Modeling Offensive Play Calls and Defensive Reactions

Jayden Cruz Berdecia, Colin Montie,
Nicolas Thomas, Nathan Wright

Overview

Our project focused on modeling offensive and defensive behaviors in NFL plays by leveraging player tracking and situational data provided by the Big Data Bowl.

The ultimate goal was to develop a "best plan of attack" framework—offering predictive insights for teams based on expected in-game scenarios.

We analyzed player movement, alignments, and matchups to capture offensive and defensive tendencies, incorporating game theory to model strategic interactions.

Background

Objective
To develop a robust predictive model for defensive reactions in NFL plays—including the number of rushers, coverage schemes, and post-snap adjustments—based on pre-snap alignment, motion, and game context.

Motivation

- **Tactical Adaptation**: Defenses constantly evolve in response to offensive trends. Anticipating these adjustments equips teams with a strategic edge in play design and in-game decision-making.
- **Unlocking Big Data**: The rise of player tracking technology provides an unprecedented opportunity to model defensive behavior with high granularity.
- **Filling the Gaps**: Many existing approaches simplify defensive logic or rely solely on static alignment. Our work addresses the fluid, reactive nature of NFL defenses under pressure.
- **Team Impact**: Accurate defensive predictions empower teams to optimize play calling, reduce uncertainty, and execute with greater confidence in key situations.

Data Handling

Data Handling

- We integrated multiple data sources to enrich our modeling process:
  - **Tracking Data** (frame-by-frame player movement)
  - **Game Metadata** (contextual information from Big Data Bowl)
  - **Play-by-Play Data** via nfl_data_py (event-level details and game situations)

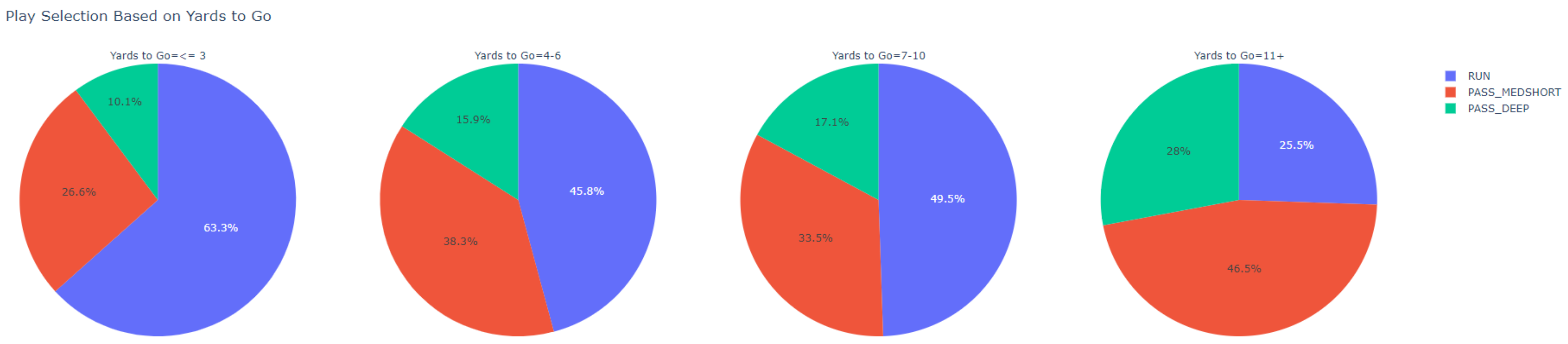
Preprocessing focused on isolating high-value moments in each play. We filtered for critical events like line_set, man_in_motion, and ball_snap to trim retain the most relevant pre- and post-snap actions—laying the foundation for accurate modeling of defensive behavior.

Methodology

- We engineered features like distance from the ball, line of scrimmage, and nearest offensive player to capture spatial context.
- Categorical variables (e.g., personnel, alignment, motion, down & distance) were one-hot encoded for modeling.
- Tracking data was merged with play-level context and outcomes (yards gained, result) to support analysis.
- From there, we built models to capture offensive tendencies and predict defensive adjustments

