Predicting This Year's NFL MVP: Insights from Historical Data*

Jacob Gilbert

2024-11-27

"This paper aims to predict the NFL MVP for the 2024 season by analyzing historical quarterback performance data. Using an XGBoost classifier, we identify the key performance metrics, such as passing touchdowns, win percentage, and efficiency, that significantly influence a quarterback's likelihood of receiving MVP votes. The model highlights that both individual performance and team success are critical to a player's MVP candidacy, with win percentage being the most influential factor. Understanding these key characteristics helps fans, analysts, and stakeholders gain insights into what drives MVP-level success, enriching the broader conversation around evaluating elite player performance."

Table of contents

*code available at https://github.com/JfpGilbert0/NFL

Introduction

The NFL MVP dates back to 1957, and since its introduction by the associated press, it has been the league's most coveted individual award. It recognizes a player's individual success but also their impact on team success, and sometimes the sport as a whole, in a given season. The quarterback position has dominated the award's winners, with the biggest names in the sport today such as Tom Brady, Aaron Rogers, Patrick Mahomes and Lamar Jackson all having won the award in the past decade. In the process setting records in major statistical categories such as passing yards and touchdowns. The quarterback positions focal point in the offence and statistical contributions to teams lead to its frequency at the top of mvp voting. This paper thus focuses on what makes a qb an mvp recipient, and which variables must a player excel at in order to qualify, according to historical data.

Just winning is often not enough to show who wins an mvp, especially for quarterbacks. Players must show excellence in many key aspects of their position. This paper analyzes the determinants from statistical measures that are key to football's most important position , including passing yards, touchdowns, interceptions, completion percentage. The model evaluates whether a player is in the top 15% for a given statistical category in that season impacting their mvp candidacy. Helping us determine how excelling at which specific metrics are most strongly linked to MVP success.

The 2024 season is more than half way through and as such the "race" for the mvp is heating up. Player performances this season have elevated some above the rest, as such the question of who has been the most valuable is becoming more and more relevant to all stakeholders of the league. By identifying mvp characteristics, the ongoing seasons' most valuable players can be predicted, which holds great value for most NFL stakeholders.The model in this paper predicts 2 likely candidates for mvp in this current season, as well as understanding why some of the league's best players are not in the conversation.

The paper is structured as follows: Section introduces the NFL MVP award, focusing on its history and the significance of individual performance metrics for quarterbacks. Section section provides details on the data used, including its compilation, features, and key variables, as well as measurement practices and limitations. Section 3 presents the XGBoost model used to predict MVP candidacy, including feature selection, model assumptions, and implementation details. Section 4 contains the results from the model, discussing feature importance, model accuracy, and the predicted MVP votes for the 2024 season. Finally, Section 5 discusses the implications of the results, providing a deeper analysis of what makes a quarterback an MVP candidate, and concludes with the model's predicted outcome for this year's MVP race.

Data

The dataset used for this analysis was compiled using the NFLverse (Contributors (2024)) package in R (Team (2024)) and supplemented by publicly available sources (League (2024)), covering player statistics since 2006. Data cleaning and figure generation is done using Python (P. S. Foundation (2024)) specifically utilizing the following packages: pandas (al. (2024)) for data manipulation and analysis, pyarrow (A. S. Foundation (2024)) for working with Apache Arrow and parquet file formats, and numpy (Community (2024)) for numerical operations and calculations.

The data spans multiple aspects of player performance, including passing, rushing, and receiving metrics, showing player performance in each game they participated in each season. The data used includes records from both regular-season and playoff games, providing insight into players performances in every game they participated over the available timeframe.

Key variables in the dataset include passing and rushing touchdowns and yards, passing attempts and completions, as well as negatively impacting plays such as fumbles and interceptions. Further details are included in the nature of plays made in a game such as passing and rushing first downs. More advanced statistics are included in the dataset such as Passing Air Conversion Ratio (PACR) which is a ratio of passing yards/ passing air yards, reflecting a passers ability to find receivers in space and thus resulting in more yards after the catch. Each datapoint in the dataset provides insights into a players performance in a game.

