Predicting Rookie Season Performance Based on National Football League (NFL) Scouting Combine Movement Analysis

A Major Qualifying Project Proposal
Submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
In partial fulfillment of the requirements for the
Degree of Bachelor of Science

Submitted By:
Matthew Dzwil
Mia Hopman
Katie Houskeeper

April 27, 2023

Advisors:
Dr. Kenny Ching & Dr. Yanhua Li

Sponsor:
National Football League (NFL)

WORCESTER POLYTECHNIC INSTITUTE

Table of Contents

List of Figures

List of Tables

Abstract

This Major Qualifying Project seeks to advance player evaluations in the National Football League (NFL) by transforming wide receiver positional tracking data during the 2022 NFL Scouting Combine into meaningful metrics. To achieve this, our team developed a systematic approach to preprocessing data that utilizes specific thresholds to identify and segment drills. A similar process was applied to routes throughout the 2022-2023 NFL season. Following this, metrics were created on the drills and routes to represent players' movements. Finally, several techniques, such as correlation and regression, were performed alongside the creation of a route prediction model.

Acknowledgments

The success of our Major Qualifying Project results from the contributions of numerous individuals who generously dedicated their time, expertise, and resources to support our team. Their cooperation enabled us to overcome challenges and achieve our goals. We extend thanks to everyone involved and wish to express our appreciation by acknowledging their efforts.

Firstly, we would like to thank our project advisors, Dr. Yanhua Li and Dr. Kenny Ching, who provided many hours of guidance and constructive criticism throughout our project.

We extend our appreciation to the National Football League for partnering with Worcester Polytechnic Institute (WPI) and offering us a valuable opportunity to gain insight into the field of sports. This partnership has given us a deeper understanding of the industry and the available career opportunities. Specifically, we would like to thank Dr. Michael Lopez, Senior Director of Football Data and Analytics, and Thompson Bliss, Football Operations Data Science Manager, for their expertise in tackling this project's complex problems. Your efforts to help our team did not go unnoticed.

We want to recognize Dr. Elke Rundensteiner, the Data Science (DS) Department Head, Dr. Kelsey Briggs, DS Project Coordinator, and Dr. Anna Maria Young, DS Program Manager. Each played an instrumental role in solidifying and bettering this project experience.

Thank you to Dr. David Ortendahl, Executive Director of Corporate Relations, and Lisa Drexhage, Associate Director of Corporate Relations, for attending our presentation at the NFL Headquarters in Manhattan, NY. Your kind words and presence were greatly appreciated, and we acknowledge your dedication to fostering meaningful partnerships with the NFL.

Lastly, we are grateful to the staff and faculty at WPI, who provided us with the necessary resources and facilities to carry out this project. We also thank our fellow students and other faculty members of the DS department for their support over the past four years. Our team is fortunate to have been a part of the collaborative environment that WPI and the DS department foster.

1.0 Introduction

The NFL Scouting Combine has been held annually since it was founded in 1982. It provides top draft eligible prospects a chance to display their skills and abilities to NFL teams. A player's performance at the combine is not the main factor of determining if a player gets drafted. However, it can impact a player's draft stock (NFL Scouting Combine History, n.d.).

Traditionally, the evaluation of a player's combine performance drills has relied on broad quantitative evaluation, such as the total time taken by a player to complete a drill, among other similar metrics. However, the 2022 introduction of Zebra's MotionWorks wearable device, which utilizes radio-frequency identification technology, has enabled more precise and comprehensive data collection during the event (Petrus, 2022).

This device can measure a player's position, velocity, and acceleration in two directions every tenth of a second. 2022 was the first-year players ran with these devices. This meant our team was among one of the first teams to evaluate this data. We focused not only on improving the NFL's ability to utilize this data but using this data to help NFL teams make more informed decisions (Zebra Technologies, n.d.).

We established a reliable process to transform raw combine files, as well as 2022-2023 NFL season game data, into insightful metrics, which first involved identifying drills or routes and segmenting them based on player movements. Once metrics were created for both the combine drills and in-game routes, our focus shifted to exploring the extent to which player movements during the NFL Scouting Combine could predict future performance in the NFL through correlation.

We also developed a CatBoost predictive model by first establishing a baseline model that served as a reference point to evaluate future models' performance. We applied machine learning techniques, such as hyperparameter tuning, class weighing, and early stopping to optimize the model.

2.0 Background

2.1 The National Football League

2.1.1 Overview

The National Football League (NFL) was formed in 1920 through the merger of two competing professional football leagues, bringing organization to the sport (Britannica, The Editors of Encyclopedia, 2023). Despite early criticisms, the NFL has grown in popularity and revenue and now stands as one of the most followed sports leagues in the world. It garners passionate fans of all ages, and backgrounds tuning in to watch their favorite teams and players compete (Gough, 2023). In the 2022-2023 season alone, the league amassed 18 billion dollars in total revenue (Moore, 2022).

As of the 2022-2023 NFL season, the NFL comprises 32 teams from all over the United States, shown in Figure 1. Teams are divided into two conferences (the American Football Conference and National Football Conference) and four divisions per conference (National Football League, 2023). The league operates on an 18-week regular season schedule, followed by a playoff tournament culminating in a championship game referred to as the Super Bowl (NBC Sports, 2023). It attracts millions of viewers each year as it is America's highest-viewed annual sporting event. For example, over 56 million people worldwide tuned in to watch the 2023 Super Bowl LVII, where Kansas City Chiefs secured a 38-35 win against the Philadelphia Eagles (National Football League, 2023).

Figure 1: Map of NFL teams and approximate locations (NFL map, n.d.)

A football game consists of four 15-minute quarters. The dimensions of the playing field are 100 yards in length, with an extra 10-yards end zone on each side, and 53.3 yards in width.

The game commences with an initial kickoff, where one team kicks the football to the opposing team. From there, each team has four downs to advance at least 10 yards toward the opponent's end zone. Four more downs are granted if the necessary distance is covered. A change of possession ensures that the 10 yards are not achieved within four downs. Points can be scored in five ways: a touchdown, one or two-point conversion, a field goal, or a safety. The team with the most points at the game's conclusion wins (Adler, 2019).

For each team 11 players concurrently take the field, with roles designated among offense, defense, and special teams. The responsibility of stopping the opposing team from scoring falls on defensive positions such as linebackers, defensive backs, and safeties. Special teams, composed of kickers, punters, and return specialists, handle various plays, including kickoffs, punts, and field goal attempts. Offensive players, including quarterbacks, tight ends, running backs, and wide receivers, are primarily responsible for advancing the ball up field toward the opponent's end zone. (Rookie Road, n.d.). Our report will focus on the wide receiver position.

2.1.2 The Role of Wide Receivers

In football, the wide receiver position involves running pass routes and catching passes from the quarterback to progress towards the end zone. Wide receivers are often among the field's fastest and most agile players. They are responsible for various duties, including running specific pass routes, such as slants, corners, and comebacks, illustrated in Figure 2, to get open and create separation from defenders. They can also be responsible for blocking for their fellow team members when necessary (Rookie Road, n.d.).

A wide receiver's route is determined based on the play called and the defensive coverage. Additionally, their skill set, and abilities often determine the routes they specialize in. For example, Julian Edelman, a former wide receiver for the New England Patriots, was not the fastest player but could turn quickly. Therefore, he preferred running shorter routes such as an in or out route, compared to deep routes like go or post.

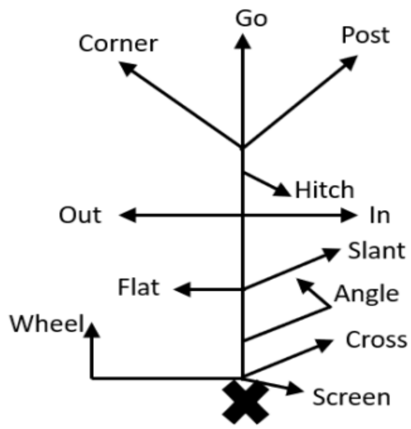


Figure 2: Wide receiver routes (Champion & Varanasi, 2022)

Figure 2 depicts the general routes run by wide receivers. A description of each is shown below (Bowen, 2017; Nogle, 2016).

- *Screen route:* A route where a short pass is quickly thrown to a receiver behind the line of scrimmage.
- *Cross route:* A route where the receiver makes a sharp cut across the field in the direction of the opposite sideline.
- *Flat route:* A route where the receiver initially runs straight and then quickly turns towards the sideline.
- *Wheel route:* A route where the receiver initially runs toward the sideline before turning up field.
- *Angle*: A route where the receiver initially runs at a 45-degree angle and then quickly cuts back 90 degrees toward the center of the field.
- *Slant route:* A route where the receiver runs diagonally across the field from the outside to the center of the field.
- *Out route*: A route where a receiver initially runs straight and then cuts toward the sideline.
- *In route:* A route where a receiver initially runs straight and then cuts toward the center of the field.
- *Hitch route:* A route where the receiver initially runs straight and then cuts back toward the quarterback.
- *Post route:* A route where the receiver initially runs straight and then cuts towards the center of the field at an angle.
- *Corner route:* A route where the receiver initially runs straight and then cuts towards the corner of the endzone at an angle.
- *Go route:* A route where the receiver runs straight as fast as possible.

2.2 The NFL Scouting Combine

2.2.1 Purpose and Overview

Each year, roughly 259 players of all positions are selected by one of the 32 teams through the NFL Draft (National Football League, 2023). This four-day event is typically held in late April or early May, allowing teams to acquire talented players to contribute to the team's upcoming year's success (NFL Scouting combine, n.d.).

The NFL Combine is an annual event held before the NFL Draft. It is an opportunity for approximately 310 of the top draft-eligible players to impress teams. Players' NFL Combine performance can impact their draft position as they endure mental and physical tests, including the 40-yard dash, vertical jump, position-specific drills, and team interviews.

2.2.2 Drill Specifics

As shown in Figure 3, in the 40-yard dash, the player begins in a three-point stance and sprints in a straight line for 40 yards as fast as they can. The drill hopes to obtain a benchmark for a player's speed (Davenport, 2013). While a typical 40-yard dash time for college players is 4.6 - 4.9 seconds, the time does vary by position. For instance, a 2000-2022 study showed a typical 40-yard dash time for a defensive tackle was around 5.06 seconds. However, for wide receivers, the average 40-yard dash time in the NFL is 4.48 seconds (Doll, 2013).

