

445TestPresentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

445TestPresentation

Zach Goldstein Sam Rolsten Andrew Davenport Ryan
Lang

2025-12-08

This makes a new slide (one #)

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

- These dashes make bullet points for the slide
 - ① You can also make further indented bulleted points
 - ② They can be numbered
 - Or they can be dashed

You can also use ">" to make italicized text

This makes a new slide (one #)

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

You can also make subheadings in the slides that make almost a “table”

- You can use this to add more organized info
- This is helpful to add tables/graphics that you want to title

Problem Description

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

Data Insight

- Wanted to predict what NFL teams should do on 4th down
- Can either punt, kick a field goal, or go for it
- From those options, if they go for it, can either run or pass the ball
- Selected the **nflfastR** data set
- Filtered down to only 4th down plays from 1999-2025
- Split the data into training set (1999-2019) and test set (2021-2025)

Data Pre-processing

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

Overview of Cleaning & Variable Selection

- Loaded full play-by-play dataset (1999–2025) using **nflfastR**.
- Selected key predictors:
- Game context: season, week, quarter, game clock, score differential
- Field context: yardline_100, ydstogo, posteam/defteam timeouts
- Play descriptors: rush, pass, penalty, EPA/WPA, success
- Created a unified **decision** variable:

Data Pre-processing

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

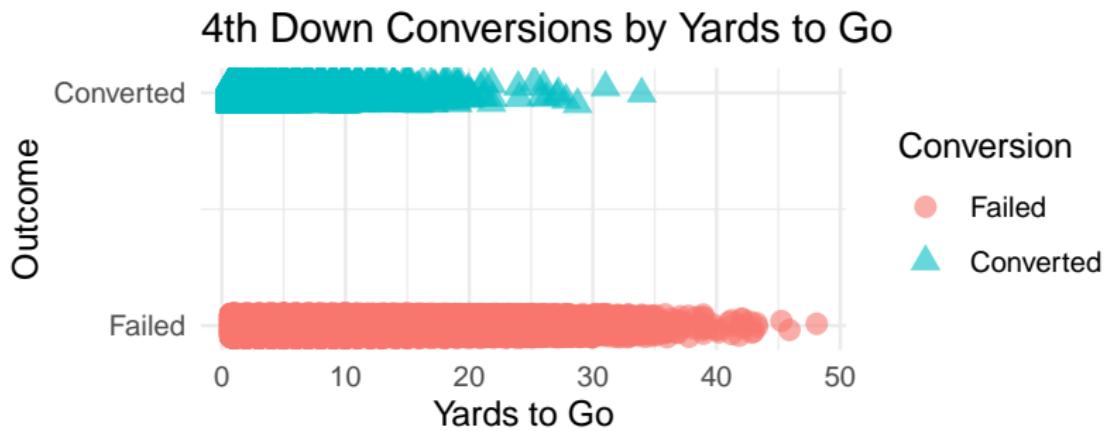
Additional Feature Engineering

- Added intuitive indicators:
 - close_game (within 7 points)
 - short_to_go (<= 3 yards)
 - must_go (trailing, < 5 minutes in 4Q)
- Binned variables to reduce model memorization:
 - yardline_zone: red_zone_own / mid_field / red_zone_opponent
 - ydstogo_bin: short / medium / long / very_long