The MVP is a season award and thus is focused on the regular season performance, as such we restrict our data to data in these games. Focusing on quarterbacks allows for direct comparison so only those players' statistical performances are kept for analysis. Season stats are aggregated from the game data provided by nflverse(Contributors (2024)). Since 1984 no player has won the mvp award while playing less than 14 games in a season. The last being "insert name here", thus analysis of players that did not play more than thirteen games is deemed unnecessary, especially since the 17 game season was introduced in "insert year" players who play less games become more irrelevant in our dataset.

Team success is gathered from an alternate dataset provided by nflverse (Contributors (2024)) that provides game outcomes throughout each season. From this each team's win record with each QB is calculated for a given season. This is used to provide insight into how much a team's success relates to their quarterback's individual accolades. The NFL (League (2024)) provides nfl mvp voting records since the awards inception. 32 votes decide the mvp with 16 resulting in winning the award. Receiving votes signifies that a player has performed at an mvp level in a given season,and thus the analysis dataset will include information on who won in a season but also the vote count for all players to allow for comprehensive analysis.

Measurement

The data collection process for NFL statistics is highly standardized and conducted by the NFL's official statisticians during each game. The NFL uses an advanced tracking system known as Next Gen Stats, which captures player movements using sensors embedded in players' shoulder pads and around the stadium. This system provides detailed information about player performance, such as speed, distance traveled, and player positioning, in addition to traditional stats like passing yards, rushing yards, touchdowns, and completions. Player statistics are recorded in real time during games by official NFL statisticians, who work to ensure accuracy through a rigorous validation process. These statistics are then cross-referenced and reviewed to correct any discrepancies. The data is made available publicly through the NFL's official platforms, which ensures consistency across sources used for analysis. For this paper, variables such as passing yards, touchdowns, and interceptions were chosen because they are key indicators of quarterback performance and directly affect a player's likelihood of being considered for the MVP award. The derived variables, such as touchdown-to-interception ratio and win percentage, reflect more nuanced aspects of a quarterback's contribution to their team's success. The data thus represents a combination of both raw performance metrics and situational metrics, providing a comprehensive view of a player's overall value to their team. From table 2 we can see how the key statistics we are looking at vary across the seasons. As the game changes and years differ the meaning that values of these statistics hold change. For example in 2002 the 28 was the most passing touchdowns achieved, whereas this would be below the average in 2020. In the table 1, this type of variablility is shown by the large standard deviations for the max values. Thus comparing across the seasons stats directly would incorrectly represent the importance of passing touchdowns. To avoid this issue we will measure how a player performed in a given statistic by if they fall in the 90th percentile for that statistic. This allows us to compare across seasons where the benchmark for greatness in a statistic may differ. This way we can identify which attributes are more attributed with MVP consideration when players excel at them.

Summary Statistics for Key Variables{#tbl:summary_stats}

Statistic	85th Pct Mean	85th Pct std	Max Value Mean	Max Value std
passing_yards	4531.35	246.401	4988.28	315.902
passing_tds	32.5917	3.84315	40.7222	6.80662
passing_first_downs	222.583	11.1717	251.222	18.5427
interceptions	16.4444	2.52061	20.7778	4.46629
pacr	1.06952	0.0596346	1.21848	0.116848
sacks	42.4333	4.06868	52.2222	6.28308
sack_yards	281.675	31.9917	351.056	51.2669
total_fumbles_lost	5.04444	0.638216	7.05556	1.10997
rushing_yards	360.597	149.114	668.444	285.88
rushing_tds	4.07778	1.40401	7.88889	3.67646
rushing_first_downs	25.9833	10.5236	44.8889	15.9259
win_pct	0.749451	0.0396244	0.867758	0.058604
td:int	3.51242	0.872895	6.3491	2.48576

Table 1: Summary of 85th percentile lower and upper bounds from 2006 - 2023, mean, standard deviation, for key metrics used in predicting MVP candidates. This provides insight into how the levels of these key metrics cary from season to season.