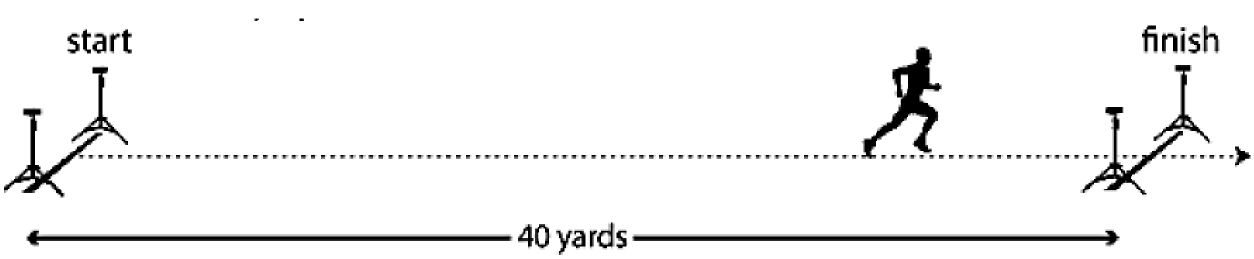


Figure 3: 40-yard dash visualization (40 yard dash, n.d.)

The 3-cone drill, also known as the "L-drill", is a test that measures a player's agility and athleticism. During this drill, cones are set up in an L-shape pattern, as shown in Figure 4 on page 13, with each cone placed five yards away from the next. The player begins at the first cone and sprints forward to the second cone, then changes direction quickly and returns to the first cone. After arriving at the first cone, the player must promptly change direction again and sprint around the second cone toward the third cone. At the third cone, the player must run a half circle around it before turning around the second cone to finish back where they started (Rogust, 2022). Wide receivers and defensive backs typically complete the drill in about 6.5-7.0 seconds, while other positions sometimes complete it within 7.0-7.5 seconds (NFL Combine Trainer, n.d.).

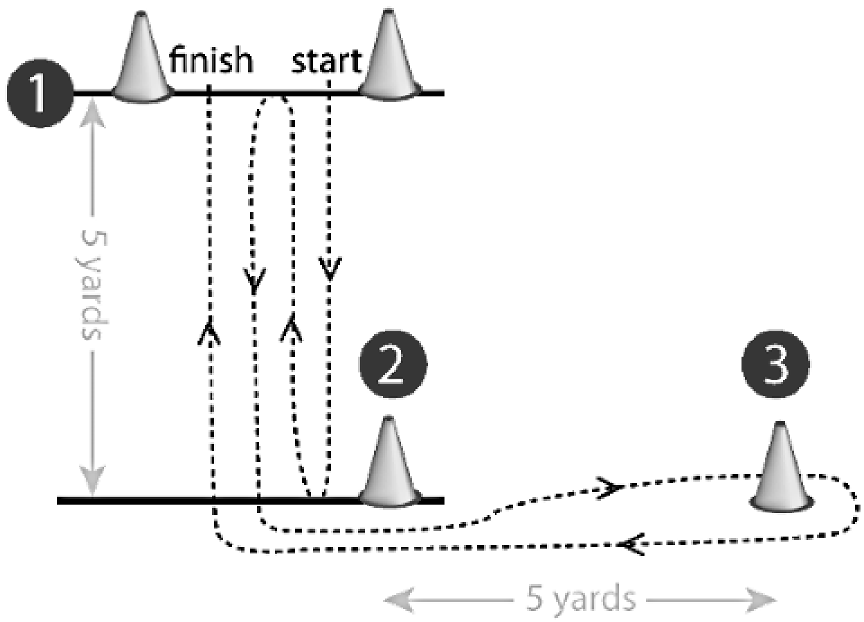


Figure 4: 3-cone drill visualization (3-cone shuttle drill test, n.d.)

The shuttle drill is designed to measure an athlete's change-of-direction ability and lateral quickness. As shown in Figure 5, to begin the drill, the player stands behind cone one and sprints five yards to touch the second cone. They then pivot and sprint ten yards in the opposite direction to touch the third cone before quickly changing direction and running five yards back to the first cone line (Sidhu, 2022). A shuttle run time of 4.5 seconds or less is generally considered good, with the fastest shuttle attempts clocking in at around 3.9 seconds (BlazePod, n.d.; Hope, 2023).

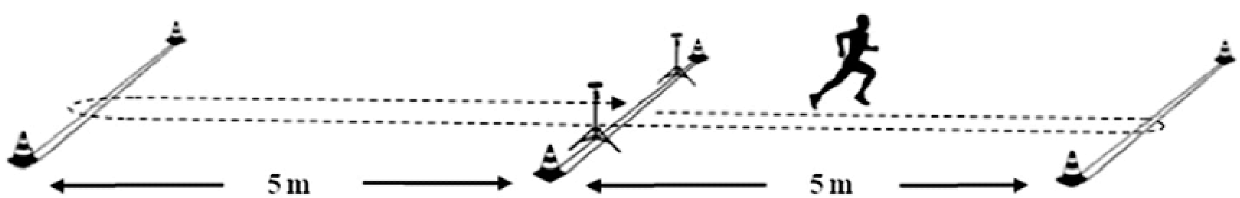


Figure 5: 20-yard shuttle visualization (20 yard shuttle test, n.d.)

For the remainder of this report, these three combine drills, the 40-yard dash, 3-cone drill, and shuttle drill, will be referred to as standard drills (Grimes, 2022). Regardless of their position, all players can run these standard drills in the combine. They serve to emulate players' movements during a game.

The NFL combine also features position-specific drills that evaluate a player's ability to execute the movements required for their position (Gridiron Elite Training, n.d.). We will refer to these as skill drills. The purpose of these drills is to replicate the specific actions and movements that players would perform during a game. Wide receivers, for instance, run various passing drills during the combine, including the in, out, and corner drills, which closely resemble their in-game routes (Landow & Jarmon, 2022). NFL teams use the results of these drills to assess a player's skill level and suitability for a specific position.

2.2.3 Metrics and Limitations

Historically, NFL Combine player analysis has relied on two sources: qualitative and quantitative evaluation of a player's Combine performance by teams. For instance, coaches and team representatives often stand at the four 10-yard intervals during the forty-yard dash, carefully noting players' split speeds. Team personnel also value interviews, which can reveal the personality, aptitude, and organizational fit. Although careful analysis can be effective, this method may be prone to subjectivity and bias (Starkey, 2016). Additionally, general metrics like the time taken to complete each standard drill are used for analysis. However, no consistent statistical relationship has been found between Combine drills and future NFL performance (Kuzmits & Adams, 2008). A separate study focused on quarterbacks, running backs, and wide receivers from 1999-2004 showed similar results, with sprint tests for running backs being the only exception. The study also acknowledged that unreported combine aspects, such as physical measurements, injury evaluations, and urine tests, may still be beneficial for drafting players (Kuzmits & Adams, 2008).

2.2.4 Current Advancements

Zebra's MotionWorks technology was recently introduced in the 2022 Combine, utilizing wearable devices on players' shoulders to collect more precise and comprehensive data during the event (Petrus, 2022). This radio-frequency identification device gathers real-time location data from all players on the field, including their speed, orientation, and positional data (NFL, n.d.). For the past nine years, this technology has served as the NFL's official on-field player tracking provider (Zebra Technologies, n.d.). Apart from tracking players, it has also been utilized for ball-tracking purposes, such as capturing data on throwing and kicking, which includes ball spin and height.

3.0 Methodology

This chapter summarizes the methodology used by our team to achieve our project goal. Specifically, we describe the data sets provided to our team and how we leveraged them to determine our project scope. Following this, we outlined the various processes we developed and followed to accomplish our two project objectives.

3.1 Project Goal & Objectives

The goal of this project is to enhance the NFL's utilization of Scouting Combine data by developing insightful metrics through data manipulation techniques, with the aim of assisting teams in making informed player evaluation decisions. In order to achieve our project goal, we made the following objectives:

1. Establish a structured approach to processing the 2022 NFL Scouting Combine data into meaningful metrics that can then be used for further analysis.
2. Determine the extent to which player movements during the 2022 NFL Scouting Combine serve as a reliable predictor of future performance in the 2022-2023 NFL season.

The first objective was centered around developing a method for inputting prepossessed drill files and outputting the metrics of the drill segments from the Scouting Combine. The second objective leverages the metrics generated in the first objective to evaluate a relationship between these metrics and NFL performance.

3.2 Determining the Scope of the Project

During the early stages of the project, our team realized we would have to develop a methodology from scratch as we were among one of the first groups to analyze NFL Scouting Combine positional tracking data. As a result, we had to carefully select which data to include and exclude based on our own research and analysis as well as advice from our sponsor.

As discussed in the background section of our report, comparing players' movements across different football positions can be challenging. Each position has unique responsibilities on the field, which requires specific skill sets, abilities, and physical attributes. For example, a defender will likely need to be more powerful than an attacking player, who may need to be more agile and reach a greater top speed.

Therefore, our team narrowed our focus to solely analyzing drill and route data for players in the wide receiver position. While this approach limited our data set to a specific position, it allowed us to have a consistent data set with consistent physical demands and movement patterns. We believed that having a uniform data set outweighed the potential loss of additional positions in our analysis.

3.3 Data Overview

The NFL provided our team with access to data and supplementary information related to the 2022 NFL Scouting Combine held at Lucas Oil Stadium in Indianapolis, Indiana. Due to the limited space on the field and the large number of players, among other reasons, the combine split players into different groups to complete the events. As a result, we were provided with data on two separate wide receiver groups, who completed each event at different times.

Three data files, each containing different types of data outlining aspects of 30 college prospects' combine performance, were provided to our team. First, the Player Info file included players' names and college information. Next, the Combine Results data file contained a player's height and weight measurements and results from the standard drills, vertical jump, and broad jump. Finally, the Combine Tracking Data file contained players' x-y coordinates, velocity, and acceleration at the timestamp. Each row within the file had a unique player ID and timestamp at a 0.1-second interval during the players' completion of the NFL Scouting Combine.

After the completion of the 2022-2023 NFL season, our team was provided with two game data files containing play-level information for wide receivers who participated in the 2022 Combine. The Game Tracking file contained in-game tracking details such as players' x and y coordinates, speed, acceleration, and the route being performed. The Plays file offered supplementary play-level information, including a textual description of the play, yards to go, the team in possession of the ball, and the absolute yard line number.

Along with the datafiles, we were given additional supplementary material that allowed our team better understanding of our project, particularly in familiarizing ourselves with the drills a prospect undergoes during the NFL Scouting Combine. A master schedule for the day was provided in a pdf form which outlined each event a prospect would be undergoing in 30-minute segments. Videos from the combine of two wide receiver prospects completing the drills aided our team in verifying drill positioning. Finally, a comprehensive README file accompanied the provided materials, offering guidance on effectively utilizing and interpreting each data set.

3.4 Objective 1: Data Preprocessing

3.4.1 Drill Detection

Since our team was only interested in focusing on the critical movements and actions performed by the players during the drills, a drill locator was developed to extract and filter the relevant data for our analysis. This function identified specific areas of drill movement during

the Combine, incorporating various relevant thresholds. This removed information such as warm-ups and water breaks, as well as drills that did not directly include horizontal movements, such as the vertical jump and bench press. The drill location process involved creating a function to parse through the tracking data set to detect movements that met the unique thresholds for each drill. This section will discuss the drill detection process for the combine drills.
First, to determine the threshold of the physical position of drills, we closely examined the provided player combine videos. By doing so, we identified the area in which drills were performed and the arrangement of cones, which varied depending on the drill. This allowed us to hone into specific areas where the drills would be located.

