Fantasy Football Draft Assistant

Chris Ramos
EECS Department
University of Tennessee
Knoxville, United States of America

Braeden Watkins
EECS Department
University of Tennessee
Knoxville, United States of America

***Abstract* - Fantasy Football is a game in which people gather to compete against each other based on the players they draft to be on their team. Although Fantasy Football has been around for a while, many new players are unfamiliar with the art of drafting a winning team. The goal of this project is to enable newer players, or players with less experience, to be able to compete against those who are experienced.**

I. OBJECTIVE

The primary objective of the fantasy football draft assistant is to provide users with a player-specific analysis to improve the outcome of their fantasy football team. As the current fantasy football season has already begun, this assistant will be focused on generating a mock team for the 2025 season. A mock draft will be simulated that will assist in creating an accurate team generation environment. Once the team using the fantasy football draft assistant has been created, the performance can be analyzed based on how each player is currently performing in the active season. This validation metric will provide evidence of how useful the draft assistant actually is, and while ignoring inevitable injuries, can provide information about how to improve the draft assistant for future use. In terms of future use, the system will be designed such that minimal changes will be required to update the assistant for future years (i.e. 2026 season using 2025 data).

II. DATASET ANALYSIS

To assist us in reaching our goal, we used a Kaggle dataset that contains statistics of NFL players [1]. The portion of the dataset that this draft assistant utilizes is the player and team stats from years 2021 to 2024 only. The reason for this decision is that statistics from earlier years, such as 2012, provide zero benefits as most of the stars from that era will be either retired or no longer in their prime. We also took the Covid-19 season into consideration, which is why we chose to use data from the following years.

As we looked deeper into the data, we found that some of the most valuable metrics were the season passing yards, season receiving yards, season pass attempts, season complete pass, season rush attempts, season rushing yards, season fantasy points PPR (points per reception), etc.

Instead of using these statistics directly, we calculated z-scores for each metric within each position. We decided to use z-scores as it standardized the data, making it easier to compare players to each other more fairly. For example, for a quarterback's completion percentage, the average would be at 0. Anything greater than 0 is considered above average, and anything below, is below average, with both extremes highlighting the best and worst players respectively.

We calculated z-scores for metrics such as completion percentage, yards per attempt, touchdown percentage, interception percentage, fumbles lost, and fantasy points. For stats that would be considered "negative" like interceptions and fumbles, we flipped the sign because fewer is better. After calculating these individual scores, we averaged them to create an overall rating for each offensive position, as well as for defense. We did these calculations separately for each year (2021-2024).

Our model uses these ratings along with games played as features, with fantasy points PPR being the label we are trying to predict. To make the data usable by the model, we had to calculate the global maximum and minimum values for games played and fantasy points across all years. We used this to normalize the features for each year using min-max scaling keeping everything between 0 and 1. We computed these values to ensure that larger ranges didn't affect the model's prediction. Additionally, we one-hot encoded the player positions, to create separate binary columns for each position type (QB, RB, WR, TE, DEF).

III. METHODS

After the preprocessing steps were completed, the model could be created and trained. The model we chose to implement is Scikit-Learn's RandomForestRegressor. We chose a regressor rather than a classifier because the rankings of players are based on point metrics rather than classes. A random forest approach uses an ensemble of decision trees internally, which we felt would adequately represent the interactions between features. Using one decision tree would have almost certainly led to overfitting. For that reason, we chose to use an ensemble approach that would employ the wisdom of the crowds principle to provide us with the most accurate player rankings.

To produce the best results from our Random Forest Regressor, we had to tune the hyperparameters. This was not done manually, but performed instead using Scikit-Learn's RandomizedSearchCV. The following hyperparameters were selected using this method:

- n_estimatres: 700
- random_state: 42
- bootstrap: True
- criterion: squared_error
- max_depth: 10
- min_samples_leaf: 5
- min_samples_split: 15

The model was trained on the following years' feature-label pairs. 2021 features with 2022 labels, 2022 features with 2023 labels, and 2023 features with 2024 labels. Along with these features and labels, we fit the model with sample weights. The sample weights were created to add significance to players more recently active, as recent years are more indicative of next year's performance. Our test set and final predictor used this trained model to make 2025 predictions based on 2024 features.

