THE MANY MODELS IN PRODUCTION @ NFL NEXT GEN STATS

MIKE BAND
SR. MGR, RESEARCH & ANALYTICS
NFL NEXT GEN STATS

The Many Models in Production at NFL Next Gen Stats

A brief history of five seasons of machine learning work applied to player tracking data in professional football.



I What is NFL Next Gen Stats?

II The Evolution of the NGS Toolbox

III The Many Machine Learning Models in PROD

What is Next Gen Stats?



Next Gen Stats captures NFL player tracking data in real time.

ZEBRA

RFID Tags

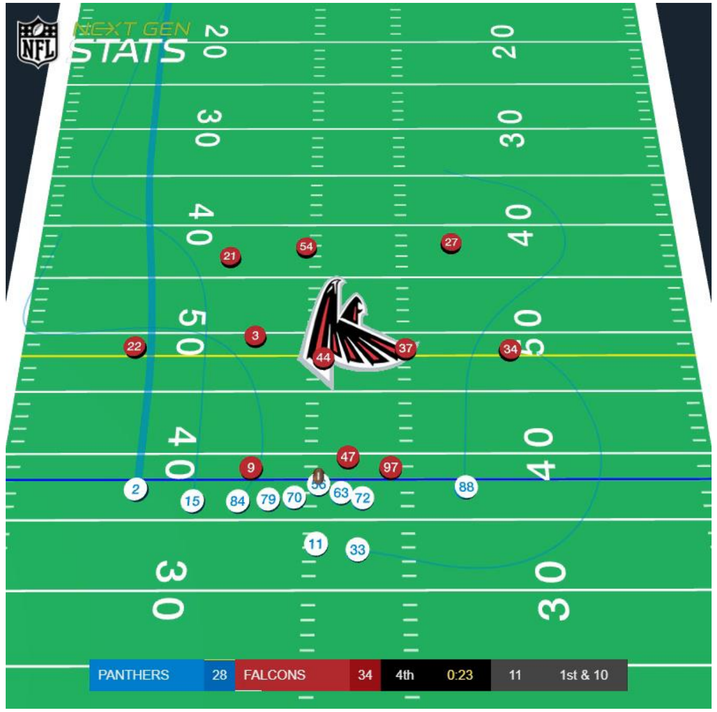


NFL player tracking, also known as **Next Gen Stats**, is the capture of real time location data, speed, and acceleration for every player on every play for every inch of the field. Sensors throughout the stadium track tags placed on players' shoulder pads, charting individual movements measured within inches.

Next Gen Stats currently creates **over 200 different stats** each play in real-time and post play used live during NFL broadcasts, by coaches, scouts and front office executives to make decisions, and the league itself to make the game better, both competitively and in the interest of player health and safety.

NFL NEXT GEN STATS

Next Gen Stats captures NFL player tracking data in real time.



NFL player tracking, also known as **Next Gen Stats**, is the capture of real time location data, speed, and acceleration for every player on every play for every inch of the field. Sensors throughout the stadium track tags placed on players' shoulder pads, charting individual movements measured within inches.

Next Gen Stats currently creates **over 200 different stats** each play in real-time and post play used live during NFL broadcasts, by coaches, scouts and front office executives to make decisions, and the league itself to make the game better, both competitively and in the interest of player health and safety.

The Primary Users of NFL Player Tracking Data

Media & Fan Engagement

All 32 NFL Clubs



The NFL League Office

"Storytelling Beyond the Box Score"

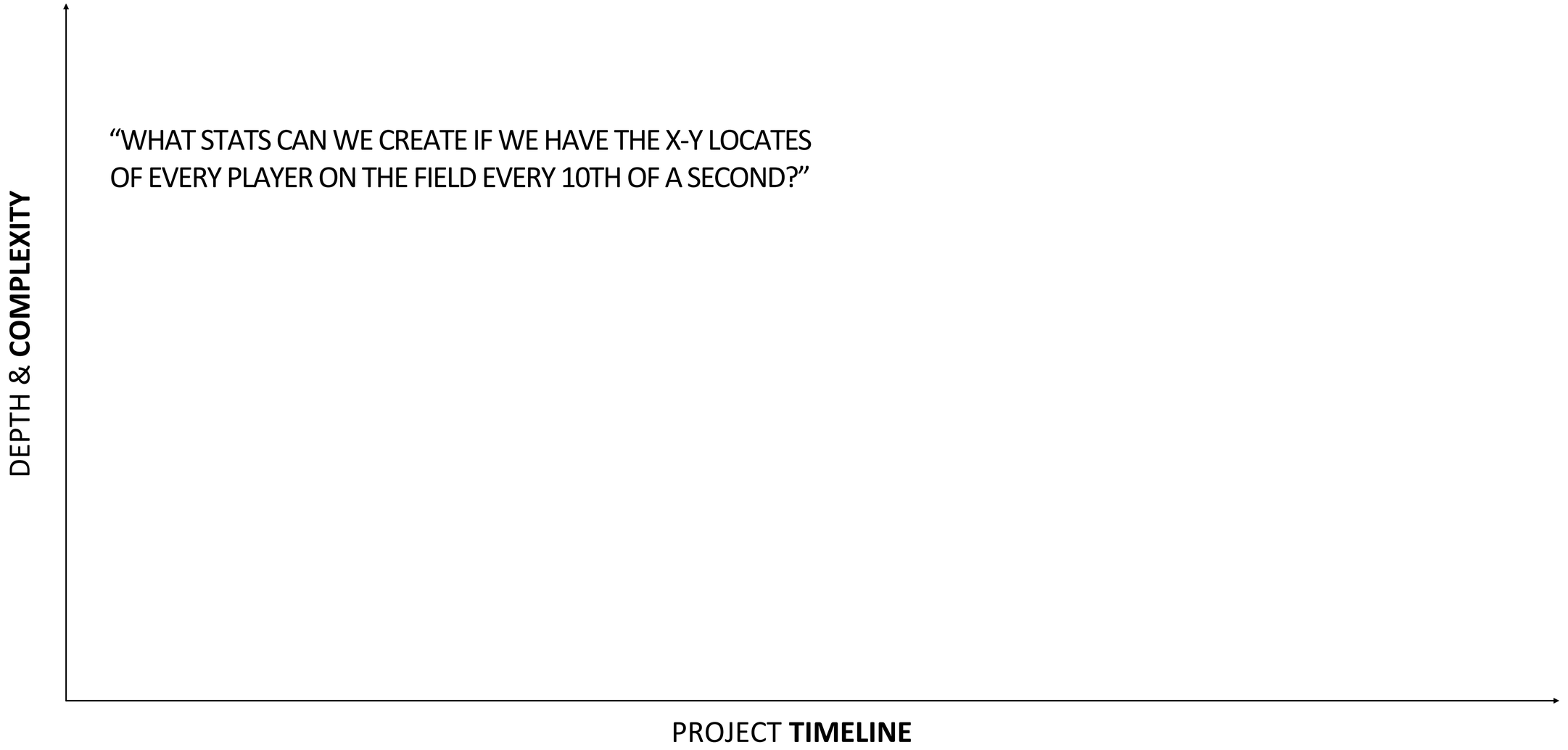
"The Race for Competitive Advantages"

"How Can We Improve the Game? & How Can We Make the Game Safer?"

The Evolution of the Next Gen Stats Toolbox



The evolution of the Next Gen Stats toolbox



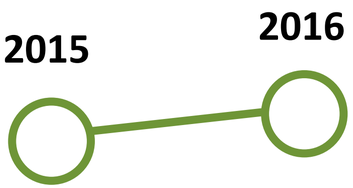
The evolution of the Next Gen Stats toolbox

DEPTH & **COMPLEXITY**

"WHAT STATS CAN WE CREATE IF WE HAVE THE X-Y LOCATES OF EVERY PLAYER ON THE FIELD EVERY 10TH OF A SECOND?"