For the context of this paper, MVP candidacy by predicting whether a player will receive any MVP votes. This approach allows us to capture those players who were significant enough in their performance to be recognized, even if they did not ultimately win the award. By modeling MVP candidates through this lens, we gain a nuanced understanding of what it takes to be considered among the league's elite players. This gives a greater sample size of mvp quality players for us to extract key characteristics from. Moreover, looking at mvp votes allow us to still gain insights from years that don't have a QB mvp winner, still informing us of the makeup of the best quarterback in the league.

For the current season, statistics are presented in a per-game format, if they are not already aggregated stats, to avoid bias related to differences in the number of games played by different teams due to the season schedule. This ensures that players on teams with bye weeks or delayed games are not unfairly disadvantaged in the analysis, allowing for a fair comparison of performance across all players.

Model

In this section, we present the XGBoost Classifier model used to predict which NFL quarterbacks are likely to receive MVP votes based on their performance metrics for a given season. XGBoost (Extreme Gradient Boosting) is an advanced, efficient implementation of gradient boosting that is well-suited for structured/tabular data. The choice of XGBoost reflects a need to capture complex, non-linear relationships among the input features while maintaining high predictive accuracy.

XGBoost Model

The XGBoost Classifier can be represented mathematically as an ensemble of decision trees:

$$F(x) = \sum_{k=1}^{K} f_k(x), \quad f_k \in \mathcal{F}$$

Where:

- $F(x)$: The final model that is used to predict the target variable, aggregating all the decision trees.
- f_k : Represents an individual tree in the ensemble.
- $\mathcal{F}$: The space of regression trees.

The XGBoost model uses a gradient boosting approach to iteratively minimize the loss function:

$$L(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Where:

- $L(\phi)$: The total loss function, composed of two parts.
- $l(y_i, \hat{y}_i)$: The loss function measuring the difference between the predicted value ($\hat{y}_i$) and the actual value (y_i). In our case, the "logloss" metric is used to measure this for classification.
- $\Omega(f_k)$: The regularization term, controlling the complexity of each individual tree to prevent overfitting.

The output is a probability, which is then thresholded to predict whether a quarterback will receive MVP votes (1) or not (0).

- **`eval_metric='logloss'`**: Log loss is used as the evaluation metric. It measures the accuracy of probability predictions in a logistic regression context.

Feature Selection

The predictor variables (X) used in the model are chosen based on their importance in determining quarterback performance and team success. They include:

1. **`win_pct`**: The win percentage of the team, representing team success and thus, the player's contribution to winning.
2. **`td:int`**: Ratio of touchdowns to interceptions, indicating efficiency.
3. **`passing_tds_85pctile`, `rushing_tds_85pctile`**: Indicators of whether the player is in the top 15% for touchdowns—both passing and rushing—during a season.
4. **`passing_yards_85pctile`**: High passing yardage, representing offensive production.
5. **`passing_first_downs_85pctile`, `rushing_first_downs_85pctile`**: First downs gained via passing and rushing, important for keeping drives alive.
6. **`pacr_85pctile`**: Passing Air Conversion Ratio, representing a quarterback's efficiency in gaining yards after the catch.
7. **`total_fumbles_lost_85pctile`, `sacks_85pctile`, `sack_yards_85pctile`**: Measures for minimizing negative plays such as fumbles and sacks.
8. **`interceptions_85pctile`**: Indicates a player's interception count relative to others.

The selection of features is based on their ability to capture different aspects of both team and individual performance that contribute to MVP candidacy.

Rationale for Using XGBoost

XGBoost was chosen for its ability to: - **Capture Non-linearity**: Unlike logistic regression, XGBoost is capable of capturing non-linear relationships between variables, which is crucial when dealing with complex datasets like those representing quarterback performance. - **Feature Importance**: XGBoost allows us to identify which features are most influential, providing insight into what makes a quarterback an MVP candidate. - **Regularization**: It includes parameters for regularization to control overfitting, which is particularly useful given the complexity of the dataset.

