

NFL-Combine-Draft-Classification

Eric Au



The Business Problem

Every year, the National Football League (NFL) holds a week long showcase where college football players, otherwise known as prospects, perform physical drills and tests in front of team coaches, scouts, and general managers. These drills are intended to measure a player's physical ability such as speed, quickness, strength, and overall athleticism.

- *But what can NFL teams learn from these workouts?*
- *What exactly do non-football athletic testing measurements contribute to prospect evaluation? Is there any value to the combine?*

These are questions that many fans ask to this day and NFL teams try to interpret to make the best decision possible when drafting their players.

The Task: Create a model that can determine whether a player was **Drafted** or **Undrafted** based on the available data and information provided in the NFL Combine.

- How significant are the measurables in determining draft status?
- Which measurables are good indicators of draft status?

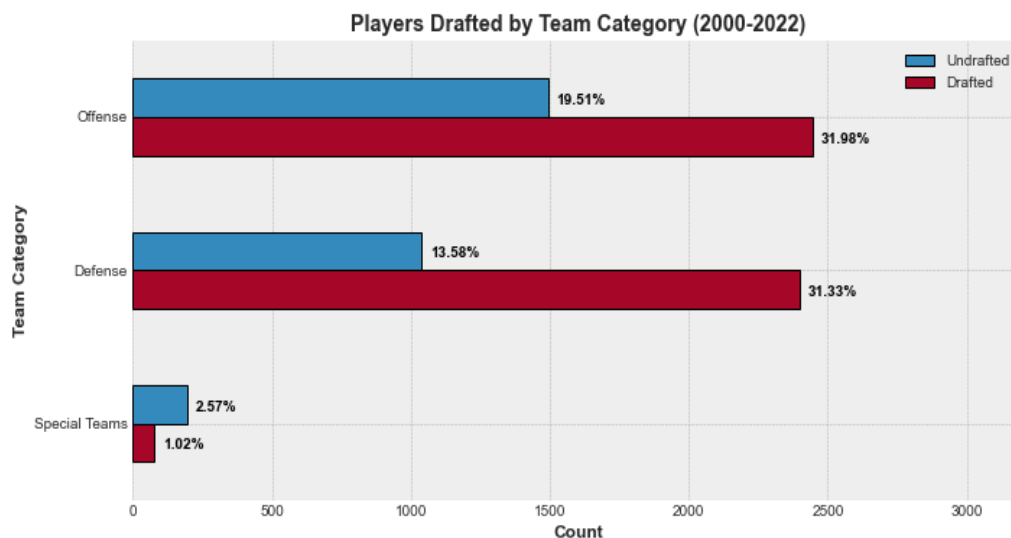
Stakeholders: The New York Giants front office (General Manager, President, Scouting Department).

The Dataset

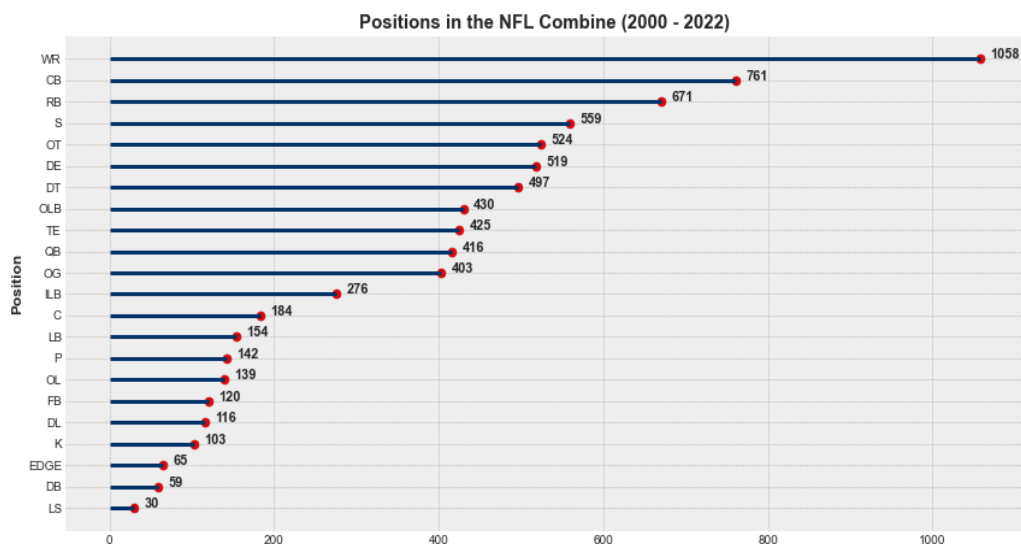
In this analysis, player combine data was scraped from [Pro-Football Reference](#) over the last 22 years (2000-2022). For this analysis, players are classified as 1 for Drafted and 0 for Undrafted.

