Predicting Draft Round Based on NFL Combine Results*

Jonathan Adams[1] and Julian Swart[2]

Abstract— **The purpose of this analysis is to determine if it is possible to predict a National Football League (NFL) prospect's draft round based on his physical attributes, height and weight, and his performance at the NFL combine. The analysis is broken down by player position due to the varying nature of available variables at each position. The NFL combine test results include metrics such as bench press repetitions, forty yard dash time, vertical jump, and more. The NFL draft pick is ranked based on the round in which a player was selected by an NFL team. The resulting outcome of this analysis will be to provide college football athletes a clearer picture of what metrics matter the most to increase their draft equity. This will allow the athletes to focus their time into appropriate training.**

I. INTRODUCTION

The NFL is the pinnacle of success for an American football player. Thousands of college football players in America work hard towards the dream of making it to the NFL and playing for their favorite team. However, an extremely small percentage of collegiate level athletes make it to the professional level. The NFL draft is yearly event in which each of the thirty two current NFL teams come together and decide which college athletes to add to their roster for the upcoming football season. Several factors play into a team's decision on which players to draft. A few of these factors include: current team roster needs (at which positions are they in need), player success on the field in college, NFL combine test results, and player attitude. This analysis will focus on tangible results obtained from NFL combine test results as well as the physical attributes of the athlete, specifically height and weight. The NFL combine is also a yearly event in which top performing college players are invited to come participate in workouts and drills while recruiters from NFL teams watch and evaluate their performance. A few of these tests which we will be analyzing are the 225 pound bench press repetition test, the forty yard dash, the vertical jump, the three cone drill, and the shuttle. These tests are all designed to determine a player's physical abilities. While watching the NFL combine, the commentators often talk about how a certain player has increased his draft ranking based on his performance on these tests. The following analysis will attempt to determine if a player can affect his draft stock (positively or negatively) based on NFL combine performance metrics.

*This work was not supported by any organization

[1]J. Adams is a Graduate Student in the Business Analytics Program, University of Tennessee, 201 Stokely Management Center, Knoxville, TN 37966

[2]J. Swart is a Graduate Student in the Business Analytics Program, University of Tennessee, 201 Stokely Management Center, Knoxville, TN 37966

II. DATA COLLECTION AND CLEANING

Data for this analysis was collected from two different sources. The NFL combine test results will need to be obtained from This website has historical combine results dating back to 1987. The NFL draft pick ranking will be obtained from The draft results website contain archives dating back to 1936. This analysis will only include data from 2004 - 2017.

After scraping the data and exporting the results to comma separated (CSV) files using Python, it was read into R as individual data tables. The next step was to merge the two data tables together to have draft ranking line up with the correct player. This merge was done by player name since both tables included this field. One issue encountered while merging was that there were several names that were the same among multiple athletes. This issue affected approximately fifty athletes out of the rough forty-five hundred in our data set. Due to the small number of issues and the difficulty in sorting out the merges correctly, the decision was made to exclude all athletes with matching names.

Once the tables were merged, it was discovered that multiple tags existed for similar positions. A table of the merged positions can be seen below.

Original Position Labels	New Position Label
ILB, MLB, OLB	LB
C, G, OG, OG	OL
DE, DT, NT	DL
DB, FS, SS	DB

Along with combining the above positions to create similar groups, the draft ranking was converted into draft round. The number of players selected in the NFL draft is typically around two hundred fifty with some variation from year to year. However, for this paper, draft round is the response variable. Therefore a conversion was made from draft rank to draft round. The number of players selected per round is also subject to the yearly variation. The table below describes the numeric conversion. Note: due to the variation, it is possible that a few athletes' draft round was wrong but the number of incorrect rounds was most likely negligible.

Draft Rank	Draft Round
1 - 32	1
33 - 63	2
64 - 99	3
100 - 134	4
135 - 169	5
170 - 205	6
> 205	7

After handling the above issues, some data exploration led to the discovery of unrealistic values for some of the combine test results. For the timed drills (forty yard dash, three cone drill, and shuttle), there were several entries of 9.9 seconds. These values are certainly beyond reason for drills in question and was most likely a data entry error or a place holder that was never corrected. This error was only present for a small number of athletes so these athletes were removed from the data set since there was no way to obtain accurate estimates for the actual values.