Then, the master schedule was used to identify the time of day each drill was conducted. It should be noted that the time slots for the two different wide receiver groups differed. Additionally, we created a duration threshold, taking into consideration the expected length of each drill. By doing so, these two thresholds helped us to exclude any unofficial attempts, such as a player practicing beforehand or a disqualification, further refining our data analysis.
Initially, we established a minimum velocity and acceleration threshold to ensure that a player met a certain performance level. Any instance of a player falling below this number would deem the drill incomplete. This approach proved effective when identifying drills exclusively involving unidirectional motion, such as the 40-yard dash. However, it became apparent for drills that involved turns, such as the in, we needed to implement a momentary stop threshold to identify when a player's acceleration or velocity dropped below the established minimum velocity and acceleration threshold. Therefore, we allowed players to fall below this momentary stop threshold for a drill-specific duration before the data collection was terminated.
In certain drills, a more comprehensive threshold was necessary to ensure accurate capture of the drill. These thresholds were designed on a drill-by-drill basis. An example of utilizing several of the drill-specific thresholds we created to pinpoint a drill is demonstrated in the case of the 3-cone drill. The following explanation details this process:

First, an initial set of thresholds is designed to capture all the movements typically associated with a three-cone drill. To isolate the location of the drill on the field, the data set was filtered only to include data within an 11x11 yard square where the drill was conducted.

To further refine the movement, collection began only after a player's acceleration exceeded 2 yards per second squared or their velocity exceeded 3 yards per second. Once the starting point was identified, the drill collection process continued until:

1. The player's velocity and acceleration fell below the momentary stop threshold.
2. The player is no longer to the right of their start position.

Following the initial capture, a second set of thresholds was applied to determine if the movement was an official attempt. To be officially collected, an isolated drill must be completed between six and eight seconds, and the starting point must be at a distance of more than 50 yards from the right side of the field. Figure 6 displays the final output of this drill location process.

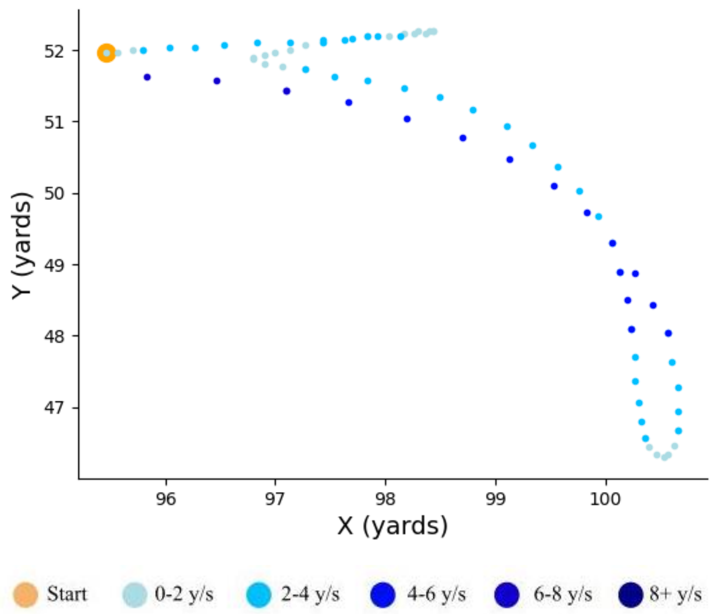


Figure 6: Isolated 3-cone drill

3.4.2 Drill Segmentation

To conduct a precise analysis of the drills, our team developed a method to segment each drill based on its movements. We identified two distinct types of movements within drills: turns and straights, each of which require different approaches to analysis. It is worth noting that some drills, such as the go and 40-yard dash (Figure 7, page 19), only incorporate straight movements, as minimal y-direction movement and no turns involved. However, all other drills involve both a straight and turn component.

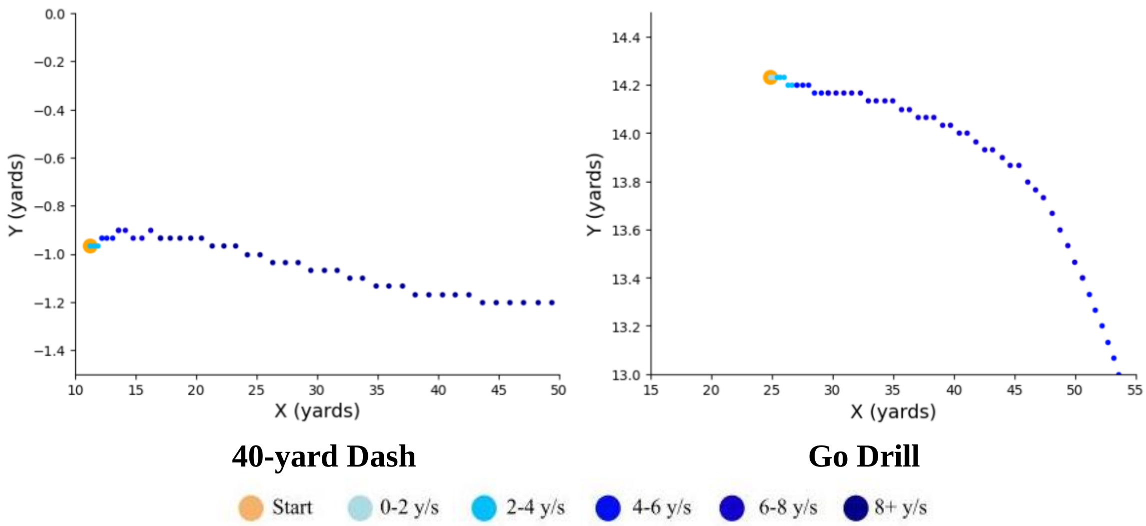


Figure 7: Straight movement-only drills

Straights refer to movements made by a player in a single direction without any turns or changes in direction. Wide receivers use their speed and agility to beat their opponents, making movements critical to assess. Similar thresholds to those employed in the drill locator were used to locate the start and end of the straight, allowing us to extract the necessary data for further analysis.

Turns, however, involve a change in direction, such as a player maneuvering around a cone. These movements are critical for evaluating a player's route-running abilities as they are performed to evade defenders. To analyze turns accurately, we implemented thresholds similar to the drill locator and the straight to evaluate the location of the turn.

The process outlined below describes how we isolated different segments of the combine in (Figure 8, page 20). For this drill, the first leg is a straight. It is captured until the velocity in the y direction is less than -1 yards per second. Once this threshold is met, turn segmentation begins. To be classified as part of the turn, the player's movement must not fall below 2.5 yards per second in the x direction and -5 yards per second in the y direction. Once these thresholds are met, any subsequent motion was considered post-turn and disregarded. A pass was typically made after the initial turn and is data that was irrelevant to our project's objectives.

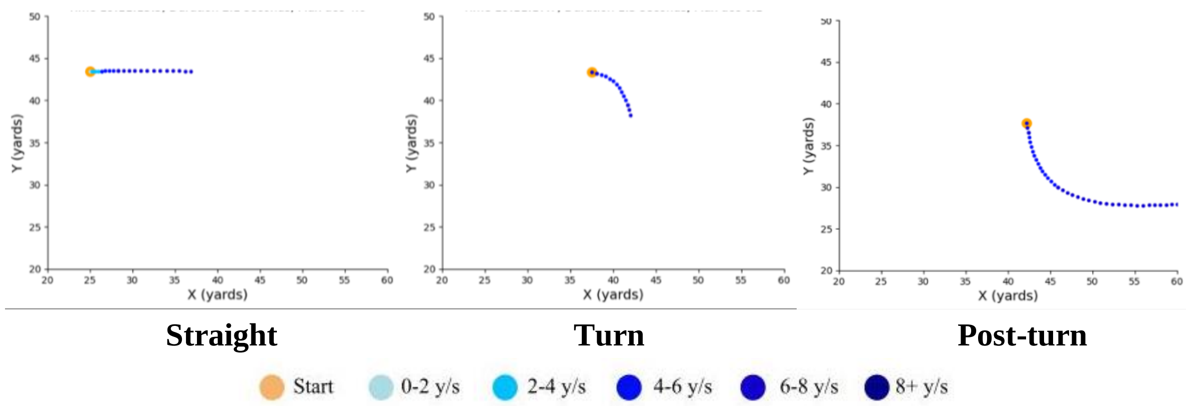


Figure 8: Three segments of the in drill obtained after segmentation

3.4.3 Metric Creation

For the next stage in our pipeline, we worked to ensure accurate comparisons between different drills. Our team developed a process of creating metrics for both straight and turn segments.

3.4.3.1 Straight Segment Metric Creation

To evaluate the performance of players in different drills accurately, we implemented a process that involved breaking straights into 5-yard intervals. This process was completed for the entire 40-yard dash and go drill, as well as the straight portions of all skill drills. By dividing each drill into intervals, we were able to compare matching segment lengths even if the overall drill lengths and overall movements were dissimilar. Segments at the end of a straight that were not an entire 5 yards were disregarded.

Once 5-yard intervals were established, we created three metrics categories to analyze a player's straights: burst, speed, and ratio. Burst metrics evaluated a player's ability to accelerate at the beginning of the movement, including average and maximum acceleration. Speed aimed to assess a player's ability to increase or maintain velocity, therefore, include movement measurements such as average and maximum velocity. Finally, we also created ratio metrics to facilitate comparisons of metric relationships across straights. An example of a ratio use case would be to compare the performance of the first 10 yards versus the last 10 yards of the 40-yard dash. This metric aimed to analyze how a player's metrics evolved throughout a drill. Table 1 on page 21 shows the value of the metrics of a straight for the first 0-5 yards of the 40-yard dash in Figure 9, shown on page 21.