For further information on the drills/measurables, this [USA Today article](#) provides an in depth explanation.

Data Understanding & Visualization



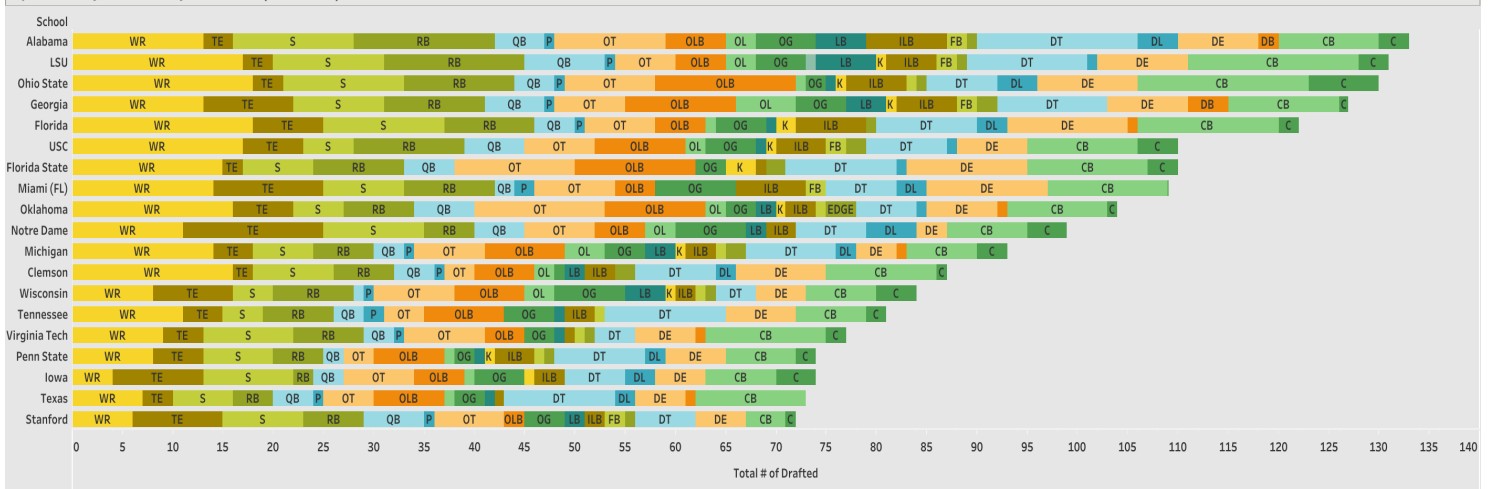
Players Drafted By Team Category: Of the 7600+ entries of NFL combine data, approximately 2/3 of the data consists of a drafted player.



Positions in the NFL Combine: When breaking down the distribution of player positions in the draft, we notice that the skill positions such as wide receiver, cornerback, and running back, dominate the dataset. While special teams and more generic positions such as "defensive lineman", "defensive back", and "edge" are less frequent.

