Analysing the Success of NFL Offensive Plays

Toby Cheng, Emil Sebastian Jino, Nikil Chandrasekar, Jack Bellamy, Alex Bott, Xiaojun Zhang

***Abstract*—This report explores the offensive performance of NFL teams during the 2018-2023 seasons. Machine learning and statistical models are applied to the four main play types (running, passing, punting and field goals), aiming to identify the factors that contribute to the success of a play. The findings and models from this study can be implemented by players to plan more effective training, and by coaches to improve their strategic decision making in real game scenarios.**

I. Introduction

In the past 50 years, the average points scored by a National Football League (NFL) team has steadily increased from 19.5 per game in 1973 up to 21.9 per game in 2022 [1], as shown in Figure 1.

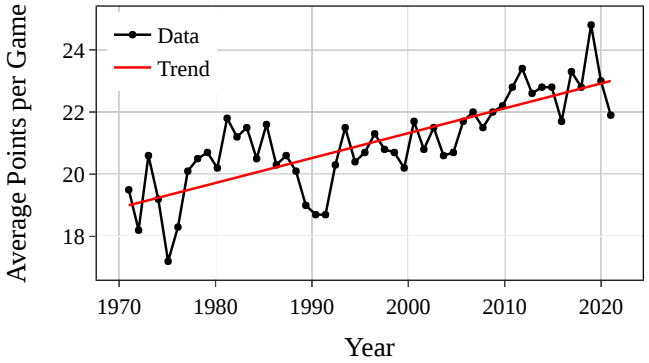

Fig. 1: *Data scraped from [1] showing the average points scored by the offense per game for each season, from 1973 to 2022.*

A likely factor contributing to this positive trend is a better understanding of strategies before and during a play. By analyzing the results of previous game scenarios, insights can be gained into the movements and key decisions players and coaches need to make in order to maximize their offensive success. A study conducted in 1988 by Carroll *et al.* [2] revolutionised the sporting world by stating teams should pass more, kick less and attempt more two-point conversions. While this study was controversial when first released, it is now considered the standard approach that modern teams take. This highlights the significant impact sport analytics can have on American Football, improving the quality of training and the effectiveness of strategical decisions.

Our project aim is to therefore contribute to the ongoing research into NFL offensive success, using a data driven approach in order to supplement the information available to players and coaches. On an individual level, we aim to form models that can be used in player training to evaluate decisions made and suggest improvements for future games. More generally, these models could be used to predict the outcome of plays, such that coaches are more informed on what the best play type should be for a given game scenario.

The NFL hosts an annual competition [3] where data analysts create new insights into American Football using real-time tracking data. In 2022, the winners of the competition created a GUI [4] to optimise punt returns. Using delaunay triangulation to find gaps created by the punting team, they were able to form an A^* algorithm [5] to find the optimal path to the endzone. With this information, it was possible to rank returners and advise coaches on best returners in a given situation. The dataset used in this competition [6] was used as the primary source of information for exploration performed in this report.

II. Data Preparation

Due to the dataset [6] being used for a high profile competition, the data is already mostly cleaned and set as a relational column store, meaning the majority of the variables are already formatted in an appropriate layout. To better understand the components in the dataset, two external notebooks [7] and [8] gave a brief description of each column key. Figure 2 shows the dimensional variables used to describe positional data, such as ball locations and player velocity directions.

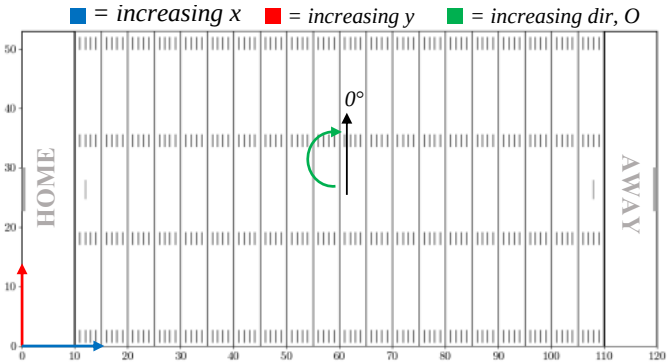

Fig. 2: *The dimensions and variables of the NFL field of play for the tracking data [6].*

Using the `python` library `Pandas`, the data can be read into a 2 dimensional dataframe through the function `read_csv()` and, using downcast, compressed by 36%. This allows much more efficient data manipulation, especially for the larger tracking datasets which inherently take up a lot of memory. For some data values we altered the datatyping such as 'height' to `int` (cm) from `str` (feet/inches) to allow for more efficient visualisations, training and testing.

`Pandas` data frames are ideal for concatenating and merging data as they can be combined using shared key variables. Through the use of left merging, groupby and aggregation functions, we we were able to reduce, compare and visualise the merged data.

III. Data Exploration

A. DataFrames

Following the data preparation phase, there were five different Dataframes to analyse. A brief description of each Dataframe and key observations are as follows:

- Games: Contains the teams playing in each game.
  - Dataframe shape: (764,7)
  - Key variables: 'gameId'
  - missing values: None
  - Key findings: 33 unique teams, 16 unique game times and 151 unique game dates.
- Plays: Contains special team plays that occurred.
  - Dataframe shape: (19979,25)
  - Key variables: 'gameId','playId'
  - missing values: 10 columns
  - Key findings: 4 unique pass results, 4 unique special team plays
- Players: Contains individual player statistics.
  - Dataframe shape: (2732,7)
  - Key variables: 'nflId'
  - missing values: 2 columns
  - Key findings: 26 unique player roles/positions
- Tracking2018-20: Contains positional data for each tenth of a second from all games.
  - Dataframe shape: (12777351,18), (12170933,18), (11821701,18)
  - Key variables: 'gameId','playId','nflId'
  - missing values: 5 columns - only 5% are NaN
  - Key findings: positions for each play, descriptions at each time stamp and which 'team' each player was on. This allowed us to simulate specific plays (can be seen via [9]).
- PFFScoutingData: contains play-level scouting information
  - Dataframe shape: (19979, 20)
  - Key variables: 'gameId','playId'
  - missing values: all except key variables.
  - Key findings: 11 unique kick types - majority deep-kicks - and 14 unique contact types.

