Analyzing 2022 NFL 4th Down Success by Team

Jack Miller and Alex Jackson

Introduction

Teams across the National Football League have been incorporating analytics into their decision making more and more over the past 20 years. The emergence of analytics has been increasingly dragged by those who think football and numbers should stay separate. Recently, the decision to go for 2 points after scoring a touchdown to go down 6 instead of 7 points late in the game has become a controversy, yet it is an analytically-correct move. We believe the next enhancement of the game in regards to the incorporation of analytics comes one 4th-down decision making. Our project examines NFL 4th down decision making using play-by-play data dating back to 1999. Our goal is to build three models that predict expected win probability added for each of the three decisions coaches have on fourth downs: kick a field goal, punt it, or go for it. Personally, we believe teams should be going for it more and punting less than how they operate currently, but we wish ti support this claim with models.

For the purpose of our study, we simplified 4th down decision making, meaning we did not differentiate between run or pass plays, and any trick plays were included as going for it (i.e. lining up in punt formation does not classify it as a punt). To train our data, we used everything that was available to us. Our data comes from the `nflfastR` package which contains data on every play dating back to 1999. After filtering for 4th down plays, we used everything from this season all the way back to 1999. Although we considered basing our project on more recent data, we figured the more data the better and decided to use it all.

Exploratory Data Analysis

Count of NFL 4th Down Plays
4th Down Plays Between 1999 and 2022
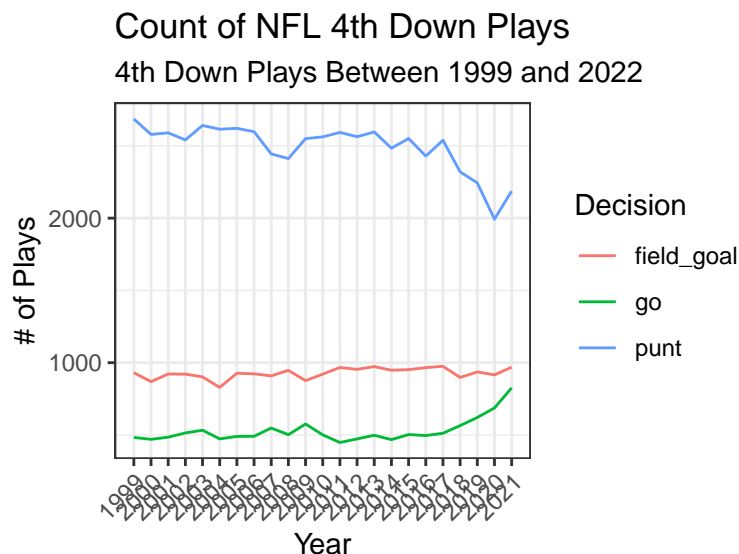
play_type	count	mean_wpa
field_goal	22075	-0.00035
go	12713	0.00393
punt	59018	-0.00043

To get a better idea of what our data looks like, we first wanted to see how much data of each type of play we have. Most 4th down plays are punts, of course, and only a small piece of our dataset is go for it plays, which also makes sense. Looking at the graph above, however, we can see an uptick in go attempts and a steady decrease in punts over the last 4 full seasons. This hints at the greater picture we are trying to discover: coaches haven't found the right go/punt/kick balance. But they are trending in the right direction.
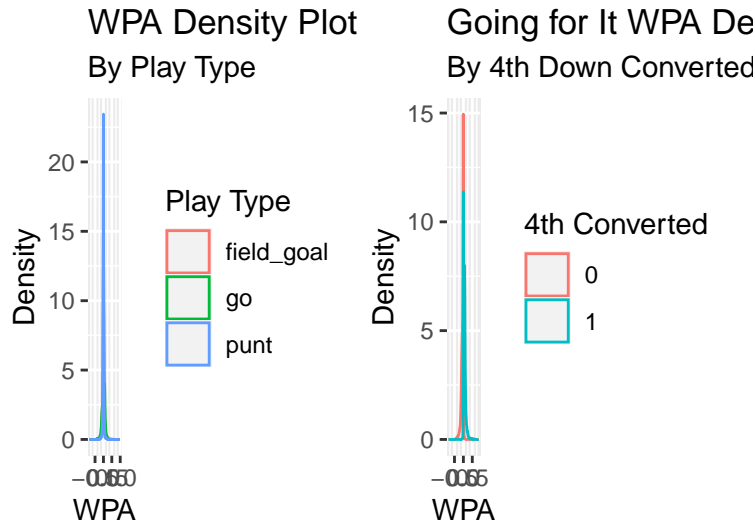
We can also see that while most of the density plots for WPA are relatively normal, the density plot for going for it on 4th down appears slightly bimodal. This is because a successful 4th down conversion usually leads to a bigger change in WPA than a 4th down failed attempt. While we weren't sure if this was worthy of transforming WPA for the go for it model or all the models, it is something to beware of moving forward.