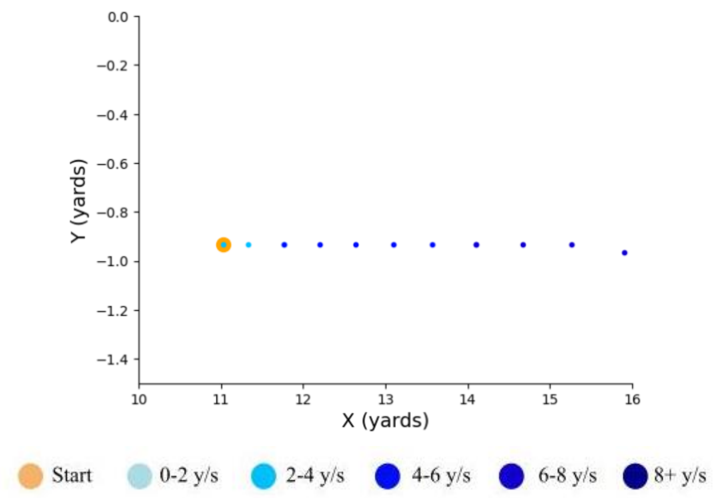


Figure 9: First 5 yards of the 40-yard dash

X Velocity (y/s)	X Max Velocity (y/s)	Position of Max Velocity (yd)	X Acc. (y/s²)	X Max Acc. (y/s²)	Position of Max Acc. (yd)	Y Position Mean (yd)	Y Position Variance (yd)	Total Time (s)
5.8	7.3	5.9	2.4	2.9	4.7	-0.9	1 E-4	1.0

Table 1: Example metrics of the first 5 yards of the 40-yard dash

3.4.3.2 Turn Segment Metric Creation

Creating metrics for turn segments posed a greater challenge compared to the straight metrics, as turns exhibited a greater variability in length and direction. As a result, we divided turns into two categories: right running turns and left running turns, to prevent the comparison of velocities of turns that pointed in opposite directions.

Then, we began creating metrics aimed at evaluating the movements a player performed during the process of their turn. These metrics were similar to the burst metrics created for the straight segments and were designed to assess a player's ability to accelerate during a turn. For instance, we collected the average and maximum acceleration metrics at the entry and exit points of a turn. Moreover, additional metrics that measured players turning ability by noting the angle of turn, change of position throughout the turn, as well as the average and maximum velocity achieved throughout the turn’s duration. Additionally, we generated ratio metrics to enable comparison of the same turn metrics across the same or different drills. Table 2 on page 22 shows some of the metrics values for the turn segment of the out drill in Figure 10 on page 22.

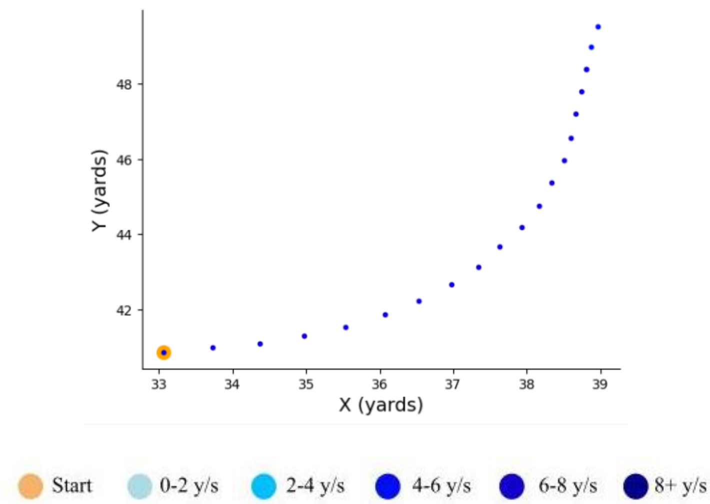


Figure 10: Turn segment of the out drill

Velocity Enter (y/s)	Velocity Exit (y/s)	Acc. Enter (y/s^2)	Acc. Exit (y/s^2)	Change in Position (yd)	Angle (°)
5.8	5.9	7.3	2.4	2.9	1.0

Table 2: Example metrics of an out drill turn

Overall, the development of these turn and straight metrics hold significant value in assessing a player's performance and potential as they closely resemble the typical movements executed by wide receivers in-game. Providing these metrics of evaluation of a receiver's performance, it allows for easier to compare their in-game performance to their combine results, leading to more informed decisions regarding their potential and development.

3.5 Objective 2

3.5.1 Game Data Preprocessing

The first step toward accomplishing objective two was to implement a metric creation pipeline from in-game data using a methodology similar to that of objective one with combine data. This was necessary as it enabled direct comparisons between in-game and combine metrics and provided a more comprehensive evaluation of player performance.

The data we received for in-game analysis was comparable to the combine data. It contained information on the entire play, starting before the quarterback snapped the ball and continuing until the play was completed.

The combine data contained information regarding the velocity and acceleration in the x and y directions, whereas the in-game data did not differentiate between the x and y directions. However, the in-game positional data did contain two-dimensional coordinates. Therefore, to draw a direct comparison between metrics, the first step in preprocessing was to calculate velocity and acceleration by using general physics equations. In doing so, the velocity and acceleration data did not have the same smoothing algorithms applied to them as the tracking data supplied directly from Zebra Technologies and was rounded to one decimal.

From there, we worked to isolate only the portion of the players' movement while they were running a route from the larger file that encompassed all movements. For our project's purposes, we define a route as the movement from the moment the ball is snapped until the event of a pass outcome was ruled to be caught or incomplete. This process was undergone to ensure consistency between the in-game data and combine data, as well as to avoid potential external factors that could affect pre-route and post-catch movements and not accurately reflect route running ability.

Once the prior steps were completed to produce a data set that contains only relevant in-game route-specific data, we needed to ensure that all the data was oriented in the same direction to facilitate a standardized baseline for comparison. To do this, we first established that all routes begin at the origin. Next, unlike the combine where routes are only run left to right, players run routes towards either end zone during games. Therefore, to ensure consistency, we inverted all routes ran from right to left.

After implementing these changes to the original data set, a more uniform and systematic approach to analyzing and exploring the players' in-game movement data became possible. Following this preprocessing, we applied a similar pipeline of segmenting and creating metrics as outlined in objective one. This allowed us to compare the performance of wide receivers during the combine drills with their in-game performance metrics.

3.5.2 Predicting Combine Performance based on Game Data

After completing the preprocessing of both the combine drills and in-game routes, our focus shifted to exploring the extent to which player movements during the NFL Scouting Combine can predict future performance in the NFL.

We initially performed correlations as they represent the linear relationships between different metrics. We opted to perform this analysis on drills with similar movement patterns, such as the go drill and 40-yard dash. Our team also implemented correlation heat maps to compare the multiple metric correlations against each other at once and detect trends within different groups of metrics. Using this visual representation of the data made it simpler to pinpoint areas of interest.

While it was possible to compare all segments of the same type, such as a turn to another turn, we limited our comparisons to only compared segments with similar movements. For example, in the 3-cone drill, a 180-degree turn is completed around cone three. This movement is different from the turn performed during out, as shown in Figure 11. Although both are considered turns, we did not compare them. When areas of high correlations between metrics were found, scatterplots were used to represent the data visually and to identify trends and patterns.

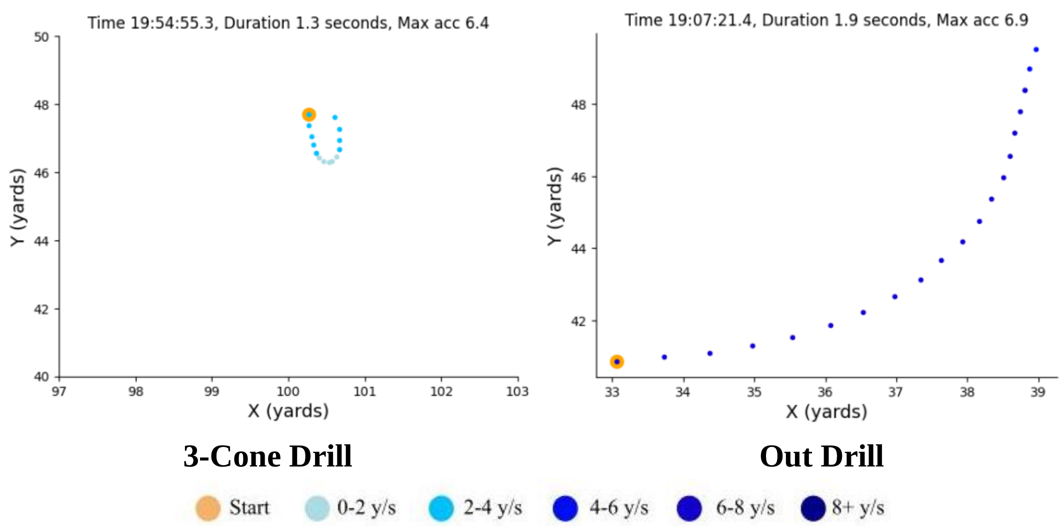


Figure 11: 3-cone turn vs. out drill turn

Following this initial assessment of the relationship between the in-game and combine metrics, we began to explore other techniques that could be used to increase our sample size. This was necessary as throughout the completion of our correlation, we were only able to use data from a maximum of 14 players. Although we could not change the sample size of the players, we realized that these players ran anywhere from three to over 100 in-game routes. By adjusting the data structure to make the route the target variable, we could subsequently restructure our data set. As a result, our new approach was to merge our movement analysis with in-game data, ultimately working to determine if movement analysis enhanced route prediction.

As depicted in Table 3 on page 25, our new data set structure comprised in-game data, movement analysis data, and additional information extracted from the 2022 combine published on the NFL Next Gen website. This information included the player's height, weight, and athleticism scores. As previously stated, the target variable of this data set would be the route the player ran. We limited routes to only include the go, corner, in, out, and hitch. We decided to use all instances of a player running in-game routes, even if they only ran a limited number.

Player	Route	Down	Yards To Go	Absolute Yardline Number	Y	Side of Field	Combine Statistics	Burst Metrics	Speed Metrics	Left Turn Metrics	Right Turn Metrics
1	GO/ CORNER	1	10	40	29	Left	…	…	…	…	…
2	HITCH	1	10	60	4	Right	…	…	…	…	…

Table 3: Revised data set structure

3.5.3 Official Baseline model

Initially, we created a baseline model to serve as a reference point. This model was used to assess the performance of other models after incorporating movement analysis data. The data set for the model was created by merging features from two sources. From the in-game data, we included the y-position of the player, side of field, down, yards to go, and absolute yard line. These features allowed our model to understand the circumstances surrounding the play, therefore assisting it in predicting what route a player may run. For example, if a team is on 4th down and 20 yards away from the first down line as opposed to a team on 2nd down and 1 yard away from the first down line, the wide receivers will run different routes accordingly. Additionally, 11 features scraped from the NFL Next Gen website were included in the baseline model. These features provided the model with additional insight into the non-positional aspects of the players' combine performance, such as their achieved vertical jump height.

3.5.4 Movement Analysis Model Iterations

After our baseline model was established, we incorporated our movement analysis data and implemented many different techniques to optimize our models and maximize its prediction power. For example, a fundamental step in any machine learning prediction is to divide the data into training and testing sets. Our reshaped model contained over 1800 instances of data; however, this data was from only 16 players. This meant, for each unique in-game play, the same combine data was associated with it, as players had only one instance of combine data available for them. Therefore, to prevent data snooping, our team initially split the data into 80% training and 20% testing, ensuring no overlap between players within these two subsets.

Next, our team merged the go and corner routes into one target variable due to the class imbalance in our data set. The corner route had a small number of routes run in-game, and since both the go and corner tend to be deeper infield routes, combining the two makes sense for this use case.