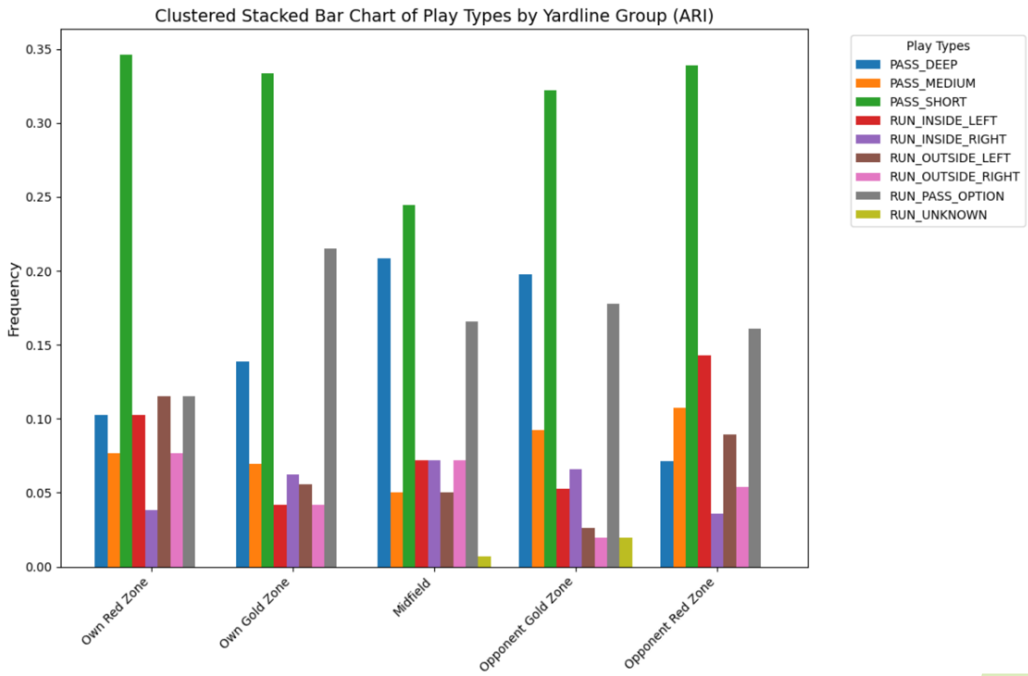
Play Type by Yards to Go

- The following pie charts display leaguewide offensive play-calling tendencies grouped by yards to go
- Plays are sorted into runs, short or medium passes, and deep passes
- Run plays are most common when there are fewer than 11 yards to go while short and medium passes are the most common when not running
- Teams rarely throw deep unless they must



Play Type by Yards to Go (based on downs)

- These pie charts once again depict leaguewide play-calling by yardage to go, however they are also grouped by down
- An interesting takeaway from these is the lack of deep passes on 1st and short situations
- The idea of taking a shot on 2nd and short can be seen with 8% of plays on 2nd and <=3 being deep passes and 11% on 2nd and 4-6
- This idea has not carried over to 1st downs despite having an extra down to work with, leading to teams potentially missing out on big plays



Offensive Modeling

- We started by classifying team tendencies across play types: Inside/Outside Runs (Left/Right), RPOs, and Passes (Short, Intermediate, Deep), benchmarking each team against league averages.
- Plays with undefined or irrelevant types (e.g., QB kneels) were removed to reduce noise.
- We then modeled expected play type probabilities based on pre-snap and snap-time alignment, incorporating contextual features like motion and formation shifts.
- Scouting visuals were created from features including:
 Personnel, Field Position, Game State (Time/Down/Distance), Scoring Probability, Drive Yards, and Custom RB Metrics.

Offensive Play Prediction

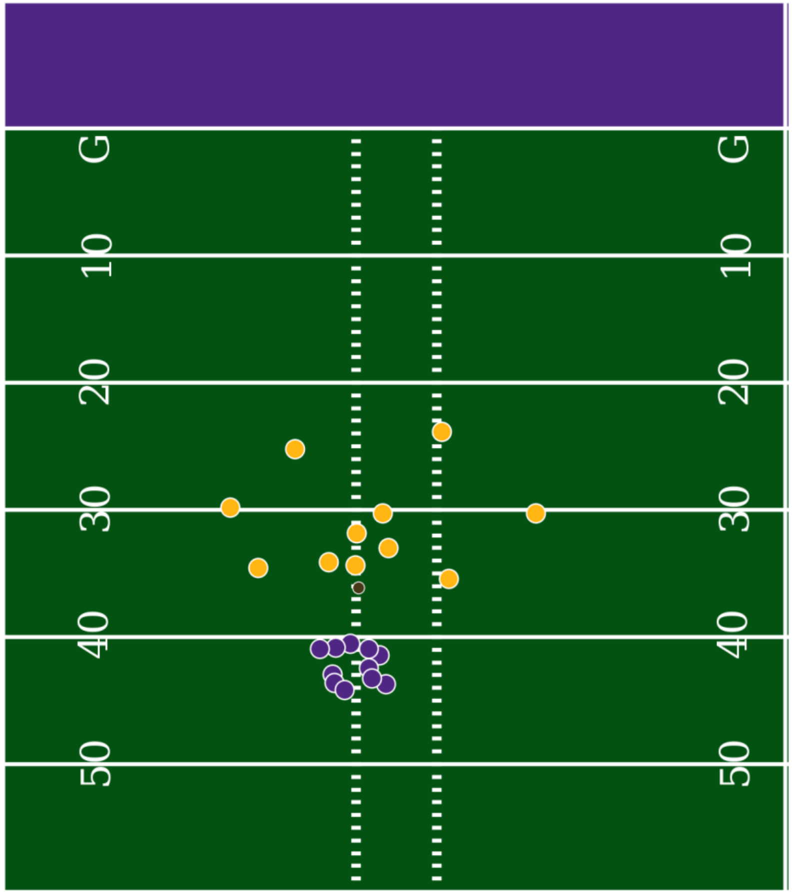
This example highlights our model's prediction of the play type on a 36-yard touchdown pass from Kirk Cousins to Justin Jefferson.

