**X Marks The Spot: Assessing NFL Offenses with the Formation Performance Score (FPS)**

MRBL Systems

*Overview:*

Football analytics has evolved into a sophisticated field, leveraging advanced statistical models to dissect player performance and team strategies. In this project, we delve into the realm of offensive plays, aiming to assess and quantify the performance of NFL teams using a novel metric – the Formation Performance Score. This score, rooted in the analysis of player tracking data, specifically focuses on predicting the outcome coordinates of offensive plays based on the offensive formation (such as shotgun, pistol, wildcat, etc.) and play type (run or pass). The objective is to provide a nuanced understanding of an offense’s strengths and weaknesses as well how said strategies unfold on-field. Using machine learning, we can uncover nuances that revolutionize how we evaluate offensive strength in the NFL.

The Formation Performance Score (FPS) introduces a novel approach to assessing NFL offenses. Unlike traditional player-centric metrics, FPS employs machine learning on player tracking data to predict the outcome of offensive plays. This fills a crucial gap in traditional statistics by considering the spatial dynamics and coordination inherit to the game of football. The FPS uses a Regression Multilayer Perceptron (RegMLP) model to achieve actionable insights for NFL teams on a seasonal or weekly basis. It transcends numerical evaluation and serves as a tool for teams to optimize offensive strategies based on data-driven assessments. This section sets the stage for a holistic understanding of the FPS’s significance and ability to revolutionize offensive strategy evaluation in the NFL.

This project was conducted as a part of the NFL Big Data Bowl 2024 competition and will be submitted to the metric track.

*Data Preparation:*

The foundation of our analysis lies in the comprehensive dataset, merged\_df, a compilation of player tracking data (from tracking\_week\_[1-9].csv) and play-specific information (from plays.csv). The dataset captures the details of each play such as offensive formations (offenseFormation, type: string, example: ‘SHOTGUN’), play types (playType, type: string, example: ‘run’, analysis generated), and the spatial coordinates of the football at the beginning of the play (x\_start, y\_start, type: float), tackle location (x\_tackle, y\_tackle, type: float), and the distance between initial and final positions of the football (delta\_x, delta\_y, type: float, analysis generated).

*Key Features:*

*offenseFormation:* Drives the analysis by categorizing plays based on the offensive team’s formation.

*playType:* Further differentiates analysis between run and pass plays, determined by the function determine\_play\_type, seen here below:

A screen shot of a computer program

Description automatically generated

*delta\_x* and *delta\_y*: serve as target variables for the RegMLP model (we will run 2 instances of the model, one predicts the delta\_x coordinate and the other predicts the delta\_y coordinate) representing the spatial outcomes of offensive plays. Statistical code for both features is included here below:

A computer screen with colorful text

Description automatically generated

*delta\_x:* horizontal distance covered over the course of the play from x\_start to x\_tackle where the X axis represents the distance downfield. This is calculated as the absolute distance difference between the final horizontal position (x\_tackle) and the initial position (x\_start). Increasingly positive values (0 to 120) represent horizontal distance covered from the line of scrimmage to the play’s tackle location.

*delta\_y:* vertical distance covered over the course of the play from y\_start to y\_tackle where the Y axis represents the lateral position on the field. Increasingly positive values (0 to 26.65) represent lateral positions to the left of midfield. Increasingly negative values (0 to -26.65) represent lateral positions to the right of midfield (left and right are determined if an individual was standing at quarterback position at line of scrimmage).

*Preprocessing Steps:*

The merged\_df dataset undergoes several crucial preprocessing steps before being fed to the RegMLP model for analysis and prediction. First, we identify any missing values in the dataset and address them via mean imputation. After doing this and calculating delta\_x and delta\_y, we utilize One Hot Encoding to categorize our qualitative features in the dataset such as offenseFormation and playType. Finally, we divide the processed dataset into training and test sub-sets to facilitate model training and evaluation – thus ensuring model generalizability (maybe change this word) to previously unseen data.

*The Regression Multilayer Perceptron (RegMLP) Model:*

The Regression Multilayer Perceptron is a neural network model designed for predicting a continuous target variable, in this case, the delta\_x and delta\_y coordinates. The model is implemented using the TensorFlow library for neural network architecture and the scikit-learn library for data preprocessing.

The architecture of the RegMLP is structured as follows. The input layer accommodates the preprocessed features including the one-hot encoded offenseFormation and playType features. The model consists of three hidden layers which respectively and subsequently contain 128, 64, and 32 neurons. Each layer utilizes the Rectified Linear Unit (ReLU) activation function to introduce non-linearity and capture complex relationships within the data. Finally, the output layer consists of a single neuron representing the delta\_x or delta\_y coordinate, depending on the set target variable or model. The choice of layers and neurons aims to strike a balance between model complexity, computational efficiency, and performance optimization. In addition, the ReLU activation function is chosen for its ability to handle non-linearities effectively.