Using these 2025 predictions, we developed a mock snake draft to illustrate how the player rankings could be used to create a team. A snake draft is one of the most popular fantasy football drafting formats; therefore, it is highly representative of the real-world use case this draft assistor could provide. The order in which drafting occurs is the following: Drafter 0 - Drafter N, Drafter N - Drafter 0, for R number of rounds. Our implementation consists of 15 draft rounds with the option of performing the draft with either 4, 6, 8, or 10 teams. The results of a mock draft and their comparison to real-life performance will be explored in the results section.

The snake draft took the ideal team composition that we decided on into account. We determined the composition based on Braeden's current team, which consists of the following:

- QB: 2
- RB: 4
- WR: 6
- TE: 2
- DEF: 1

This baseline composition was then used to determine the starter and bench teams. Typically, this is just split in half, with the single defensive team getting placed in starters. With these compositions in mind, we adapted the snake draft to fill the starter team first using a counter, and then once those positions are full, switching to the bench. This would put better players on the starters and those worse off as bench warmers.

We also developed a website to display the model predictions in a format that could be usable by fantasy football players needing draft assistance. The website lists all of the highest-ranked players in descending order, organized by position type. This provides users with a priority list that could be referenced when drafting players for their team. The website also includes the snake draft simulator to provide users with an example of potential draft outcomes.

IV. RESULTS

Our model predictions can not be analyzed using simple accuracy scores, as this is not a traditional classification problem. The best metric for analyzing the performance of our model is to compare the 2025 performance predictions with how the players are currently doing in real life. This can be even better illustrated by comparing the generated model teams with Braeden's own fantasy football team, along with one of his opponents. First, a 6-team snake draft was performed, where two of the generated teams were used in the comparisons. Braeden's draft position in his draft was 6, and his opponent's was 2. For this reason, the two model-generated teams compared were Team 2 and Team 6.

The first comparison was performed using only data from the first week of the 2025 NFL season.

Braeden's Current Team:

Starters:
QB: Jayden Daniels - 20.12
RB: De'von Achane - 16.5
RB: Jahmyr Gibbs - 15.0
WR: Nico Collins - 5.5
WR: Justin Jefferson - 14.8
WR: Mike Evans - 10.1
TE: Brock Bowers - 15.3
DEF: MIN Defense - 4.0

Bench:
QB: Kyler Murray - 18.32
RB: Kyren Williams - 13.9
RB: Kenneth Walker III - 5.4
WR: Courtland Sutton - 18.1
WR: DJ Moore - 8.6
WR: Tee Higgins - 6.3
TE: Mark Andrews - 1.5

Starters Total: 101.32
Bench Total: 72.12

Overall Total: 173.44

Simulated Model Team:

Starters:
QB: Jared Goff - 10.9
RB: Zach Charbonnet - 10.7
RB: Rico Dowdle - 3.6
WR: Ja'Marr Chase - 4.6
WR: Amon-Ra St. Brown - 8.5
WR: Ladd McConkey - 13.4
TE: Brock Bowers - 15.3
DEF: DET Defense - 2.0

Bench:
QB: Baker Mayfield - 22.58
RB: Tony Pollard - 7.9
RB: Kareem Hunt - 4.6
WR: Terry McLaurin - 4.7
WR: Drake London - 13.5
WR: Josh Downs - 3.2
TE: Jonnu Smith - 12.5

Starters Total: 69.0
Bench Total: 68.98

Overall Total: 137.98

Figure 1: Braeden vs Model (Team 6) Week 1

Opponent's Current Team:

Starters:
QB: Bo Nix - 6.84
RB: Chuba Hubbard - 17.9
RB: Jonathan Taylor - 12.8
RB: Josh Jacobs - 14.0
WR: A.J. Brown - 1.8
WR: Jaxon Smith-Njigba - 19.4
TE: Sam Laporta - 13.9
DEF: PIT Defense - 1.0