Handling this class imbalance proved challenging, causing us to experiment with various models, such as multinomial logistic regression and support vector classification. After careful consideration, we employed the CatBoost model as our primary choice model due to its gradient boosting and a known ability to handle imbalanced data sets well. To further increase the prediction power, many common machine learning techniques were applied to our CatBoost model, as described below:

- **Hyperparameter tuning**: Involves selecting user-defined parameters to optimize a given learning algorithm. In the context of our model, we experimented with different parameters, including but not limited to the learning rate, decision tree depth, and regularization technique used.
- **Class weighting**: Penalizing the model for misclassifying the minority class. Due to our imbalanced data set, we believed introducing this technique would help prevent our model from heavily predicting the majority class (go/corner).
- **Early stopping**: Stopping the learning algorithm prior to the completion of all iterations. This technique was implemented to prevent overfitting.
- **Model Structure**: Split our data set into two distinct classification models: the majority class (go/corner) and all others. Initially, the model predicted if a route was a go/corner. Then all routes not predicted to be a go/corner were fed into a separate model to classify the remaining minority classes.

Despite implementing these techniques, our model's performance did not significantly improve compared to the baseline. It continued to predict the majority class at a high proportion compared to its overall percentage of the data set. As a result, we re-evaluated our training set and created a new one consisting of 100 go/corner routes, 100 in routes, 100 out routes, and 100 post routes completed by 4 out of the 16 players in our data set. Upon implementation, our team developed our final movement analysis model based on the model with the best weighted accuracy in our testing data.

Our team utilized data analysis and machine learning techniques to develop models that predict in-game routes. A CatBoost model was chosen as the primary analysis choice, with the route as the target variable. The baseline model combined features from two sources, while the movement analysis model incorporated the addition of combine drill metrics as features. Feature analysis was performed to select the optimal combination of metrics, yet there was a limited increase in the weighted accuracy of the model.

4.0 Results

This section will discuss our findings for our two objectives. First, our ability to develop a comprehensive 2022 NFL Scouting Combine data analysis pipeline, and second, the degree to which a player's combine performance correlates with their in-game performance. In doing so, we will determine the impact of our metrics on the predictions, utilizing a multi-classification CatBoost model, which is a powerful gradient-boosting learning model.

4.1 Data Pipeline

The pipeline we developed was assessed on its ability to complete the three fundamental tasks: drill finding, drill segmentation, and drill metric creation.

Our pipeline excelled in drill finding, where it was able to extract the relevant data frames containing only the desired drill information from the raw data of the combine. It was able to identify and extract every standard drill that was performed, showcasing its capability to handle large datasets and filter out irrelevant data. However, we did not have access to records of which players performed which skills drills. This limited our ability to perform a comprehensive analysis as it is difficult to verify the correctness of the results without knowing a player's drill attempts. Table 4 presents the number of players detected by our pipeline for the standard drill

Drill	**Number of Players Ran**	**Number of Players Found**
40-yard dash	29	29
3-cone drill	13	13
Shuttle	12	12

Table 4: Drill location accuracy of standard drills

We were able to identify the drills within a few tenths of a second of the official recorded time from the Combine. This level of accuracy, shown in Table 5, is impressive, considering the complexity and variability of the NFL Combine data.

	Forty Yard Dash			Shuttle			Three Cone		
Player	Our Time	Official Time	Difference	Our Time	Official Time	Difference	Our Time	Official Time	Difference
1	4.4	4.43	0.03	-	-	-	7.3	7.28	-0.02
2	4.5	4.55	0.05	-	-	-	7.4	7.28	-0.12
3	4.7	4.41	-0.29	4.7	4.32	-0.38	7.2	7.13	-0.07
4	4.4	4.31	-0.09	-	-	-	-	-	-
5	4.7	4.65	-0.05	5.5	4.57	-0.93	7.1	7.14	0.04
6	4.3	4.33	0.03	-	-	-	-	-	-
7	-	-	-	4.7	4.38	-0.32	-	-	-
8	4.4	4.32	-0.08	4.6	4.07	-0.53	6.6	6.65	0.05
9	4.6	4.5	-0.1	4.7	4.28	-0.42	7.0	7.03	0.03
10	4.4	4.34	-0.06	-	-	-	7.0	6.98	-0.02
11	4.8	4.62	-0.18	4.9	4.42	-0.48	7.7	7.28	-0.42
12	4.8	4.66	-0.14	-	-	-	-	-	-
13	4.5	4.43	-0.07	4.6	4.15	-0.45	6.9	6.71	-0.19
14	4.6	4.6	0	4.9	4.54	-0.36	7.4	7.21	-0.19
15	4.9	4.65	-0.25	-	-	-	-	-	-
16	4.6	4.53	-0.07	-	-	-	-	-	-
17	4.6	4.38	-0.22	4.4	4.36	-0.04	-	-	-
18	4.6	4.39	-0.21	-	-	-	-	-	-
19	4.6	4.36	-0.24	-	-	-	-	-	-
20	4.4	4.44	0.04	-	-	-	-	-	-
21	4.3	4.28	-0.02	-	-	-	-	-	-
22	4.4	4.47	0.07	-	-	-	-	-	-
23	4.7	4.49	-0.21	-	-	-	7.2	7.08	-0.12
24	4.6	4.49	-0.11	4.8	4.21	-0.59	7.3	7.28	-0.02
25	4.9	4.58	-0.32	-	-	-	-	-	-
26	4.5	4.55	0.05	-	-	-	-	-	-
27	4.5	4.59	0.09	3.8	4.36	0.56	-	-	-

	Forty Yard Dash			Shuttle			Three Cone		
Player	Our Time	Official Time	Difference	Our Time	Official Time	Difference	Our Time	Official Time	Difference
28	4.8	4.65	-0.15	-	-	-	-	-	-
29	4.4	4.57	0.17	-	-	-	-	-	-
30	4.6	4.51	-0.09	5.0	4.53	-0.47	7.4	7.54	0.14
Average	4.56	4.48	-0.08 (-1.8%)	4.71	4.35	-0.37 (-8.5%)	7.19	7.12	-0.07 (-1.0%)

Table 5: Drill locator time accuracy

In addition to filtering data for the drills, the locator programs also provided visualizations, two examples are displayed in Figure 12.

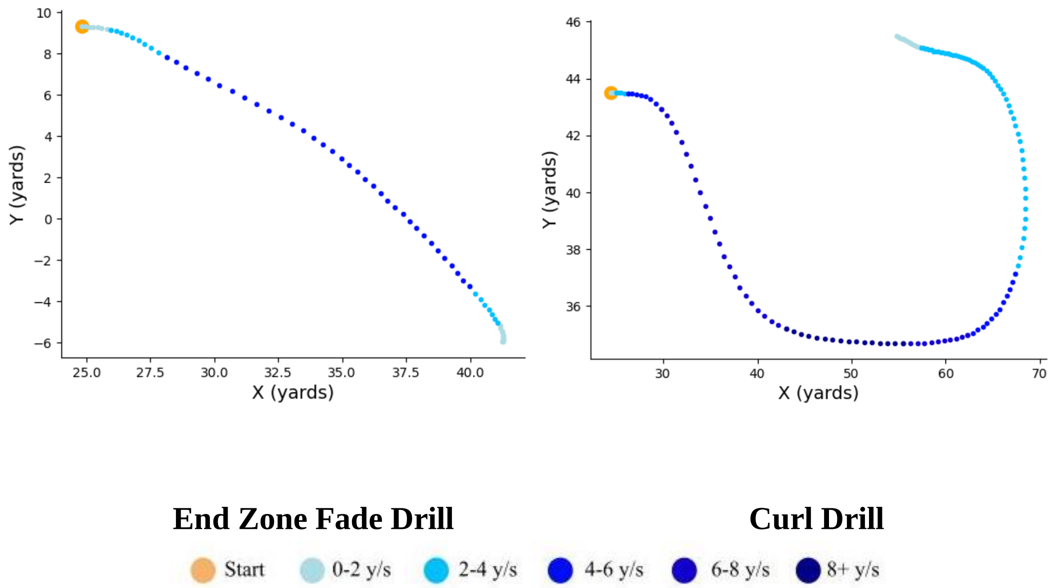


Figure 12: Example locator outputs

Our pipeline was able to segment every straight and turn for all standard and skill drills required for the scope of our analysis. For example, the 3-cone drill is a relatively complex drill to segment as it involves multiple turns compared to other drills such as the in and out, which only have one straight followed by a turn. Despite this added complexity, our pipeline was able to segment it to our expectations as seen in Figures 13 & 14, shown on page 30.