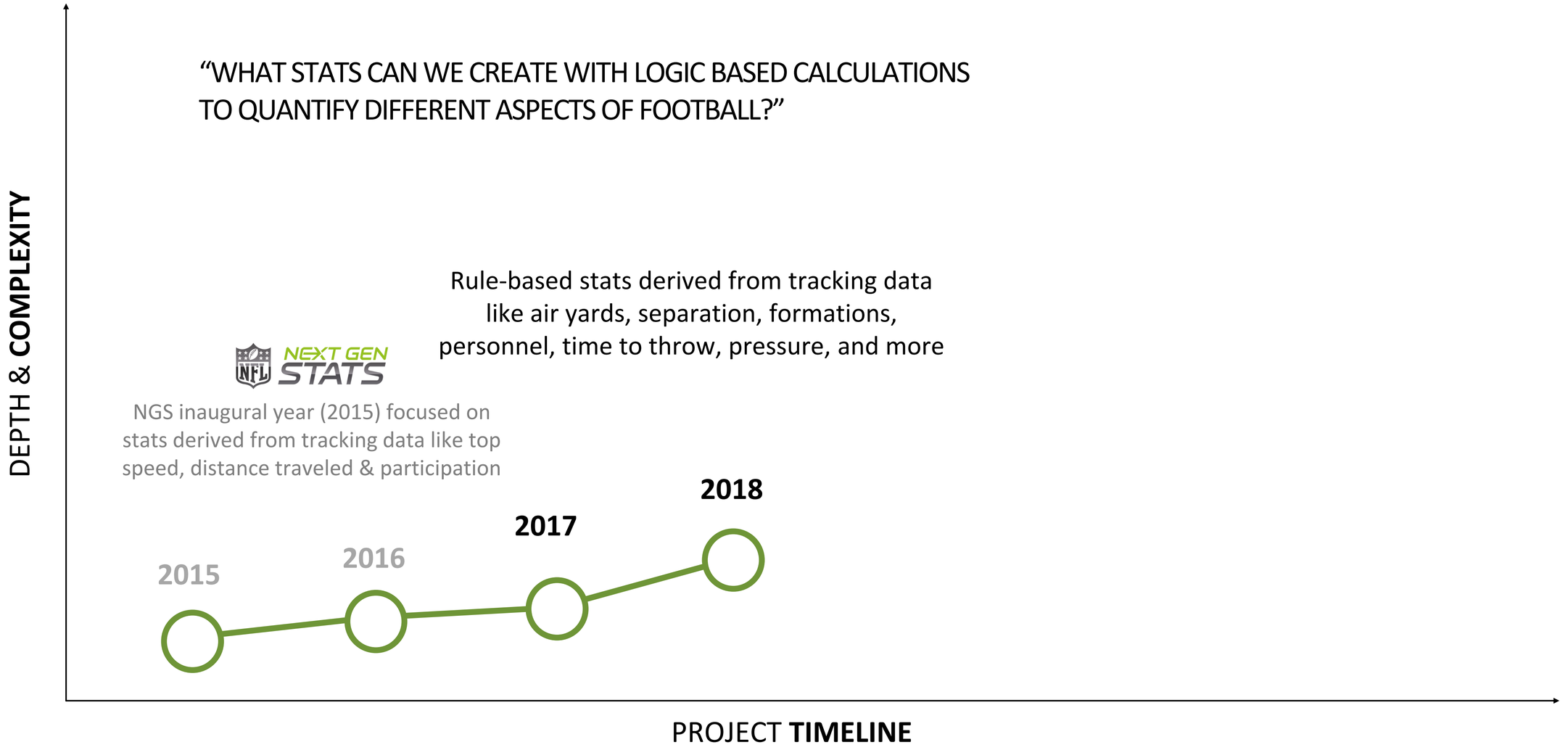
NEXT GEN STATS

NGS inaugural year (2015) focused on stats derived from tracking data like top speed, distance traveled & participation