Assumptions and Limitations

- **Assumption of Data Quality**: The model assumes that the features are representative of player performance and are free from significant biases. Poor data quality could impact the model's performance. As
- **Complexity and Interpretability**: XGBoost is a more complex model compared to simpler alternatives such as logistic regression. While this complexity helps in accuracy, it also makes interpretation more challenging.

Implementation and Validation

The model was implemented using Python (P. S. Foundation (2024)) with the `xgboost` library (Chen and Guestrin (2024)). To validate the model, the dataset was split into training (90%) and testing (10%) sets using sklearn packages (Pedregosa et al. (2024)), ensuring that the model's performance was evaluated on unseen data. The following steps were taken:

1. **Training and Testing Split**: The data was split into 90% for training and 10% for testing, allowing us to assess out-of-sample predictive performance.
2. **Standardization**: Predictor variables were standardized using `StandardScaler` to ensure they contribute equally to the model.
3. **Metrics for Validation**:

    - **Accuracy Score**: The overall accuracy of the model was calculated for the test set.
    - **Confusion Matrix**: This matrix showed the true positive, false positive, true negative, and false negative counts, providing insight into where the model performed well or struggled.
    - **Classification Report**: The report included metrics like precision, recall, and F1-score, giving a holistic view of the model's classification performance. ## Alternative Models Considered An alternative model considered was **logistic regression**, which provides more interpretability. However, logistic regression does not adequately capture non-linear relationships, which are likely important in this context. We ultimately chose XGBoost due to its superior performance in capturing complex interactions among features and improving accuracy.

Conclusion

The XGBoost Classifier was used to predict which quarterbacks would receive MVP votes based on season statistics. The model's choice was guided by its ability to handle non-linear relationships, provide feature importance, and regularize effectively to prevent overfitting. The resulting model achieved high accuracy and provided insightful conclusions about what makes a quarterback an MVP candidate. While XGBoost offers increased accuracy, it also introduces complexity that limits interpretability, a trade-off that was carefully considered in our analysis.

Results

Regression Results

Feature importances for the XGBoost model

Variable	Importance
win_pct	0.322893
td:int	0.0552517
passing_tds_85pctile	0.146538
rushing_tds_85pctile	0.0596878
passing_yards_85pctile	0.0257456
passing_first_downs_85pctile	0.0804869
pacr_85pctile	0.0830801
rushing_yards_85pctile	0.0168114
rushing_first_downs_85pctile	0.00959435
total_fumbles_lost_85pctile	0.0237403
sacks_85pctile	0.0470513
sack_yards_85pctile	0.0864964
interceptions_85pctile	0.0426228

Table 2: Feature importance values for the XGBoost model predicting binary value for MVP votes. These metrics highlight the significance of each variable, with 'win percentage' playing a dominant role in determining a quarterback's MVP candidacy.

The regression results from the XGBoost model reveal that win_pct was the most influential predictor, with an importance score of 0.3229. This underscores that team success, as reflected by a player's win percentage, is the most critical determinant of MVP candidacy. The model also highlighted passing_tds_90pctile (importance = 0.1465) and td:int (importance = 0.0553) as significant features, indicating that players who excel in passing touchdowns and maintain a strong touchdown-to-interception ratio are more likely to receive MVP votes. Other features such as pacr_90pctile (importance = 0.0831) and passing_first_downs_90pctile (importance = 0.0805) contributed moderately to the model. Conversely, variables like rushing_yards_90pctile (importance = 0.0168) and rushing_first_downs_90pctile (importance = 0.0096) had lower importance, suggesting that while these metrics influence MVP candidacy, they are less crucial compared to standout offensive achievements in the passing game. Key negative play indicators like sacks and interceptions are lower than others implying that strong positive play outshines the impact of reducing negative plays. Despite this being in the 90th percentile of sack yards does have a decent impact on mvp consideration.