The play took place on 1st down from the opponent's 36-yard line with 45 seconds left in the half.

The Vikings initially aligned in a 3x1 singleback set, then motioned the tight end across to create a balanced 2x2 look.

A second motion to the left occurred just before the snap

Given these factors along with the other variables taken into consideration, our model correctly predicted this play to be a **deep pass**



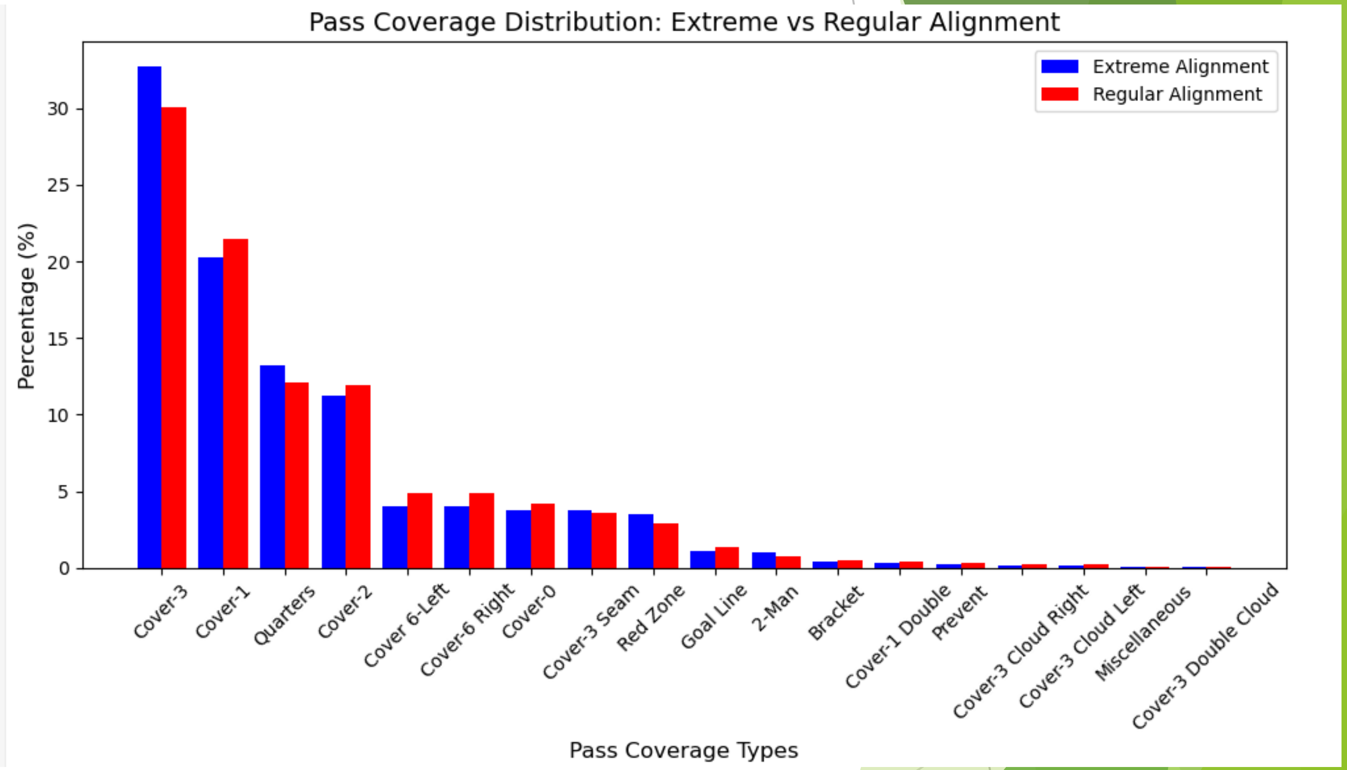
Defensive Modeling

Coverage	Precision	Recall	F1-score	Support
2-Man	1.000000	1.000000	1.000000	2608
Bracket	1.000000	1.000000	1.000000	968
Cover 6-Left	1.000000	0.944865	0.971651	10447
Cover-0	0.999324	1.000000	0.999662	8873
Cover-1	0.867812	0.763583	0.812368	49015
Cover-1 Double	1.000000	1.000000	1.000000	762
Cover-2	0.956246	0.697107	0.806368	27683
Cover-3	0.729967	0.952746	0.826609	79062
Cover-3 Cloud Left	1.000000	1.000000	1.000000	426
Cover-3 Cloud Right	1.000000	1.000000	1.000000	462
Cover-3 Double Cloud	1.000000	1.000000	1.000000	134
Cover-3 Seam	0.998705	0.969611	0.983943	9543
Cover-6 Right	0.997220	0.933758	0.964446	10371
Goal Line	1.000000	1.000000	1.000000	2305
Misc.	1.000000	1.000000	1.000000	141
Prevent	1.000000	1.000000	1.000000	523
Quarters	0.935546	0.735548	0.823579	31692
Red Zone	0.997851	1.000000	0.998924	7892
Accuracy	0.852813	0.852813	0.852813	0.852813

- We began by identifying pre-snap defensive alignment using each player's location (X, Y), Speed, Acceleration, Orientation, Direction, and Distance Traveled.
- Next, we analyzed how defenses respond to expected offensive actions—focusing on number of rushers, coverage type (man/zone), and overall scheme.