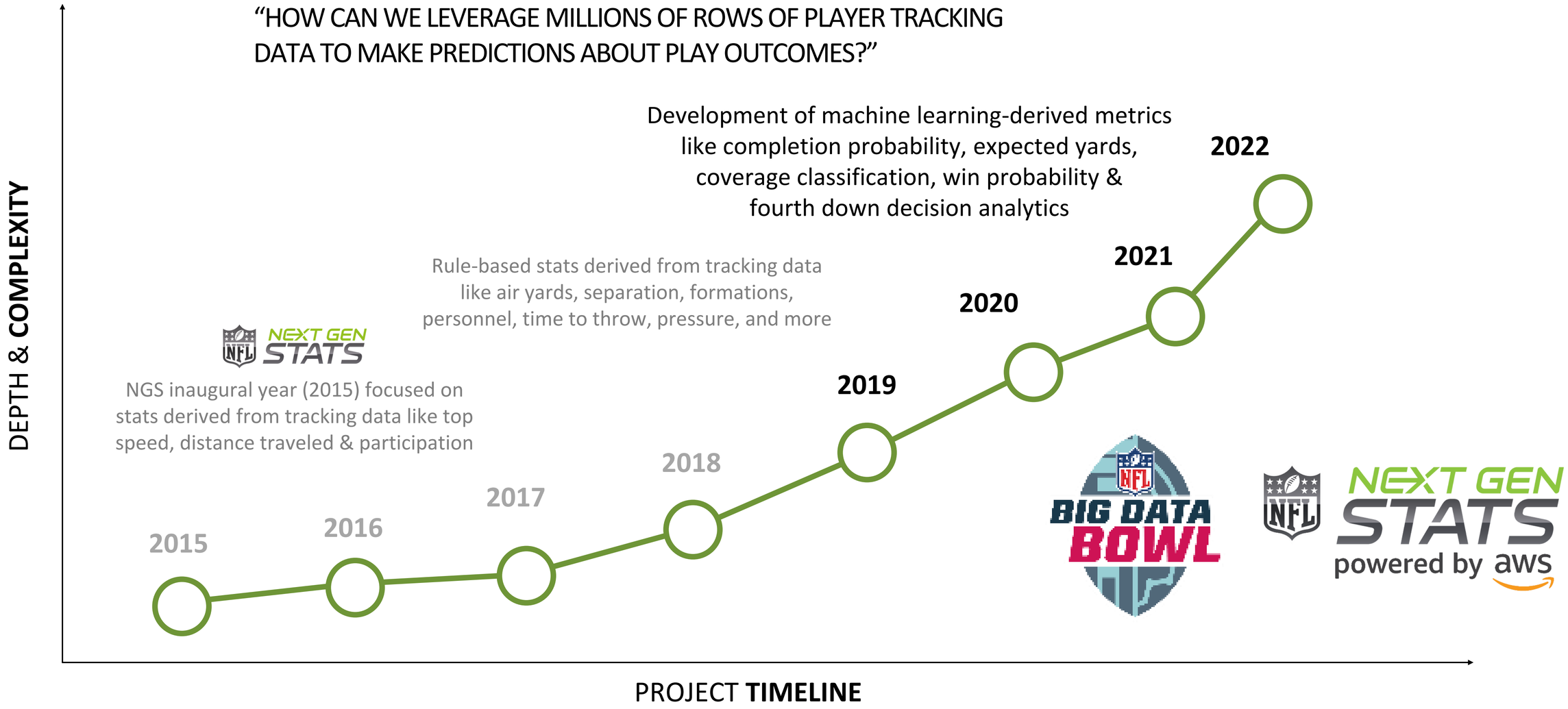


PROJECT **TIMELINE**

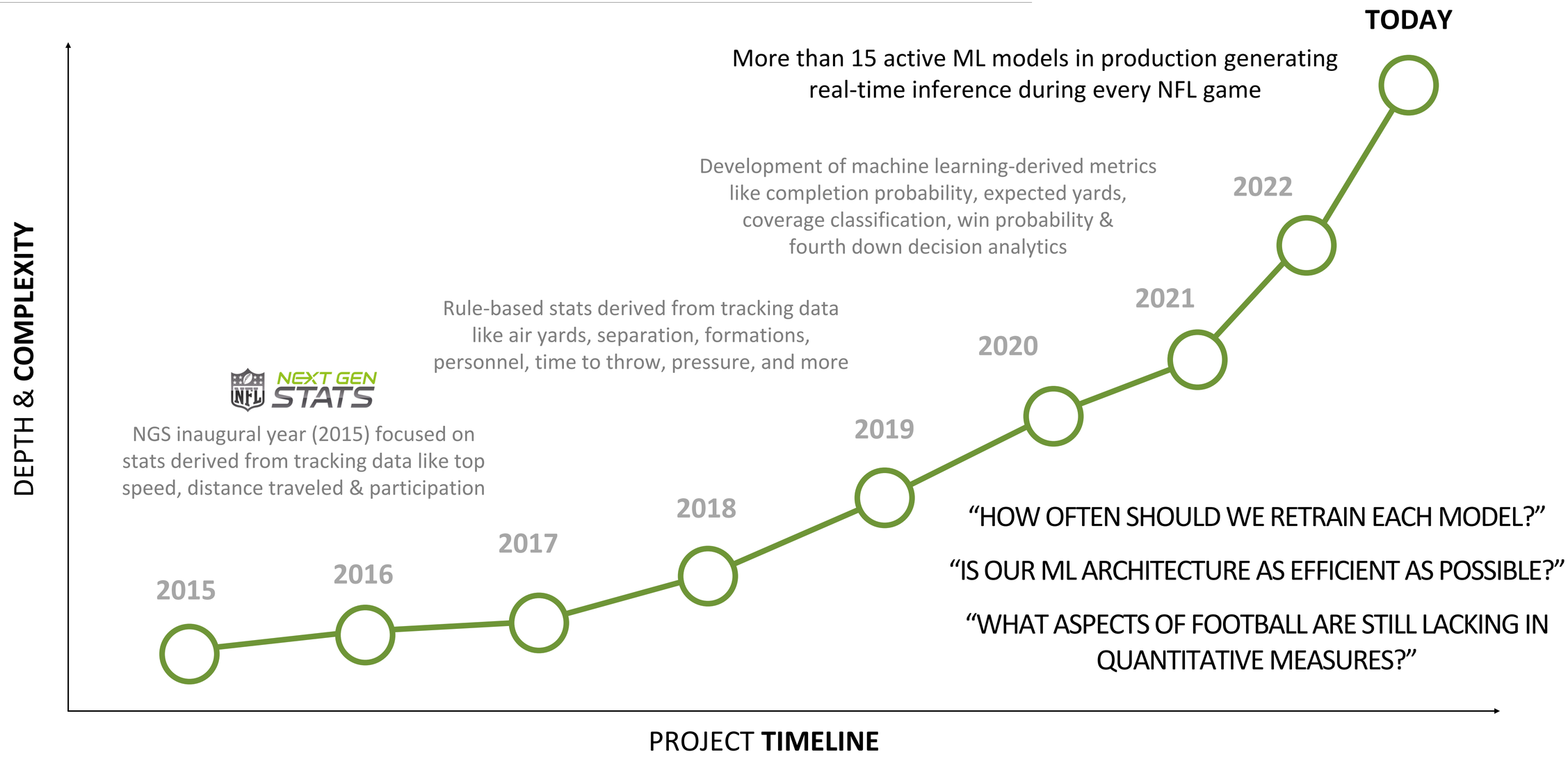
The evolution of the Next Gen Stats toolbox



The evolution of the Next Gen Stats toolbox



The evolution of the Next Gen Stats toolbox



The Many Models in Production at Next Gen Stats



The many models in PROD at Next Gen Stats

Completion Probability
What are the chances a given pass will be completed based on the difficulty of the throw?

QB Dropback Classification
Each dropback is categorized as one of: Standard Dropback, Designed Rollout, or Scramble.

NGS Passing Score (Expected EPA)
An estimate of the value of a pass to be attributed to a quarterback by controlling for yards after catch.

Expected Yards After Catch
Estimates the number of yards after the catch a receiver expects to gain on a reception.

Receiver Route Classification
Which route type did an eligible receiver run on a given pass play? (ex. Post, Corner, Go, …)

Expected Rushing Yards
Estimates the number of yards a ball carrier is expected to gain on a rushing play.

Coverage Classification
What coverage type did a defense run on a given play? Man or zone? Cover 1 or cover 2?

Run vs Pass Probability
What is the probability the next play will be a pass or run?

Field Goal Probability
What are the chances a field goal will be made based on the level of difficulty of the kick?

Expected Return Yards
Estimates the number of return yards a returner expects to gain after fielding a punt or kickoff.

Expected Points
Estimates the number of points a team will eventually score on the drive based on the situation.

Win Probability
What is the likelihood of either team winning at any moment in the game?

Fourth Down Decision Guide
Should a team go for it on fourth down, or kick?

2-Pt Decision Guide
Should a team go for two or kick the extra point?

NGS Draft Scores
Based on a prospect's athleticism, college production and size, what is their pro projection?

A few of the models in PROD at Next Gen Stats

Completion Probability

What are the chances a given pass will be completed based on the difficulty of the throw?

Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.

Receiver Route Classification

Which route type did an eligible receiver run on a given pass play? (ex. Post, Corner, Go)

Coverage Classification

What coverage type did a defense run on a given play? Man or zone? Cover 1 or cover 2?

Win Probability

What is the likelihood of either team winning at any moment in the game?

KEY ASSUMPTIONS & CONSIDERATIONS

- Each model must be able to generate inference predictions in fractions of a second before & after every play.
- The more explainable the model, the better. If deep learning is required, the model must align with the problem.
- The results of the model must pass the eye test! The validation of each model goes beyond model scoring techniques.

Completion Probability

What are the chances a given pass will be completed based on the difficulty of the throw?



Completion Probability

What are the chances a given pass will be completed based on the difficulty of the throw?

Model Type: xgboost

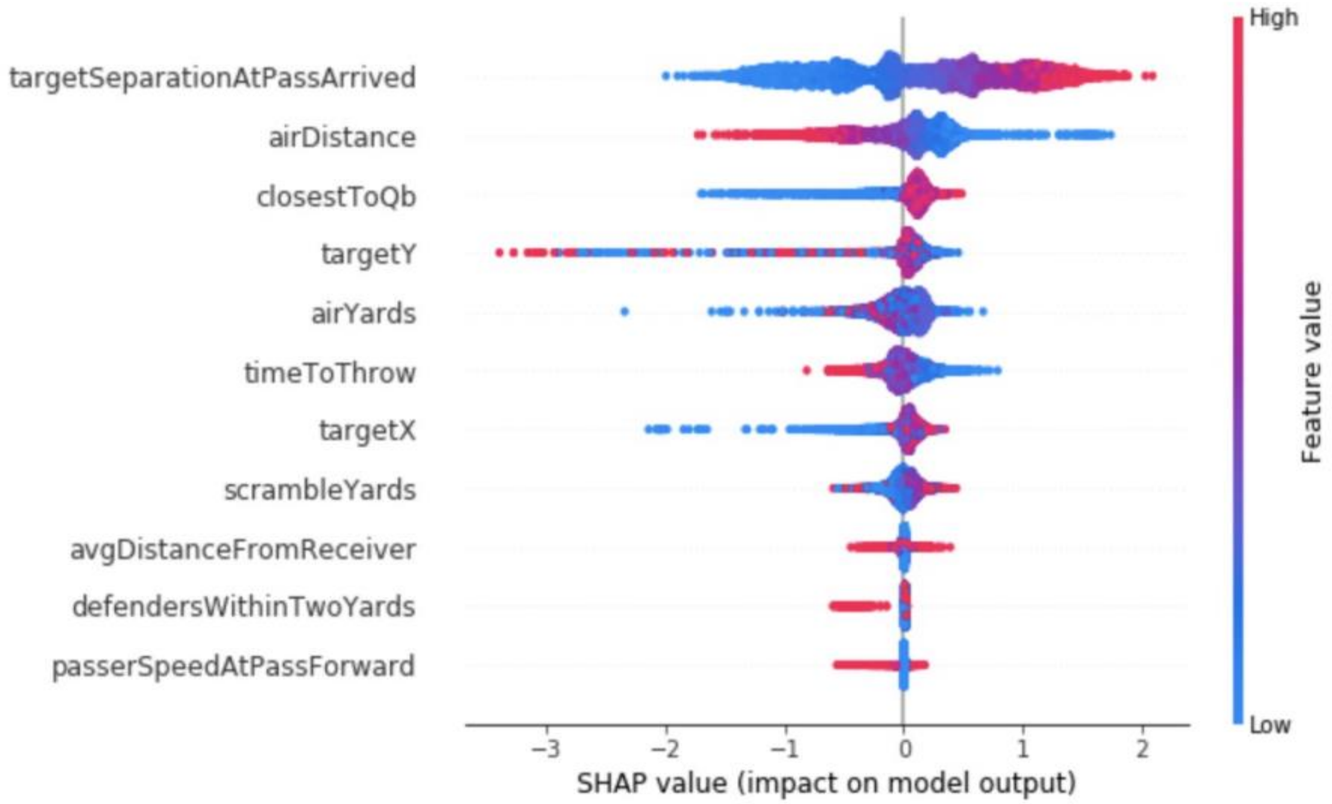
Features: How far the ball traveled in the air, receiver separation, pass rush pressure, QB & receiver speed, location on the field, and more.

