

NFL Offensive Play Calling using Prescriptive Trees

Jonathan Chan
Joe Zaghrini

Problem Statement

- American Football is a complex sport with many facets to decision-making. Coaches must decide when to run, when to pass, and when to kick.
- Poor decisions can be extremely costly: two recent Super Bowls have been lost due to questionable play calls
- Play calls are incredibly complex decisions from an “input” standpoint: coaches have many different sources of information to consider when selecting a play
- The complexity and time-sensitivity (40 second play clock) decisions result in most decisions being made on “intuition,” which is subject to human biases
- Coaches often employ a greedy as opposed to global decision-making, choosing the best choice for a play, not for the drive or the game

American Football - Overview

- Scoring:**
 - A **Touchdown (TD)** is when a player on the offensive team runs to the End Zone (at the end of the field). This is worth 6 points. After a touchdown, a team can kick a short field goal for one point (Extra Point Attempt) or go for another touchdown for another 2 points (Two-Point Conversion). In the vast majority of cases, NFL teams opt for the extra point.
 - A **Field Goal (FG)** is when a team opts to kick the ball between the vertical yellow bars. It is worth 3 points
 - A **Safety** is when the offensive team ends a drive in the End Zone, resulting in 2 points for the defense
 - A “**Pick Six**” is when the defense intercepts a pass and returns it for a touchdown. It is worth 7 points for the defensive team
 - No Score: This is possible if neither team has any of the scoring plays above
- A **drive** begins when the offense starts with the ball, and ends with a Touchdown, Field Goal, Punt, or a Turnover
- The start of each drive is a 1st Down. The team then has four attempts to gain 10 yards. If they gain 10 yards within these attempts, they get another 1st Down. This continues until they fail to get to a 1st Down, kick the ball (on 4th Down), or score a touchdown.

Data

Two Sources:

- Kaggle: Play by Play between 2009 and 2018 with 250+ features for over 450k plays
 - Features selected: game_id, yardline_100, down, ydstogo, goal_to_go, drive, passer_name, year
 - Decision variable: play_type
- Sports Reference: Quarterback Ratings (QBR) between 2014 and 2018
 - Features selected: Quarterback Rating (QBR) and Quarterback Name

Approach & Methods

Simplifications:

- Extra Point Conversions: Converted all touchdowns to 7 points as this is the most likely scenario
- Time Considerations: All plays within four minutes of halves were removed to avoid special scenarios with time remaining, score differential and timeouts remaining
- Penalties, QB Kneels, QB Spike and Third Down Kicks: Omitted as they are special cases

Models:

- Optimal Classification Tree (OCT): Used to determine and value intermediate plays, that do not terminate a drive
- Optimal Prescriptive Tree (OPT): Used to recommend decisions for play calling based on data, cost functions, and OCT valuations

Cost Functions:

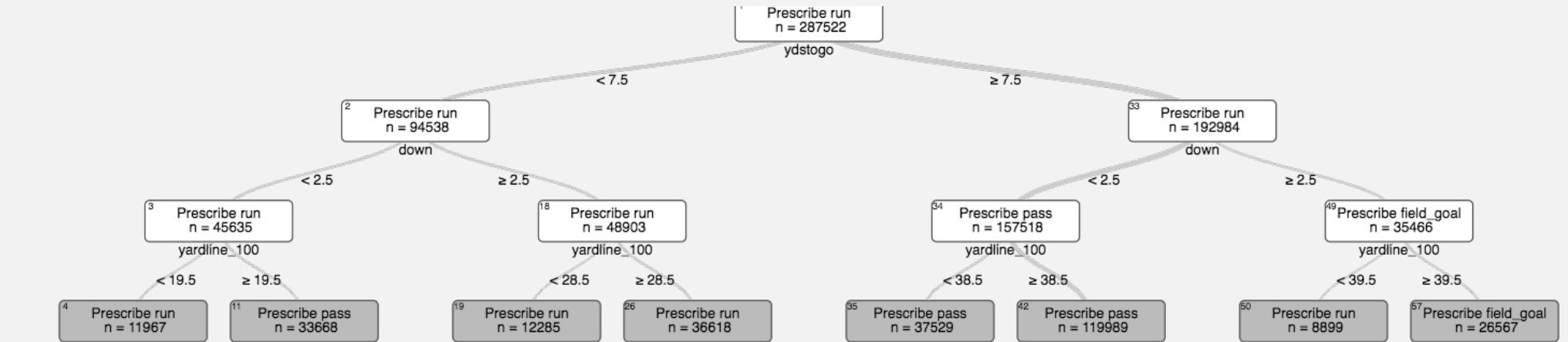
	Touchdown	Field Goal	No Score	Safety	Pick Six
Base	7	3	0	-2	-7
Down by Four	7	0	0	0	0
Any Score	1	1	0	0	0
Down by Three	0.9	0.5	0	0	0