B. Visualisation

Visualisation techniques are an essential part of exploratory data analysis (EDA) as they enable quick insights into the data and interpret general patterns, trends, and outliers that might not be apparent from raw data alone. Additionally, with such a large dataset, visualization provided a better understanding of the contents in each Dataframe.

Initially, simple visualisation methods were used to gather descriptive statistics, such as the distribution of play types, shown in Figure 3. This initally identified four play types: Extra Points, Field Goals, Kickoffs and Punts.

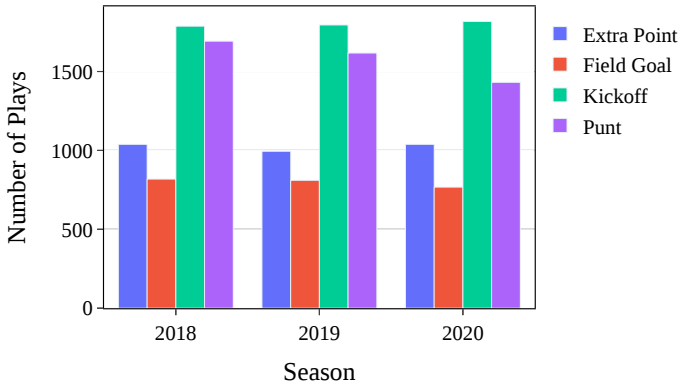

Fig. 3: *Bar plot showing the counts for different Special Team play types.*

A key observation from this data visualization was that the simulation of the play was not always consistent with the play type. This introduced the concept of fake plays, where the formation of the offense indicates one type of play, only to execute a different one to surprise the defense. Through further investigation of faking, two additional play types were identified; passing plays and running plays.

C. Web Scraping

As the 2022 competition dataset was focused around players specialised for kicking, the dataset only included passing and running plays when the offensive team was faking another option. This resulted in only 45 passing and 30 running plays to analyse, which is insufficient to draw reliable conclusions. Other NFL competitions have focused on these aspects of the game, thus the datasets from the 2023 competition [10] and 2020 competition [11] were used for further analysis.

IV. Kickoffs

Kickoffs initialize every half of the game, and additionally follow each scoring play. From the 2022 dataset, they were the most frequently occurring play type of of specialists kicking teams. Kickoffs are set plays meaning that they must be performed, and are thus excluded from the strategic decision making of coaches.

There exists strict rules for Kickoffs in terms of player and ball positioning. This makes this play difficult to optimize, as there are few parameters that can be explored when determining the outcome success rate.

The most notable variable is the kick type, as it is less volatile measure than other variables such as the player kicking the ball, whilst maintaining a significant impact in the variation of absolute yards gained. Figure 4 shows the frequency of various kick types, along with the respective average yardage gained.

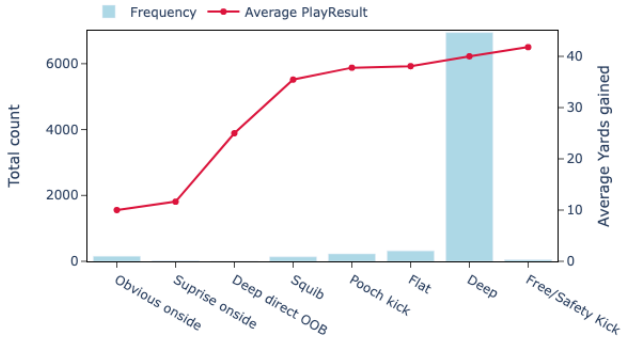

Fig. 4: *Plot showing Frequency and average yards gained for each Kickoff kicktype.*

As can be seen in Figure 4, a large majority of kicks are 'Deep' kicks and therefore most kicks will have a play result of around 40 yards excluding onside and 'out of bounds' kickoffs.

Because of this majority, it makes training a model impractical, as proven by attempts at fitting a logistic regression through one-hot encoding; the predicted result is heavily skewed by the multitude of 40 yard values.

V. Touchdown - Extra Points

Extra Point plays occur after a touchdown has been scored. The offensive team place the ball behind the 15 yard line and have a choice between kicking between the uprights or attempting another touchdown, for 1 or 2 points respectively. It is worth noting that in 2015, a rule was changed such that extra points were snapped from 15 yards (previously 2) and therefore the percentage conversion rate for 1pt dropped from an all time high of roughly 99%.

From the Dataset, the number of successes and failures for each point attempt was recorded, as shown in Figure 5.

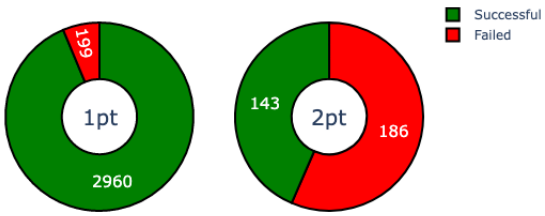

Fig. 5: *Conversion rates for 1 and 2 point extra point attempts.*

It can be seen that the success rate for both are approximately 94% and 43% respectively. The expected points are 0.86 for a 2pt attempt, and 0.94 for a 1pt attempt, indicating that the latter is the statistically better option. This observation does not account for the chance that the defense can block the kick and return a touchdown - which would be +2 for the defending team and the least desirable outcome. However, through EDA, only 2 out of over 3000 plays had this outcome and therefore has little impact on the above findings. Despite the expected points of a 2pt conversion being lower, it may be necessary to attempt a touchdown when a losing team is in immediate need of more than 1 point.

VI. Field Goals

A field goal occurs when a player kicks the ball between the uprights and above the crossbar of the goal, resulting in 3 points. Field goals can be attempted by the offensive team from anywhere on the pitch, and require a holder to hold the ball in place after the snap and a kicker to attempt the shot. A visualization of locations where field goals where scored and missed are shown in Figure 6.

