

# UNDERSTANDING NFL Fourth Down Decisions

BRETT LADERMAN



# Table of Contents

<b>Overview</b>	<b>1</b>
<b>Methodology and Analysis</b>	<b>1</b>
Data Used	1
Trends Over Time	1
Team Aggressiveness	2
Fourth Down Success	3
XGBoost Model	4
<b>Results and Recommendations</b>	<b>5</b>
Predictive Metrics	5
Insights	6
<b>Conclusion</b>	<b>7</b>

# UNDERSTANDING NFL FOURTH DOWN DECISIONS

## Overview

This report presents an analysis of NFL fourth down decision-making from 2021 to 2024 using historical play-by-play data. I developed a prediction model to estimate whether a team would go for it on fourth down based on in-game factors. The methodology combines statistical feature selection, domain expertise, and machine learning to identify key drivers of fourth down decisions. The model achieved a prediction accuracy of 0.86, offering strong predictive performance and providing data-driven insights to support decision-making in the NFL.

## Methodology and Analysis

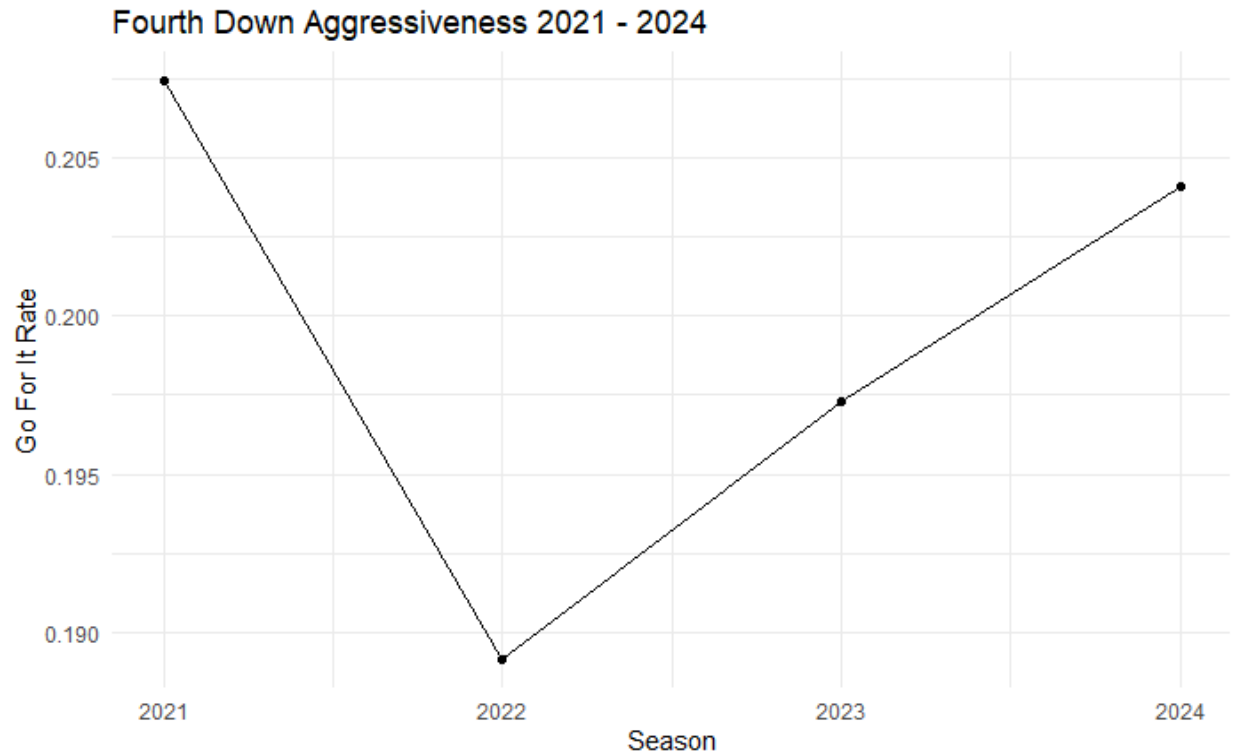
This section outlines the data sources, historical team trends, and model selection process used to analyze NFL fourth down decision-making.

### *Data Used*

I used NFL play-by-play data from the 2021 through 2024 seasons, sourced from the nflfastR package. I filtered the dataset to include only fourth down plays and created a Fourth Down Decision variable indicating whether a team chose to go for it or opted to punt or attempt a field goal.

### *Trends Over Time*

Over the last four NFL seasons, teams have elected to go for it on fourth down between 18.9% and 20.7% of the time league-wide. Figure 1 displays the year-over-year rate at which teams chose to go for it on fourth down across the league.