Target/Output: Completed Pass (0 or 1)

Training Set: 72,000+ Regular & Postseason Pass Attempts All Pass Attempts (2018-2021)

Test Set: 2022 Regular Season Pass Attempts

Validation Set: 2022 Postseason Pass Attempts



Completion Probability

What are the chances a given pass will be completed based on the difficulty of the throw?

Model Type: xgboost

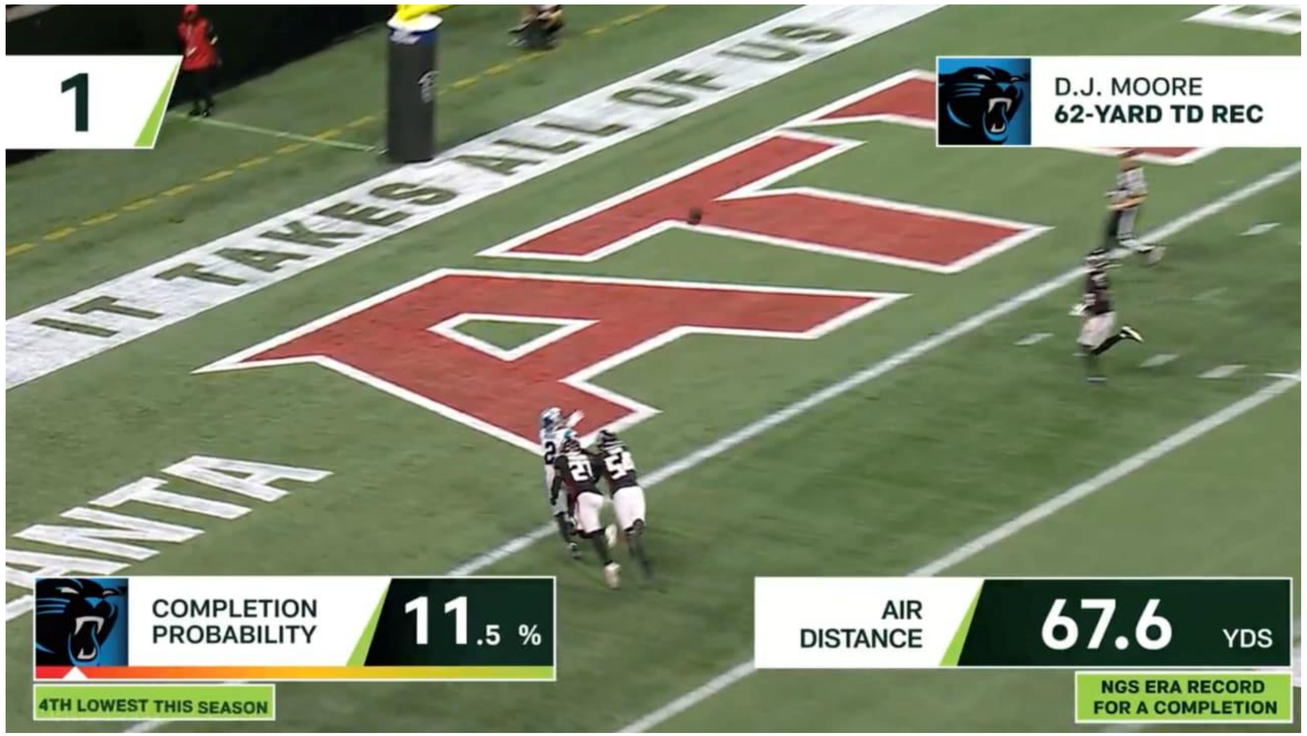
Features: How far the ball traveled in the air, receiver separation, pass rush pressure, QB & receiver speed, location on the field, and more.

Target/Output: Completed Pass (0 or 1)

Training Set: 72,000+ Regular & Postseason Pass Attempts All Pass Attempts (2018-2021)

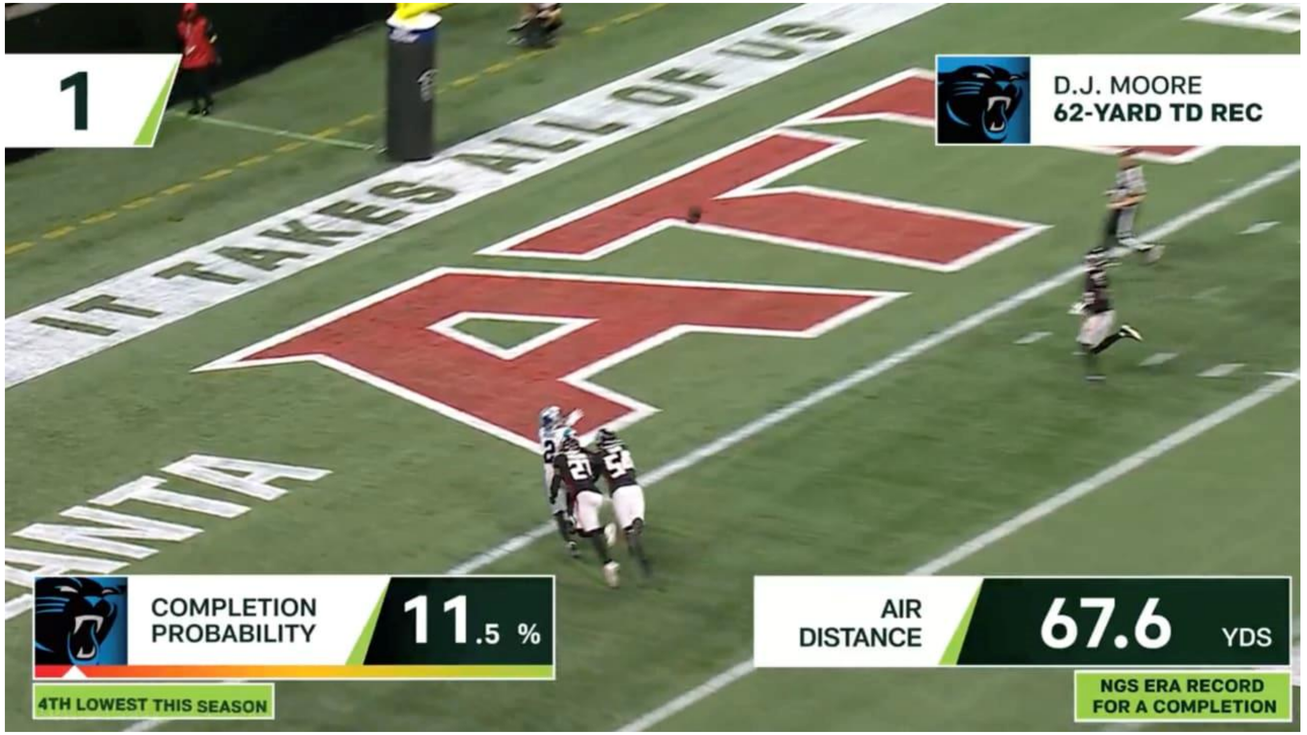
Test Set: 2022 Regular Season Pass Attempts

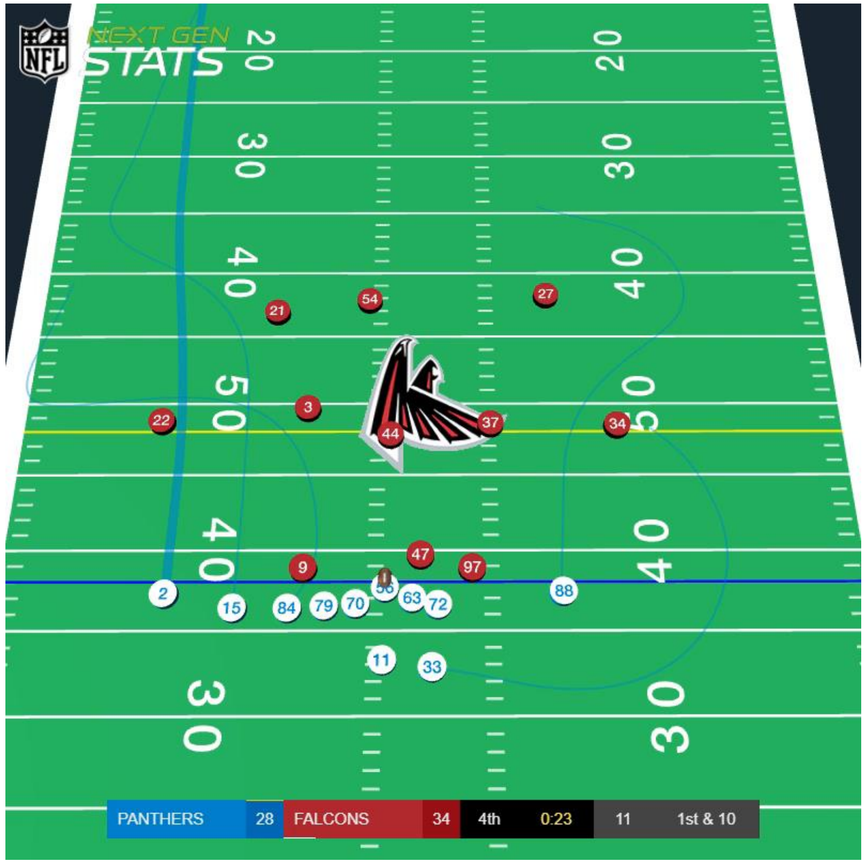
Validation Set: 2022 Postseason Pass Attempts



Completion Probability

What are the chances a given pass will be completed based on the difficulty of the throw?





Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.



Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.

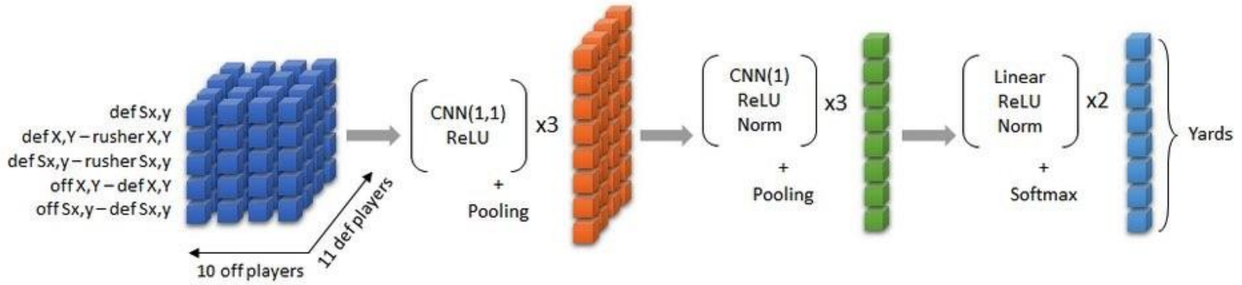
Model Type: Convolutional Neural Network

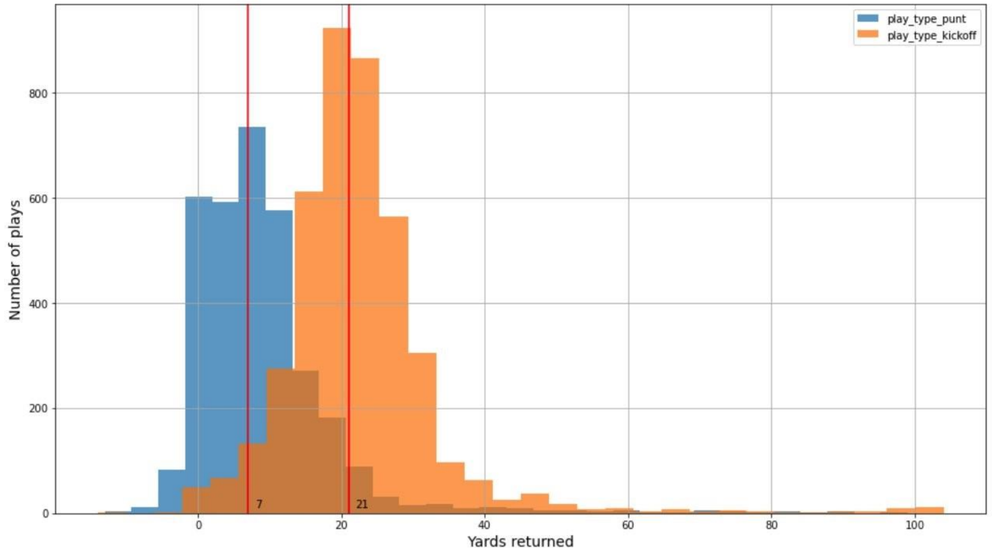
Models (5): Expected Rushing Yards (handoff & scramble), Expected Yards After Catch, Expected Return Yards (kickoff & punt return)

Features: Relative X-Y coordinate locations, speed and direction of all 22 players on the field at a single timestamp

Target/Output: The number of yards gained

Derivative Analytics: Expected Yards, Yards Over Expected, First Down Probability, Touchdown Probability, NGS Rushing Score





Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.

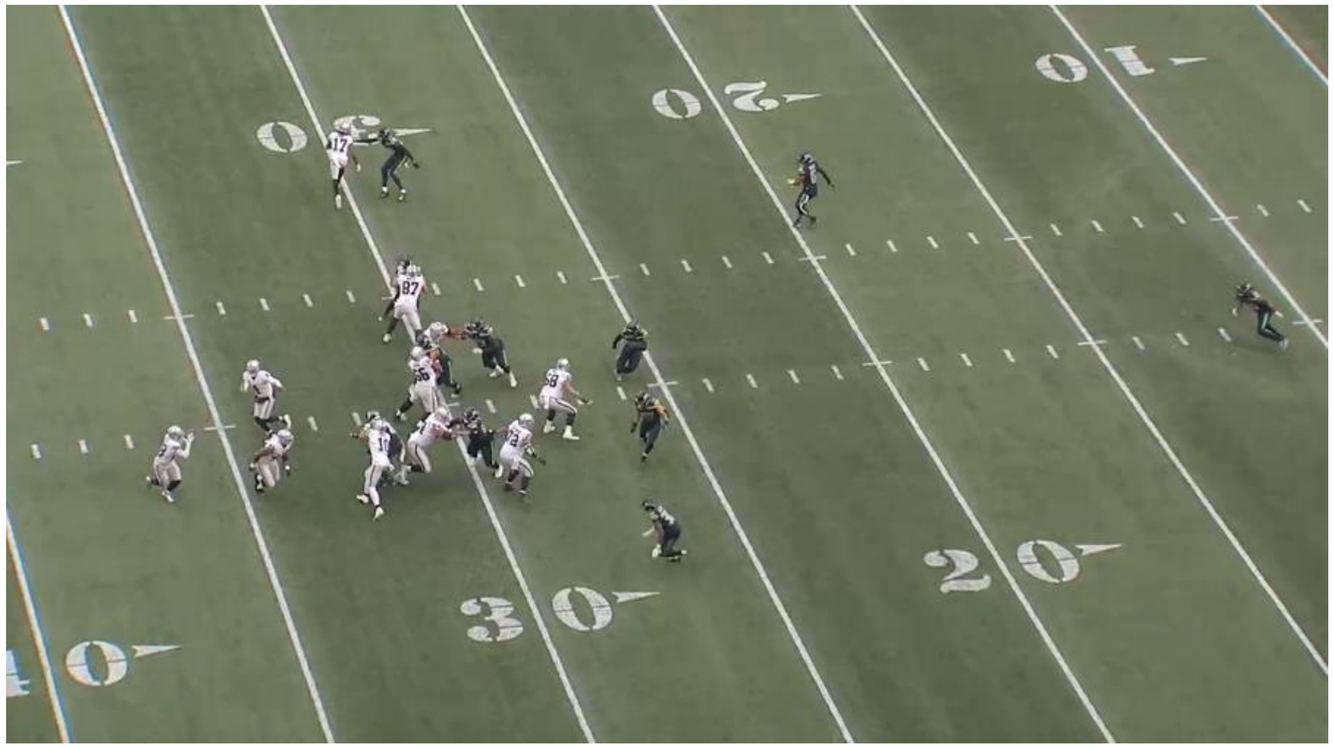
Model Type: Convolutional Neural Network

Models (5): Expected Rushing Yards (handoff & scramble), Expected Yards After Catch, Expected Return Yards (kickoff & punt return)

Features: Relative X-Y coordinate locations, speed and direction of all 22 players on the field at a single timestamp