Model Evaluation

Accuracy Score: 0.97

Confusion Matrix:

	Predicted 0	Predicted 1
Actual 0	30	0
Actual 1	1	3

Classification Report:

Class	Precision	Recall	F1-Score	Support
0	0.97	1.00	0.98	30
1	1.00	0.75	0.86	4
Accuracy			**0.97**	34
Macro Avg	0.98	0.88	0.92	34
Weighted Avg	0.97	0.97	0.97	34

The model achieved an accuracy score of 0.97, demonstrating a strong predictive capability. The confusion matrix shows that out of 34 players, 30 were correctly identified as not receiving MVP votes, while 3 were accurately identified as receiving votes. There was 1 false negative, indicating a player who received MVP votes but was not identified by the model. The high precision of 1.00 for players who received votes means that all positive predictions made by the model were correct, although the recall of 0.75 for this class suggests that not all players who received votes were successfully predicted

MVP Votes Predictions for 2024 Season

Prediction and probability for 2024 MVP vote recipients

Player Display Name	MVP Votes Prediction	MVP Votes Probability
Aaron Rodgers	0	0.000336
Matthew Stafford	0	0.007637
Kirk Cousins	0	0.000169
Geno Smith	0	0.001000
Jared Goff	1	0.603842
Patrick Mahomes	0	0.032164
Lamar Jackson	1	0.604000
Baker Mayfield	0	0.014920
Josh Allen	1	0.976419
Sam Darnold	0	0.084216
Kyler Murray	0	0.002409
Justin Herbert	0	0.007366
Jalen Hurts	0	0.036282
Joe Burrow	0	0.386225
C.J. Stroud	0	0.000133
Bo Nix	0	0.000997
Jayden Daniels	0	0.001547
Caleb Williams	0	0.001353

Table 3: Predictions for MVP votes based on the model analysis for the 2024 season. binary percentile variables are calculated using per game statistics. This table presents which quarterbacks are predicted to receive votes and their respective probabilities. 3 quarterbacks are expected to recieve at least some votes for MVP.

Using the XG boost model we can extract who is predicted this year, based on their current rankings of key statistics per game, which quarterbacks will receive mvp votes. This list of predictions only includes players who have at least 10 games played, to display only those deemed to be contenders, however 36 quarterbacks percentages were predicted in total. This model predicts that based on their performance in key statistics this year only 3 players are expected to receive votes, Jared Goff, Lamar Jackson and Josh allen. Josh Allen boasts the highest probability of 97.6% chance of receiving votes. Both Jackson and Goff are sitting with approximately 60.4% chance of receiving votes. Only one other player has above 10% chance of receiving votes according to the model, that is Joe Burrow with a predicted 38.6% chance. In the Discussion section we will look into the stats to understand why this is. But compared to the rest of the quarterbacks, he is much closer to the mvp than he is to any of them.

Discussion

Model discussion

The regression results highlight several interesting implications about how MVP candidacy is influenced by efficiency versus production, as well as the importance of team success. The high feature importance of win_pct indicates that team success plays a pivotal role in determining MVP votes. This means that even if a player has a statistical season below exceptional, their chances of receiving MVP votes can be greatly inflated by the team's success. In some context this could lead to unfair favouritism to great teams rather than great players. The inverse is also true, that players playing at an mvp level could be passed over due to team success which is not always in control of the individual player. This bias towards team success suggests that the MVP is not purely a measure of individual excellence but also considers a player's contribution to the team's overall achievements.

The model shows that being in the 85th percentile of touchdowns has the second highest importance towards mvp vote. As the primary way quarterbacks impact points scored this is expected, especially as point scored impact winning which the model has identified as a key factor towards impact mvp votes. Being one of the best Rushing scorers is less impactful as an indicator of candidacy. One reason this may be is that not all quarterbacks are mobile and thus do not rely on rushing touchdowns to score. However the model suggests that scoring the most rushing touchdowns will indeed aid an mvp campaign.

Some variables are more associated with high impact plays and others relate to the consistency. Excelling at passing first downs and pacr, which is a measure of large plays that occur after the catch, are more representative of a player that makes important high impact plays. Being at the top of the league in these categories are significantly positive towards mvp consideration. We see that although the number of sacks is deemed relatively unimportant, it is considered important to be one of the best at keeping sack yards low. Highlighting again that explosive plays are more significant in the mvp conversation, even if that means reducing negative ones. In contrast with this, keeping fumbles lost to a minimum is not considered as important as other factors, this could be due to the fact that these often result from poor team play. However other variables that could result from the same still hold great weight in the mvp conversation. Another possibility is variables like sacks and fumbles occur more when players are in command of the football for longer, and thus the best players are not often in the top percentile for these statistics, in figure 5 we will see some stats of mvp contenders that offers some insight.