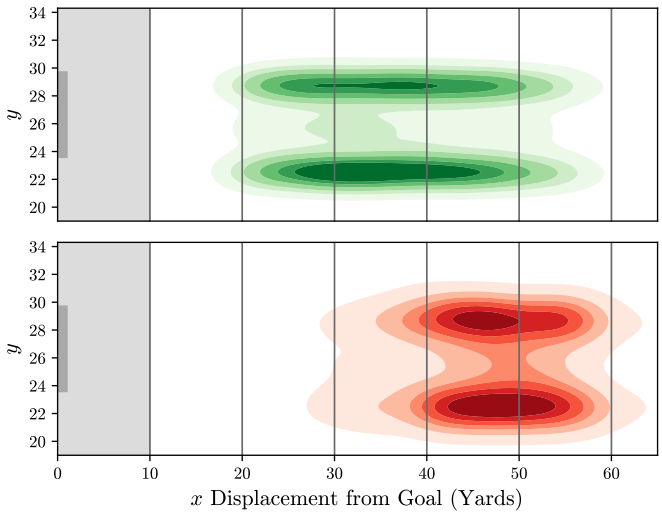

Fig. 6: *Field locations of where a field goal attempt was successful (green) or missed (red). Each outcome is rotated to show home and away attempts aiming in the same direction. Note that the aspect ratio is not to scale.*

An interesting note from this figure is the clustered y locations of left and right footed attempts. As a field goal attempt can be from any point along the line of scrimmage, the kicker can receive the ball at the y location of their choice. This results in a clear segregation between kicking foot preference.

This density plot indicates that there may be a correlation that the distance from the goal line and the likelihood of the attempt being successful, as the misses are closely clustered to the centre of the pitch. A plot showing probability of success against yardage to the goal is shown in Figure 7.

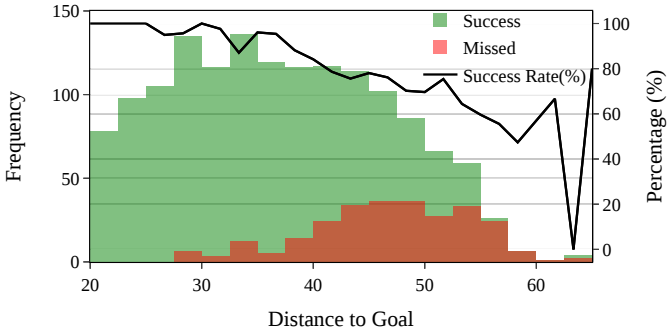

Fig. 7: *Probability of a successful field goal attempt at increasing distance from the goal.*

Due to there few data points for attempts that were over 60 yards away, the probability of success is highly inaccurate. As such, without further data, conclusions on score probability beyond this point are unreliable.

Predicting the Success Rate of a Field Goal Attempt

The model used to predict the success rate of a kick attempt can represent both field goals and extra point attempts. A logistic regression model was implemented, which is a type of supervised learning technique commonly used for binary classification problems. This model uses 'kick success' as the dependent variable and 'kick length' as the independent variable. Kick success was represented as a binary result where a miss $= 0$ and a success $= 1$. The kick distance is defined as the distance of the kick from the goal posts. The model forms the following logistic equation [12],

$$\log\left(\frac{P(success=1)}{1-P(success=1)}\right) = \beta_0 + \beta_1 \times Distance \quad (1)$$

where β_0 is the intercept and β_1 is the gradient. Applying equation (2) to find the different β values of the seasons forms the table,

Season	β_0	β_1
2018	6.7863	-0.1102
2019	6.3304	-0.1060
2020	5.7440	-0.0915
All	6.2822	-0.1031

TABLE I: *β values for the different seasons.*

For ease of use, rearranging Equation (1) allows for the β values and a specific distance to be inputted, outputting the probability of success.

$$P(success=1) = \frac{e^{\beta_0+\beta_1\times Distance}}{1+e^{\beta_0+\beta_1\times Distance}} \quad (2)$$

A graph showing the various fitted logistic regression models from each years data is shown in Figure 8.

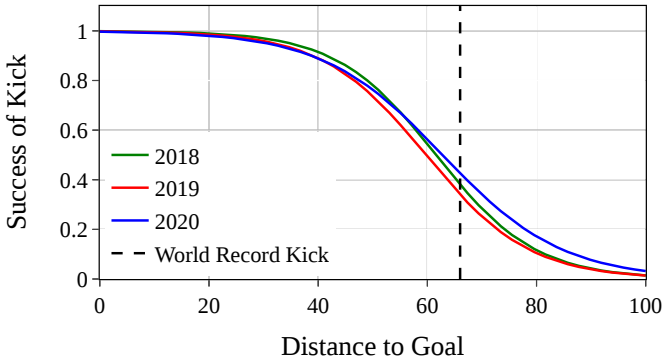

Fig. 8: *The fitted logistic regressions for data from each of the three different seasons, using the β values from Table I.*

The longest ever successful field goal attempt was at a distance of 66 yards by J. Tucker in the 2021 season [13]. This is shown in the graph to highlight a limitation of the model; predictions are still given for the probability of success even though it is effectively impossible. A logistic regression model is hindered by its ability to converge to zero as it will always return a probability even if it is minute. Therefore this model should only be used to interpolate data for attempts less than 66 yards away.

The model will always return the same output for a consistent input, which is not necessarily true. It is therefore necessary to introduce confidence intervals, which is a statistical measure used to help estimate the range of values that the output could be.

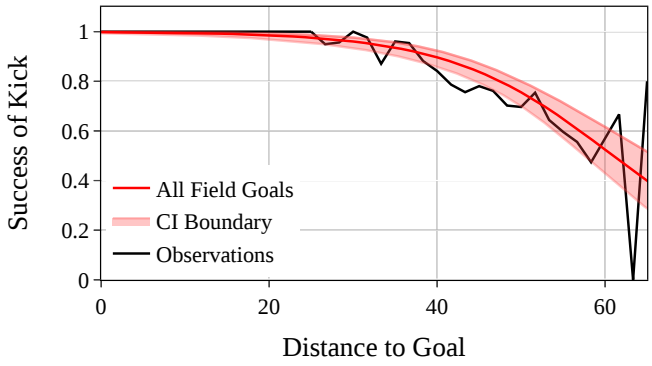

Fig. 9: *The fitted logistic regression to field goal attempts across all three seasons. The 95% confidence intervals is shown by the shaded red area.*

Implementing confidence intervals of 95% means when repeatedly interpolating data from the model 95% of the time, the true probability mean weight would be expected to lie within the confidence interval boundaries. The confidence intervals separate from the fitted mean as distance increase, suggesting that greater the distance the greater the uncertainty. This makes attempts further away from the goal line more risky, giving additional information to the coaches.