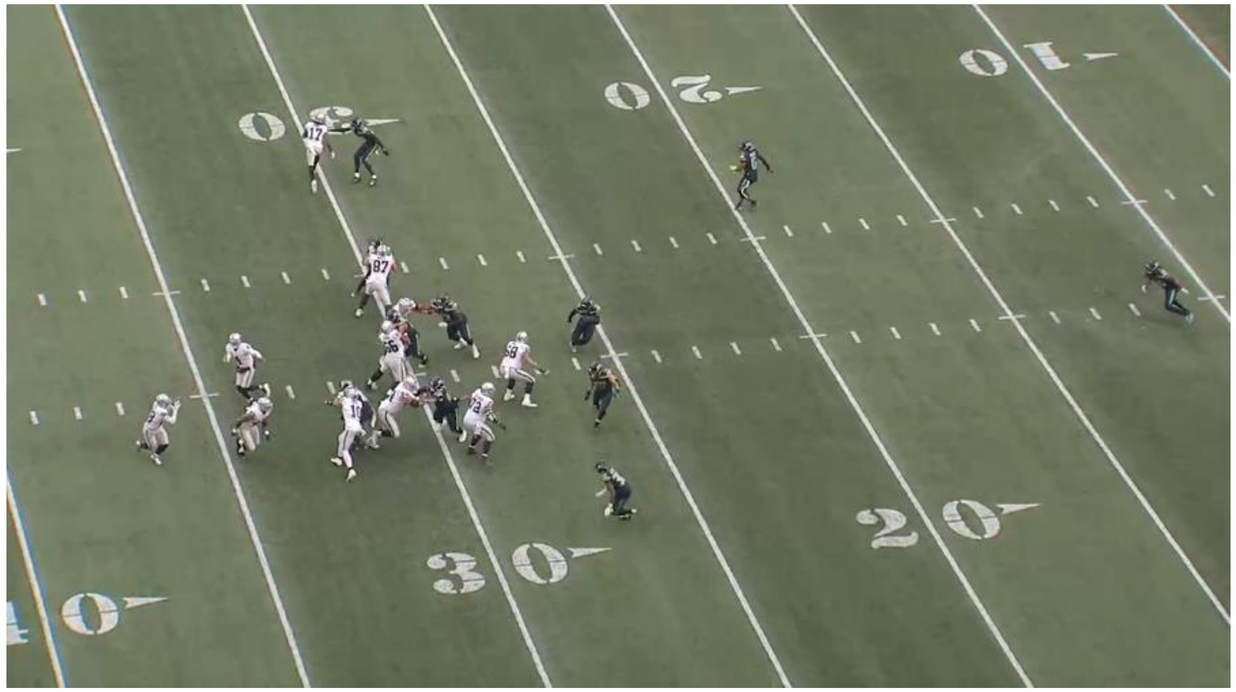
Target/Output: The number of yards gained

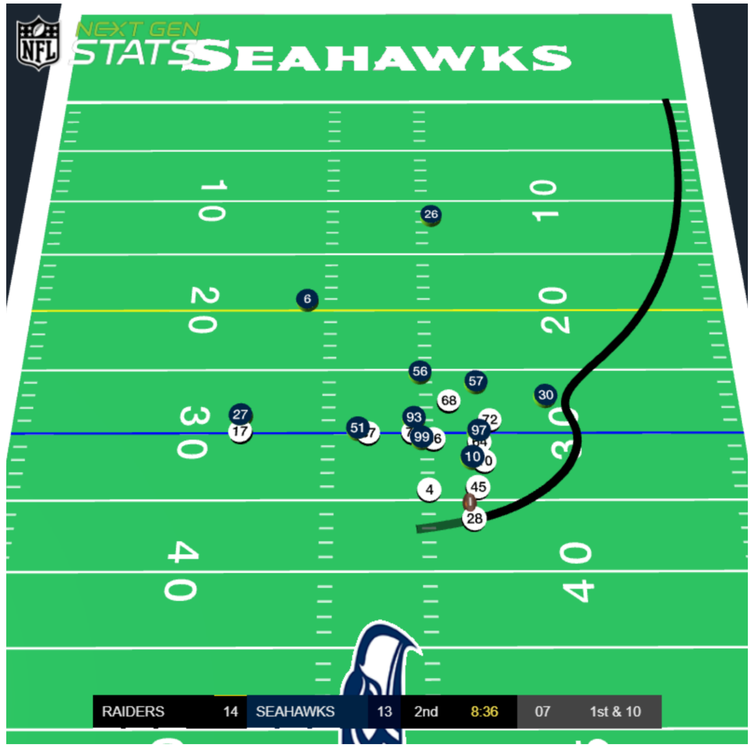
Derivative Analytics: Expected Yards, Yards Over Expected, First Down Probability, Touchdown Probability, NGS Rushing Score



Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.

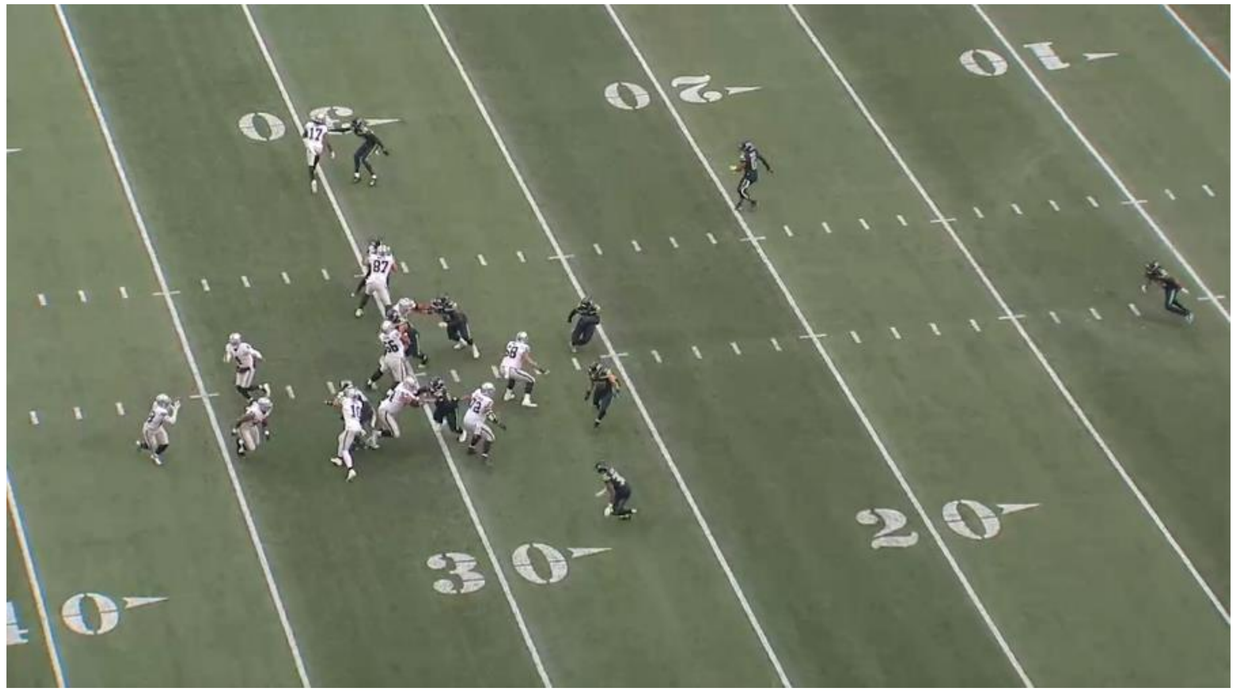


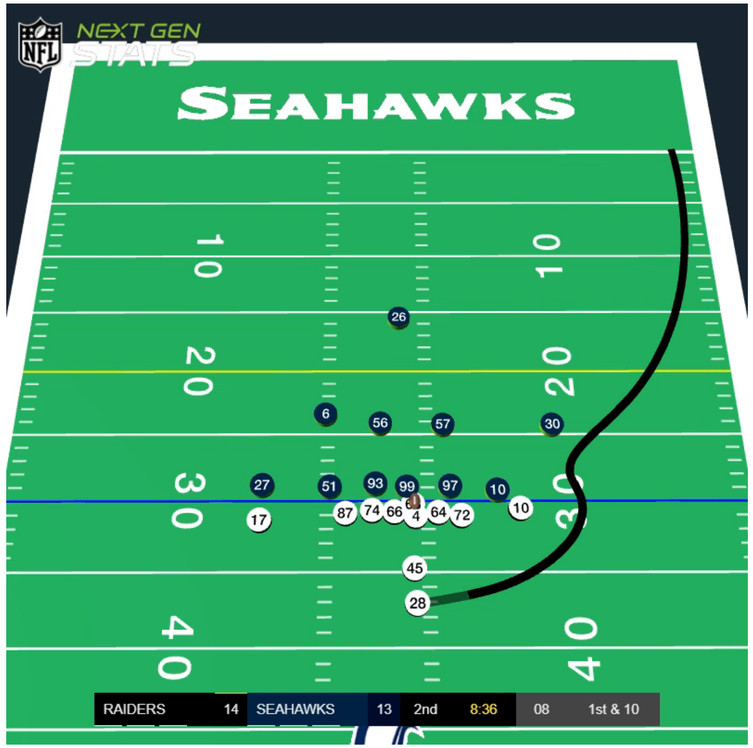


Expected Rushing Yards: 3 | Actual: 30 | Yards Over Expected: +27 | First Down Probability: 9.0% | TD Probability: 0.6%

Expected Yards

Estimates the number of yards a ball carrier is expected to gain from the moment they receive a handoff, catch a pass, or return a kick.





