

NFL 4th down Analytics

Morgan Abbitt



Contents

1. Goals
2. Introduction / Overview
3. Data
4. EDA
5. Modeling
6. Conclusions

Goals

Predict whether a team is going to **Punt**, kick a **Field Goal**, or **Go for it** on 4th down.

1. Motivation
2. Explore and Understand the data
3. What affects their choice on 4th down?
4. Develop a Model

Introduction / Overview

Does anyone know the different decisions a team has on 4th down?

Introduction / Overview

Does anyone know the different decisions a team has on 4th down?



Punt

Introduction / Overview

Does anyone know the different decisions a team has on 4th down?

Punt

**Field
Goal**

Introduction / Overview

Does anyone know the different decisions a team has on 4th down?

Punt

**Field
Goal**

Go for it

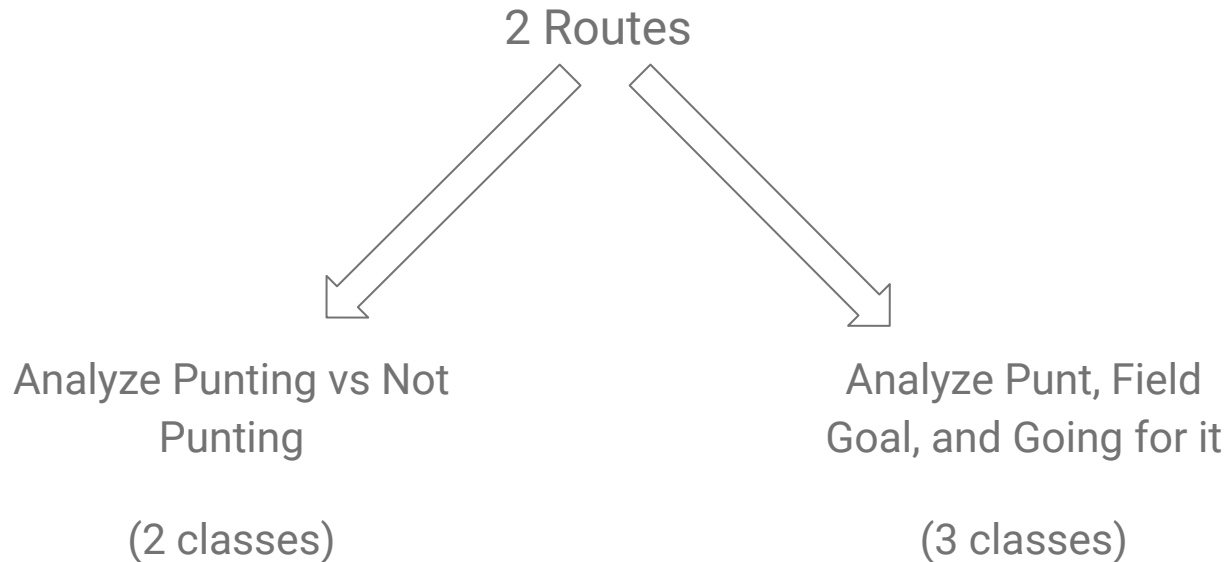
Data

Play-by-play data webscraped from the NFL, provided on Kaggle

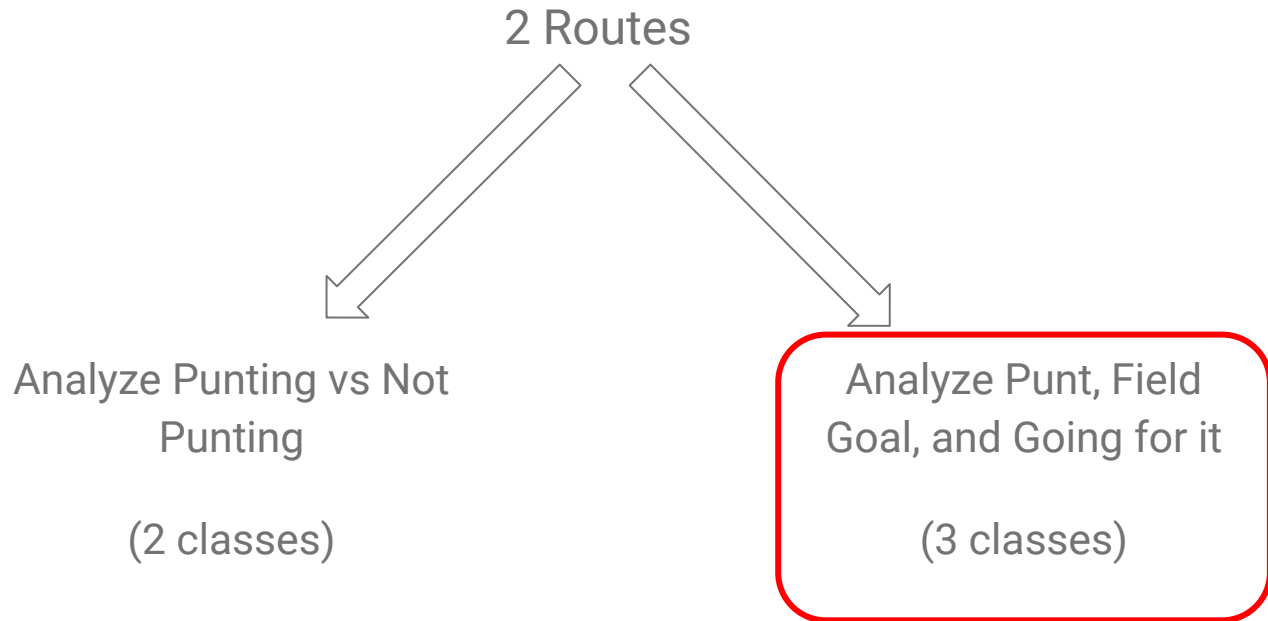
Includes:

- 449371 individual plays
- 255 features
- 2009 - 2018 regular seasons
- 2,526 games
- **37,769 4th down plays**

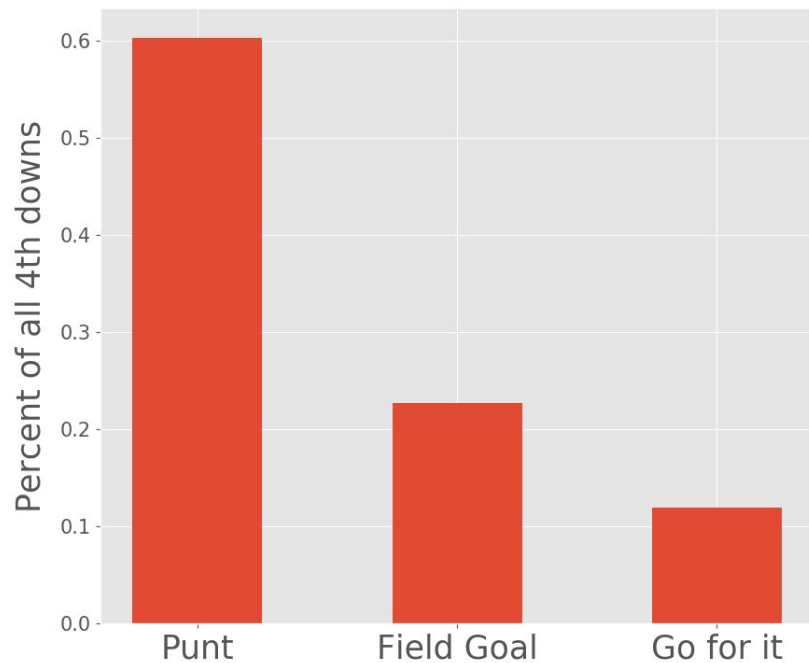
Exploratory Data Analysis



Exploratory Data Analysis

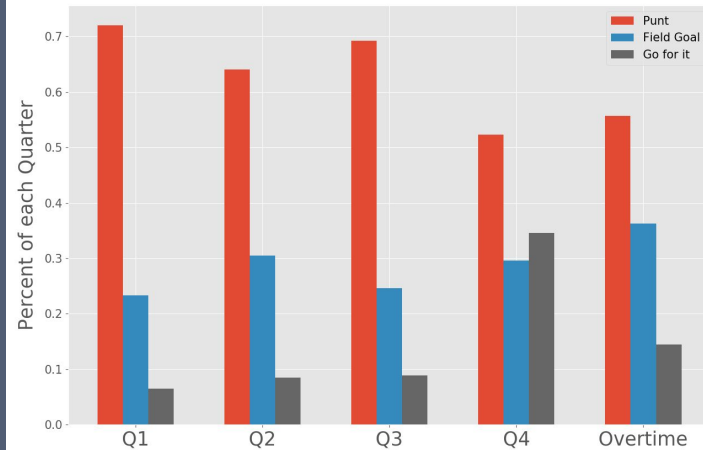


Percent of Plays on 4th down

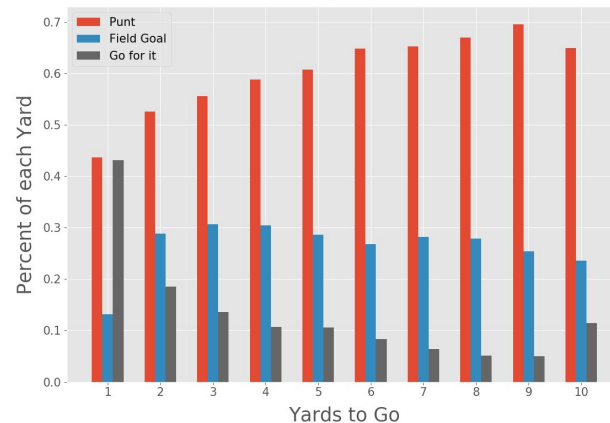


Exploratory Data Analysis on what features (columns) affect the decision on 4th down