Bench:
QB: Matt Barkley - RETIRED
RB: James Conner - 14.4
RB: Breece Hall - 16.5
RB: Treveyon Henderson - 11.1
WR: Terry Mclaurin - 4.7
WR: Devonta Smith - 4.6
WR: Jaylen Waddle - 7.0

Starters Total: 87.64
Bench Total: 58.3

Overall Total: 145.94

Simulated Draft Team:

Starters:
QB: Jayden Daniels - 20.12
RB: James Conner - 14.4
RB: J.K Dobbins - 14.8
WR: Mike Evans - 10.1
WR: Jaxon Smith-Njigba - 19.4
WR: D.J. Moore - 8.6
TE: George Kittle - 12.5
DEF: MIN Defense - 4.0

Bench:
QB: Kyler Murray - 18.32
RB: Jahmyr Gibbs - 15.0
RB: Kyren Williams - 13.9
WR: Jauan Jennings - 3.6
WR: Devaughn Vele - 2.3
WR: Tyreek Hill - 8.0
TE: David Njoku - 6.7

Starters Total: 103.92
Bench Total: 67.82

Overall Total: 171.74

Figure 2: Opponent vs Model (Team 2) Week 1

Braeden's Current Team:
Starters:
QB: Jayden Daniels - 111.6 (Injured 5 weeks)
RB: De'von Achane - 235.0
RB: Jahmyr Gibbs - 259.0
WR: Nico Collins - 146.5 (Injured 1 week)
WR: Justin Jefferson - 151.9
WR: Mike Evans - 34.0 (Injured 7 weeks)
TE: Brock Bowers - 116.2 (Injured 4 weeks)
DEF: MIN Defense - 63.0

Bench:
QB: Kyler Murray - 77.8 (Injured 6 weeks)
RB: Kyren Williams - 172.0
RB: Kenneth Walker III - 124.4
WR: Courtland Sutton - 135.9
WR: DJ Moore - 116.3
WR: Tee Higgins - 141.5
TE: Mark Andrews - 102.0

 # Injured Players: 5
Current Starters Season Total: 1,117.2
Current Bench Season Total: 869.9

Current Overall Total: 1,987.1

Simulated Model Team:
Starters:
QB: Jared Goff - 191.0
RB: Zach Charbonnet - 96.2 (Injured 1 week)
RB: Rico Dowdle - 162.0
WR: Ja'Marr Chase - 193.7
WR: Amon-Ra St. Brown - 218.3
WR: Ladd McConkey - 142.4
TE: Brock Bowers - 116.2 (Injured 4 weeks)
DEF: DET Defense - 76.0

Bench:
QB: Baker Mayfield - 183.8
RB: Tony Pollard - 103.0
RB: Kareem Hunt - 113.9
WR: Terry McLaurin - 39.3 (Injured 7 weeks)
WR: Drake London - 177.0 (Injured 2 weeks)
WR: Josh Downs - 86.6 (Injured 1 week)
TE: Jonnu Smith - 61.1

 # Injured Players: 5
Current Starters Season Total: 1,195.8
Current Bench Season Total: 764.7

Current Overall Total: 1,960.5

Figure 3: Braeden vs Model (Team 6) Total Accumulation

Opponent's Current Team:
Starters:
QB: Bo Nix - 198.1
RB: Chuba Hubbard - 81.0 (Injured 2 weeks)
RB: Jonathan Taylor - 282.5
RB: Josh Jacobs - 180.5 (Injured 1 week)
WR: A.J. Brown - 126.7 (Injured 1 week)
WR: Jaxon Smith-Njigba - 255.0
TE: Sam Laporta - 106.9 (Injured 2 weeks)
DEF: PIT Defense - 81.0

Bench:
QB: Matt Barkley - RETIRED
RB: James Conner - 33.3 (Injured 8 weeks)
RB: Breece Hall - 152.9
RB: Treveyon Henderson - 136.8
WR: Terry Mclaurin - 39.3 (Injured 7 weeks)
WR: Devonta Smith - 148.4
WR: Jaylen Waddle - 151.5