Expected Rushing Yards: 3 | Actual: 30 | Yards Over Expected: +27 | First Down Probability: 9.0% | TD Probability: 0.6%

Receiver Route Classification

Which route type did an eligible receiver run on a given pass play?



Receiver Route Classification

Which route type did an eligible receiver run on a given pass play?

Model Type: Convolutional Neural Network (CNN) & Long Short-Term Memory (LSTM)

Features: Temporal player location data of each eligible receiver from snap to pass (with a maximum of 4.6 seconds); Includes x-y coordinate locations, speed, and direction.

Target/Output: 15 Unique Route Types

Wideout Routes (10): Screen, flat, slant, cross, out, in, hitch, corner, post, go

Backfield Routes (5): Screen, flat, angle, out, wheel



Big Data Bowl winner, and now Cleveland Browns data scientist, Nate Sterkin is joined by the Big Data Bowl emcee Cynthia Frelund and Michael Lopez, the NFL's Director of Football Data and Analytics.

Receiver Route Classification

Which route type did an eligible receiver run on a given pass play?

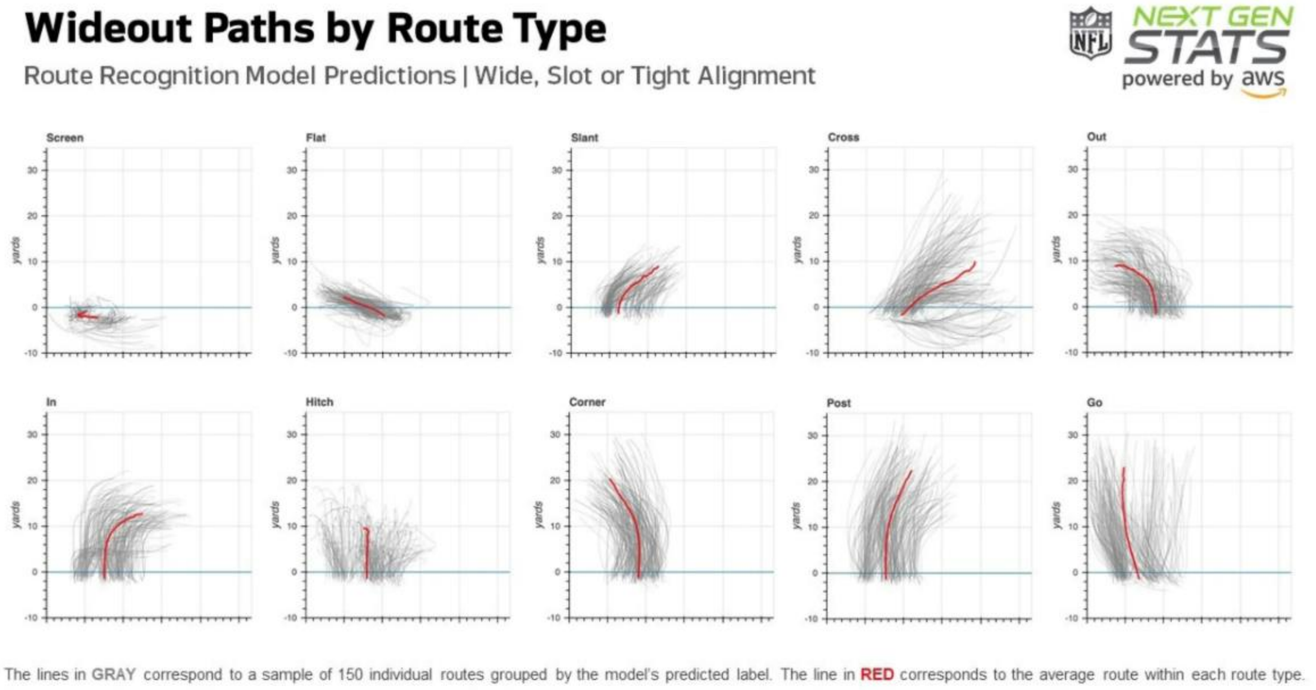
Model Type: Convolutional Neural Network (CNN) & Long Short-Term Memory (LSTM)

Features: Temporal player location data of each eligible receiver from snap to pass (with a maximum of 4.6 seconds); Includes x-y coordinate locations, speed, and direction

Target/Output: 15 Unique Route Types

Wideout Routes (10): Screen, flat, slant, cross, out, in, hitch, corner, post, go

Backfield Routes (5): Screen, flat, angle, out, wheel



Defensive Coverage Classification

What coverage type did a defense run on a given play?



Coverage Classification

What coverage type did a defense run on a given play?

Model Type: CNN-Attention Model

Features: Temporal player location data of every defensive player from snap to pass (with a maximum of 4.6 seconds); Includes relative x-y coordinate locations, speed, and direction.

Target/Output: 8 Unique Team Coverage Types

Coverage Types (8): Cover 0, 1, 2 (Man Coverage); Cover 2, 3, 4, 6 & Prevent (Zone Coverage)



NFL Data Scientist Tom Bliss speaking at the 2022 New York R Conference. Tom created the baseline coverage classification model for Next Gen Stats.

Coverage Classification

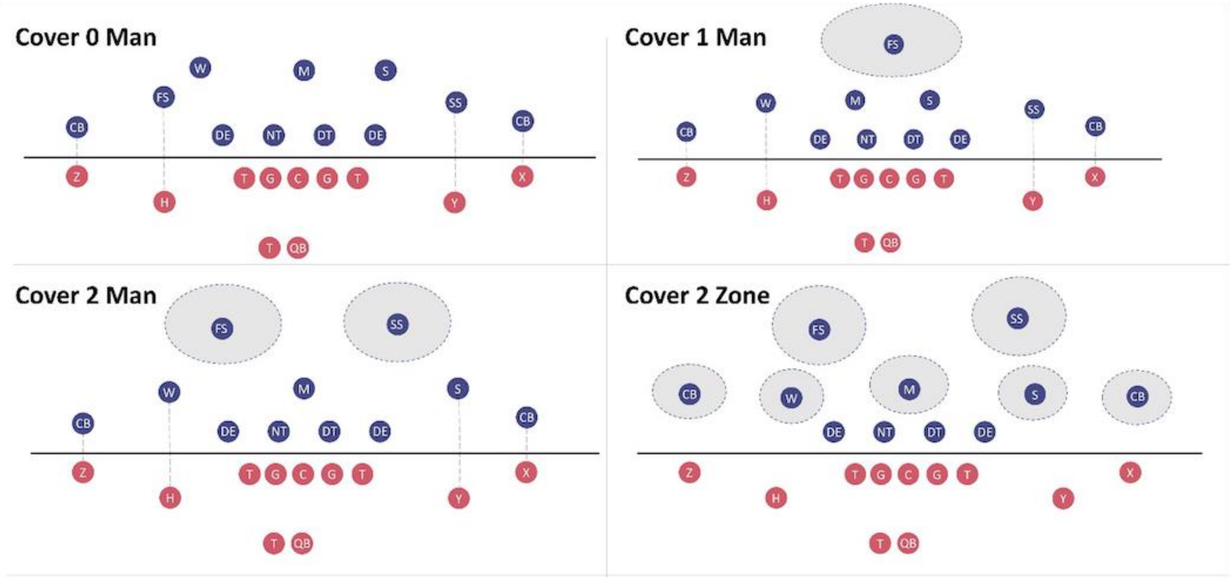
What coverage type did a defense run on a given play?

Model Type: CNN-Attention Model

Features: Temporal player location data of every defensive player from snap to pass (with a maximum of 4.6 seconds); Includes relative x-y coordinate locations, speed, and direction.