Decision by Quarter

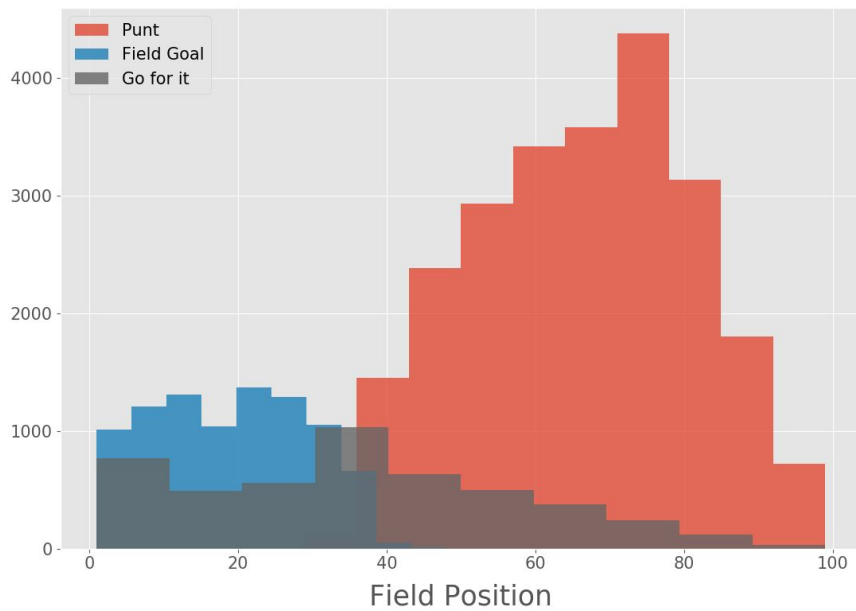


Decision by Yards to Go

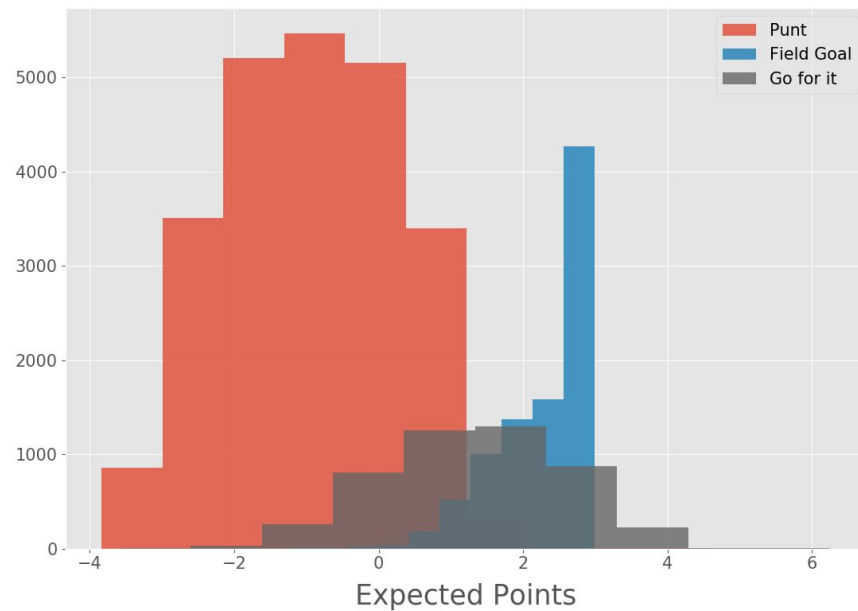


Field Position and Expected Points

Distribution of Decision by Field Position

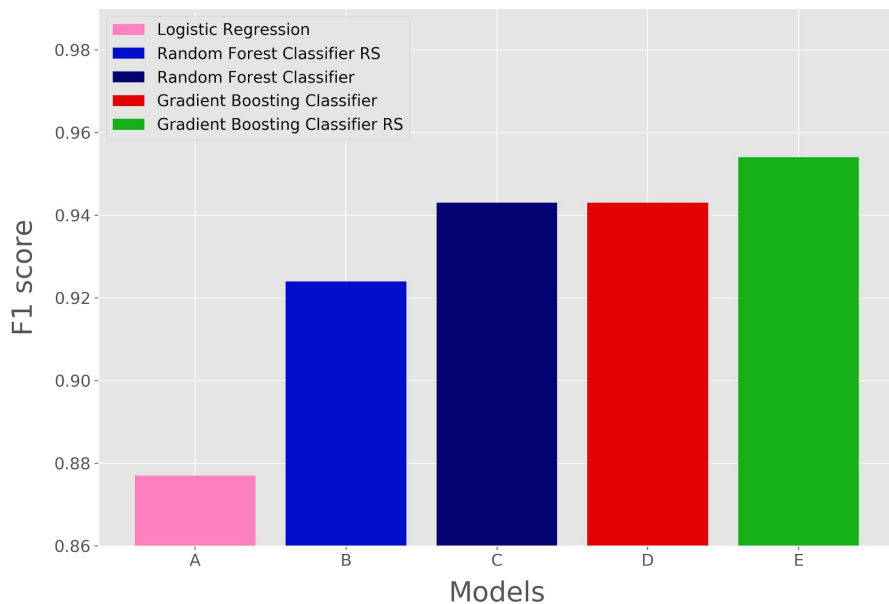


Distribution of Decision by Expected Points



Modeling

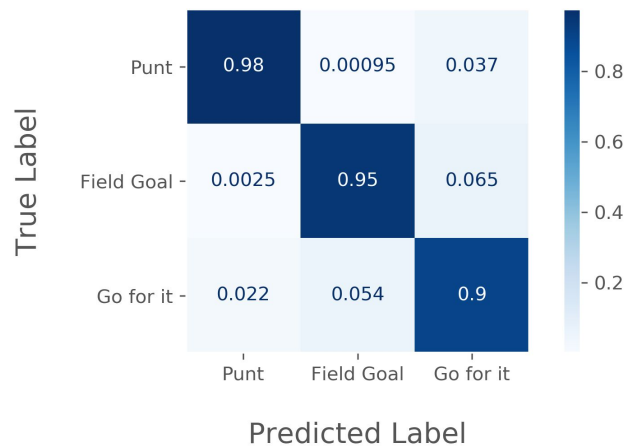
CV Model Selection



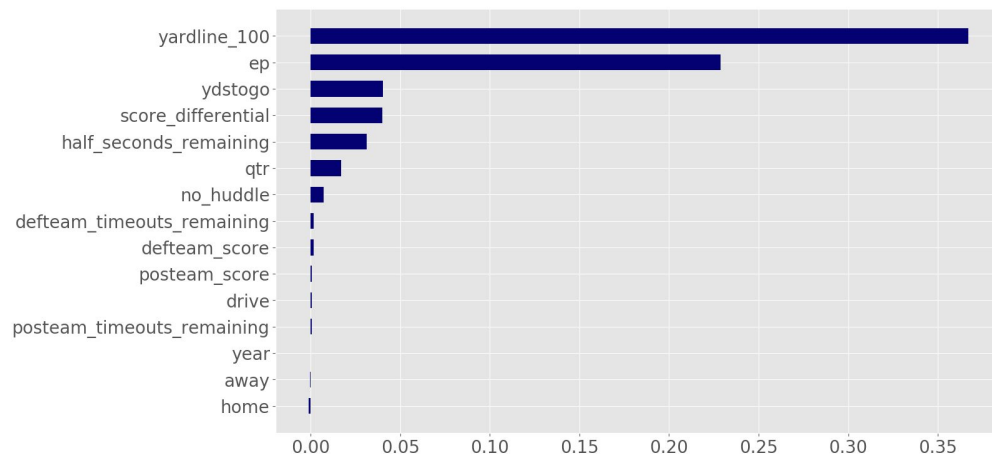
- Multi-Classification
- Imbalanced Classes
- Scoring Metric

Final Model

GradientBoosting RS Confusion Matrix



Permutation Importance of Gradient Boosting RS:
Testing Dataset



Gradient Boosting RS weighted F1 score: 0.944

Conclusions

Coaching Decisions

- Which unit to have out on the field
- Defense on 3rd down
 - Yards to Go
 - Field Position

Education Tool

Football Leagues