The model is compiled with an Adam optimizer and Mean-Squared Error (MSE) loss function – both typical choices for regression tasks. Mean Absolute Error (MAE) is chosen as a metric to provide a measure of the average prediction error. Finally, the model is trained using the training dataset consisting of offense formation, play type, and corresponding delta\_x or delta\_y values. Of the entire parent data set (merged\_df), the training subset is comprised of 80% of said parent dataset. Model performance is evaluated on the test subset (unseen by the model). Training is performed for a specified number of epochs with a defined batch size.

*Model Evaluation Metrics:*

During and after training the model’s performance is evaluated using the following metrics. These metrics collectively offer insight into the model’s predictive capabilities and how well it captures the nuances of offensive plays in football. The evaluation process ensures that the model generalizes well to unseen data, contributing to its utility for assessing offensive performance. It is important to note that said evaluation metrics can exhibit variation each time the model is run due to several factors such as random initialization, data splitting, and stochastic elements.

*Model Evaluation Metrics:*

*Mean Squared Error (MSE)*: Measures the average squared distance between the predicted delta\_x or delta\_y value and the actual delta\_x or delta\_y value, thus emphasizing larger errors.

*Mean Absolute Error (MAE)*: Measures the average absolute difference between the predicted delta\_x or delta\_y value and the actual delta\_x or delta\_y value, thus offering a more interpretable measure of prediction accuracy.

*Coefficient of Determination (R2 Score)*: Represents the proportion of variance in the target variable (delta\_x or delta\_y) explained by the model. Here R2 values closer to 1 indicate a good fit.

*Interpreting the Formation Performance Score (FPS):*

The RegMLP model’s prediction, which can be understood as an ordered pair of X and Y coordinates (delta\_x and delta\_y), gives us a regression prediction of how the entire NFL performs on a given play formation and play type. By comparing this prediction with individual teams’ performances using the same play formations and play types, we can assess an offense’s ability to be “better than” or “worse than” the rest of the league when it comes to said formations and play types. The most pragmatic way to understand the FPS would be through a case study.

To set the scene, we will predict the regression outcome (i.e. tackle locations, delta\_x and delta\_y) of all NFL shotgun run plays from 2022. To do so, we execute the code found in the attached appendix. Running the functions i) PredictPlayOutcome\_MLP\_X and ii) PredictPlayOutcome\_MLP\_Y will result in a predicted delta\_x and delta\_y, respectively, that is representative of the whole league’s performance for shotgun run plays. This prediction can be illustrated using the map\_to\_field function where we see our prediction demarcated by a red ‘x’ and all of the other shotgun run play outcomes from the 2022 season marked as blue circles.

With this prediction, relative to the rest of the league’s performance, we can now move on to the map\_FPS function. This code will output a similar map as map\_to\_field, however the user has the ability to select which team they would like to measure FPS for via the parameter offenseTeam. For our purposes we will use BUF (Buffalo Bills) as the argument for offenseTeam. Once the function has run, we will now see that all shotgun run plays that the Bills executed are highlighted in green while the rest of the graphic remains unchanged. This gives us a visual representation of the FPS, however in a more compact form, the function tells us that Buffalo outperformed the league regression prediction 46.25% of the time – this is the Formation Performance Score. Of course, these functions can be run with any combination of offensive formation, play type, and team of interest with up to 448 unique combinations available.

With this specific FPS in mind, lets discuss what it means and how it can be interpreted from an offensive and defensive perspective. Offensively, running this analysis on every combination of formation, play type, and team of interest you can get a holistic view of your offense’s strengths and weaknesses relative to the rest of the NFL. In reference to our case study, the shotgun run is one of Buffalo’s best offensive plays in their arsenal according to our FPS metric. Conversely, Buffalo’s singleback pass play had an FPS score of 30.77% indicating that more offensive refinement may be needed when it comes to this play combination. The FPS metric can also have significant implications in regard to game planning for defenses. For example, if a team’s offensive FPS is relatively high for a given play combination, defensive plays can be more effectively prepared or called in-game. It is important to note that this FPS metric is inherently tailored towards offensive predictions given that the data only includes offensive play formations, however if similar data were available for defenses (such as defensive formations and/or man vs. zone coverage), the same interpretations and analyses could be made on the other side of the ball.

*Conclusion(s):*

Our exploration of the Formation Performance Score (FPS) has illuminated a novel and powerful approach to evaluating NFL offenses. The FPS, driven by a Regression Multilayer Perceptron (RegMLP) model, stands at the intersection of advanced analytics and football strategy – providing teams with a tool that transcends traditional metrics.   
 Interpreting the FPS unveils its practical significance for offensive and defensive strategies. Via our small case study, we can see how teams can leverage the FPS to identify strengths and weaknesses in their offensive playbook. Beyond the offensive implications, and with more data, the FPS can also be repurposed as a valuable tool for defensive coordinators allowing them to make data-driven decisions when planning against specific offensive play combinations. The visual representation of FPS (from map\_FPS), offers a concise yet powerful way to assess team performance.

Looking ahead, the FPS opens up further avenues of exploration and expansion as previously discussed in regard to the defensive side of the ball. As football analytics continue to evolve, the Formation Performance Score stands as a pioneering metric, empowering teams to navigate the complexities of football with data-driven precision. The synergy between machine learning and football strategy showcased in this project marks a significant step forward in the search for a deeper understanding of America’s game.