Good vs. Bad Quarterbacks:

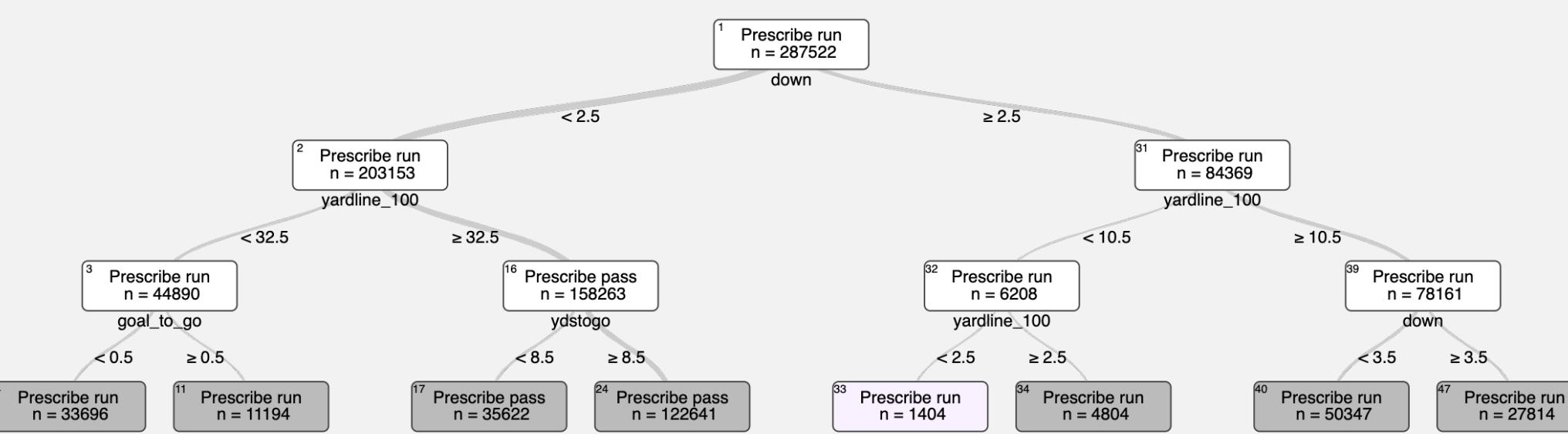
- Top and bottom 10 QBs for each year were identified by QBR with at least 100 pass attempts
- Plays separated with 26,000 and 12,000 rows in each group to see if model prescribes different strategy

Results

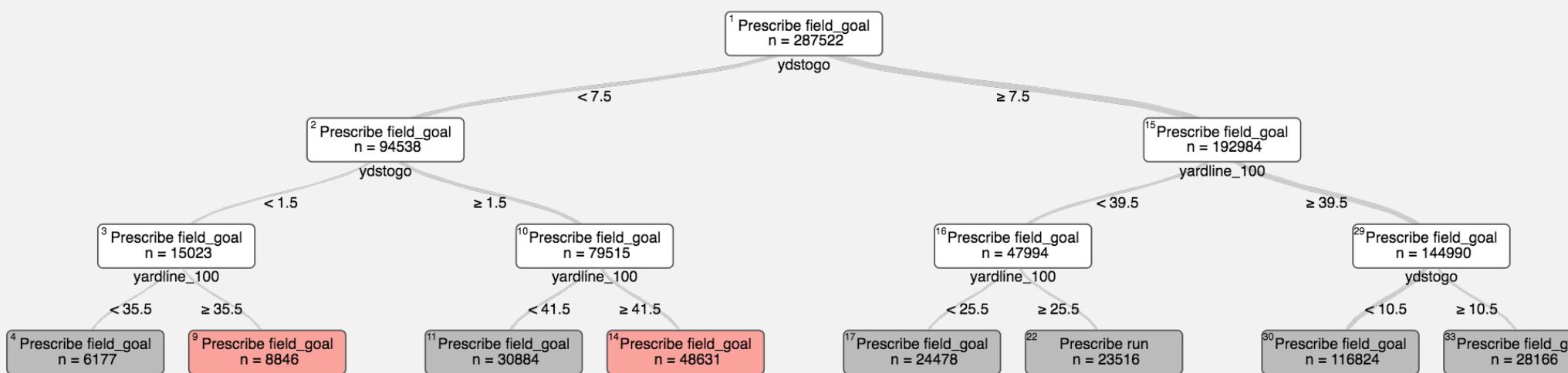
OPT: Base Scenario (Overall)