College Powerhouses

Top Schools By Count of Players Drafted (2000-2022)



College Powerhouses: There is no surprise that the college powerhouses like Alabama, LSU, and Ohio State produce the most NFL drafted talent. However, it is interesting to note the variation of quantity of drafted players at each position for each college. This can likely be attributed to each school's unique football program structure, team personnel strategy during games, and importance of coaching the position.

The Measurables At A Glimpse

Median Performance in Drills By Position & Draft Status



Color shows details about draft status (1 - Drafted, 0 - Undrafted). The view is filtered on Pos, which excludes K and P.

The Measurables at a Glimpse: We further examine the median performance in drills sorted by position and draft status (where green represents drafted and red represents undrafted). When categorizing by positions, we notice an obvious trend where drafted players tend to perform better at the combine drills than undrafted players.

Modeling

Following a train-test split whereby 25% of the entire dataset was held out, an iterative modeling process was implemented to determine the most accurate model for the testing set. For reference, we want to have a better baseline accuracy score of about 63% (those who were drafted) if our model was just simply randomly guessing.

F1-Score was the main metric used to determine model accuracy which takes into consideration a balance between the False Positives and False Negatives. In other words, we want to have a 'harmonized' balance between precision and recall and take into consideration the False Positives and False Negatives.

- False Positives are players who were labeled as 'Drafted' but they were actually 'Undrafted'

- False Negatives are players who were labeled as 'Undrafted' but they were actually 'Drafted'.

A baseline logistic regression model was implemented first with a Testing F1-Score of 79%. Additional models utilizing various classification techniques were implemented thereafter with a grid-search performed on each model to optimize the F1-Score. 5-fold cross validation was also performed with each iterative model.

The following models were implemented on the testing set:

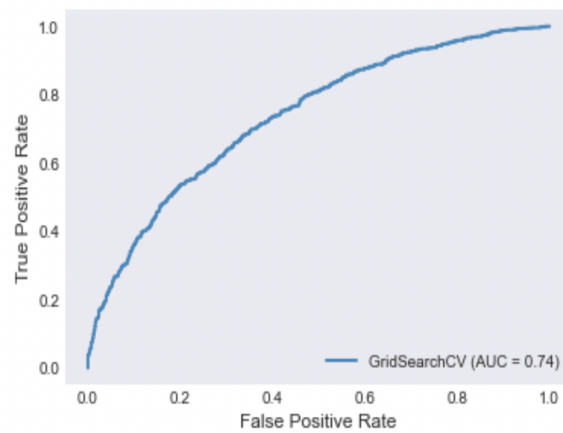
- Baseline Logistic Regression
- Best Parameters Logistic Regression
- Liblinear Solver Regression
- Decision Tree
 - Tuned Decision Tree
- KNN (K-Nearest Neighbors)
- Adaboost
- Gradient Boost
- XGBoost
- Bagging Decision Tree Classifier
- Random Forest Classifier

Evaluation

The following details each individual model's training and test F1-Scores and sorted by the highest AUC (Area Under the Curve - measuring how well the model was able to classify drafted/undrafted).

