

Analysis and Predictive Modeling for Play-calling in the NFL

Michael Flores
Flatiron School



Introduction

- ❖ “Defense wins championships” - Paul “Bear” Bryant
- ❖ In recent years, the game of football has largely shifted towards an offense-focused mindset.
- ❖ If a defense can pinpoint a play-call given the game situation, they can anticipate it ahead of time.
- ❖ I wanted to design a predictive model that could predict whether a team would run or pass on a play.



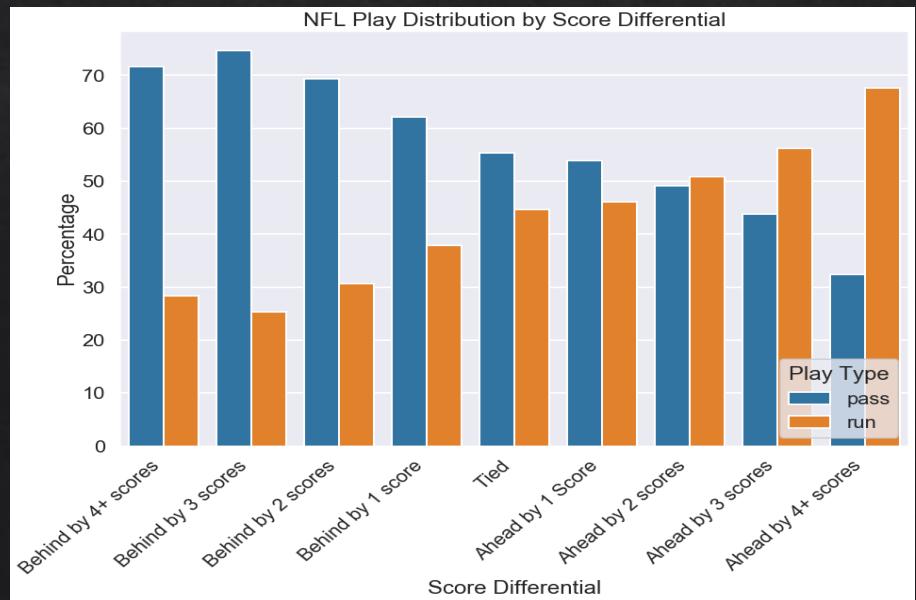
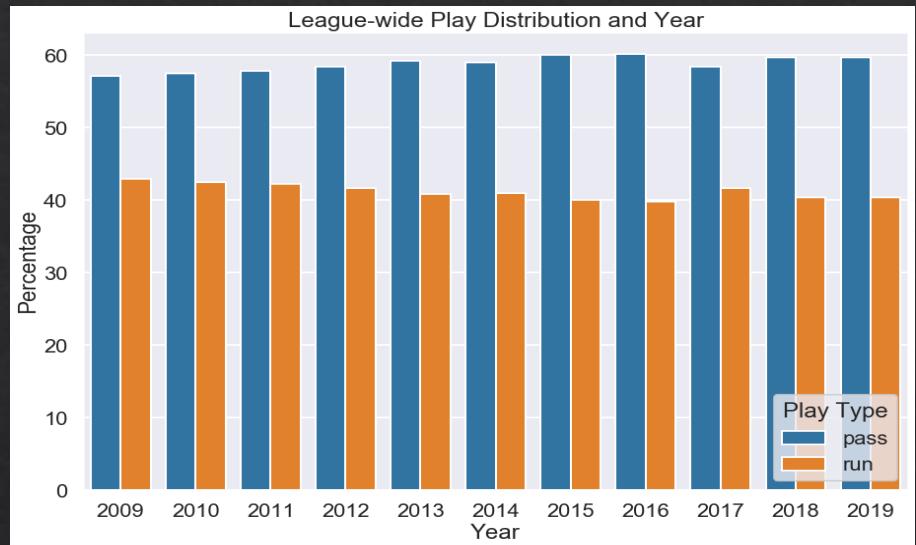
Data Information

- ❖ The data consisted of nearly 350,000 plays that occurred between the 2009 and 2019 regular seasons.
- ❖ Each play contained details about the play's description/outcome as well as game details.
- ❖ Each of the models were trained on details that an NFL defense would have in-game.