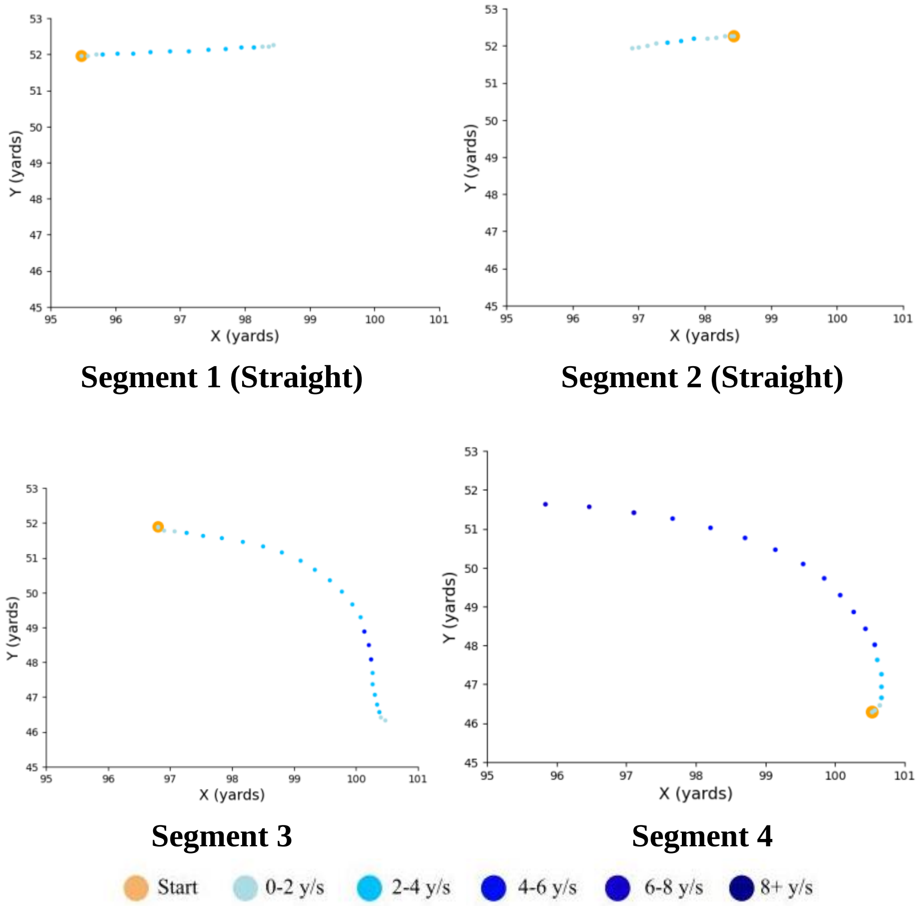


Figure 13: 3-cone drill segmentation

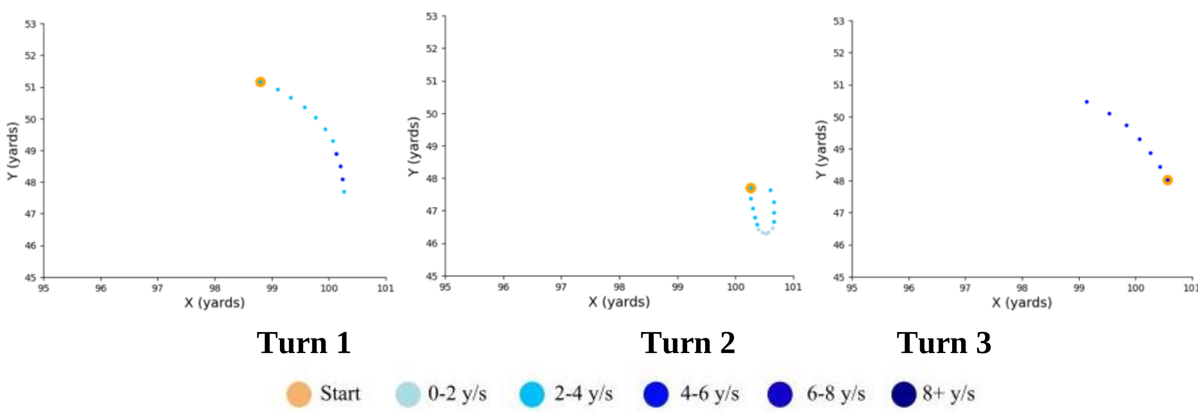


Figure 14: 3-cone drill turn segmentation

Additionally, our segmentation model performed well on drills without a strong turn, such as the slant. This further underscore the pipeline's flexibility and potential for handling a wide range of NFL Combine drills. The metric creation section of the pipeline takes in either a straight or turn segment and returns the proper metrics.

Overall, the pipeline created effectively fulfilled the three primary tasks it was designed for, as well as overarchingly accomplishing Objective One. It has also been thoroughly documented to easily facilitate modifications that may arise to suit the unique requirements of future combines.

4.2 Describing NFL Performance using Combine Metrics

The initial data set contained thirty wide receivers from the combine, but various factors decreased the sample size when comparing the combine and in-game data. For instance, some combine wide receivers did not run any in-game routes during the 2022-2023 NFL season, resulting in the absence of a target variable for these players in our analysis.

Out of the sixteen total players who did run in-game routes, two players of these ran less than five total routes. Along with this, of the fourteen players who ran more than five in-game routes, three players did not run any of the combine skill drills. As a sample size of fourteen would be too small for any advanced modeling or prediction, our team had to use simple methods to evaluate the relationship between in-game performance and combine performance.

Although limited by a small sample size, correlation analysis was still able to be performed, serving as an initial assessment of the relationship between the same metric types. For example, the go drill and 40-yard dash, which both involve players running in a straight line at a high speed can be compared to players running a go route in a game. Examples of the results of correlation analysis are presented in Tables 6 & 7 on page 32, where a value close to 1 or -1, displayed in dark red, indicates a strong linear relationship between the pair of metrics. It is important to understand that the sample size hinders the validity of these relationships.

		Velocity			Acceleration			Y Position		
No. of Players	Yards	Average Velocity	Max Velocity	Position of Max Velocity	Average Acceleration	Max Acceleration	Position of Max Acceleration	Average Y Position	Y Position Variance	Time
14	0-2.5	0.06	0.55	0.41	0.53	0.32	0.51	0.21	-0.03	-0.18
14	0-5	0.25	0.81	0.49	0.58	0.15	0.46	0.21	-0.13	-0.21
14	5-10	0.69	0.67	0.03	-0.25	-0.3	-0.45	-0.02	0.14	0.36
14	0-10	0.66	0.91	-0.38	0.78	0.28	-0.17	0.04	-0.16	-0.07
14	10-15	0.83	0.75	-0.39	-0.11	-0.07	-0.13	-0.24	-0.26	0.32
13	15-20	0.43	0.37	-0.19	0.34	0.07	0.06	-0.63	-0.04	0.32
13	10-20	0.53	0.4	-0.42	0.36	0.34	-0.19	-0.6	-0.17	0.62
13	20-25	0.42	0.36	0.09	-0.05	-0.21	0.19	0.04	-0.11	0.00
10	25-30	-0.07	0.29	-0.07	-0.07	-0.13	0.63	-0.09	-0.09	0.06
10	20-30	0.19	0.43	-0.28	-0.21	-0.08	-0.49	-0.15	0.23	0.25
8	30-35	0.44	0.49	0.46	0.32	0.10	-0.18	0.05	-0.51	-0.53

Table 6: 40-yard dash vs. go route correlations

		Velocity			Acceleration			Y Position		
No. of Players	Yards	Average Velocity	Max Velocity	Position of Max Velocity	Average Acceleration	Max Acceleration	Position of Max Acceleration	Average Y Position	Y Position Variance	Time
11	0-2.5	0.03	0.12	-0.14	0.08	0.00	-0.69	0.62	-0.67	-0.44
11	0-5	0.19	0.49	-0.11	0.17	0.14	-0.37	0.61	-0.7	-0.23
11	5-10	0.11	0.46	0.12	0.33	-0.22	-0.38	0.33	0.45	-0.27
11	0-10	-0.06	0.24	-0.37	-0.25	-0.06	-0.25	0.67	-0.59	-0.47
11	10-15	0.01	0.48	0.11	-0.07	0.28	0.34	0.73	0.32	0.00
10	15-20	0.16	0.40	-0.13	0.24	0.10	0.05	0.00	0.03	-0.21
10	10-20	0.17	0.35	-0.37	0.01	-0.16	-0.31	0.04	-0.39	-0.11
10	20-25	0.31	0.40	0.30	-0.06	0.23	-0.08	0.33	-0.37	0.09
7	25-30	0.6	0.40	-0.11	-0.41	-0.45	0.15	0.03	-0.54	0.43
7	20-30	0.62	0.40	-0.18	-0.33	0.12	-0.21	0.03	-0.47	0.42

Table 7: Go drill vs. go route correlations

The heatmaps in both figures above reveal interesting correlations between certain metrics, such as the moderate to strong correlation between the average velocity and max velocity columns in Table 6 for about the first 20 yards of the drill. To investigate these relationships, scatterplots were created for metrics of moderate correlation. Figure 15 outlines the relationship between the max velocity of the 40-yard dash and the go route in five-yard segments for the first twenty yards.

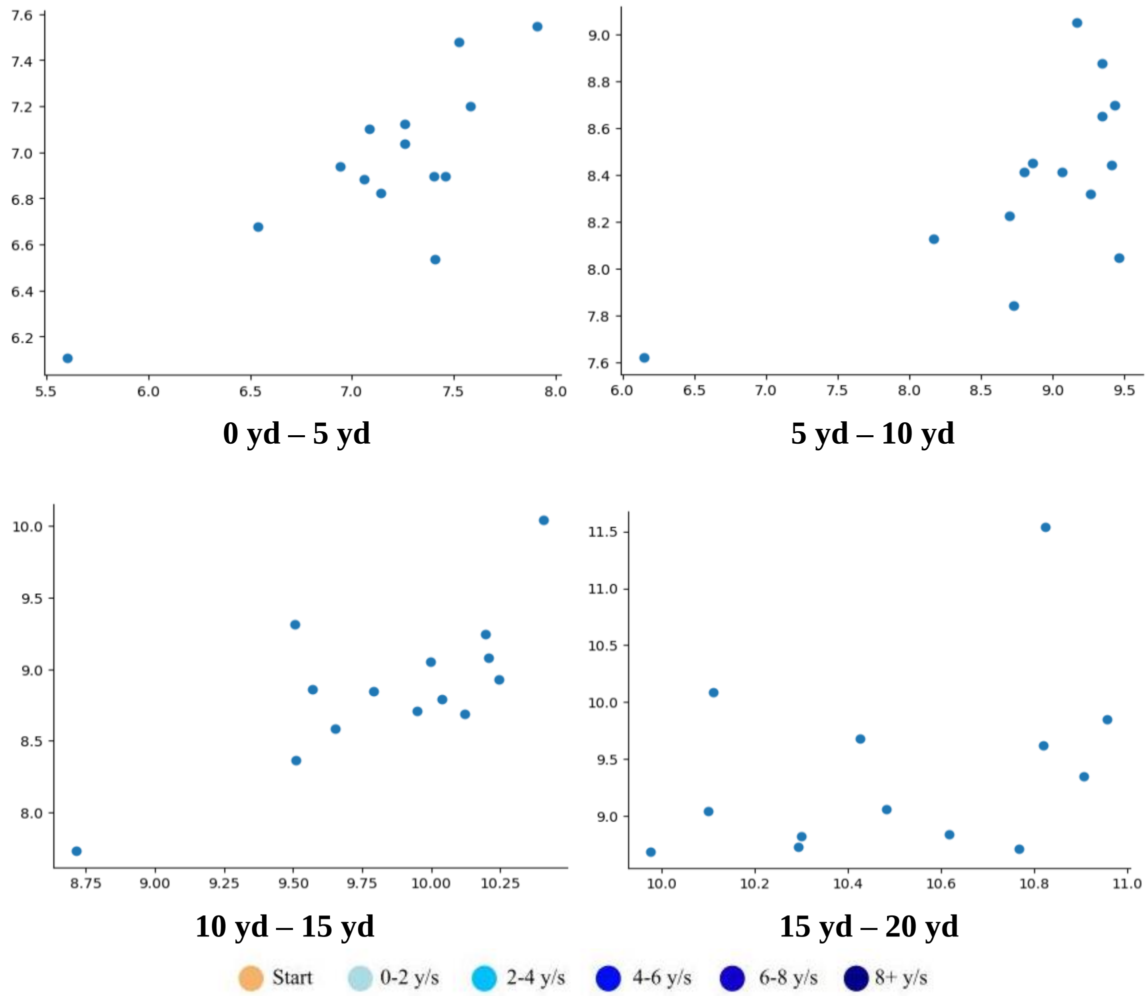


Figure 15: Correlation scatter plots of 40-yard dash vs. go route max velocity

Despite outliers shown in the scatter plots, there does seem to be a moderate correlation between these combine and in-game metrics, indicating a potential relationship. However, it is crucial to remember our sample size was at most 14 players and we, therefore, cannot make any conclusions based on the results.