Modeling

The goal with our models is to create the best possible models for predicting WPA (win probability added) given a certain game state. To do this, we trained 3 separate xGBoost models, each on their own data of one play type. We began by training a baseline xgboost model as well as a Lasso and Ridge regression model for each play type. After examining the MSEs of the models, the xgboost baseline model performed much better than both the ridge and lasso optimal models. After we decided to use xgboost for our final model, we played around with the parameters of the model to try and achieve the best predictive model possible.

[section on the reason we chose xgboost] We split the data for each model into train and test sets using 70-30 splits and then attempted to minimize the RMSE through adjusting model parameters. We also tested lasso and ridge regression models, just to make sure we were fitting the best possible type of model.

Punt Model

Our optimal punt XG Boost model has a RMSE of 0.0216, which is the lowest RMSE we obtained throughout our XG Boost, Lasso, and Ridge Regression modelling process. Therefore, we will continue with this as our

punt WPA model.

Kick Model

Our optimal kick XG Boost model has a RMSE of 0.3921, which is the lowest RMSE we obtained throughout our XG Boost, Lasso, and Ridge Regression modelling process. Therefore, we will continue with this as our kick WPA model.
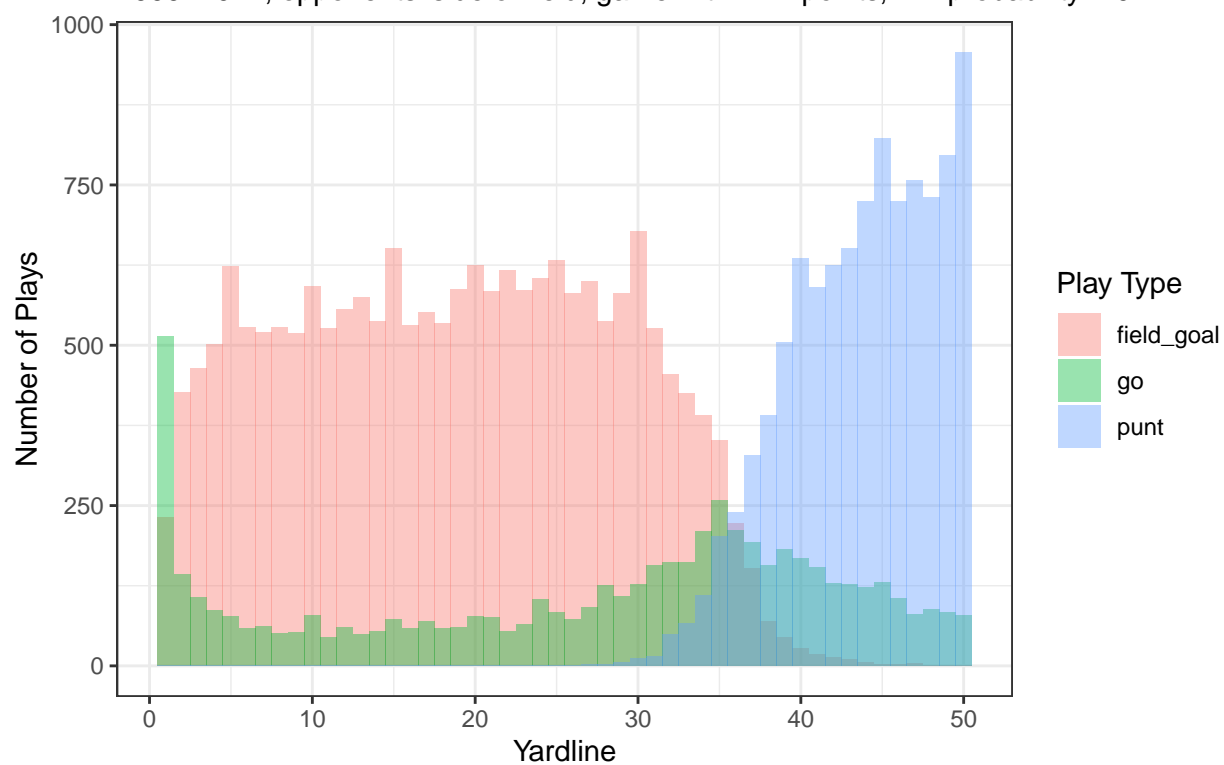
Go Model