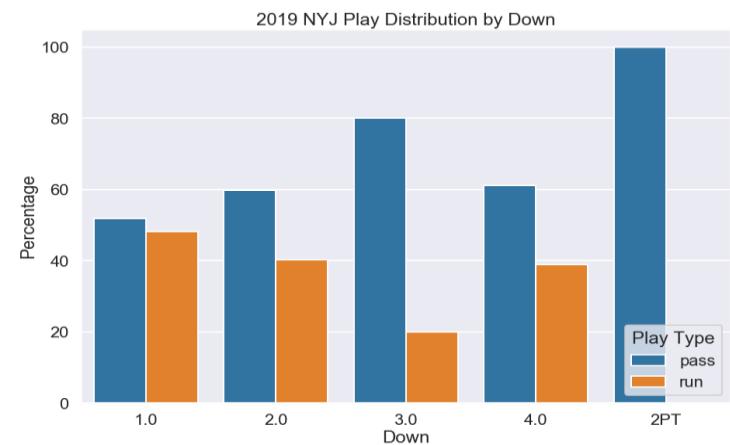
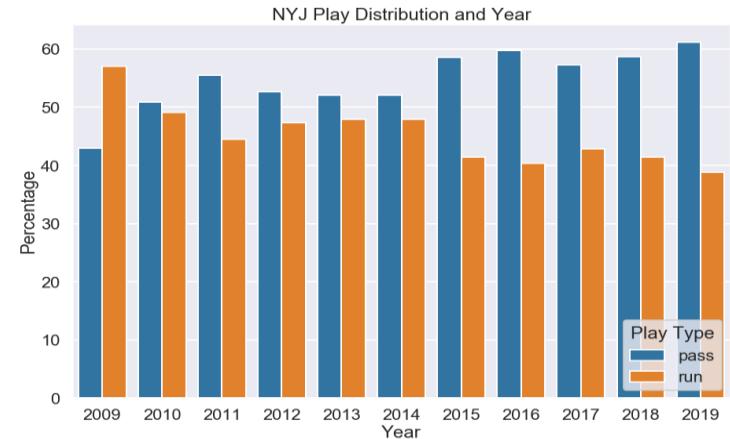
NFL General Trends

- ❖ I first wanted to observe general trends across all the teams in the NFL.
- ❖ Percentage of passing plays has stayed consistent over the years despite claims stating otherwise.
- ❖ Passing being quickly more favored when down by 1+ scores and runs being favored when ahead.



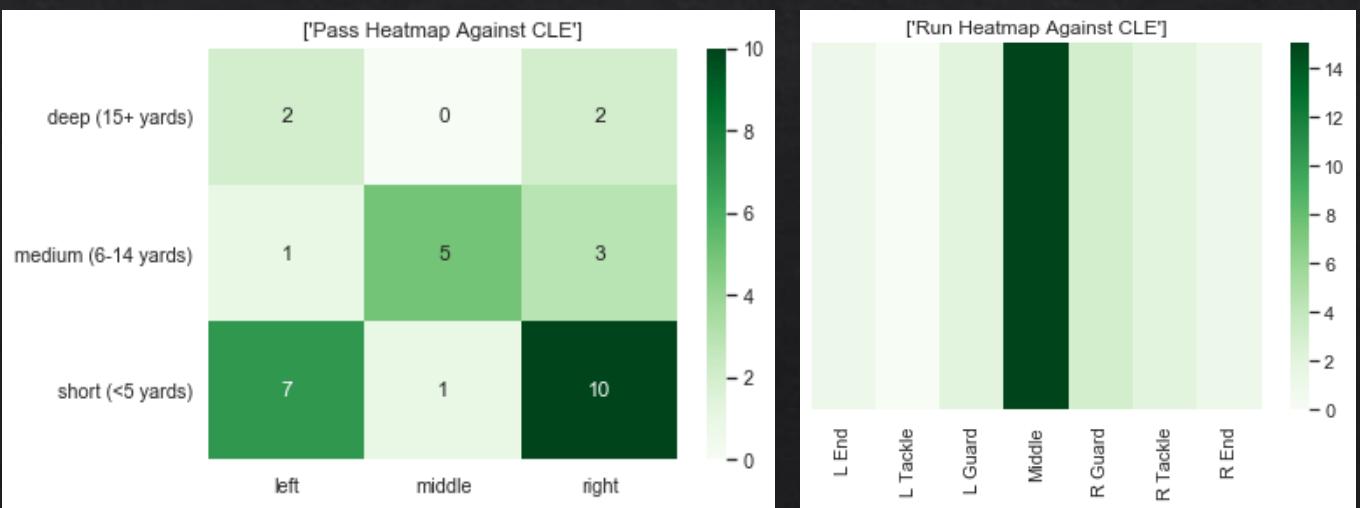
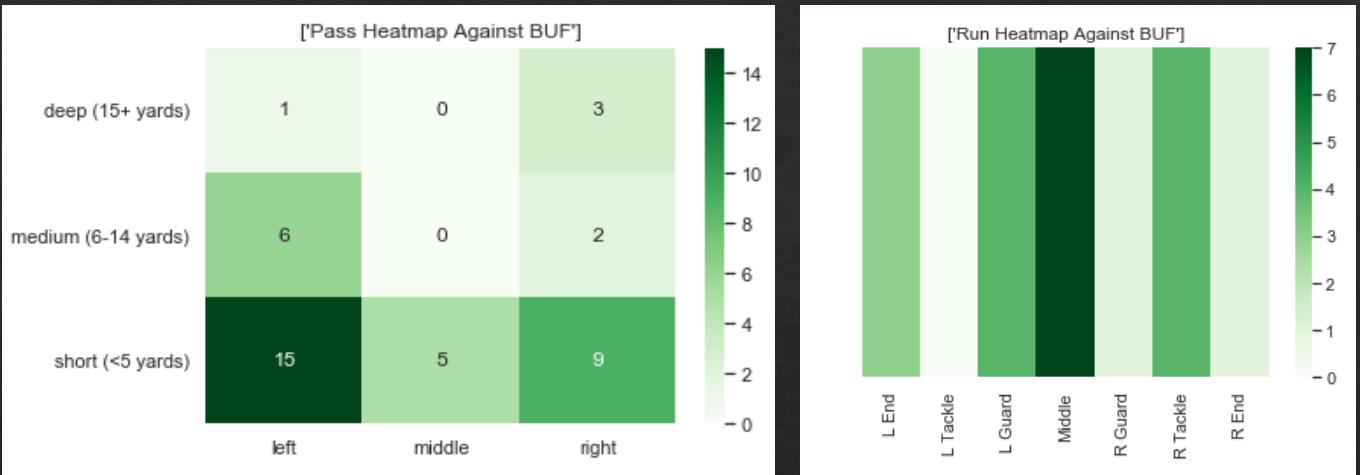
NFL Team Dashboard

