The Role of Sports Statistics in Evaluating College Quarterbacks

UCLA Department of Statistics

Research Questions & Motivations

Research Questions

- **Question 1**: Is it possible to utilize common, widely available college football quarterback statistics to predict whether a given quarterback will be drafted into the NFL or not?
- **Question 2**: Given that it is possible to predict whether a certain college quarterback will be drafted with sufficient accuracy, what are the features that most significantly influence whether a quarterback is drafted into the NFL or not?

Motivations

- **Motivation 1**: To provide a quick way to narrow down potential quarterback candidates for scouting teams of NFL organizations.
- **Motivation 2**: To provide valuable insight for graduating high school quarterbacks into the skills to develop, schools to attend, and more to increase their odds of being drafted into the NFL.

Data Sources

Before exploring the conclusions made by our research, it is important to explain how the data has been sourced and constructed:

- **QB Statistics Source**: All of the quarterback data analyzed is sourced from the summary of 2010 to 2023 player passing statistics from the college football Sports Reference page.[1]
  - *Note*: The quarterback data was transformed such that each statistic was normalized within the player's conference, each player played at least three games, each player threw at least ten pass attempts, and, in the instance of multiple seasons, a player's best season was kept.
  - *Note*: The two classes, drafted vs undrafted quarterbacks, are heavily imbalanced. There are just 151 drafted quarterbacks in the data to 1,488 undrafted.
- **Draft Data Source**: All of the draft data is sourced from the drafted players data from 2011 to 2024 on the Pro Sports Reference page.[2]

Normal and Weighted Models

How can the response class imbalance be managed while keeping all of the training data points?

- **Normal Models**: Train the models on the data without any changes to account for response class imbalance.
- **Weighted Models**: Provide a greater weight to data points in the minority class to account for the class imbalance.

Model	Accuracy	Recall (Drafted QBs)	F1-Score (Drafted QBs)
Normal RFC	0.93	0.40	0.50
Normal Boost	0.91	0.30	0.38
Weighted RFC	0.92	0.67	0.62
Weighted Boost	0.91	0.50	0.52

Table 1: Tree Model Performance Predicting Drafted Status of QBs

Undersampled Models

The normal and weighted models have a major problem, they have low recall for the minority class. How can model accuracy be improved when predicting quarterbacks who will be drafted?

- **Undersampling**: A specified procedure for eliminating training cases from the majority class to reduce, or completely eliminate, the imbalance between the response classes.[3]
  - *Random*: Randomly selected data points from the training data are deleted until a desired class balance is reached
  - *Near Miss*: Each minority class data point and its closest majority class data point are kept.
  - *Condensed Nearest Neighbors (CNN)*: The combination of all minority class data points and only those from the majority class that cannot be classified correctly based on a subset of other majority class points.

Model	Accuracy	Recall (Drafted QBs)	F1-Score (Drafted QBs)
Random (2:1)	0.86	0.77	0.50
Near Miss (1:1)	0.78	0.90	0.43
CNN (1.75:1)	0.92	0.53	0.55

Table 2: Random Forest Model Performance Using Undersampling Methods

Feature Importances

Now that a model has been trained, how can it be utilized to determine which features of the data provide the most information as to whether a college quarterback gets drafted into the NFL or not? Although tree models have built in methods for determining model importance, they have some issues:[4]

- **Issue 1**: There are various metrics that measure feature importance, which can give inconsistent results.
- **Issue 2**: Feature importances are aggregated, making it difficult to observe the effect on specific classes