Explosive plays importance is supported by the model as it is not important to have the fewest interceptions in the league but instead it is more important to have one of the better td to interception ratios. And as highlighted previously, even more key to score more touchdowns. The small factor that interceptions play is interesting however and could be slightly skewed by players who play less games, despite the methodology mentioned in the data section to diminish this. Further interceptions here reflect the season as a whole, it is possible that many interceptions occur in a few games, which is not a characteristic of the best players in the league. The model would not be able to interpret the distribution of interceptions, the effect of this would be to underrepresented the importance of interceptions.

This is a problem with the model that does extend across the variables, and that is the granularity of the data is the season. Stand-out performances both good and bad often define an mvp season. For example it would be hard to vote for a player that threw 8 interceptions in a game, however they would still be in the upper echelon of the league if they threw no more that season. WHat quantifies a standout performance is however fairly subjective, and as the mvp is a season award, using statistics from the whole season is still very functional. Other methods of data collection such as surveys may offer more insight into what variables outside of the data may influence mvp votes. Further there may be some bias present by using mvp votes as our variable of interest, this could result from players with 1 or few votes being included with those that won unanimous mvp's. These players are clearly very different but treated as the same here. In future iterations of this model it is recommended to weight by mvp votes.

In summary, the regression analysis underscores that while individual performance is crucial for MVP candidacy, team success remains a major determinant. The model indicates that standout offensive achievements, particularly in passing metrics, are highly valued. Being the best in the league in other facets does positively impact mvp odds however complementing winning with offensive production is more valued than limiting negative production. The model also does have some shortfalls as it regresses a season worth of visual spectacles to one number, people are often emotional about sports and thus there is more understanding to be had as to what makes and MVP outside of the number.

2024 prediction

Inputting this years current standings in key statistics offers a prediction from the model. Looking at the 4 quarterbacks that are leading the MVP conversation in greater detail gives an informative view at why they are contenders. The model analyses being in the 85th percentile in most statistics, in the 2024 season among qualifying quarterbacks this means ranking in the top 5 in a statistic.

2024 MVP candidates per game statistics and ranking {#tbl:summary_statistics}

Player	Win %	MVP prop	TD: INT	Pass TD	Rush TD	Pass Yrds	Rush Yrds	Pass 1sts	Rush 1sts	PAC R	Fumbles Lost	Sacks	Sack Yrds	Ints
Allen	0.82 (3)	97.6 (1)	3.6 (5)	1.6 (8)	0.5 (3)	231.2 (15)	28.7 (9)	10.6 (17)	2.7 (4)	1.08 (12)	0.2 (17)	1.2 (4)	5.5 (2)	0.5 (10)
Goff	0.91 (1)	60.4 (3)	2.22 (14)	1.8 (7)	0.0 (25)	251.0 (9)	3.5 (33)	12.2 (6)	0.5 (28)	1.59 (1)	0.0 (1)	1.9 (12)	14.8 (19)	0.8 (21)
Jackson	0.67 (6)	60.4 (2)	9.0 (2)	2.2 (2)	0.2 (11)	254.4 (7)	49.9 (1)	11.9 (9)	2.5 (7)	1.34 (3)	0.3 (27)	1.3 (6)	8.0 (6)	0.2 (3)
Burrow	0.36 (24)	38.6 (4)	6.75 (3)	2.5 (1)	0.1 (15)	275.3 (2)	13.7 (21)	14.1 (1)	1.2 (20)	1.04 (14)	0.1 (10)	2.4 (20)	12.5 (13)	0.4 (6)
Mahomes	0.91 (1)	3.2 (8)	1.64 (24)	1.6 (8)	0.1 (15)	243.0 (13)	20.6 (15)	13.5 (3)	1.5 (14)	1.24 (5)	0.0 (1)	2.5 (23)	16.5 (24)	1.0 (30)

Table 4: Comparison of key metrics for quarterbacks leading the MVP race. The table illustrates differences in these key variables among those likely to win mvp, from this it adds context as to why some players have higher odds than others.