- ❖ I wanted to create visualizations that would help to identify trends and tendencies for each NFL team.
- ❖ Visualizations for not only the entire span of the dataset, but each individual season as well.



NFL Team Dashboard (cont.)

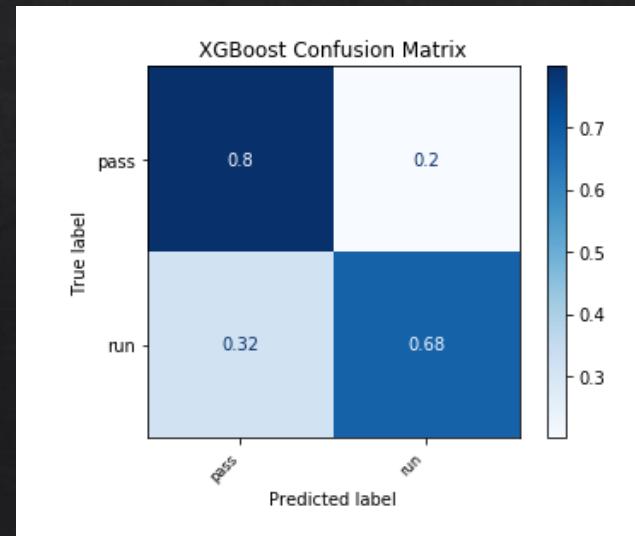
- ❖ Visualize each team's play-calling for both runs and passes for all their games.
- ❖ Heatmaps showing which sides of the field a team's offense favored for both runs and passes.
- ❖ Helps to highlight what offenses liked to run against certain opponents.



Predictive Modeling Results

- ❖ Multiple machine learning models were tested such as Logistic Regression and Random Forests.
- ❖ Metrics for measuring performance were accuracy and passing plays recall.
- ❖ XGBoost performed the best out of all models tested on accuracy and recall.

Model Type:	Overall Accuracy:	Pass Recall:	Run Recall:
Logistic Regression	0.720146	0.76	0.66
Decision Tree	0.65432	0.71	0.58
Random Forest	0.71647	0.73	0.7
Gridsearch	0.74138	0.75	0.73
XGBoost	0.75049	0.8	0.68



Recommendations

- ❖ Utilizing the model for general game situations.
 - ❖ Limitations include not having full details of each play and accounting for audibles.
 - ❖ Maintaining a level of unpredictability on play-calls for given situations.
 - ❖ Utilizing offense heatmaps to see favored areas of the field in order to anticipate a play.



Future Work

- ❖ Exploring how different personnel sets would influence the model's accuracy.
- ❖ Testing model performance across individual teams as opposed to the entire league.
- ❖ Incorporating playoff situations and seeding into the model.
- ❖ Deployment of an accessible web dashboard of each team's play-calls and tendencies.

Personnel Packages		
	RB	TE
	WR	
23	2	3
22	2	2
21	2	1
20	2	2
13	1	3
12	1	2
11	1	3
10	1	4
02	2	2
01	1	4
00	3	4



Thank you!

Michael Flores
Flatiron School