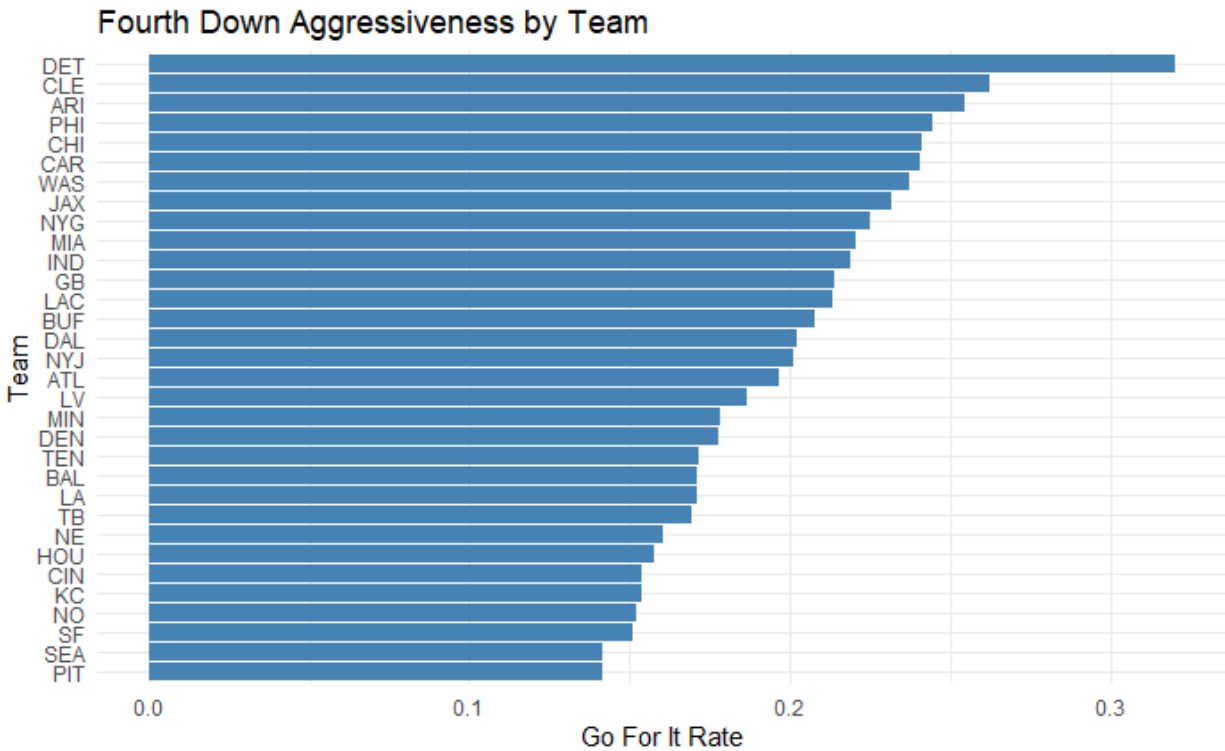


**Figure 1:** NFL Fourth Down Go-For-It Rate by Season

As shown in Figure 1, the go-for-it rate peaked in 2021 and hit its lowest point in 2022, but has steadily increased in each season since.

### *Team Aggressiveness*

Each of the 32 NFL teams has a different level of comfort when it comes to how often they choose to go for it on fourth down. Figure 2 shows each team's go-for-it rate over the past four seasons, ranked in order of frequency.