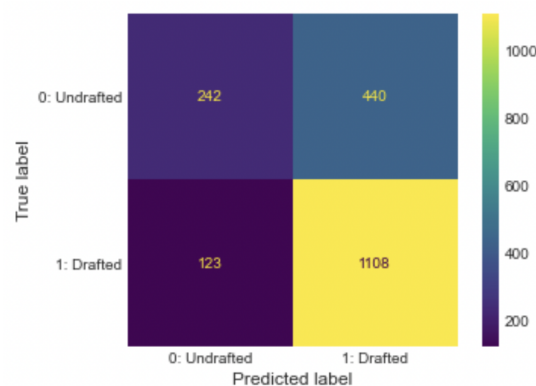
	Model	Training - F1 Score	Testing - F1 Score	AUC
8	8th Model - XGBoost	0.8348	0.7974	0.737
10	10th Model - Random Forest Classifier	0.8582	0.7960	0.735
9	9th Model - Bagging Decision Tree Classifier	0.8508	0.7918	0.733
7	7th Model - Gradient Boost	0.8198	0.7980	0.724
0	1st Model - Baseline Logistic Regression	0.8106	0.7914	0.724
1	2nd Model - Best Params (lbfgs) Logistic Regre...	0.8065	0.7916	0.718
2	3rd Model - Liblinear	0.8067	0.7911	0.718
6	6th Model - Adaboost	0.7958	0.7813	0.708
5	5th Model - KNN	0.8042	0.7922	0.689
4	4B Model - Decision Tree	0.8411	0.7681	0.662
3	4th Model - Decision Tree	1.0000	0.7156	0.582

Ultimately, the 8th Model - XGBoost performed the best when taking into account the AUC and is slightly higher than the Random Forest and Bagging Decision Tree Classifiers.



However, the 7th Model - Gradient Boost performed the best overall in terms of F1 Score, but only slightly better than the XGBoost model.

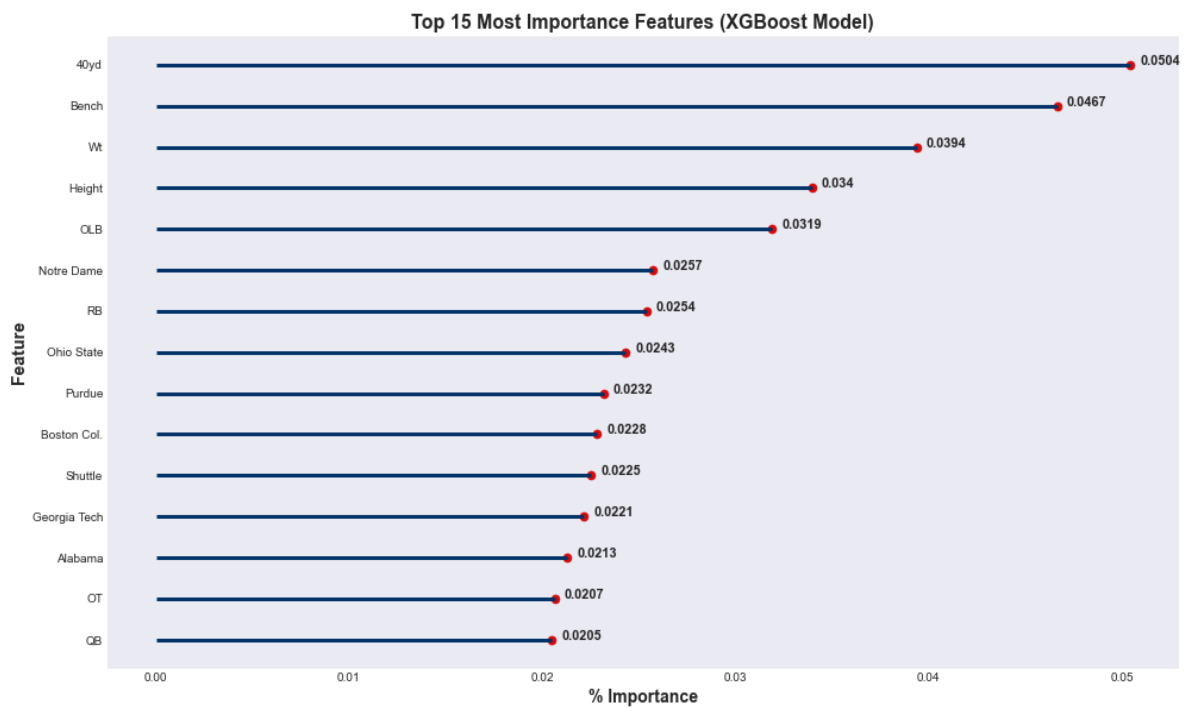
Based on these metrics, the XGBoost model was deemed best overall when it comes to classifying between players who are Drafted and Undrafted.