4.3 Route Prediction using Combine Data

After initial comparisons, our team shifted focus determining if a player's combine movements are indicative of the routes they run in-game. Weighted accuracy was used as the main metric of comparison between the baseline model (Figure 16) and the movement analysis model (Figure 17), as there is a significant class imbalance.

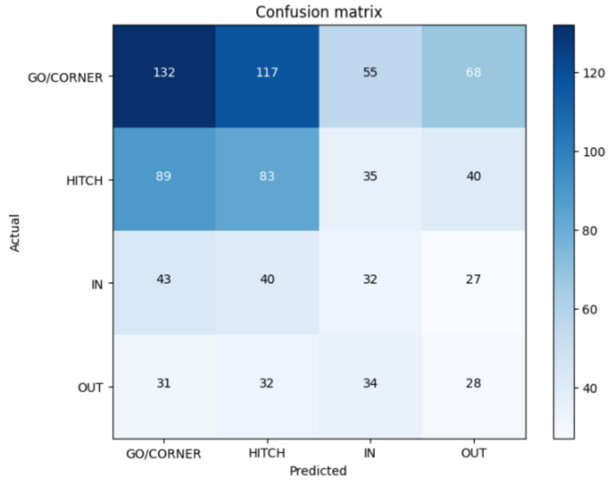


Figure 16: Baseline model

The baseline model has a weighted accuracy of 32%, and subsequent models were compared to this to evaluate their performance. Despite experimenting with various techniques such as early stopping, hyperparameter tuning, loss function creation, and balanced sampling, the model displayed in Figure 17 was ultimately determined to be the best.

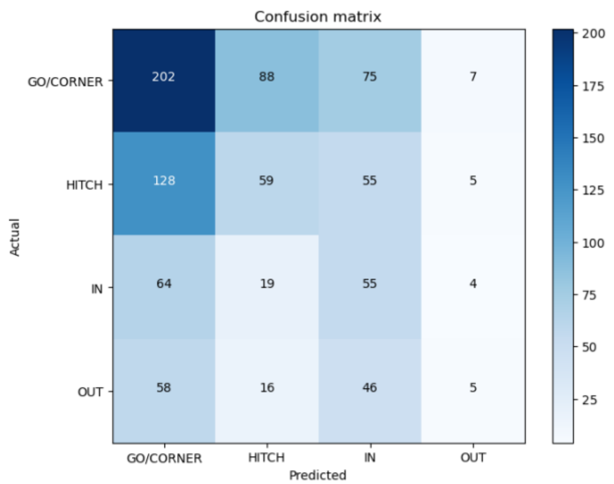


Figure 17: Movement analysis model

The chosen movement analysis model increased by two percentage points to 34%, indicating limited or no improvement in the prediction power of our movement analysis model.

As seen throughout this results section, our limited sample size had a large impact on our team's ability to make definitive results. Without the input of more data, any conclusions drawn based on movement analysis data would be limited to this same obstacle.

5.0 Conclusion

Our project aimed to enhance the NFL's utilization of 2022 Scouting Combine data by developing insightful metrics through data manipulation techniques to assist teams in making informed player evaluation decisions. This was the first-year players' movements were tracked using positional data in the combine. The NFL had not created a way to analyze this data and use it for analysis, so our goal was evaluated by the completion of two objectives:

1. Establish a systematic approach to processing the NFL Scouting Combine data into meaningful metrics that can then be used for further analysis.
2. Determine the extent to which player movements during the NFL Scouting Combine serve as a reliable predictor of future performance in the NFL.

Once these objectives were identified, our team proposed innovative ways of completing both objectives. We implemented an NFL Scouting Combine drill preprocessing pipeline that takes the four-hour data set of all of the player's combine movements and outputs a drill the user requests. This pipeline contains adjustable metrics so that it can preprocess future combines. After preprocessing, our team performed many approaches to evaluate the relationship between a player's combine performance and their movements while running a route in a game. However, our small sample size limited us from presenting strong evidence for any of our conclusions for objective two.

Although our data could not support any conclusions for objective two, our preliminary findings raise questions about the relevance of the combine. Examples of these questions are:

1. Are the findings from the combine movement analysis consistent with how players are utilized in-game?
2. Is analyzing a player's peak performance the most effective approach to player analysis?
3. Is there a need for more comprehensive metrics beyond burst, speed, and turn to assess a player's movement and athleticism?

We recommend the NFL continue collecting this data for the combine each year. With this collection, we believe the league should focus on improving the quality and accuracy of the data. Overall, our project was successful in generating movement metrics from combine data and evaluated the relationship between these metrics and their in-game equivalents effectively, given our sample size.

Citations

Adler, J. (2019, November 4). The Basics of Football. LiveAbout. https://www.liveabout.com/football-101-the-basics-of-football-1333784

BlazePod. (n.d.). Pro agility (5-10-5) test: How to master it. https://www.blazepod.com/blogs/test/pro-agility-5-10-5-test-how-to-master-it

Britannica, T. Editors of Encyclopaedia (2023, April 25). National Football League. Encyclopedia Britannica. https://www.britannica.com/topic/National-Football-League

Bowen, M. (2017, October 03). NFL 101: Breaking down the basics of the route tree. April 20, 2023, https://bleacherreport.com/articles/2016841-nfl-101-breaking-down-the-basics-of-the-route-tree#:~:text=The%20NFL%20route%20tree%20is,receivers%20run%20on%20passing%20plays

Champion, G. & Varanasi, S. (2022, March). Next Gen Stats: Going Beyond Top Speed. University of Chicago.

Davenport, G. (2013, February 25). How Are 40-Yard Dash Times Recorded? Bleacher Report. https://bleacherreport.com/articles/1543670-how-are-40-yard-dash-times-recorded

Doll, T. (2013, February 12). Some Clarification is in Order: Average Speed by Position. Mile High Report. https://www.milehighreport.com/2013/2/12/3969128/some-clarification-is-in-order-average-speed-by-position

Gough, C. (2023). Interest level in football in the United States as of March 2022, by age group. In Statista. https://www.statista.com/statistics/1098885/interest-level-football-age/

Gridiron Elite Training. (n.d.). Football Combine Drills – Football Training. https://gridironelitetraining.com/american-football-combine-drills/

Grimes, P. J. (2022, March 3). The 3 cone drill, shuttle run and other NFL combine drills explained. USA Today For The Win. https://ftw.usatoday.com/lists/nfl-combine-drills-three-cone-shuttle-explained.

Hope, D. (2023, March 4). Jaxon Smith-Njigba Runs Fastest 3-Cone Drill and 20-Yard Shuttle at 2023 NFL Scouting Combine. Eleven Warriors. https://www.elevenwarriors.com/ohiostate-football/2023/03/137689/jaxon-smith-njigba-runs-fastest-3-cone-drill-and-20-yard-shuttle-at-2023-nfl-scouting-combine

Kuzmits, F. E., & Adams, A. J. (2008). The NFL combine: does it predict performance in the National Football League?. *Journal of strength and conditioning research, 22*(6), 1721–1727. https://doi.org/10.1519/JSC.0b013e318185f09d

Landow, L., & Jarmon, C. (2022, July 5). Wide Receivers Specific Combine Drills. CoachesInsider. https://coachesinsider.com/football/wide-receivers-specific-combine-drills/

Moore, T. (2022, October 18). NFL Revenue: $18 Billion & Keeps Going Up, But Quality Of Play (3-3 Teams) Keeps Getting Sacked. Forbes. https://www.forbes.com/sites/terencemoore/2022/10/18/nfl-revenue-18-billion-keeps-going-up-but-quality-of-play-3-3-teams-keeps-getting-sacked/?sh=262b81d06816

National Football League. (2023, February 15). Global audience of more than 56 million watch Super Bowl LVII. https://www.nfl.com/news/global-audience-of-more-than-56-million-watch-super-bowl-lvii

National Football League. (2023). Teams. Retrieved April 20, 2023, from https://www.nfl.com/teams/

NBC Sports. (2023, January 1). 2022 NFL Regular Season Schedule: How to Watch, Live Stream, Dates, Times, Matchups. https://sports.nbcsports.com/2023/01/01/2022-nfl-regular-season-schedule-how-to-watch-live-stream-dates-times-matchups/

NFL. (n.d.). NFL Next Gen Stats. Retrieved April 20, 2023, from https://operations.nfl.com/gameday/technology/nfl-next-gen-stats/

NFL Combine Trainer: 3-Cone Drill. (n.d.). Bodybuilding.com. https://www.bodybuilding.com/fun/nfl-combine-trainer-3-cone-drill.html

NFL map: Teams: Logos - sport league maps: Maps of sports leagues. (2023, March 20). Retrieved from https://sportleaguemaps.com/football/nfl/

NFL Scouting combine. (n.d.). Retrieved April 20, 2023, from https://operations.nfl.com/journey-to-the-nfl/the-next-generation-of-nfl-stars/nfl-scouting-combine/

NFL Scouting Combine History. (n.d.). Retrieved April 20, 2023, from http://www.nflcombine.net/history/

Nogle, K. (2016, June 20). Football 101: Wide receiver route tree. Retrieved April 20, 2023, https://www.thephinsider.com/2016/6/20/11975890/football-101-wide-receiver-route-tree

Petrus, A. (2022, March 30). Chalk Talk, Episode 2: How RFID Tech Used to Track Player Movements at 2022 NFL Combine. Zebra Technologies. https://www.zebra.com/us/en/blog/posts/2022/chalk-talk-episode-2-how-rfid-tech-used-to-track-player-movements-at-2022-nfl-combine.html

Rookie Road. (n.d.). Offense. In Football 101. https://www.rookieroad.com/football/101/offense/

Rookie Road. (n.d.). What does a wide receiver do? https://www.rookieroad.com/football/questions/what-does-a-wide-receiver-do/

Sidhu, D. (2022, February 17). NFL Combine: What you need to know about the shuttle drills. Houston Texans. https://www.houstontexans.com/news/nfl-combine-what-you-need-to-know-about-the-shuttle-drills

Starkey, B. S. (2016, August 16). Implicit bias and the NFL draft. Andscape. https://andscape.com/features/implicit-bias-and-the-nfl-draft/

Rogust, S. (2022). NFL Combine: What is the 3-cone drill? FanSided. https://fansided.com/2022/03/03/nfl-combine-3-cone-drill/

Zebra Technologies. (n.d.). Sports Player Tracking | RTLS in Sports | Zebra Technologies. Zebra Technologies. https://www.zebra.com/us/en/solutions/intelligent-edge-solutions/rtls/sports-player-tracking.html

20 yard shuttle test. (n.d.). Retrieved April 26, 2023, from https://www.topendsports.com/testing/tests/shuttle-20yard.htm#:~:text=20%20Yard%20Shuttle%20Test,and%20back%20again%205%20yayar.