After combining positions, changing the draft rank to draft round, and removing unrealistic data, the data was subset into individual tables by position. The main reason for this subsetting was that different positions require vastly different physical attributes, receivers tend to be light and fast while lineman are heavy and extremely strong. These differing attributes often lead to large differences in combine test results so our ability to predict draft round would have been diluted due to this fact. We also chose to exclude punters (P), kickers (K), full backs (FB), and long snappers (LS) due to the limited data points available for these positions. Players in these positions typically do not get drafted but are signed by teams as free agents after the draft which is why such limited data was available.

Once the data was subset, it was obvious there were quite a few athletes who did not participate in all of the drills which left the data riddled with NAs. Athletes will sometimes get hurt in a drill and not be able to complete the others or they will only complete the drills they think are necessary to help their chances of getting drafted. Since predictions of draft round were based on these test results, the decision was made to only use athletes who had completed all of the drills. The exception to this rule was the wide receivers (WR) and quarterbacks (QB). Both of these positions rarely completed the bench press repetition test so that test was removed for these athletes.

Finally, after handling all of the steps and issues described above, it was time to move on to modeling attempts. The modeling techniques and results will be described in the following sections.

III. MODELING TECHNIQUES

The goal of the models described in this section was to predict the draft round in which a prospect would be taken in the NFL draft based on his physical attributes and NFL combine test results. This particular problem was essentially as a multi-class classification problem with the draft round as the response class. Several different algorithms were tested and benchmarked against each other as well as the baseline of selecting a draft round at random. In this case, the baseline of randomly assigning an athlete a draft round would be one in seven or around fourteen percent. Each algorithm was tested on every position individually.

The algorithms that were tested for draft round predictions were as follows: Random Forest, Naive Bayes, Multinomial, Support Vector Machine. The data was left in its original format meaning that no scaling or normalization was needed for the algorithms chosen. Naive Bayes algorithms typically perform better when the data from the predictors follows a normal distribution. The test results and physical attributes of the athletes all followed a relatively normal distribution (no large skew). Support Vector Machines often benefit from scaling, however, the software packaged used performed scaling automatically so there was no need to manually scale the data. For the first three algorithms, Random Forest, Naive Bayes, and Multinomial, the data was split into training and testing set with a 50-50 split. For the Support Vector Machine algorithm, the data was split into training, validation, and testing sets with a 1/3 split in each set. The reason for three sets is that Support Vector Machines have several parameters that must be tuned for optimal results. The validation set allows optimization of parameters before predicting on the testing set to determine accuracy. The results of modeling will be shown in the following section.

IV. MODELING RESULTS

Each algorithm described above was tested through twenty-five runs with re-sampling of training, testing, and validation sets performed for each run. The average accuracy for each algorithm was then compared. The Random Forest algorithm was by far the highest performer of the four algorithms tested. The accuracy for this algorithm was roughly 60% for each position which is approximately four times higher than assigning random draft rounds to each athlete. The other three algorithms were not even close to competing with the accuracy of the Random Forest. Both the Naive Bayes and Multinomial algorithms had accuracies ranging from the low 20%s to the mid 30%s depending on position. The Support Vector Machine performed worse than random chance for several positions and never even made it to 20% accuracy. A plot of the accuracies for the four algorithms for all the positions can be seen in the appendix in Figure 1. The orange line in Figure 1 depicts the one in seven chance of being correct via random draft round assignment.

Another major benefit to the Random Forest algorithm is the ability to extract variable importance. Along with the accuracy, the most important variable for each position was stored for each of the twenty-five runs of the algorithm. The table below summarizes the variable that appeared on top the most often for each position.

Position	Top Importance
OL	Height
TE	Height
LB	40 yard dash
WR	Weight
DL	3 cone drill
RB	Broad Jump
QB	3 cone drill
CB	40 yard dash
DB	3 cone drill

In order to visualize the results of the above table, Figure 2 in the appendix shows the most important variable, 3 cone drill, for the QB position. The dark blue line represents the average 3 cone drill time for all QBs. As you can see, the average time is lower than average for the first two rounds and increases as the round number increases. This proves that a QB can benefit from having a lower 3 cone drill time.