- We built scouting reports showing how defenses react to specific formations, alignments, and personnel, based on historical tendencies.
- By merging defensive positioning with offensive motion, we built a dataset to train random forest models that predict coverage schemes based on pre-snap indicators.

Player Alignment

•Wide alignments and motion play a key role in shaping defensive responsibilities and coverage reactions.
•To isolate impactful cases, we flagged plays where alignment metrics exceeded the 80th percentile—capturing extreme formation and motion scenarios.
•Key features included:
  o**Player X/Y location**
  o**Speed & Acceleration**
  o**Distance Traveled**
  o**Orientation & Direction**
•These variables became critical in identifying alignment-driven shifts in defensive coverage.



Model Evaluation: Alignment

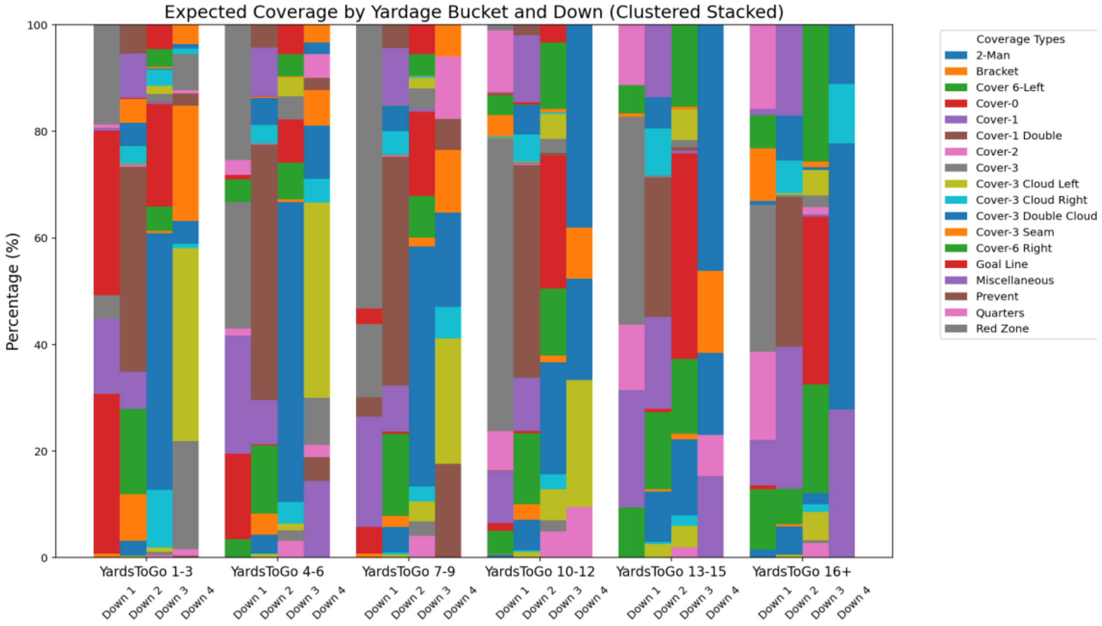
- In our research, we developed models to assess the usefulness and accuracy of player alignment features in predicting coverage types. These models incorporated key factors such as down, yards to go, and motion, with a focus on two primary variables: "pff_passCoverage" (Coverage Types) and "pff_manZone" (Man vs Zone Coverage).
- To further analyze the data, we separated the models to test the coverage variables individually. We also explored different model configurations, comparing results with and without motion, to determine the impact of motion on coverage predictions.
- All models achieved an accuracy of 84% or higher in predicting coverage, with a particular strength in identifying Man vs Zone responsibilities. Incorporating motion into the model consistently enhanced prediction accuracy for both coverage variables.