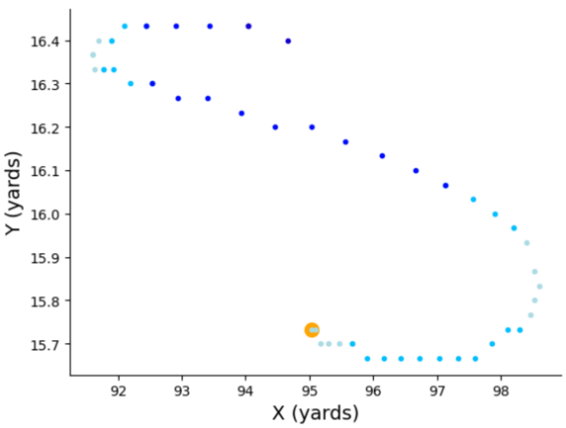
3-cone shuttle drill test (L-drill). (n.d.). Retrieved April 26, 2023, from https://www.topendsports.com/testing/tests/3-cone-drill.html

40 yard dash. (n.d.). Retrieved April 26, 2023, from https://www.topendsports.com/testing/tests/sprint-40yards.html

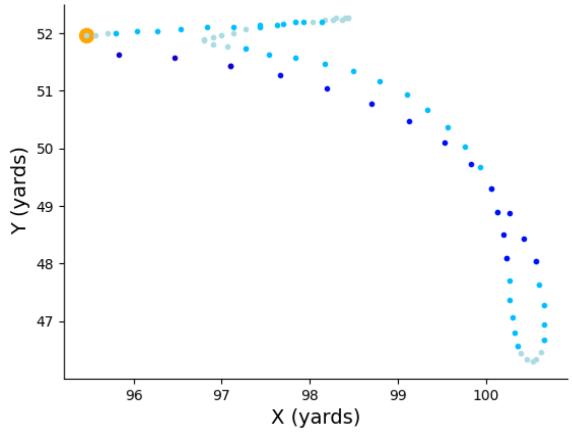
Appendix A: Movement Analysis Model Features

	Baseline Features	In/Out Burst Metrics	Forty Speed Metrics	Left Turn Metrics	Right Turn Metrics	Forty Burst Metrics
Category Description	Statistics that were previously known prior to the movement analysis of this project.	Metrics that were calculated from the first 5 yards of the in and out routes run in the combine (pre-turn).	Positional, speed and duration metrics from the 10yd - 20yd segment of the 10-yard dash.	Metrics calculated from the out drill run in the combine that reflect this type of motion.	Metrics calculated from the out drill and in drill run in the combine that reflect this type of motion.	Metrics calculated from the first 5 yards of the 40-yard dash.
Features	down	vel_x_burst	vel_x_10_20	vel_x_enter_left	vel_x_enter_right	vel_x_forty_burst
	yards _to_go	acc_x_burst	acc_x_10_20	vel_y_enter_left	vel_y_enter_right	acc_x_forty_burst
	absolute_yardline_number	time_burst	time_10_20	acc_x_enter_left	acc_x_enter_right	time_forty_burst
	y	max_vel_x_burst	max_vel_10_20	acc_y_enter_left	acc_y_enter_right	max_vel_x_forty_burst
	side_of_field	pos_vel_x_burst	pos_vel_10_20	vel_x_end_left	vel_x_end_right	pos_vel_x_forty_burst
	Height in Inches	max_acc_x_burst	max_acc_10_20	vel_y_end_left	vel_y_end_right	max_acc_x_forty_burst
	Weight	pos_acc_x_burst	pos_acc_10_20	acc_y_end_left	acc_y_end_right	pos_acc_x_forty_burst
	Prospect Grade	mean_y_burst	mean_y_10_20	change_x_left	change_x_right	mean_y_forty_burst
	NFL Stats Production Score	variance_y_burst	variance_y_10_20	change_y_left	change_y_right	variance_y_forty_burst
	NFL Stats Athleticism Score			angle_left	angle_right	
	NFL Stats Total Score					
	Forty					
	10-Yard Split (Seconds)					
	Vertical Jump					
	Broad Jump in Inches					
	Draft Round					

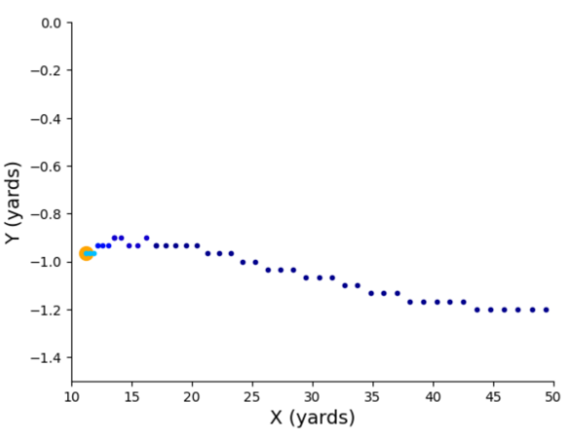
Appendix B: Combine Drill Locator Outputs



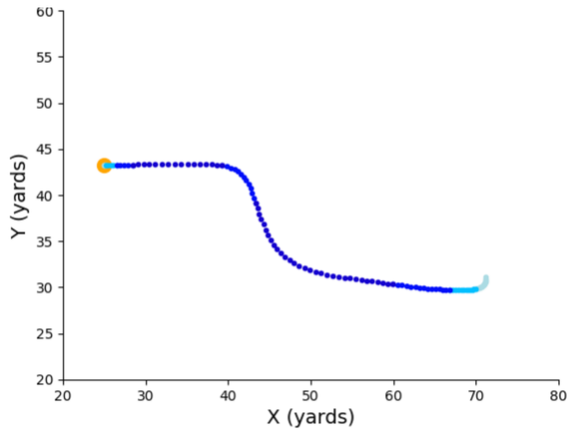
Shuttle Drill



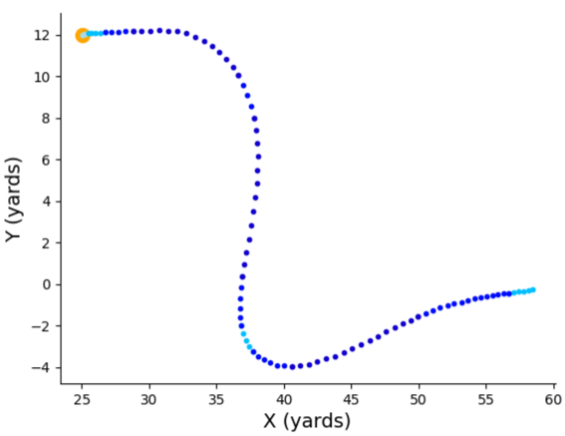
3-Cone Drill



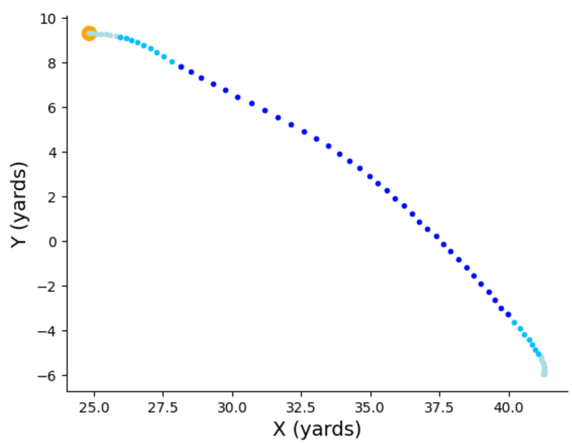
40-Yard Dash



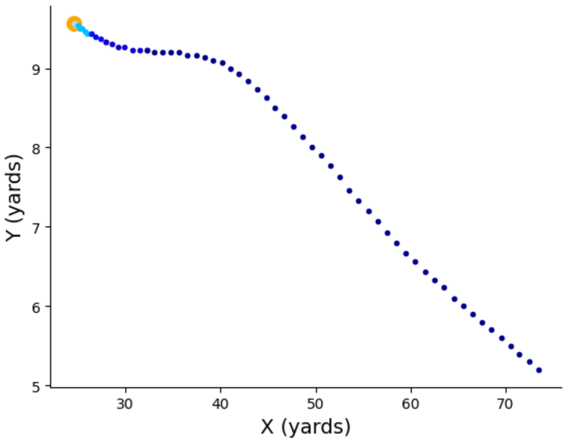
In Drill



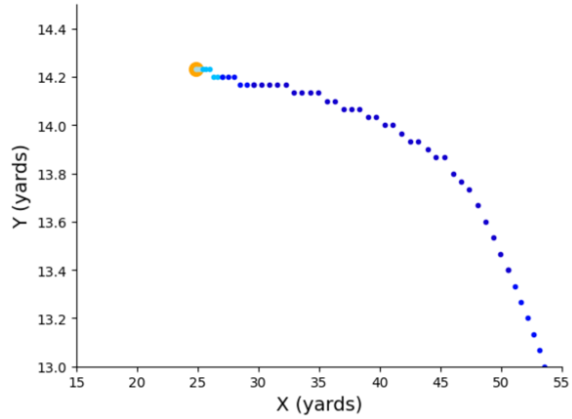
Out Drill



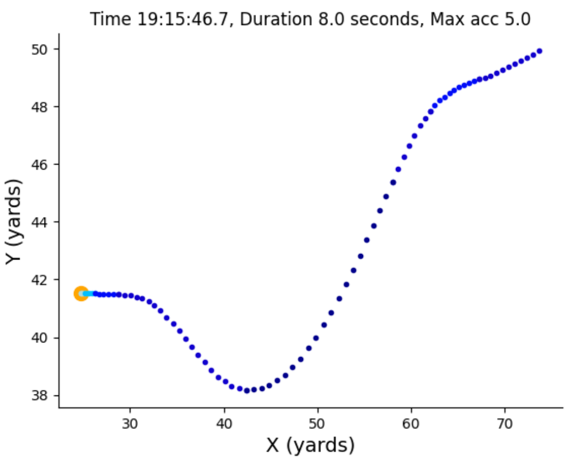
End Zone Fade Drill



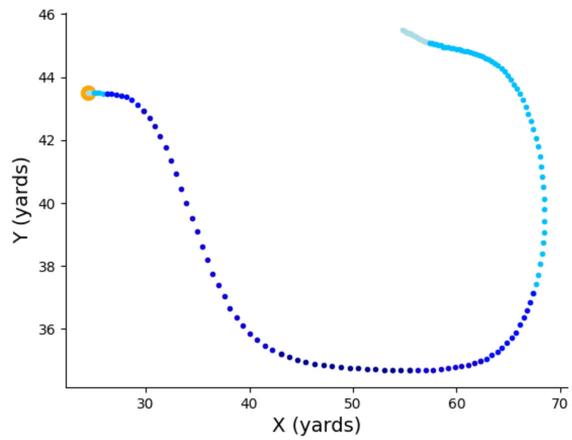
Slant Drill



Go Drill



Postcorner drill



Curl Drill