V. IMPLICATIONS

The previous results prove that our model has the ability to predict with accuracies of roughly four times greater than average which round an athlete will get drafted. It also has the capability to show which drills have the highest impact on draft round selection. The combination of these facts means that our model could be used by a prospective athlete to determine which drills he needs to focus the most time in to improving based on his position. Knowing this could increase his chances of getting drafted in a higher round. Being drafted in a higher round means that an athlete will get signed to a longer contract with a higher salary so this model has the potential to help athletes earn more money.

VI. FUTURE WORK

Going forward there are several things that could be done to potentially improve the models. One major task would be to add in college statistics for each player. Our original intent was to include this information. However, it was discovered too late that we had a data integrity issue with the college stats we had obtained. The way in which we scraped the college stats was plagued by the issue of duplicate names described earlier in this paper. We were not accounting for the fact that multiple players would have the same name so our stats were all wrong for almost every player. Therefore, a better method for obtaining the college stats would need to be discovered so that this information could be added to our models.

Another task that would be potentially beneficial would be to determine a way to how to impute the missing data from test results. This would allow a larger data set to be used since we are currently limited to athletes who completed every drill. One potential way to do this would be to consider the correlations between each combine test and calculate the missing values based on the drills the athlete actually completed. A brief look was taken at these correlations and the results for defensive lineman can be seen in Figure 3 of the appendix. As you can see from Figure 3, there definitely appears to be some strong correlation between certain drills and also between drills and physical attributes. This result is promising and hints towards the possibility of missing values imputation based on other present data for each athlete.