The model predictions for the 2024 NFL season provide interesting insights into how different quarterbacks are expected to fare in the MVP race. Based on the current season's statistics, the model identifies Jared Goff, Lamar Jackson, and Josh Allen as the top three candidates likely to receive MVP votes. Among these, Josh Allen stands out with a 97.6% probability of receiving votes, which is significantly higher than his peers. Lamar Jackson and Jared Goff also show strong probabilities at around 60.4%, indicating that their performances this season have positioned them as key contenders.

An initial shock is that the fundamental variable from the model, win percentage, Allen and the Buffalo Bills are not leading the league. In fact he is well behind Goffs, another mvp candidate. This is suppressing considering the almost certain probability that he will receive mvp votes. Looking at ALlens other stats we see why his odds are so high, he ranks in the top 5 for 5 other major statistics. Ranking as 2nd and 4th in rushing touchdowns and td:int ratio respectively, two significant variables. Interestingly he is just outside of the top 5 in a very key statistic, Passing touchdowns. This essentially should not be captured by the model implying that his MVP odds may be even higher.

The prediction results also reveal several well-known quarterbacks, such as Patrick Mahomes, Joe Burrow, with lower-than-expected probabilities of receiving votes. For instance, Joe Burrow's probability stands at 38.6%, suggesting that, despite his solid individual performance, other factors like team success look to be affecting his candidacy. This highlights how despite Burrow ranking highly in almost all the most important statistics, the poor team success is heavily holding back his candidacy. Patrick mahomes' mvp campaign looks to be hindered from the opposite affliction. The Chiefs match the Lions and Jared Goff with the highest win percentage, 0.909. However the model has Mahomes at less than 4% chance of receiving mvp votes. COmparing his stats to goff we see a very high interception rate, while not compensating with passing touchdowns. Leading to him being outside to the top of the league in key factors such as TD:INT ratio, Touchdown passes and Interceptions. Emphasising that a mvp campaign cannot rely solely on team success.

Lamar Jackson is an interesting case, as his statistics per game are very similar to our MVP favourite Josh allen. In some statistics even surpassing him, especially in offensive production and efficiency based statistics that were highlighted as key in the model. He sits in the top 5 in a majority of the key variables, However his win percentage is the second lowest in this list of contenders.

In conclusion Our model predicts a race likely to come down to three major candidates for MVP. Josh allen however has so far excelled at a majority of key statistical categories while maintaining a strong winning percentage. This balance of personal and team success puts him at the top of the MVP conversation, for now. Should this success continue in both key facets, The model expects Josh Allen to win his first MVP award.

Appendix

In this appendix, we explore how an idealized survey or observational data collection methodology could enhance the insights provided by the analysis in this paper. The dataset used to predict NFL MVP candidacy relied heavily on objective statistics collected from game performances, such as passing yards, win percentage, and touchdown counts. However, one area of interest not captured by these metrics is the public and media perception of a player's impact, which can significantly influence MVP voting outcomes. To address this gap, an observational survey or polling approach could provide supplementary insights into MVP candidacy.

Survey Design and Sampling

A well-structured survey could be conducted targeting NFL fans, analysts, and sports journalists to understand their perceptions of player performance, team success, and other less quantifiable traits that contribute to MVP consideration. This survey would aim to answer questions about which qualities fans and analysts deem most valuable in an MVP candidate, beyond the statistics. For instance, questions could explore the perceived impact of players on team morale, clutch performance in high-stakes games, and leadership—factors that might not be adequately represented in objective datasets (Kuper and Szymanski 2014). This approach is similar to survey-based studies conducted in the field of behavioral economics where public perception influences economic outcomes (Thaler and Sunstein 2008).