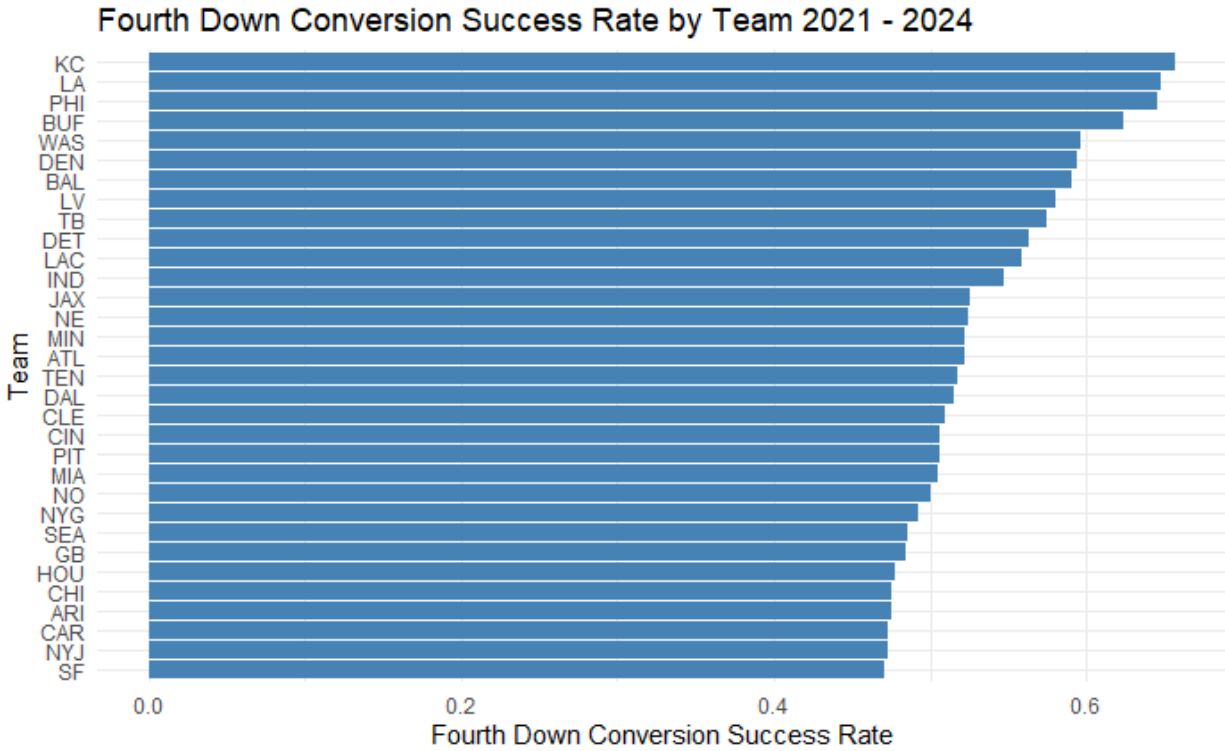


**Figure 2:** NFL Fourth Down Aggressiveness by Team

Figure 2 shows that the Detroit Lions are by far the most aggressive team when it comes to going for it on fourth down, doing so at a 32% rate. They go for it, on average, 6% more often than the next most aggressive team, the Cleveland Browns. Dan Campbell became the Lions' head coach prior to the 2021 NFL season and has developed a strong reputation as the most aggressive head coach in the league. The least aggressive team is the Pittsburgh Steelers, who go for it on fourth down just 14% of the time.

### *Fourth Down Success*

Over the past four NFL seasons, the average success rate on fourth down is 53.5%. Figure 3 highlights each NFL team's fourth down success rate during that span.



**Figure 3:** NFL Fourth Down Conversion Success by Team

Even though Figure 2 showed the Detroit Lions as the most aggressive team in the NFL when it comes to going for it on fourth down, they ranked only 10th in fourth down conversion success. The most successful team at converting fourth downs is the Kansas City Chiefs, with a conversion rate of 65.8%. On the other end, the San Francisco 49ers have been the least successful, converting just 47.1% of their attempts.

### *XGBoost Model*

I built an XGBoost model and performed hyperparameter tuning to predict when an NFL team will choose to go for it on fourth down, based on in-game factors. Table 1 displays the four variables included in my final model.

**Table 1:** Variables Included in Final Fourth Down Prediction Model

VARIABLE	DESCRIPTION
yardline_100	Total yards remaining from the line of scrimmage to the endzone
quarter_seconds_remaining	Total seconds remaining in the quarter
score_differential	Score differential in the game based on the possession team
wp	Estimated win probability for the offensive team at the time of the play

I conducted hyperparameter tuning to maximize the model's predictive power. The optimized model used 17 trees, an eta of 0.3, a gamma of 0, a max depth of 7, and a subsample of 0.9. Any

unmentioned parameters were left at their default values. After fine-tuning the model, we assessed variable importance based on gain. Gain measures the improvement in accuracy before and after a variable is used in the model. Table 2 ranks the variables based on their gain.

