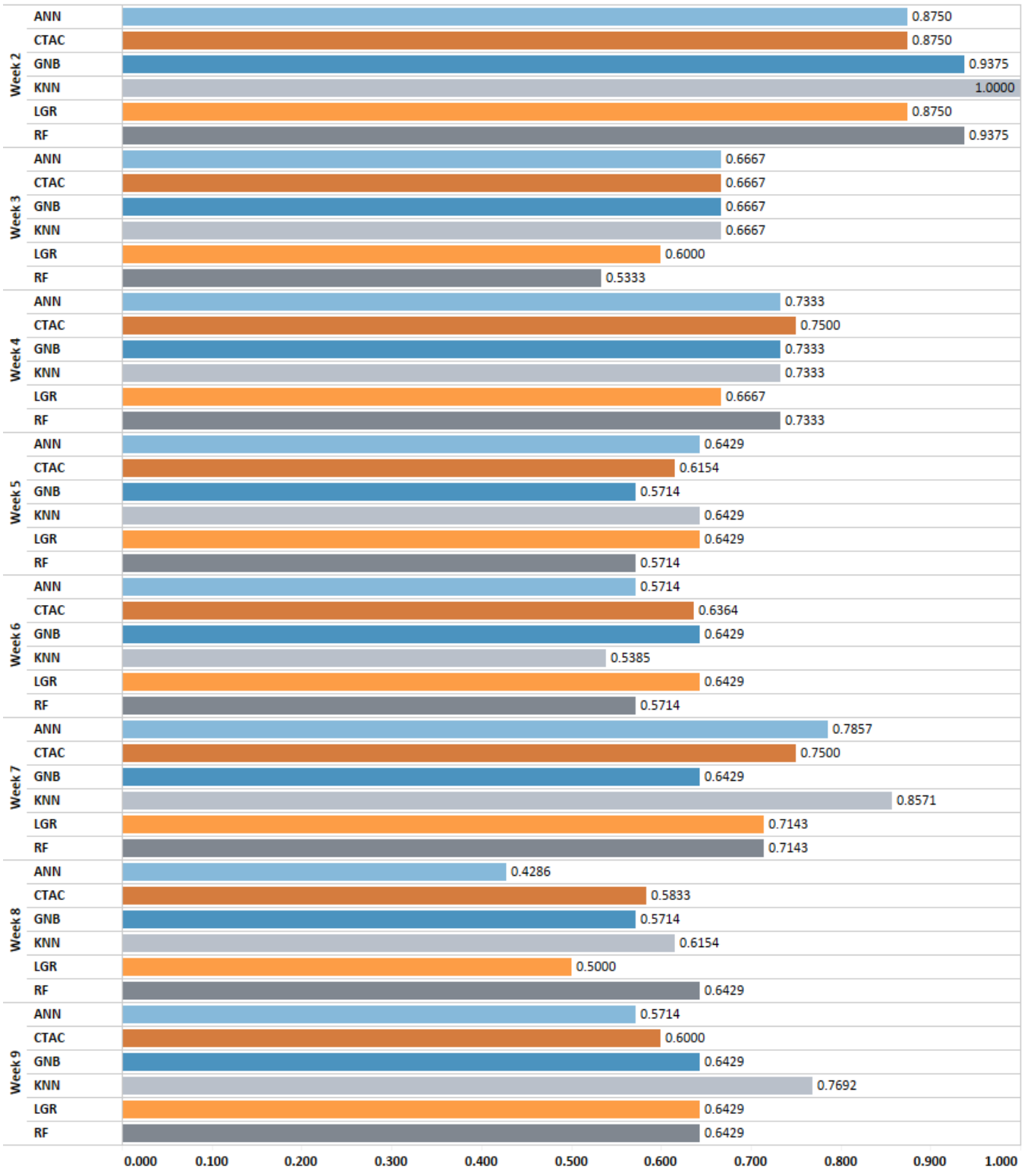
Over/Under Results 2020 Season



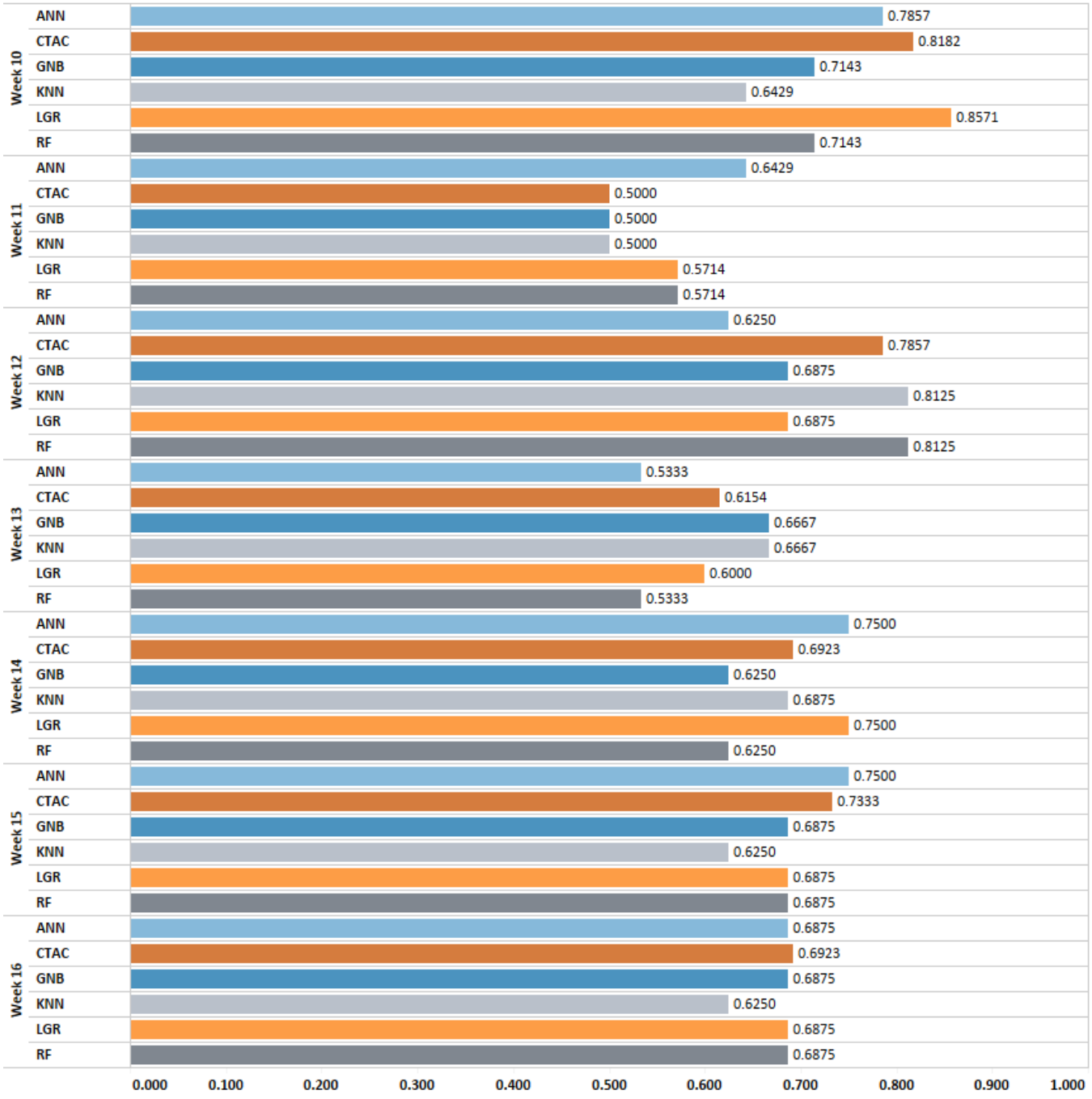