The use of a logistic regression model meant the small quantity of data points round a distance of 60 yards did not skew the predictions. From this model, coaches will be able to evaluate the current distance to the goal and identify the probability of a successful outcome. For example, inputting a distance of 55 yards using the β values from 2018 outputs a likeliness of 67%. The coach can then consider other play options, and make an informed decision on whether to attempt a field goal or not.

A study by Lopez *et al.* [12] mentions the use of grass length and other independent variables such as wind direction to form a model on field goal success rate. These variables could be experimented with on this model, to potentially further improve the accuracy of predictions.

VII. Punting

Punting is when the attacking team kicks the ball downfield in the hope of gaining yards. They are similar to Kickoffs in that the opposing team can attempt to return the ball downfield. From the dataset, it could be seen that punts only happen in the fourth down being a 'last resort' option. This is likely due to the fact there is no score reward from punting the ball; the aim is to give as little advantage to the opposition as possible. Teams often punt when deep in their own half to avoid potential turnovers of possession when dangerously close to their own endzone. It was found that accuracy of punts was 85.8%, measured by dividing the number of successful punt plays by total number of punt plays attempts.

Unsupervised Learning on Punting formations

Unsupervised learning techniques, such as clustering, can help uncover hidden patterns and relationships in the data without prior knowledge of the outcomes. By applying unsupervised learning methods to offensive punt plays, we can identify distinct player formations, strategies, and trends that contribute to the success of these plays

The following methods and functions were used for clustering offensive punt plays:

1. Data Preprocessing: We started by processing the tracking and play data and transforming it into a suitable format for clustering. For each play, we created a binary representation of the field with players' locations, We then created separate DataFrames for offense player locations.

2. Dimensionality Reduction (UMAP): To reduce the high-dimensional player location data, we employed Uniform Manifold Approximation and Projection (UMAP), a powerful dimensionality reduction technique. We created a function called `umapify()` that applies UMAP to the player location data and projects it into a 2D space.

3. Clustering (HDBSCAN): We used HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) as our clustering algorithm, which is a density-based clustering technique. We integrated HDBSCAN into our `umapify()` function to assign cluster labels to each play based on the reduced UMAP components.

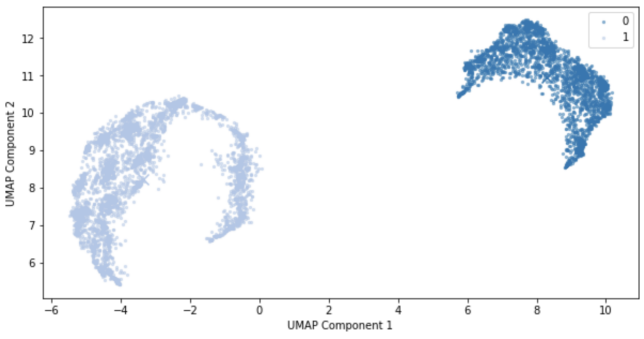

Fig. 10: *Offensive Punt Play Clusters*

The clusters in Figure 10 display clear differences in offensive player formations and punt plays. Cluster 1 shows a count of 120 downed plays compared to the higher 160 downed plays in Cluster 0.

The differences in downed play counts between the clusters could be indicative of distinct player formations, strategies, or play styles that lead to different outcomes in punt plays. Some possible explanations for the differences between the clusters in relation to downed plays are as follows:

Player Formations: The clusters might represent different player formations or positioning during the punt plays. For example, one cluster might have a more aggressive coverage formation, with players positioned to quickly move downfield and prevent the return team from advancing the ball. This could result in a higher number of downed plays for that cluster.

Punting Strategies: The clusters could also be indicative of different punting strategies employed by the teams. One cluster might consist of plays where the punting team is focusing on maximizing hang time and minimizing the returner's ability to run the ball back, thus increasing the likelihood of downing the ball. On the other hand, the other cluster might represent plays where the punting team is more focused on distance, sacrificing hang time for longer punts.

Return Team Performance: Another possible explanation for the differences between the clusters could be the performance of the return team. One cluster might represent plays where the return team struggles to field the ball cleanly or create running lanes for the returner, leading to more downed plays. In contrast, the other cluster might include plays where the return team is more successful at setting up returns and avoiding downed punts.

Carrying out this technique allows us to uncover hidden patterns in the dataset and also gives way to further analyzing the preliminary findings from the clustering process.

Predicting the Absolute Yardage Gain for a Punt

Support Vector Regression (SVR) is a machine learning algorithm that is effective in handling large datasets. It can deal with nonlinear data effectively and is less sensitive to outliers, making it a good choice for analyzing complex datasets. The model for predicting punts considers independent variables such as game clock minute, yards to go, kick length, returner average, number of gunners, num of punt rushers, num of special teams safeties, and number of vises.

The punt predictor model was created by merging the plays and scouting CSV's to allow for a full analysis of individual punts. From this, the returner's yard average was calculated, and NaN values were replaced with 0. SVR was chosen over Random Forest as it can better capture complex patterns and relationships in the data. The dependent variable selected for the analysis was 'playResult', which measures the yardage gained or lost on a play. To evaluate the model's performance, the data was split into 75-25 train and test sets. The train set was used to train the model, while the test set was used to evaluate its performance, ensuring that the model is not overfitting and can generalize well to new data.

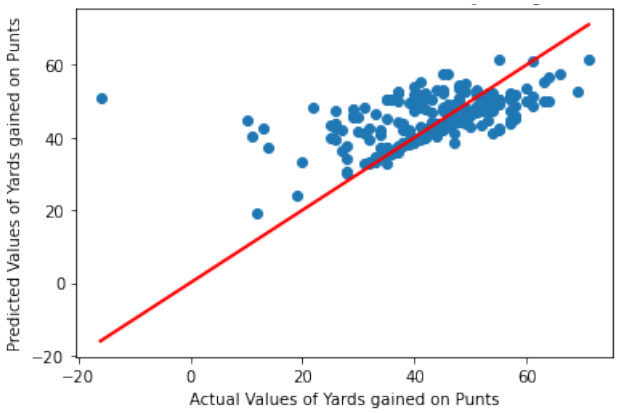

Fig. 11: *Plot of 1468 observations predicting play yardage against actual play yardage using the 8 independent variables.*

In Figure 11, the predicted yards gained have been plotted against actual yards gained where correct solutions lie on the $y = x$ line. The accuracy achieved by this model was 0.583 within a 10% interval, and showing that the model's predictions are within 10% of the actual values for just over half of the test data. Most of the data points cluster around the $y = x$ line, however, there is a spread of data points to the left and right of the line.