Our optimal go for it XG Boost model has a RMSE of 0.5492, which is the lowest RMSE we obtained throughout our XG Boost, Lasso, and Ridge Regression modelling process. Therefore, we will continue with this as our go for it WPA model.
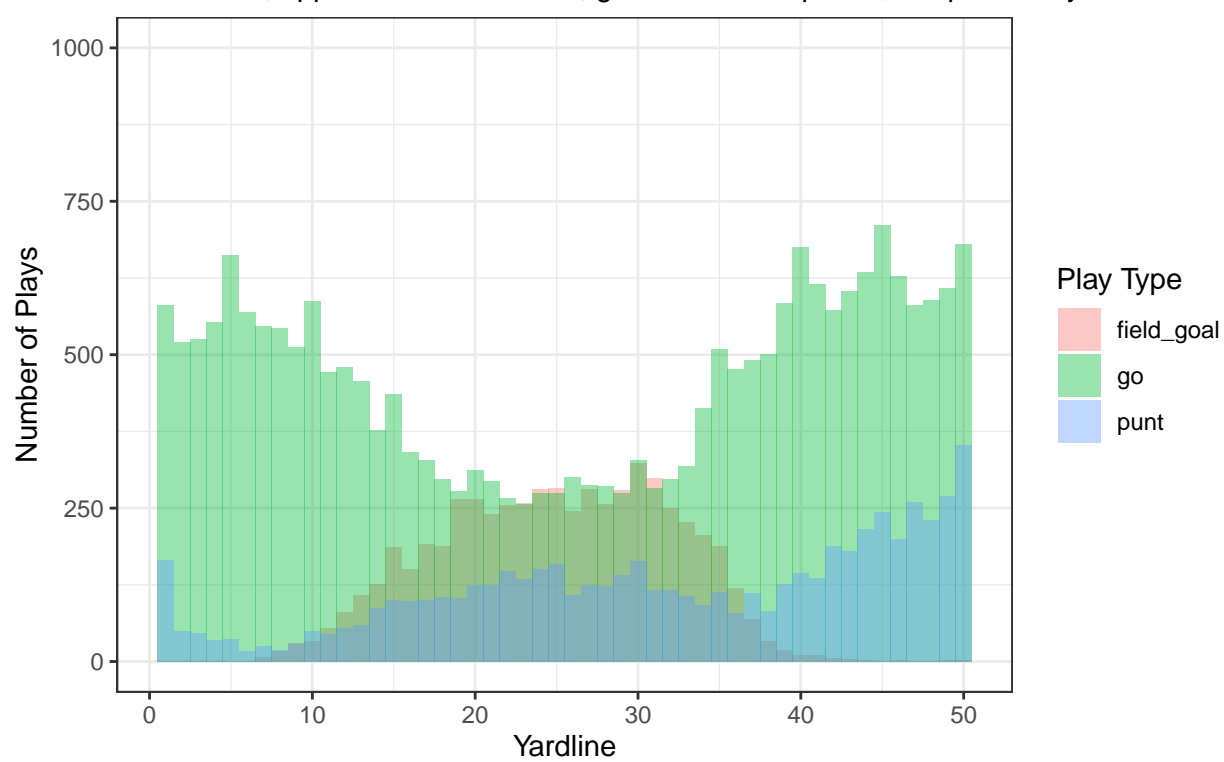
Results

Real Life Decisions

1999–2022, opponents' side of field, game within 14 points, win probability > 0.1
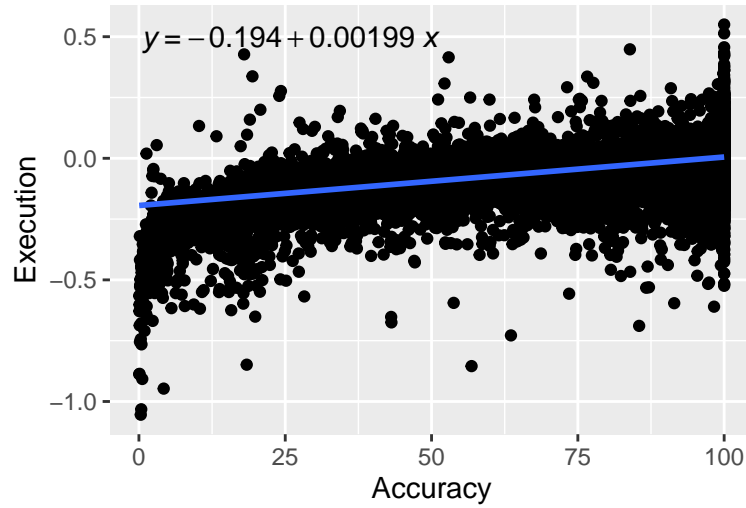

Model Recommended Decisions

1999–2022, opponents' side of field, game within 14 points, win probability > 0.1

	field_goal	go	punt
field_goal	5690	11264	2375
go	127	5225	1131
punt	106	16755	4215

Accuracy vs. Execution of NFL Teams on 4

$y = -0.194 + 0.00199\, x$

Top 5 Executed 4th Downs

ewpa_punt	ewpa_kick	ewpa_go	recommendation	play_type	wpa	execution	accuracy
-0.138	-0.122	-0.018	go	go	0.532	0.550	100.000
0.016	-0.054	0.000	punt	punt	0.529	0.514	100.000
-0.109	-0.297	0.025	go	go	0.481	0.457	100.000
0.002	-0.172	-0.095	punt	go	0.450	0.448	83.898
-0.075	-0.424	-0.016	go	go	0.418	0.434	100.000

Discussion and Conclusion

While our models and process did not turn out exactly how we wanted it to, we still ended up with a working product and learned a lot while making it.