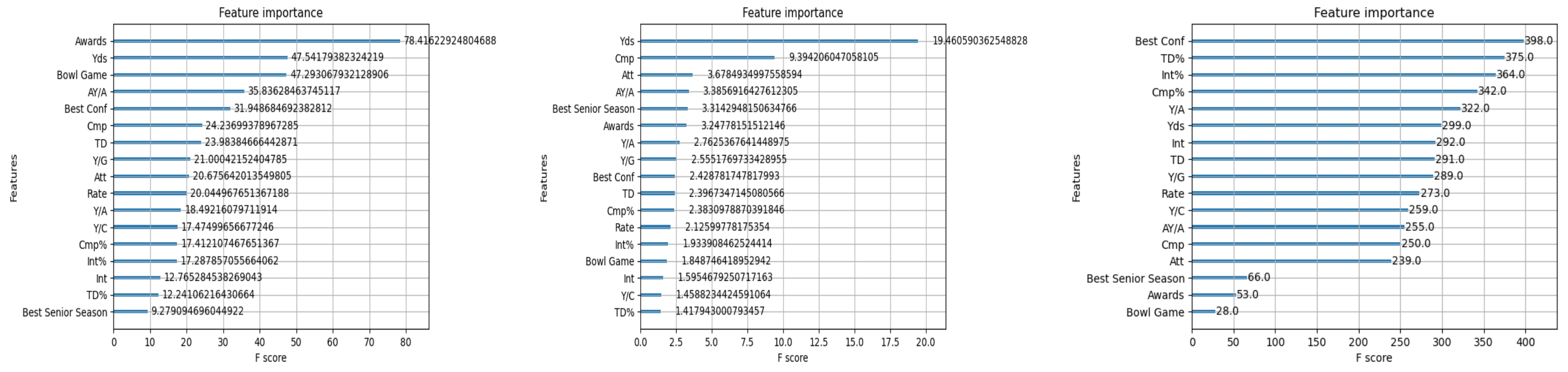

Figure 1: Feature Importances Vary Based on "Cover", "Gain", and "Weight"

References

[1] *2010-2023 College Football Passing Stats*, Sports Reference LLC, Mar. 2025.

[2] *2011-2024 NFL Draft*, Sports Reference LLC, Mar. 2025.

[3] Brownlee, Jason. "Undersampling Algorithms for Imbalanced Classification." *Machine Learning Mastery*, 27 Jan. 2021.

[4] Marsh, Emily K. "Calculating XGBoost Feature Importance." *Medium*, 31 Jan. 2023.

[5] John, Brain. "How to Use SHAP Values to Optimize and Debug ML Models." *Neptune.ai*, 23 July 2024.

Shap Values

SHAP (Shapley Additive Explanations) values can be used in place of feature importances for boost models. What are the advantages of SHAP values?[5]

- **Advantage 1**: SHAP values apply a singular, consistent method to analyze feature importance.
- **Advantage 2**: Able to visualize how each feature specifically affects the prediction of individual data points.

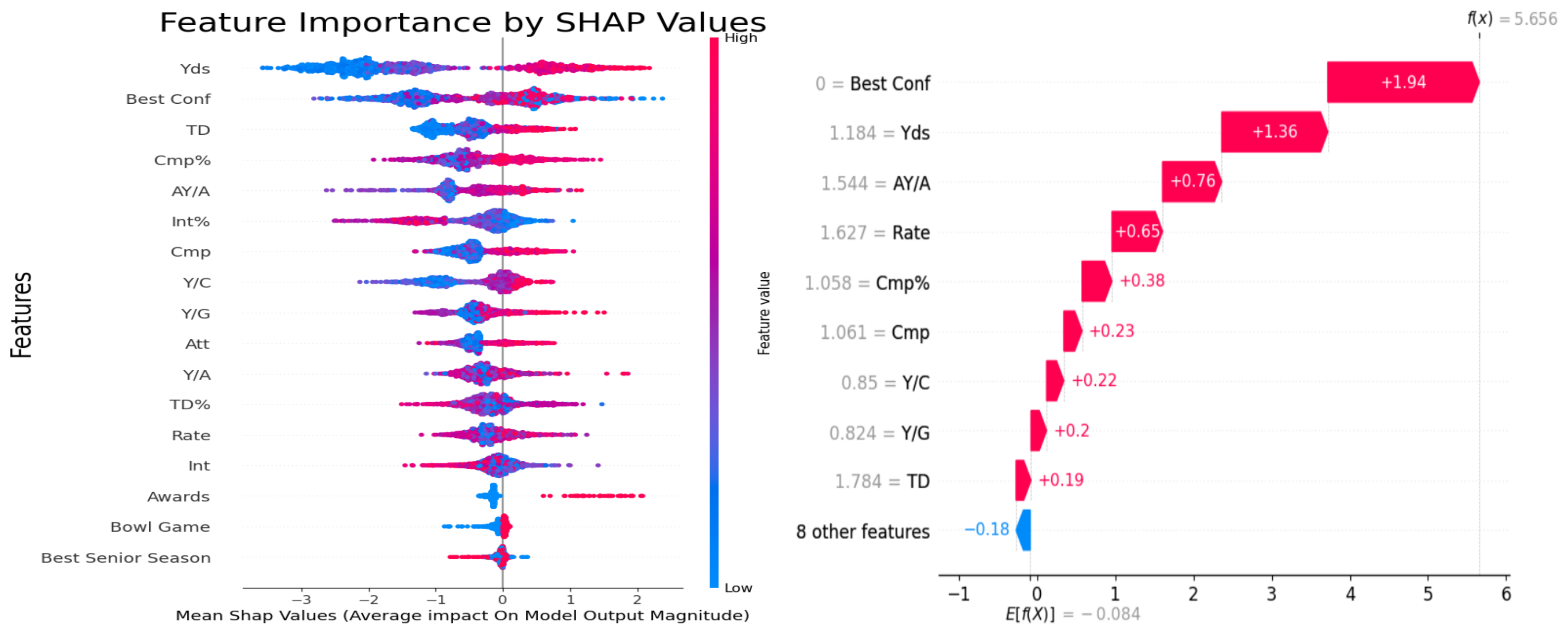

Figure 2: Feature Importances for Weighted Boost Model Based on SHAP Values

Draftability

SHAP values are powerful tools for observing the specific features that affect each individual data point. However, the more features that are present within the data, the harder they can be to interpret. How can the interpretation of these SHAP values be simplified for the quarterback data?