Since the model does not take into account the number of missed tackles, it over-predicted the play result. This is because we did not input missed tackles on the outcome variable, as it would not be mathematically correct and did not accurately capture the true relationship between the variables. The SVR model therefore relies on the relatively weak assumption that all plays were executed perfectly without any missed tackles.

The metrics mean squared error (MSE) and R-squared (R2) can be used to evaluate the performance of the SVR. MSE is the average squared difference between the actual and predicted values, implying the smaller MSE is the better the model is. R2 is a statistical measure that represents the proportion of variance, R2 ranges from 0-1 analyzing the variability in the response variable. The SVR had an MSE of 88.88 and R2 of 0.285. The R-squared score indicates that it can explain 29% of the variation in the yardage gained from punts. The mean squared error shows that the model's predictions are relatively close to the actual values of absolute yardage gained.

This model provides coaches in American football with a valuable data-driven tool that predicts the yards gained on punts. This additionally information can be used by coaches to decided whether to conduct a punt play, or whether a different play type may result in more favourable outcomes. The model can also be used for scouting opposing teams' punters and their expected yardage gained, allowing coaches to develop game strategies that account for their opponents' strengths and weaknesses.

VIII. Passing Plays

A passing play occurs when the quarter back receives a snap from behind the line of scrimmage and attempts to throw the ball either down the pitch to an eligible receiver. Passing plays can be interrupted before the quarterback can even release the ball, known as a quarterback sack. This counts as a down and play is resumed at the previous line of scrimmage. In the worst case scenario, the defensive team can gain possession of the ball during a quarterback sack, which is called a strip sack.

There are three outcomes if the quarter back is able to initiate a pass: a complete pass - an eligible receiver legally catches the ball, an incomplete pass - the ball touches the floor or the receiver drops/loses control of the ball or steps out of bounds, or an interception - the defensive team catches the ball without it touching the ground resulting in a turnover. A visualization of the special team forward passing plays between 2018-2020 can be seen in Figure 12.

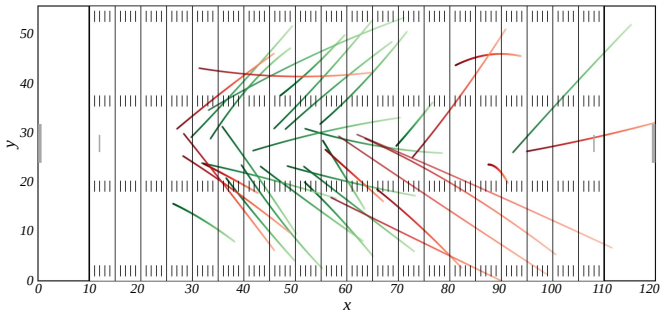

Fig. 12: *Visualization of passing plays in the 2022 competition dataset [6]. The direction of play is normalized to the right. Green lines indicates complete passes, while red lines indicate incomplete and intercepted passes.*

Unsupervised Learning for Passing Plays

By utilizing unsupervised learning algorithms, similar data points can be automatically grouped together, ultimately providing a better understanding of the underlying structure of the data and enabling data-driven decisions. In our analysis, we used K-means clustering, to group passing plays based on selected features.

We used the following methods for clustering these passing plays:

1. Data Preprocessing: We started by processing the tracking and play data and transforming it into a suitable format for clustering. For each play, we created a binary representation of the field with players' locations, We then created separate DataFrames for offense player locations.

2. Feature Selection: To carry out meaningful analysis it is imperative the features selected are done so with adequate domain knowledge of the dataset you're working with, we went with player location and supporting passing play features such as down, yards to go, quarter, and pass result. This selection would then provide underlying patterns in our dataset.

3. Clustering (Kmeans): The elbow method is a widely-used technique for determining the optimal number of clusters

in K-means clustering. When applying the elbow method, multiple K-means models with different numbers of clusters (K) are trained. For each model, the within-cluster sum of squares(WCSS) is calculated. As the number of clusters increases, the WCSS tends to decrease. The elbow point is where the decrease in WCSS becomes less pronounced and adding more clusters does not lead to a significant improvement in the model's performance. Essentially, it represents the balance between having too few clusters (high WCSS) and having too many clusters (overfitting the data). By selecting the number of clusters corresponding to the elbow point, we can achieve an optimal trade-off between model complexity and performance.

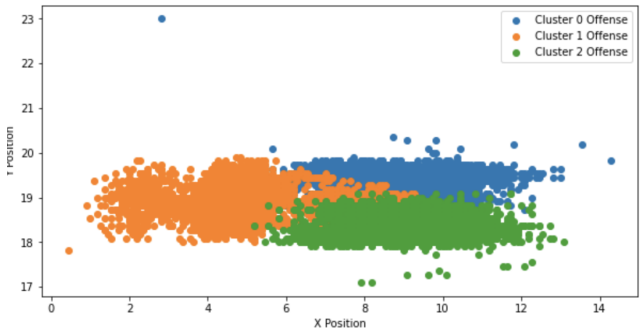

Fig. 13: *Passing Plays Clustering*

The resulting clusters provided insights into the relationship between situational and positional factors and pass results, we found three distinct clusters, each representing a different type of passing play:

Cluster 0: The offense is slightly more spread out horizontally compared to the defense. This cluster may indicate a mix of short and medium pass plays, with the offense trying to gain a moderate amount of yards.

Cluster 1: Plays at the beginning of a possession, likely kickoffs or punts. In these plays, the offense players are closer together, and the defense players are more spread out. This cluster likely represents special teams plays rather than typical offensive pass plays.