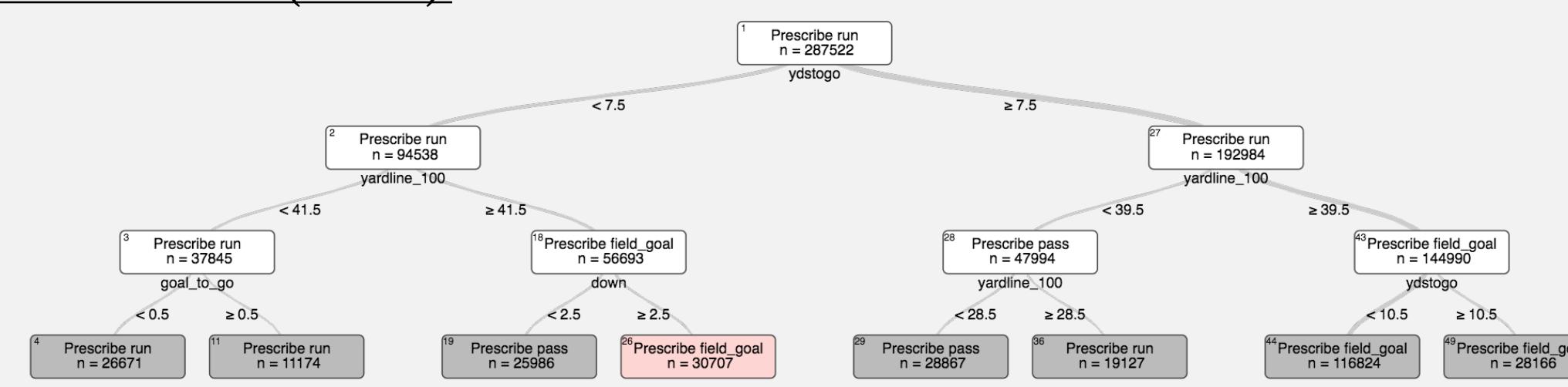
OPT: "Down 4" (Overall):



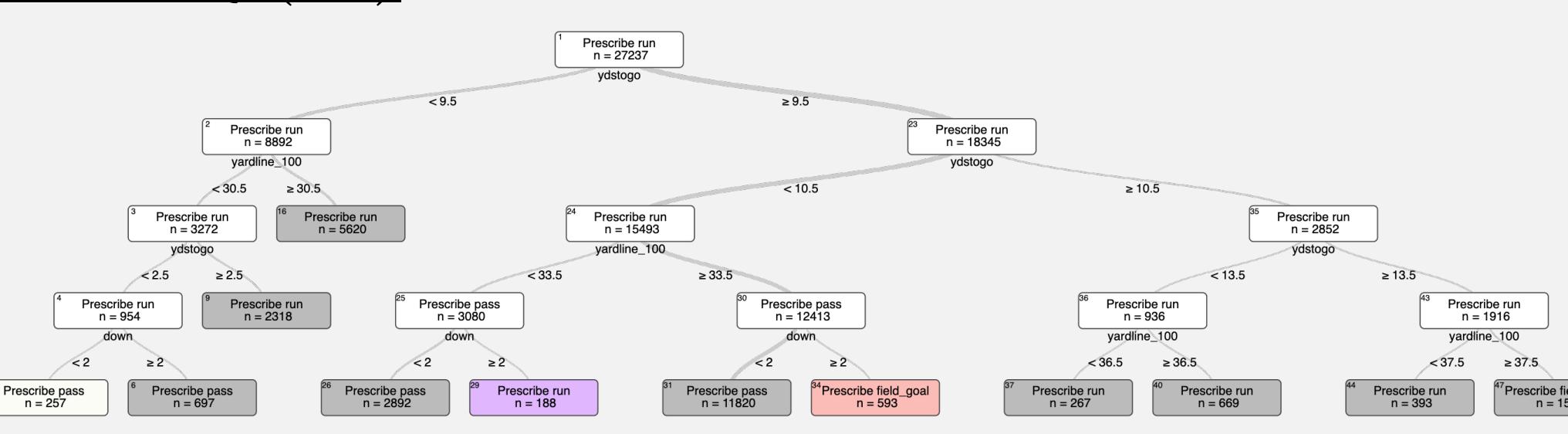
OPT: Any Score (Overall):