Over/Under Results 2020 Season



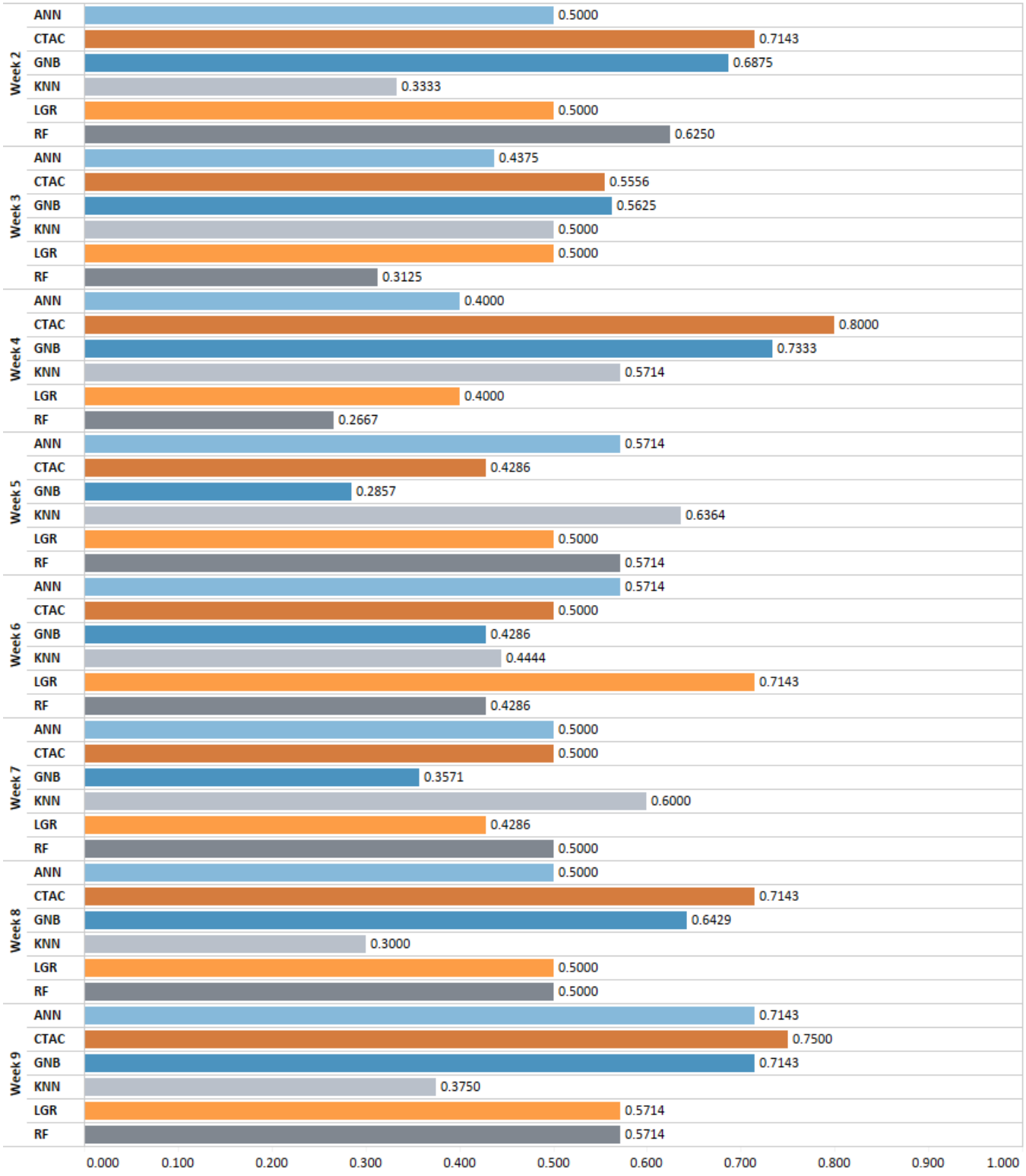
Winner Prediction Results 2020 Season



Winner Prediction Results 2020 Season



Spread Results Season 2020



Spread Results Season 2020