Cluster 2: Both the offense and defense players are relatively close together in their positions which could indicate plays where the offense is trying to exploit specific weaknesses in the defense with short passes.

While the analysis provided a degree of distinct clusters, a low number of data points in our passing play dataset can have several effects on clustering which implies the nature of the data we're working with:

1. Overfitting: With a limited number of data points, the clustering model may overfit the data. This means that it captures noise and small variations in the data rather than the underlying patterns.

2. Insufficient representation: A small dataset may not adequately represent the entire population of passing plays. This can lead to biases in the clustering results and may not reflect the true underlying structure of the data.

3. Less robust clusters: With fewer data points, the clusters formed can be less robust and more sensitive to slight changes in the data. This can result in unstable clustering results

From the above insights, we can confirm that using the current dataset to implement models based on passing plays would not inherently provide meaningful findings. We used this as an opportunity to explore the 2023 dataset which had much more data points for passing plays.

Predicting the Outcome of a Pass Attempt

The 'passResult' column in the plays files contains four unique outcomes complete, incomplete, interception and sack. A decision tree classifier works by iteratively partitioning the features into smaller regions based on the values of the input features, with the goal of separating the different classes as cleanly as possible. The algorithm builds a tree-like model of decisions and their possible consequences, where each node represents a feature or attribute, each branch represents a decision rule based on that feature, and each leaf node represents a class label.

There were a total of 45 special team passing plays making this a small dataset, therefore external data from [10] was used to train the model more extensively. The training to test data split was 75% and 25% respectively. To enhance the model further, a min-max-scaler was used to pre-process the data by setting all the features within the range of 0-1. This helps improve the models accuracy and allows for large input values to be scaled down so they do not dominant the model.

The decision tree used the stopping criterion of a max depth of 15. This value was chosen as it balances a lower complexity while not limiting the performance of the model. This avoids overfitting to the data, while saving on computational costs of simulation. Using this depth produced a tree made up of 1393 nodes; one branch of the tree can be seen in Algorithm 1.

Algorithm 1 A Branch of the Decision Tree

1: | down $<= 0.62$
2: | $-$ | absolute yardline number $<= 0.07$
3: | $-$ | $-$ | down $<= 0.38$
4: | $-$ | $-$ | $-$ | game clock minute $<= 0.23$
5: | $-$ | $-$ | $-$ | $-$ | quarter $<= 0.38$
6: | $-$ | $-$ | $-$ | $-$ | $-$ | absolute yardline number $<= 0.03$
7: | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ | class: Complete

		Actual			
		Complete	Incomplete	Intercept	Sack
Prediction	Complete	437179	16985	826	1769
	Incomplete	91132	173066	410	1723
	Intercept	6196	1496	10468	500
	Sack	17824	2674	169	40673

TABLE II: *Confusion matrix using the decision tree model predictions versus the actual outcome.*

Table II shows the predictions for the model which gave a training accuracy of 82.28% and a test accuracy of 82.36%. The high accuracy of this model provides a reliable tool that can be applied in real time. This model can be used to predict the outcome of a passing play, providing coaches with information on the likelihood of success.

IX. Running Plays

Running plays are simply when a player attempts to gain yards by carrying the ball up the pitch. Most running plays will begin with a snap, either directly to the runner or to another player who hands it off to the runner. The success of running plays are not only dependent on the skill of the player with the ball (rusher), but additionally the blockers who stop defensive players attempting to tackle the runner.

Field control is a sports analysis theory that was originally proposed by Spearman [14], developed for use in Soccer games. The concept is to visualize which parts of the pitch a team is in 'control' of, by considering the time it will take for players to reach that part of the pitch. The first step is to gather all the player locations and velocities, an example of which is shown in Figure 14.

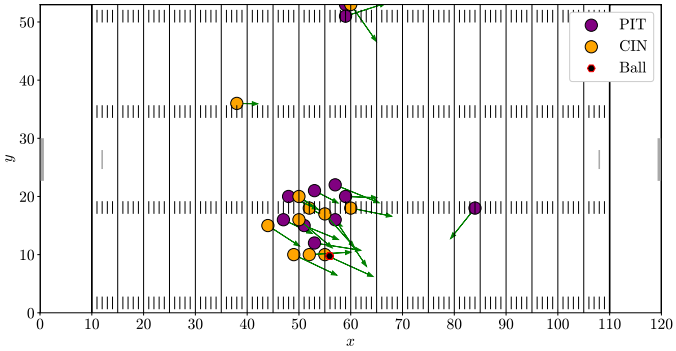

Fig. 14: *Player locations and velocities during a game between Cincinnati Bengals (Orange), and Pittsburgh Steelers (Purple). The play is a 39 yard run by Shawn Williams for the Bengals, and the frame shown is 4.1 seconds after the snap. The green arrow indicates the direction and magnitude of each players current velocity.*

The time taken for a player to reach a location on a pitch is difficult to model, as there is a high degree of uncertainty depending on multiple factors. As an approximation, it was decided that a player would continue at their current velocity for 1 second, before turning directly towards the new location at a constant speed of 8 yards s^{-1}. The 1 second delay accounts for the time taken to react, turn and accelerate in the new direction.

The field control value (FC) for a point (p) is first calculated by returning the total time taken for the closest player from the home team $T_h(p)$ and from the away team $T_a(p)$ to reach p. The difference between these two values

$$T_d = T_h(p) - T_a(p)$$

is placed inside the logistic function

$$FC(p) = \frac{1}{1 + \exp(T_d)} \tag{3}$$

to map the values from 1 - full home team control, to 0 - full away team control. The field control of S. Williams 39 yard run is shown in Figure 15.