Conversion Probability by Distance

445 Test Presentation

Zach
Goldstein,
Sam Rolsten,
Andrew
Davenport,
Ryan Lang

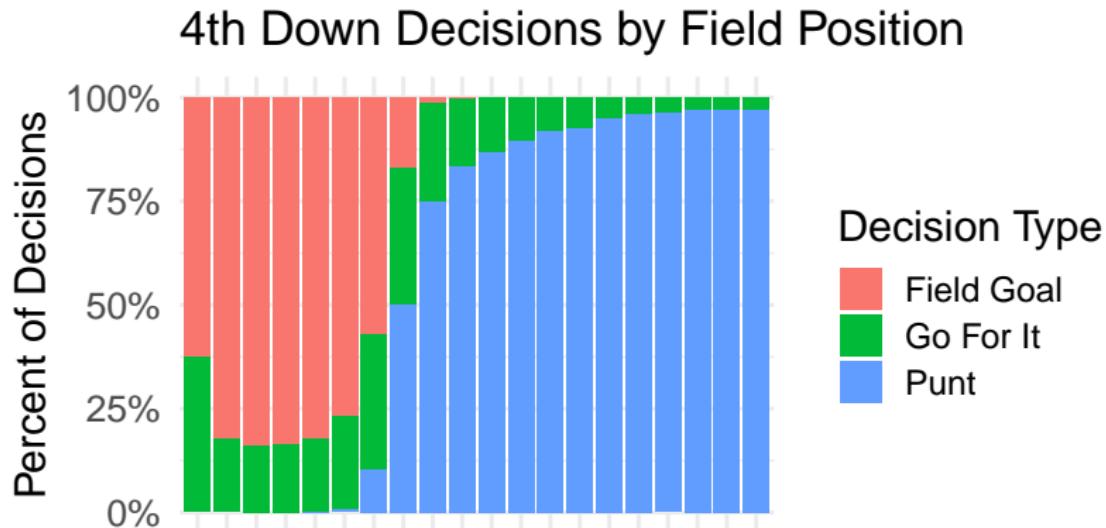


- Conversion probability decreases as distance increases.
- Noticeable separation around 4 yards.

Decision Percentage: Field Position

445 Test Presentation

Zach Goldstein,
Sam Rolsten,
Andrew Davenport,
Ryan Lang

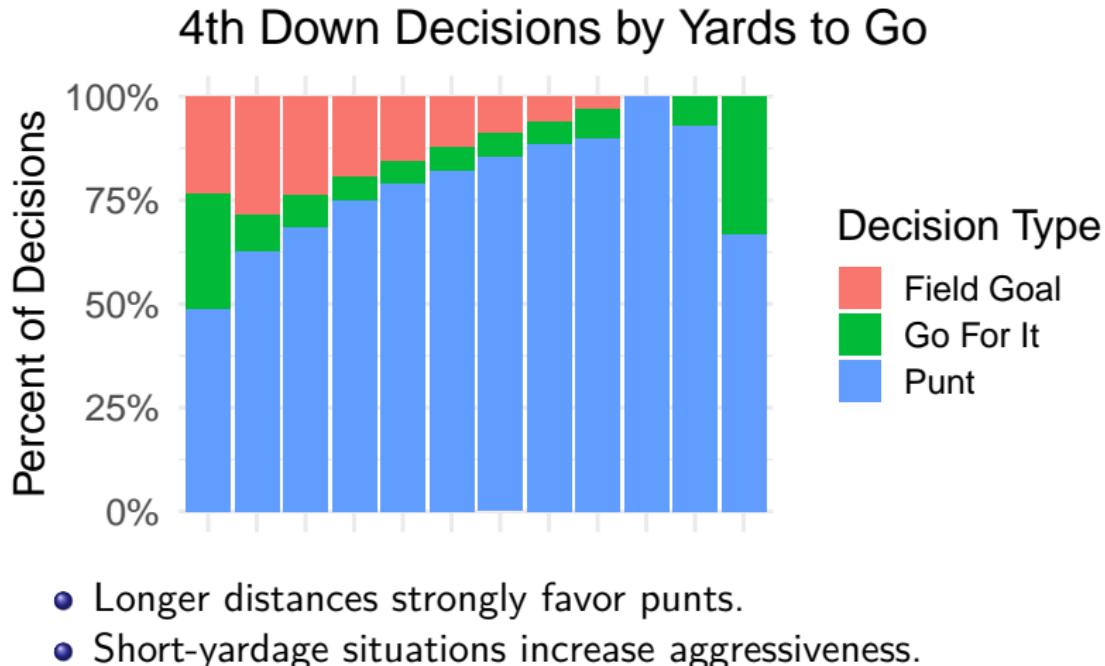


- Punt dominates deep in own territory.
- Go-for-it frequency increases past midfield.

Decision Percentage: Yards to Go

445 Test Presentation

Zach Goldstein,
Sam Rolsten,
Andrew Davenport,
Ryan Lang



Decision Percentage: Score Differential

445 Test Presentation

Zach Goldstein,
Sam Rolsten,
Andrew Davenport,
Ryan Lang



- Trailing teams are more aggressive.
- Leading teams favor safer options.