Model Accuracy: 0.9256946612550734

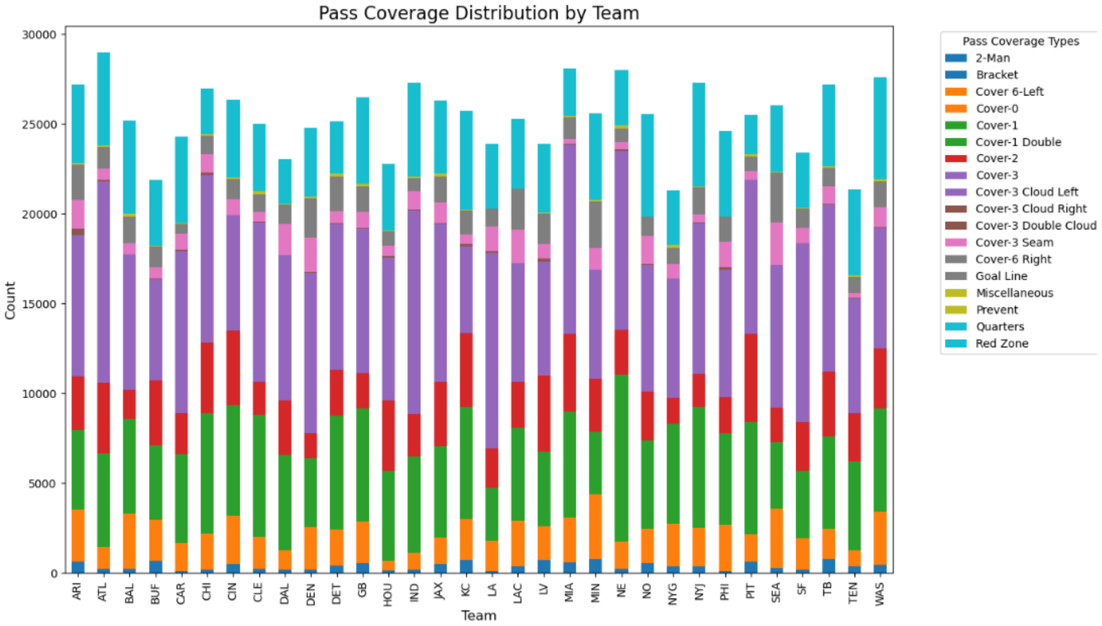
	precision	recall	f1-score	support
Man	0.97	0.74	0.84	13015
Other	0.98	0.93	0.96	2786
Zone	0.91	0.99	0.95	35447
accuracy			0.93	51248
macro avg	0.95	0.89	0.92	51248
weighted avg	0.93	0.93	0.92	51248