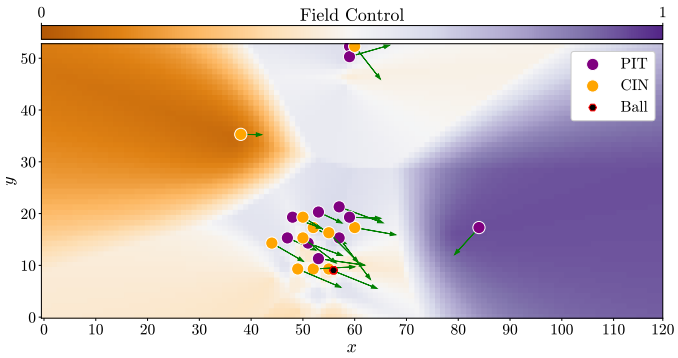

Fig. 15: *Field control for the game position shown in Figure 14. Areas with an FC value close to 0.5 appear white, and represent locations where neither team has a significant advantage in field control.*

Optimal Route Finder

Field control analysis can be used to calculate the optimal path between any two positions on the pitch. The FC value for any given point, as calculated in equation (3), can be considered as the cost of travelling over that point on the pitch. The optimal path would therefore travel through the route with the maximum sum field control of the offense.

In theory, it would be possible to calculate the cumulative cost of all possible paths from the current position of the player with the ball to the end zone, and then return the path that has the lowest cost. However, there are often several million possible paths for any given frame, thus a 'brute force' approach is computationally unfeasable.

A Dynamic programming approach was considered, which avoids the algorithm from making redundant calculations on routes that have already been evaluated, significantly reducing the time complexity of the overall solution. All plays were normalized for play direction to the right, and for field control being mapped from 1 - full defensive control, to 0 - full offensive control. The dynamic programming algorithm first calculates a cumulative map of field control using Algorithm 2.

Algorithm 2 Dynamic Programming - Cumulative FC

$x = x_{target}$
for $x > x_{pos}$ **do**
  if $x = x_{target}$ **then**
    $C[x, y] = F[x, y]$
    continue
  end if
  for $y = 1, 2, \ldots, 53$ **do**
    $min_C = \min\left(C[x, y-1],\ C[x, y],\ C[x, y+1]\right)$
    $C[x-1, y] = min_C + F[x-1, y]$
    $x = x - 1$
  end for
end for

F is a matrix containing the FC values as calculated in (3) and C is a matrix containing the newly formed cumulative FC map. The variable x_{pos} denotes the x position of the rushing player, while x_{target} describes the x point where the minimum path aims to reach. The latter can be set to reach the endzone ($x_{target} = 120$), or to reach a first down ($x_{target} =$ Line of Scrimmage + Yards Needed for First Down).

To calculate the optimal route, a secondary algorithm starts at the position of the rusher and moves in the direction of the minimal cumulative cost. Figure 16 shows the cumulative cost map and optimal path to the endzone for S. Williams run, along with the actual route taken by the rusher.

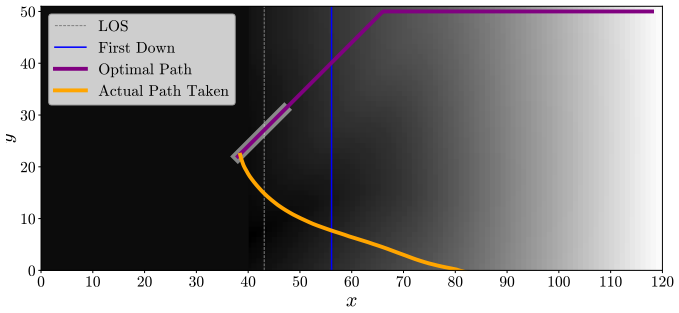

Fig. 16: *Cumulative Field control for the play shown in Figure 14. The optimal path is shown in purple, while the orange line indicates the actual route taken by S. Williams.*

The model suggests that S. Williams should have ran in the opposite direction where his team held greater field control. There does not appear to be any previous literature surrounding a dynamic approach to optimal rushing path, and the model outperforms machine learning techniques such as the A^* algorithm, in terms of computational efficiency.

The main limitation of this Dynamic Programming approach is the assumption of constant forward motion. While runners must be constantly moving in order to avoid being tackled, they will occasionally turn backwards in an attempt to open up more space.

Expected Yards using Distribution Fitting

To formulate the expected yardage gained from a given play, data from a previous big data bowl competition [11] containing 16,074 running plays was evaluated. Previous work by Brighenti [15] found that successful rushes held a greater field control in front of the path that the runner took. Using this information, the optimal path was calculated for each play, recording the average cost of the first ten steps in each path and the actual yards gained.

To perform the analysis, it must be assumed that the rushers ran the optimal route, which is not necessarily true as shown in Figure 16. From the special team dataset [6], 68.97% of rushers approximately took the optimal route. However, these are all fake plays with non-specialist rusher, thus the percentage of quarterbacks who take the optimal route will likely be much higher. Unfortunately, the 2020 dataset [11] does not contain full tracking data and thus there is no way to verify the strength of this assumption without further information.

The data from this analysis was split into four quantiles based on the average cost of the first ten steps in each path. For each quantile, a log-norm probability density function (PDF) was fit on the distribution of yards gained. The complementary cumulative distribution function (CCDF) is shown for each quantile in Figure 17.

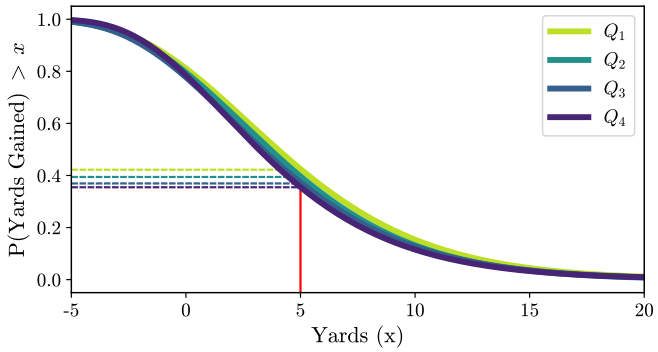

Fig. 17: *CCDF for the four average field control quantiles. The probability of gaining 5 yards or more is 42.21% for the lowest quantile (Q_1) and 35.47% for the highest quantile (Q_4).*

This graph verifies the findings in [15]; holding a greater field control along the path infront of the rusher increases the probability of success. The PDF for S. Williams, had he taken the optimal path is shown in Figure 18, with the average cost of the path falling into the second lowest quantile of data.