- **Draftability Index (Dft)**: Summation of a quarterback's statistics weighted by their respective SHAP values
- **Normalized Draftability Index (Dft+)**: Normalizes each quarterback's Dft value around a mean of 0 and standard deviation of 1.

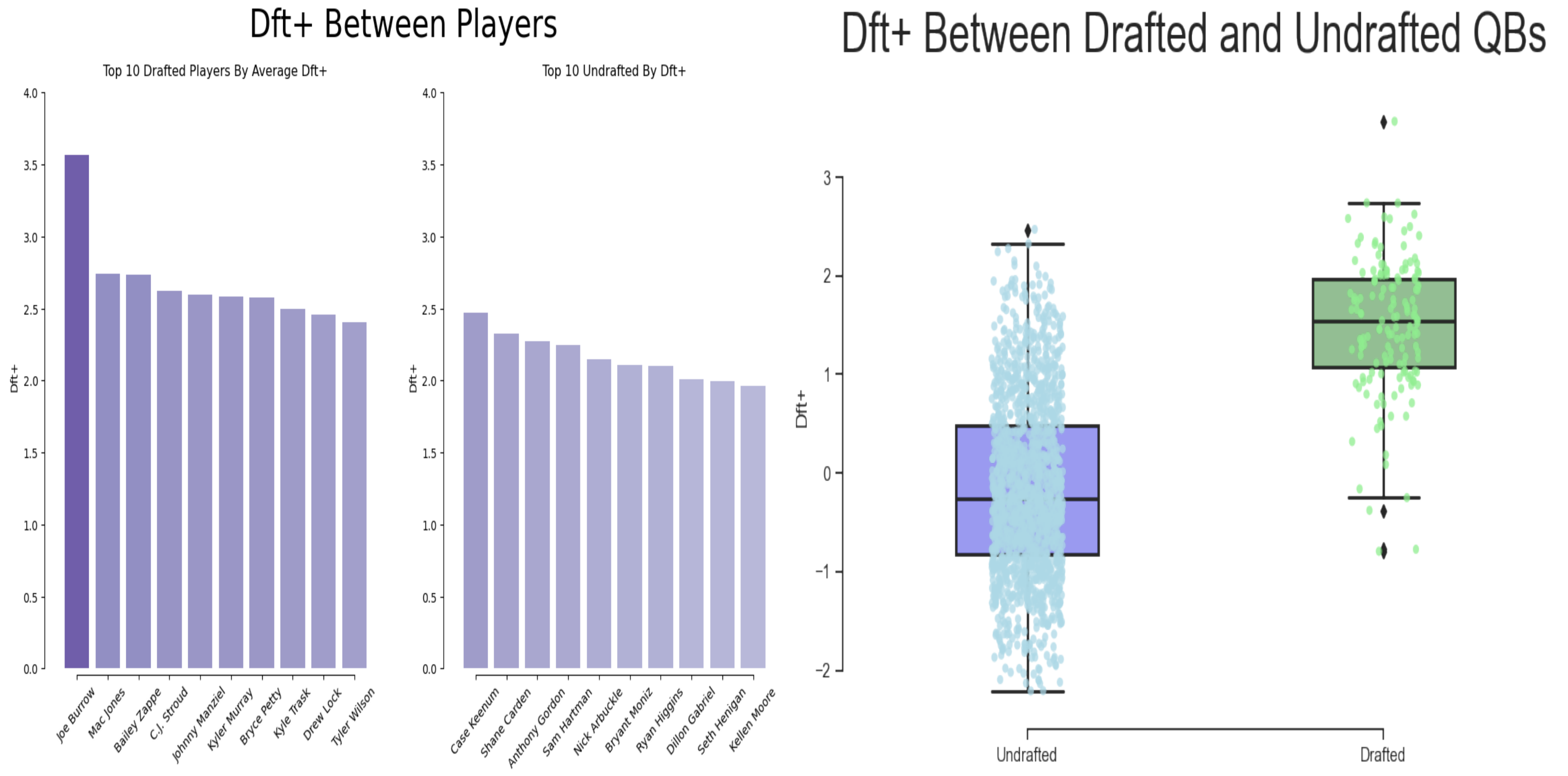

Figure 3: Use Cases of Dft+