APPENDIX

All plots described in this paper can be seen on the following pages

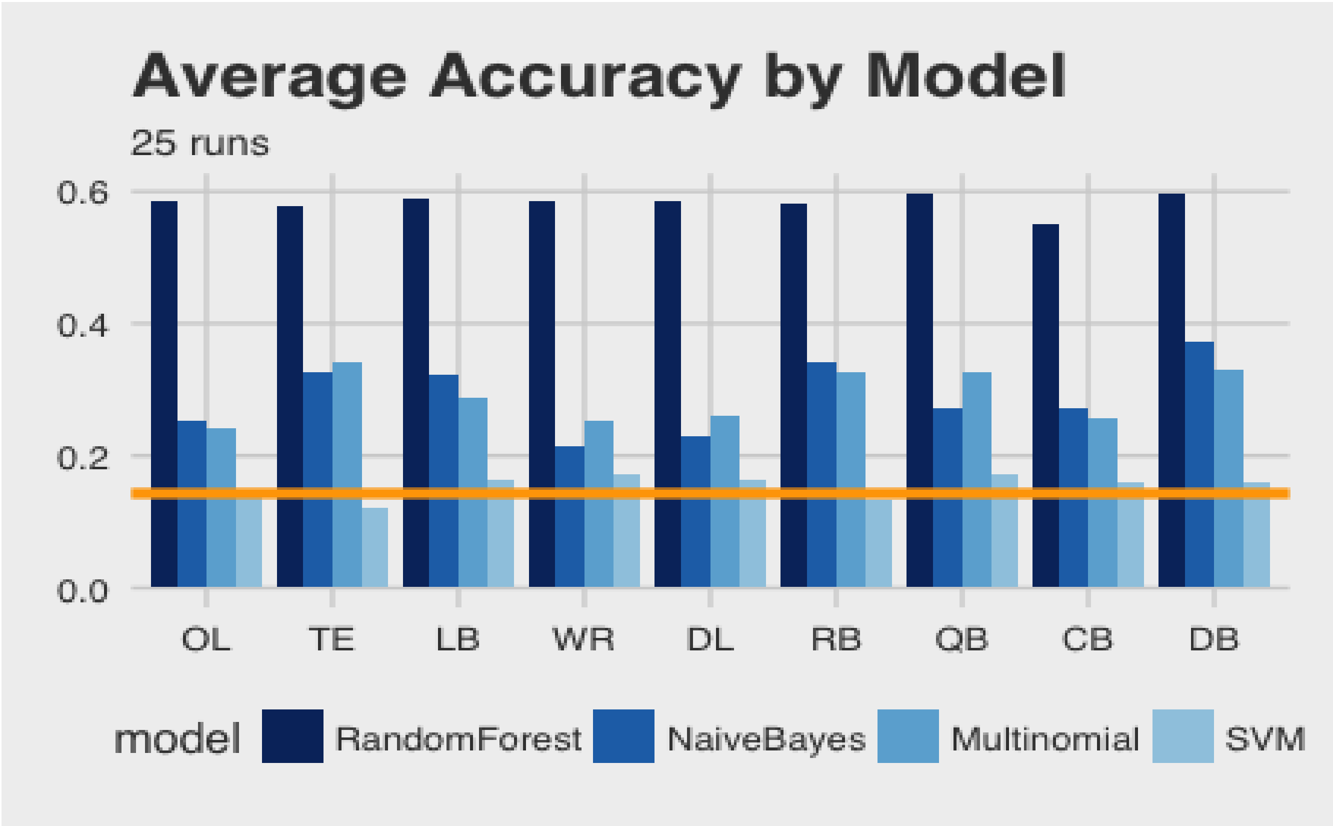

Fig. 1. Model Accuracies