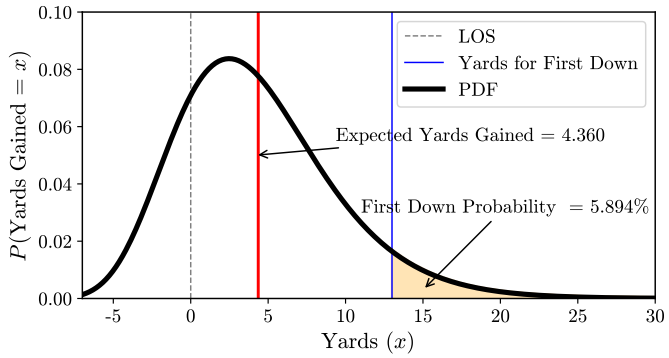

Fig. 18: *Probability Distribution Function (PDF) of yardage gain, for S. Williams run had he taken the optimal path. The probability of reaching a first down is simply the integral of the PDF over the domain between the first down yard line and the endzone. In this case, there was a 5.849% chance of achieving a first down.*

The findings from this analysis can be used at an individual level, to evaluate the movement decisions made by a rushing player. Training could be introduced to improve the rushers ability to choose the optimal path. Coaches can also use this information to predict the yards gained, and consider whether alternative play types may yield a more successful outcome.

X. Discussion

The previous section considered a specific running play by S. Williams, concluding that he did not take the optimal route. The other models can be used to evaluate whether the correct play was chosen. As the play occurred in 2020, the β values from this season (Table I) could be used in the field goal model to predict the probability of success. However, as the distance to the goal was 76 yards, extrapolation would be required, determining a field goal attempt was not an option. The punt model predicted yards gained of 14.39, which is low and would result in a turnover of possession. This was due to the lack of special team punt players, such as gunners on the pitch, determining that a punt play should not be considered. The passing model predicted the outcome of the pass to be an intercept, therefore losing possession of the ball concluding that this play should also not be considered. The expected yardage gained from a running play was 4.360 (regardless of field control) concluding that running was the only valid play. However, the likelihood of success could have been greater had he taken the optimal path, increasing the the probability of reaching a first down from 5.357% to 5.894%. This example shows how the models found in this report can be used by coaches in game to assess what the best play option should be, and also by players in training to evaluate the decisions made during a play.

A limitation of the models are that they do not consider the individual attributes of each player, such as their physical abilities, experience, or psychological factors like faking or being surprising. These individual factors can have a significant impact on the success of a play, and future research could explore ways to incorporate this information into data modeling approaches.

XI. Conclusion

This report has provided valuable insights into the four main play types of running, passing, field goals and punting. Visualization has been performed on the data such that relationships between game variables and offensive success is clearly displayed. We have created several models using a variety of advanced techniques, such as Support Vector Machining and Dynamic Programming. The models are easy and intuitive to use, and can supplement the information available to coaches and players. This will allow NFL teams to make more statistically informed decisions, and increase the effectiveness of their offensive game.

In conclusion, data modeling for NFL offensive plays is a promising area of research, with the potential to provide valuable insights for coaches and analysts. However, it is important to recognize the limitations of this approach and to continue exploring new ways of collecting and analyzing data to gain a more complete understanding of the game. By combining data modeling with traditional coaching and scouting methods, teams can develop a more comprehensive approach to strategy and game planning.

XII. Future Work

While our analysis of NFL offensive plays provides valuable insights, there are opportunities for future research to further enhance the understanding of the game. The inclusion of more granular player data; psychological factors such as faking and having a home advantage, and a wider range of play types, such as trick plays and screen passes, along with considering weather and injuries can provide deeper insights into the success of offensive plays. Coaches can use weather data to adjust strategies in response to changing weather conditions while identifying potential strategies for minimizing the impact of injuries on the field can help in improving offensive play success rates. Incorporating these factors into future research can provide a more comprehensive understanding of the success of NFL offensive plays.

In the discussion section, we evaluate a single example play using the models formed in this report. Future work could consider creating a generalised model, that would evaluate all options simultaneously, and output the best option for coaches.

References

[1] Pro football reference. https://www.pro-football-reference.com/years/NFL/scoring.htm. Accessed 2023-29-03.

[2] Bob Carroll, Pete Palmer, and John Thorn. *The Hidden Game of Football: A Revolutionary Approach to the Game and Its Statistics*. University of Chicago Press, 1988.

[3] National Football League. Big data bowl. https://operations.nfl.com/gameday/analytics/big-data-bowl/, 2023. Accessed 2023-11-03.

[4] R Ritchie, B Kumagai, R Moreau, and E Cavan. Optimal path generator. https://rr-sportsstats.shinyapps.io/path_generator/, 2022. Accessed 2023-10-03.

[5] A Patel. Introduction to a*. http://theory.stanford.edu/~amitp/GameProgramming/AStarComparison.html, 2023. Accessed 2023-10-03.

[6] National Football League. Nfl big data bowl 2022- special teams. https://www.kaggle.com/competitions/nfl-big-data-bowl-2022, 2022. Accessed 2023-10-03.

[7] Baek Kyun Shin. Nfl big data bowl basic eda for beginner. https://www.kaggle.com/code/werooring/nfl-big-data-bowl-basic-eda-for-beginner, 2022. Accessed 2023-10-03.

[8] V Sanjay. Nfl big data bowl - beginner's complete eda. https://www.kaggle.com/code/sanjayv007/nfl-big-data-bowl-beginner-s-complete-eda, 2022. Accessed 2023-10-03.

[9] Github repository. https://github.com/JackBellamy/NFL, 2023.

[10] National Football League. Nfl big data bowl 2023- passing. https://www.kaggle.com/competitions/nfl-big-data-bowl-2023/data, 2023. Accessed 2023-10-03.

[11] National Football League. Nfl big data bowl 2020- running. https://www.kaggle.com/c/nfl-big-data-bowl-2020, 2020. Accessed 2023-10-03.

[12] M Lopez. Logistic regression and nfl kickers. *WordPress*.

[13] 10 longest field goals in nfl history. https://sportsnaut.com/longest-field-goal-in-nfl-history/. Accessed 2023-05-04.

[14] William Spearman. Quantifying pitch control. 02 2016.

[15] C. Brighenti. Why successful run plays work. https://www.footballoutsiders.com/stat-analysis/2020/why-successful-run-plays-work, 2020. Accessed: May 4, 2023.