Target/Output: 8 Unique Team Coverage Types

Coverage Types (8): Cover 0, 1, 2 (Man Coverage); Cover 2, 3, 4, 6 & Prevent (Zone Coverage)



Coverage Classification

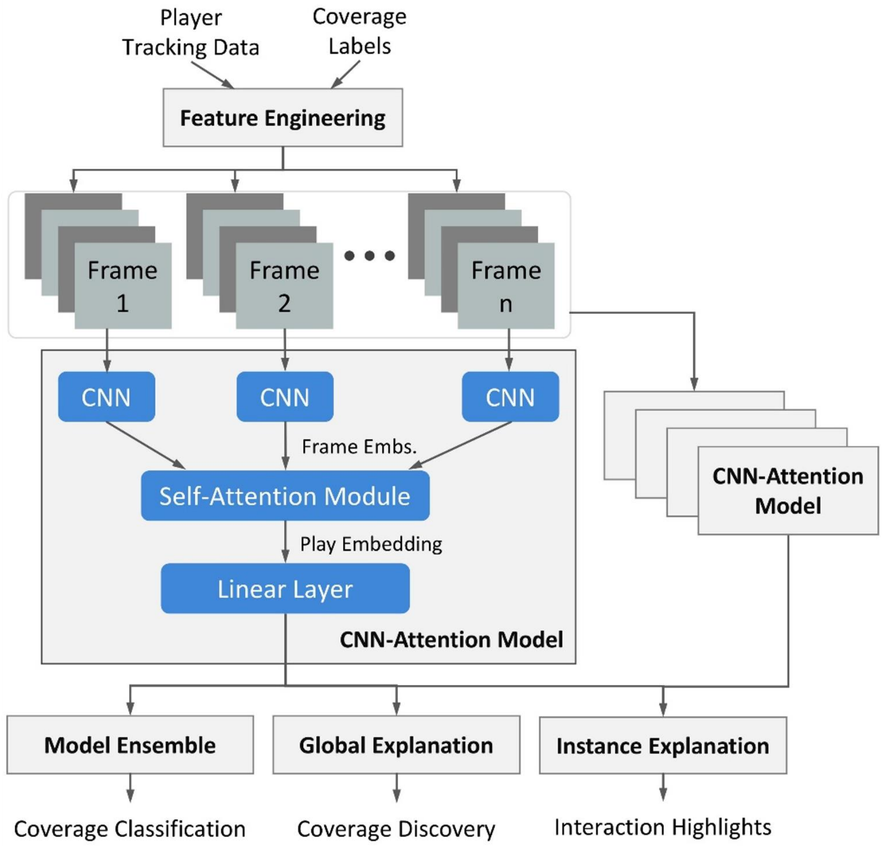
What coverage type did a defense run on a given play?

Model Type: CNN-Attention Model

Features: Temporal player location data of every defensive player from snap to pass (with a maximum of 4.6 seconds); Includes relative x-y coordinate locations, speed, and direction.

Target/Output: 8 Unique Team Coverage Types

Coverage Types (8): Cover 0, 1, 2 (Man Coverage); Cover 2, 3, 4, 6 & Prevent (Zone Coverage)



Win Probability

What is the likelihood of either team winning at any moment in the game?



Win Probability

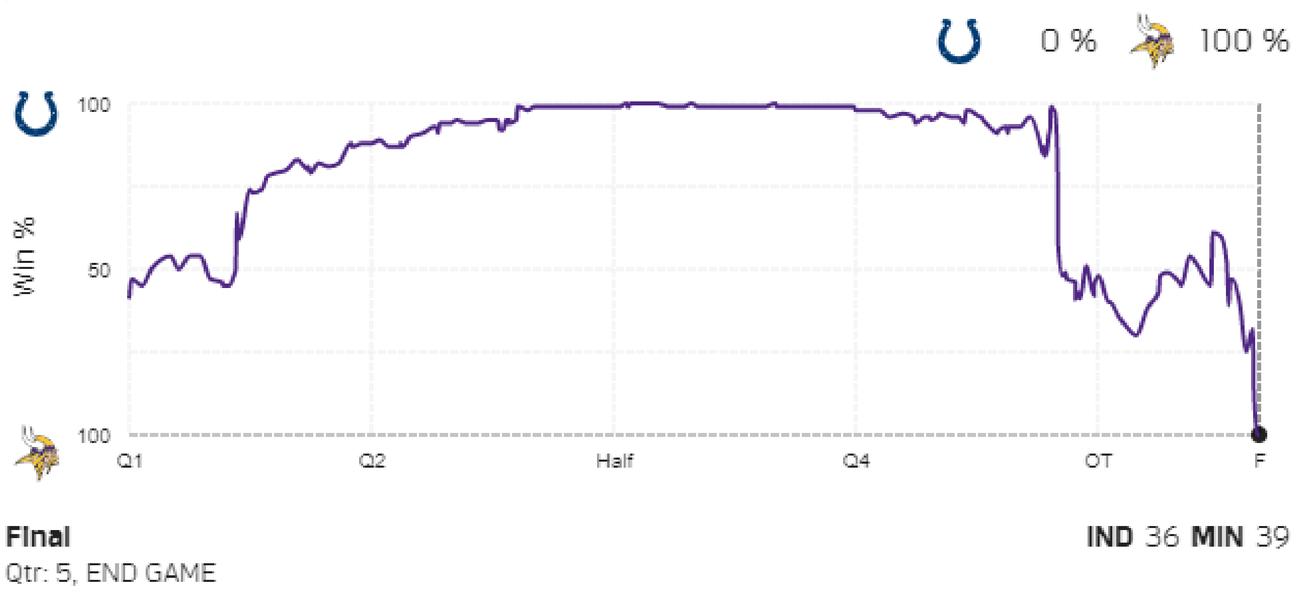
What is the likelihood of either team winning at any moment in the game?

Model Type: xgboost

Features: Trained on every historical play in the last 10 seasons looks at the score differential, down/distance, time remaining, timeouts, expected points, and the quality of the teams.

Target/Output: Game Outcome (Win = 1, Loss = 0)

Model Dependency: Expected Points that represents a composite score expectation



Win Probability

What is the likelihood of either team winning at any moment in the game?

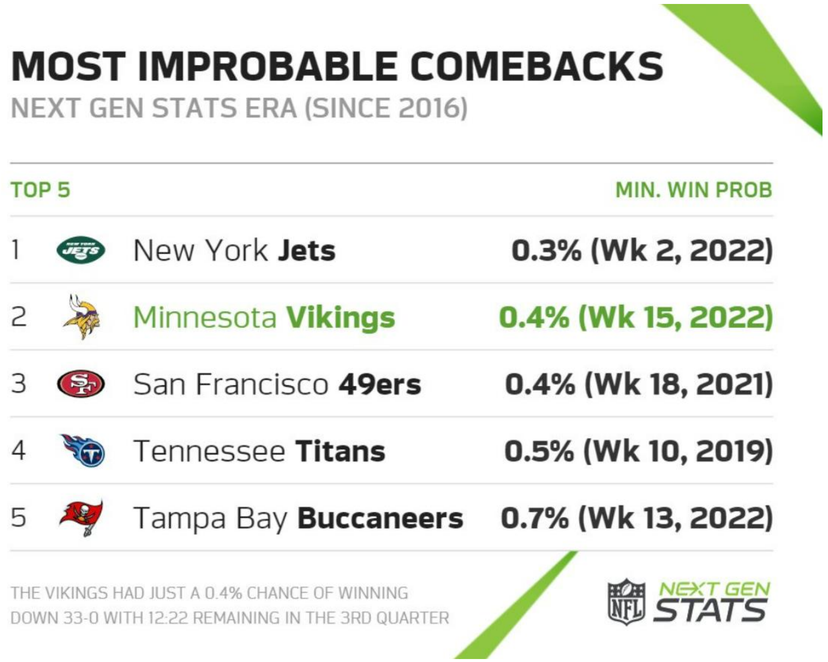
Model Type: xgboost

Features: Trained on every historical play in the last 10 seasons looks at the score differential, down/distance, time remaining, timeouts, expected points, and the quality of the teams.

Target/Output: Game Outcome (Win = 1, Loss = 0)

Model Dependency: Expected Points that represents a composite score expectation





Follow Next Gen Stats on Twitter & Instagram!

Follow @NextGenStats on Twitter

Follow @MBandNFL on Twitter

Follow @NextGenStats on Instagram



THANK YOU!