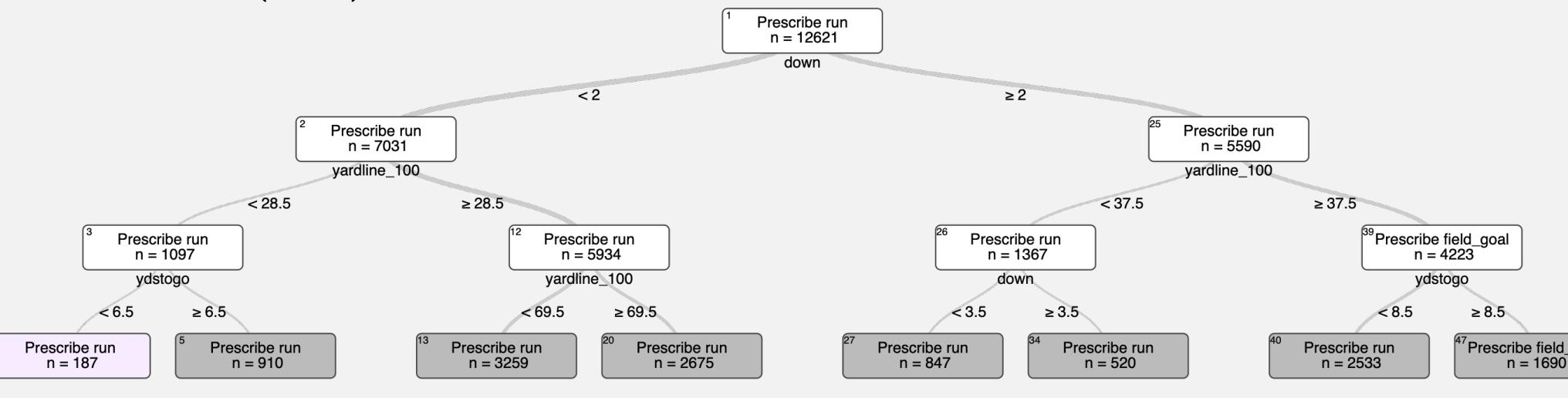
OPT: "Down 3" (Overall):



OPT: Good QB (Base):



OPT: Bad QB (Base):



Extensions

Optimal Classification Tree

- Improve accuracy: Current misclassification error of 0.65 can be improved by feature engineering and varying depth of the tree
- In-game and in-season performance metrics:
 - These can capture the effects of weather and “off-days” of the game or a player’s overall ability such as age and team
 - Quarterback: Quarterback Rating (QBR) was used, however, more metrics can be used as this is football’s most important position
 - Running Back: Best candidate since running the ball to a specific running back is a direct play call
 - Punter / Kicker: Next best candidate since kicker performance changes team’s strategy
 - Receivers: Practicality is limited since a receiver may be targeted but the ball may not be thrown to this receiver due to good defense (eg. double coverage)
- Style of Play:
 - Teams have moved from a “run-first” league to a “pass-first” league in last 5 years and opportunity exists to model based on a modern NFL
- Dynamic Programming Formulation: Replacing OCT for value function. Begin with “terminal” plays, valuing based on optimal prescription, and “walking back” to beginning of drive.

Optimal Classification Tree & Optimal Prescriptive Tree

- Objective function: Maximize likelihood of winning instead of scoring (more complex)
- Time control: Time remaining, timeouts and score differential would better simulate the practicality of play calling, but is highly subjective
- Defensive play calling: Model currently ignores defensive strategies which if known, could exploit opponent strengths and weaknesses
- Punting: Never prescribed in this model since the expected cost is ~0 and any other offensive play will be better in comparison. This can almost be viewed as a defensive play for better field position to mitigate expected points scored by opponent

Conclusion

- OPT can be a pragmatic tool for real-time decision-making
 - Practicability – running real-time models is feasible within the time constraints of a play clock
 - Interpretability – Tree structure means the decision is verifiable and defensible, and conforms to conventional wisdom
- Cost Function Effect: Reasonable to engineer cost functions to simulate strategic shifts in real-game situations
 - Example: “Any Score” OPT recommends kicking a field goal in most situations
- Reasonable and justifiable results:
 - OPT target field goal range of <37 yards matches standard beliefs in the NFL
 - Recommends more running for below-average Quarterbacks, more passing for elite Quarterbacks
 - Concentration of pass recommendations on 1st down or long yards to go
- Goes beyond a greedy approach, especially in specific situations, particularly by establishing a numerical value for non-scoring outcomes
- Interesting insights:
 - Good Quarterbacks, on average, increase a drive’s value by a point
- Appreciation for real-life play calling
 - Enormous amounts of complexity, only some of which can be captured in a single model