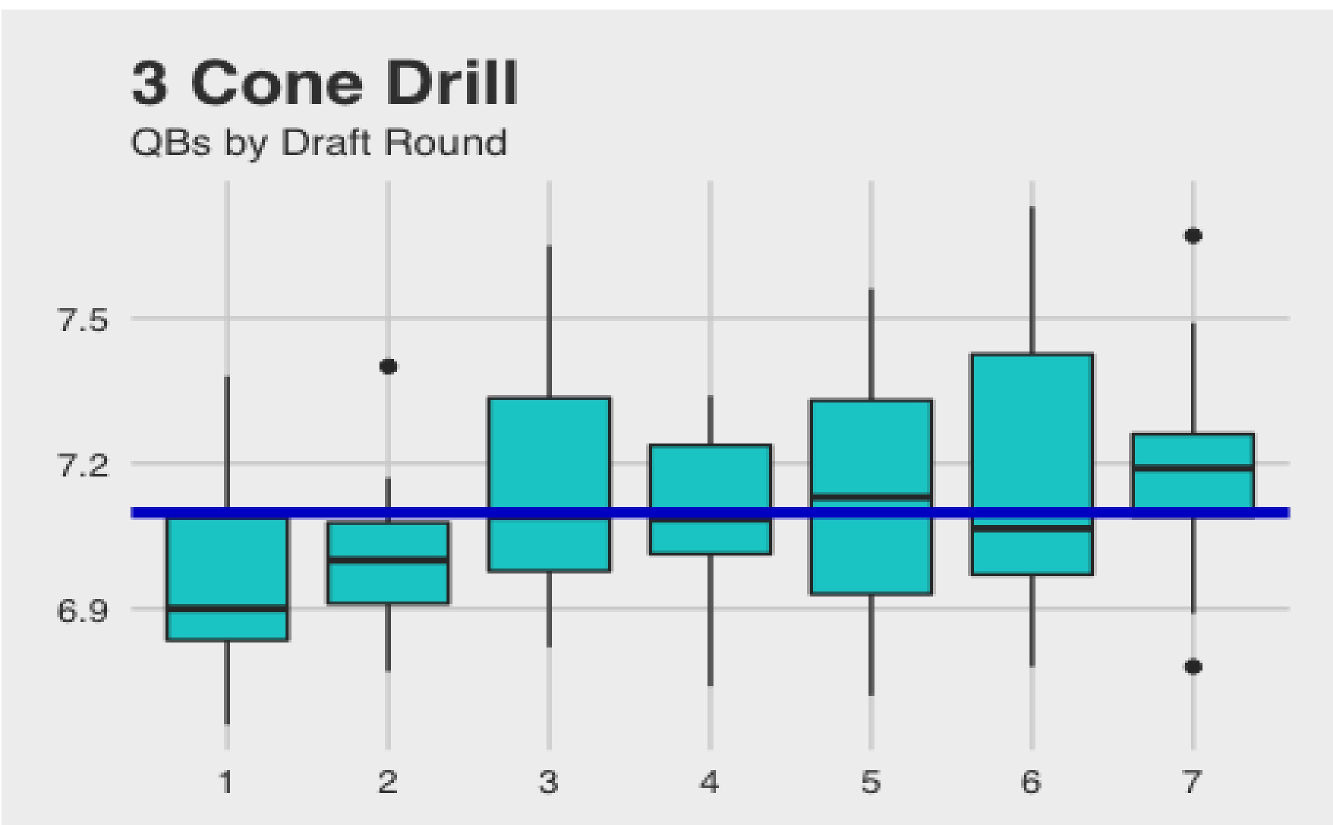

Fig. 2. QB 3 Cone Drill

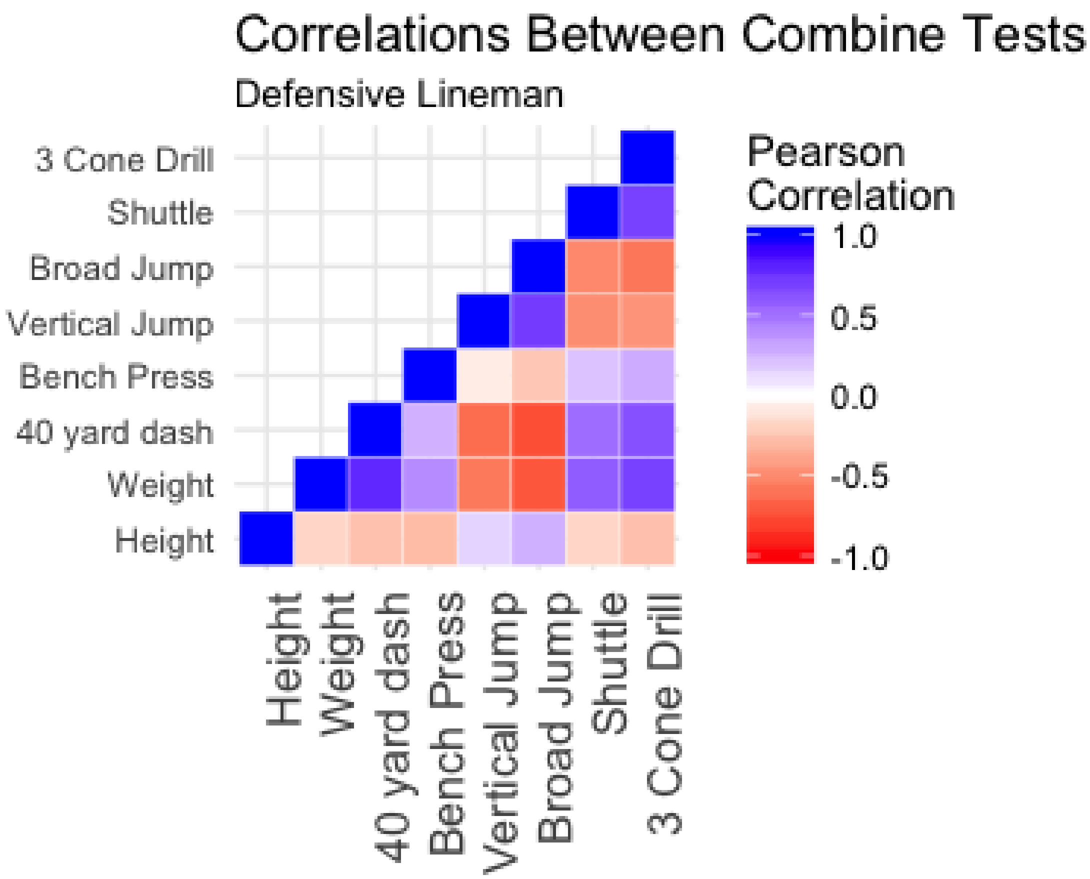

Fig. 3. DL Test Correlations