**Table 2:** Ranked Variable Importance by Gain from XGBoost Model.

RANK	VARIABLE	DESCRIPTION	GAIN
1	wp	Estimated win probability for the offensive team at the time of the play	0.3753
2	yardline_100	Total yards remaining from the line of scrimmage to the endzone	0.3518
3	quarter_seconds_remaining	Total seconds remaining in the quarter	0.1426
4	score_differential	Score differential in the game based on the possession team	0.1303

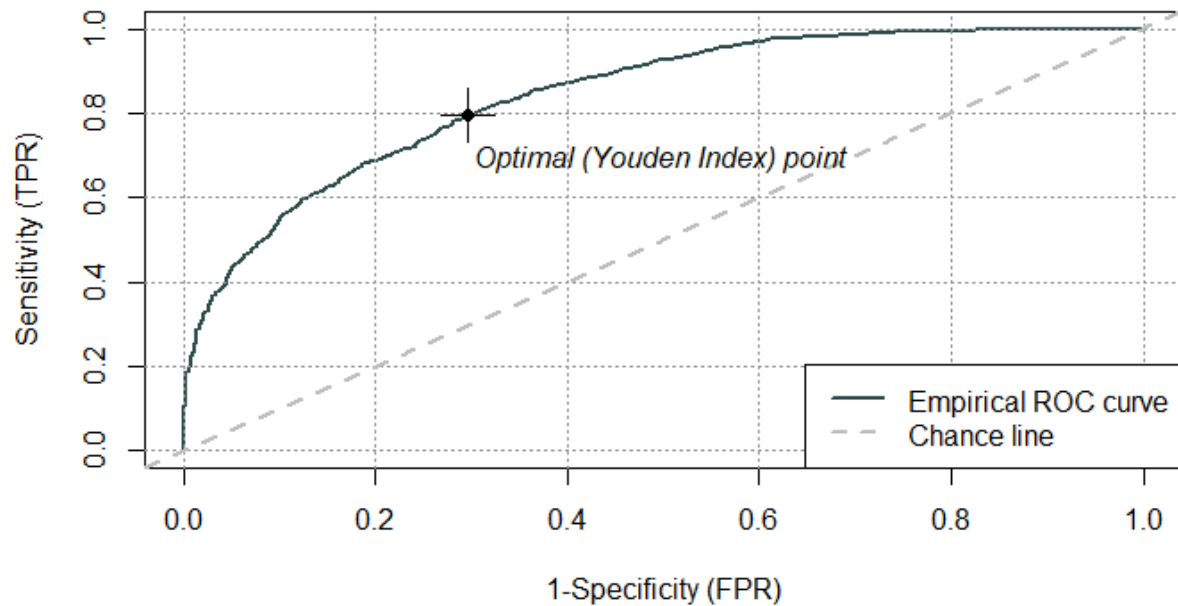
Winning Percentage is the most important variable for improving accuracy. Including this variable in my model has improved its accuracy on average by 37.53%

## Results and Recommendations

In the following sections, I present the results from my XGBoost model, along with insights that are valuable to NFL teams.

### *Predictive Metrics*

I generated the ROC curve using the XGBoost model and identified an optimal cutoff value of 0.785, as shown in Figure 4.



**Figure 4:** ROC Curve Plot of the XGBoost Model

The area under the ROC curve is 0.8414, indicating that our model correctly assigns higher predicted probabilities to plays where the team chooses to go for it 84.14% of the time.

I then generated the confusion matrix for the XGBoost model, as shown in Table 3.

**Table 3:** Confusion Matrix for XGBoost Model.

	Predicted: No	Predicted: Yes
Actual: No	3812	96
Actual: Yes	593	380

The confusion matrix shown in Table 3 shows that my XGBoost model accurately predicts a coach's decision to go for it or not go for it 86% of the time.

### *Insights*

The XGBoost model shows strong predictive performance, correctly identifying fourth down decision-making 86% of the time with an AUC of 0.8414. The most influential variable in the model is win probability, which contributes the most to prediction accuracy, followed closely by field position (yardline\_100). These results suggest that coaching decisions on fourth down are highly influenced by a team's perceived chances of winning and their proximity to the endzone. Time remaining in the quarter and current score differential also play a role but to a lesser extent.



Given these findings, NFL teams may benefit from incorporating real-time win probability models into their in-game decision-making tools. Additionally, emphasizing analytics-driven support during critical fourth down situations—especially when close to midfield or trailing late in the quarter—could lead to more consistent, data-backed choices. Teams looking to improve fourth down strategy should prioritize situational awareness dashboards that highlight win probability and field position as primary indicators.

## **Conclusion**

This analysis demonstrates the power of machine learning in uncovering the key factors that drive fourth down decision-making in the NFL. By leveraging historical play-by-play data from 2021 to 2024 and tuning an XGBoost model with carefully selected in-game variables, I achieved a high prediction accuracy of 86% and an AUC of 0.8414. Win probability and field position emerged as the most influential factors, confirming that coaches weigh both strategic context and situational urgency when making these critical calls. The model not only highlights the effectiveness of data-driven decision tools but also offers actionable insights for teams seeking to refine their fourth down strategy. Integrating such predictive models into coaching workflows could enhance real-time decision-making and provide a competitive edge in high-stakes situations.