Model Accuracy: 0.8644630034342804

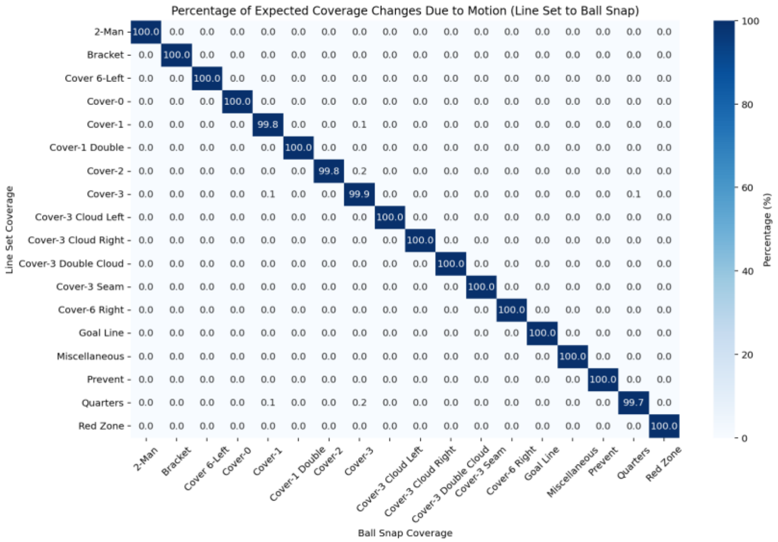
	precision	recall	f1-score	support
2-Man	0.97	0.79	0.87	570
Bracket	1.00	0.88	0.94	236
Cover 6-Left	0.95	0.71	0.82	2020
Cover-0	0.97	0.88	0.92	1996
Cover-1	0.89	0.87	0.88	10263
Cover-1 Double	0.99	0.85	0.92	186
Cover-2	0.90	0.80	0.85	6131
Cover-3	0.78	0.97	0.86	16204
Cover-3 Cloud Left	0.99	0.70	0.82	97
Cover-3 Cloud Right	0.96	0.66	0.78	125
Cover-3 Double Cloud	1.00	0.74	0.85	38
Cover-3 Seam	0.94	0.72	0.82	1967
Cover-6 Right	0.97	0.71	0.82	2128
Goal Line	0.99	0.92	0.95	454
Miscellaneous	0.94	0.85	0.89	34
Prevent	0.99	0.94	0.96	157
Quarters	0.91	0.78	0.84	6737
Red Zone	0.97	0.97	0.97	1905
accuracy			0.86	51248
macro avg	0.95	0.82	0.88	51248
weighted avg	0.88	0.86	0.86	51248

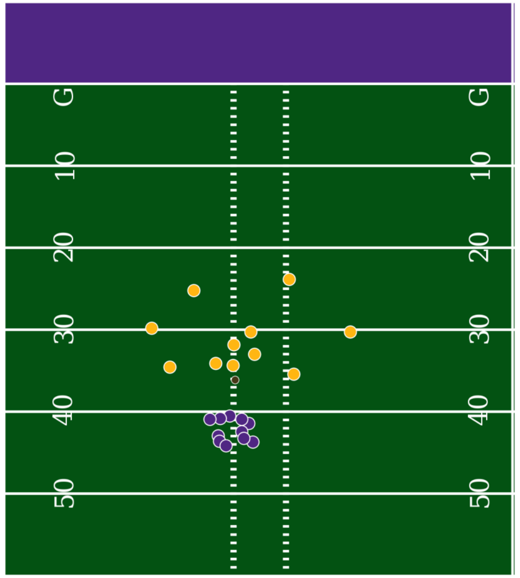


Expected Coverage Scheme by Down and Yardage Bucket



Coverage Distribution Tendencies by Team





Defensive Play Prediction

- Using the same example play, we observe that the defense initially lines up in a 2-safety look when the offense is set.
- However, after the offensive shift, one safety rotates down towards the line of scrimmage, creating a single-safety look.
- Given the expectation of a deep pass and considering other situational factors, our model accurately predicted the defense would run a Cover-3 coverage.
- Our analysis revealed that a defensive team's pre-snap shifts have minimal impact on the coverage they are likely to run.

Finalizing Research/Future Expansion

- Throughout the project, we performed Exploratory Data Analysis (EDA) to validate feature correlations by modeling NFL offensive and defensive tendencies.
- We explored a variety of features and variables, with the offensive focus on play-call decisions influenced by external factors, and the defensive focus on player alignment and how teams adjust their coverage schemes.
- Future research into offensive and defensive tendencies can be significantly expanded by incorporating additional years of data. As more games are played, our dataset will continue to grow, offering more opportunities for analysis.
- A potential future development is a machine learning model that takes user input and calculates the likelihood of specific outcomes for a given play. This would provide real-time insights, helping teams make data-driven decisions by determining the best course of action based on the current game situation.

Appendix

Link to Google Colab Notebook with code:

https://colab.research.google.com/drive/1auSNHL-PVsUHyNcSrBVrbFG_mC06bvRm?usp=sharing

Thank You!