To ensure the survey's accuracy, a stratified random sampling approach would be utilized. The population of interest would include NFL fans, analysts, and journalists, stratified based on characteristics like age, geographical location, and engagement level with the NFL (e.g., casual viewers vs. experts). This stratification would allow the survey to adequately capture diverse perspectives and biases, which is crucial given the varied ways people perceive player performance. Ideally, the survey sample size would be large enough to reduce sampling error and improve confidence in the results, similar to techniques described by Groves et al. (Groves et al. 2009) for improving survey reliability.

Incorporating Observational Data

Beyond surveys, observational data could also be gathered to assess the narrative aspects that influence MVP voting. Observational data might include media coverage volume, social media sentiment, and the frequency of players being mentioned in highlight reels or news stories. Platforms such as Twitter provide rich sources of data that could be harvested using techniques like sentiment analysis to gauge public opinion (Pak and Paroubek 2010). This approach would also align with existing literature on the impact of media coverage on award voting, such as in the field of film awards, where media visibility can significantly influence outcomes (Simonton 2004).

One possible enhancement to the study would be a simulation of MVP voting based on these qualitative factors to compare with the results of the current model. Such a simulation could use a weighted voting system that combines the statistical metrics used in this paper with survey and observational data. Each weight could be calibrated based on historical voting outcomes and validated using techniques such as cross-validation on past seasons' data. This would provide an in-depth exploration of the factors that influence MVP candidacy and enhance our understanding of both objective performance and subjective perceptions in the MVP selection process.

Potential Biases and Limitations

It is important to acknowledge potential biases inherent in survey and observational methodologies. For instance, fan surveys may be subject to response bias or over-represent fans of specific teams, particularly teams with larger fanbases. Similarly, media coverage tends to favor popular players or those in larger markets, which might skew the observational data (Entman 2007). Addressing these biases would require careful design considerations, such as weighting responses based on demographic characteristics or employing machine learning techniques to adjust for media biases.

Overall, supplementing the current statistical analysis with a survey and observational data approach would provide a more holistic view of what drives MVP voting decisions. Such a mixed-methods approach would ensure a more nuanced understanding of player value, accounting for both measurable performance metrics and the subjective perceptions held by the public and experts.

References

al., Wes McKinney et. 2024. "Pandas: Powerful Data Structures for Data Analysis in Python." https://pandas.pydata.org/.

Chen, T., and C. Guestrin. 2024. "XGBoost: Scalable and Flexible Gradient Boosting." https://xgboost.readthedocs.io/en/stable/.

Community, NumPy. 2024. "NumPy: Fundamental Package for Numerical Computation." https://numpy.org/.

Contributors, NFLverse. 2024. "NFLverse: Comprehensive Data on NFL Stats." https://nflverse.com/.

Entman, Robert M. 2007. "Framing Bias: Media in the Distribution of Power." *Journal of Communication* 57 (1): 163–73.

Foundation, Apache Software. 2024. "Apache Arrow: PyArrow Documentation." https://arrow.apache.org/docs/python/.

Foundation, Python Software. 2024. "Python: A Programming Language for Data Science." https://www.python.org/.

Groves, Robert M., Floyd J. Fowler Jr, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*. John Wiley & Sons.

Kuper, Simon, and Stefan Szymanski. 2014. *Soccernomics: Why England Loses, Why Germany and Brazil Win, and Why the US, Japan, Australia—and Even Iraq—Are Destined to Become the Kings of the World's Most Popular Sport*. Nation Books.

League, National Football. 2024. "NFL Official Statistics and Data." https://www.nfl.com/stats/.

Pak, Alexander, and Patrick Paroubek. 2010. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, et al. 2024. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research*. https://scikit-learn.org/stable/.

Simonton, Dean Keith. 2004. "Film Awards as Indicators of Cinematic Creativity and Achievement: A Quantitative Comparison of the Oscars and Six Alternatives." *Creativity Research Journal* 16 (2-3): 163–72.

Team, RStudio. 2024. "RStudio: Integrated Development Environment for r." https://www.rstudio.com/.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.