While the XGBoost Model performed with an overall F1-Score of 80%, the model still struggles when it comes to classifying whether a player will be drafted or undrafted. The above confusion matrix illustrates the difficulty of the model to classify unseen test set data of drafted players when they were actually undrafted.

There is still a significant amount of False-Positives (or players classified as Drafted when they were not) amounting to a Precision Score of 72% correctly classified Drafted players. Consequently, there is a Precision Score of 66% when classifying Undrafted players.

Feature Importance



The above chart illustrates relative scores and highlight which features are most relevant to classifying the target of Drafted . The higher the feature importance, the more of an effect it has on the model's performance.

- The model suggests that the most important combine measurables that classify Draft status are the 40 Yard and Bench. The Shuttle drill follows thereafter, though not as important of a feature.
- The schools with the most importance for Draft classification are Notre Dame, Ohio State, Purdue, and Boston College. Alabama is surprisingly behind these other schools even though they have been powerhouses recently when it comes to producing NFL talent.
- Positions with the most importance are surprisingly the OLB (Outside Linebacker), RB (Running Back), OT (Offensive Tackle), and followed by the QB (Quarterback). While the model does not suggest that these are the most important positions on the field, they are the most important when it comes to classification of draft status alone.

Conclusions & Recommendations

Model Performance

While the best model performed with an overall F1-Score of 80%, there is still evidence to suggest that draft combine metrics alone are not a strong indicator of whether a player will be drafted or undrafted.

It is clear that classification of draft status is not an exact science. Though, it is also important to note that there are potentially additional information and data that could be implemented into the model. For example, college statistics were not factored into the analysis. This would involve gathering all the data for every player in the combine and merging this information with the combine data. While this seems ideal, there is a lack of data when it comes to measuring each individual unique position in football. Further analysis would have to be split up into positional categories as not all statistics across positions are the same.

At the pro level, there are efforts to obtain more data to measure player's performance. [NFL Next Gen Stats](#) has been collecting data to assist in these efforts, though it has not been applied at the collegiate level yet.

Model Value & Limitations

Notwithstanding, there is at least some value to the combine based on our model. The best indicators from the model suggest that the 40 Yard, Bench Press, and Shuttle, are the most important metrics when it comes to classification of draft status. We also observed that on average, players who are drafted, tend to perform better at each combine drill than those who were not drafted.

We can confidently conclude that taking in consideration of combine metrics alone does not provide a sure-fire determination of whether someone should be drafted or not. However, we can still use the combine as a *guide* when it comes to predicting whether someone should be drafted or not with an 80% F1-Score accuracy using the best model. Additionally, there is still value in traditional scouting of players such as immeasurable interviews as well as accounting for a player's college career statistics.

Finally, it is also important to note the limitations of the model and that the model does not predict whether a player will be successful at the professional level. The model does not take into account any other 'intangible' measures such as the player's overall character, demeanor, or work ethic, all of which have value and factor into a draft decision.

Recommendations:

- Focus on targeting players who perform better than the average positional group for each combine drill. These players tend to be drafted players. However, due-dilligence should also be applied when assessing a player's overall value and not solely just on combine metrics.
- When drafting players, priority should be placed on the 40 Yard, Bench, and Shuttle as these drills have the highest importance in determining draft status. Height and weight of the player are also significant and should be compared against the relative average weights and heights of the position.
- Seek to incoorporate more data collection at the collegiate level; potentially partner with the NCAA Football.

Repository Structure

```
├── Data
├── Images
├── .gitignore
├── NFL Combine Classification Presentation Slides.pdf
├── NFL Combine Draft Classification.ipynb
└── README.md
```

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Languages

● Jupyter Notebook 100.0%