 # Injured Players: 6 (33 weeks)
Current Starters Season Total: 1,311.7
Current Bench Season Total: 662.2

Current Overall Total: 1,973.9

Simulated Model Team:
Starters:
QB: Jayden Daniels - 111.6 (Injured 5 weeks)
RB: James Conner - 33.3 (Injured 8 weeks)
RB: J.K Dobbins - 115.9 (Injured 1 week)
WR: Mike Evans - 34.0 (Injured 7 weeks)
WR: Jaxon Smith-Njigba - 255.0
WR: D.J. Moore - 116.3
TE: George Kittle - 81.8 (Injured 5 weeks)
DEF: MIN Defense - 63.0

Bench:
QB: Kyler Murray - 77.8 (Injured 6 weeks)
RB: Jahmyr Gibbs - 259.0
RB: Kyren Williams - 172.0
WR: Jauan Jennings - 87.8 (Injured 2 weeks)
WR: Devaughn Vele - 24.1 (Injured 1 week)
WR: Tyreek Hill - 53.5 (Injured 7 weeks)
TE: David Njoku - 76.8 (Injured 1 week)

 # Injured Players: 10 (43 weeks)
Current Starters Season Total: 810.9
Current Bench Season Total: 751.0

Current Overall Total: 1,561.9

Figure 4: Opponent vs Model (Team 2) Total Accumulation

Braeden's team was able to outperform the Model's team (Team 6) in week 1 by a margin of 35 points, as shown in Figure 1. This differential includes both starters and bench players. On the other hand, the Model's team (Team 2) was able to outperform Braeden's opponent by a margin of 26 points, as shown in Figure 2. Once again, including both starters and bench players. While this is interesting, only looking at one week of the NFL season is not very indicative of each team's performance.

Instead, it is better to look at the entire season's point accumulation. At the time of the 2025 season data collection, only 12 of the 17 weeks were complete. For that reason, the following accumulations are through the first 12 weeks of the 2025 NFL season.

As shown in Figure 3, after 12 weeks, Braeden's team accumulated a total point value of 1,987.1 points, while the model's team (Team 6) accumulated 1,960.5 points. This is actually a good indicator that our model performs quite well. Braeden is a well-established fantasy football player, so the model being 1.36% off of Braeden's overall score is rather impressive. It is worth noting as well that the unpredictable component of injured players was equal in this team scenario. Both teams had 5 injured players each through week 12, indicating they were on a relatively even playing field.

As shown in Figure 4, after 12 weeks, the Opponent's team accumulated a total point value of 1,973.9 points, while the model's team (Team 2) accumulated 1,561.9 points. At first glance, the model appears to be performing quite poorly due to the 26.28% point gap. However, this turns out not to be the case when once again considering the unpredictability of injured players. The Opponent's team had a total of 6 injured players throughout the 12 weeks, for a total injury time of 33 weeks. The model's team had a total of 10 injured players throughout the 12 weeks, for a total injury time of 43 weeks. This means that the model's team lost out on 10 additional point accumulation opportunities. So, taking this aspect into account, it can be understood that the model is not as far off from expected performance as it originally appears.

In evaluating these two comparisons, it can be stated that the model does an acceptable job at evaluating potential player performance and generating winning teams. Braeden's team's current record is 8-5, while the Opponent's team is 9-4. This places both teams at the top of their league, and the similar scores performed by the model indicate good performance.

V. ISSUES ENCOUNTERED

Throughout the project, we encountered a few issues regarding the scoring of players. One challenge was players with few games/snaps played having high average statistics. For example, Player A, a running back who only played a handful of snaps, but had very successful plays, could have a higher yards-per-carry average than Player B, a player who had a decent performance over all 17 games. While the average of Player A might look enticing, it is not indicative of what their ranking should be, due to the lack of data. The model became "confused" by these high numbers and ranked these players higher than they deserved, which had a negative impact on our predictions.

To resolve this issue, we assigned thresholds for each position to filter out bad players. Initially, we used the following thresholds for each position:

- Quarterback: minimum of 1500 passing yards
- Running back: minimum of 600 rushing yards
- Wide Receiver: minimum of 500 receiving yards and at least 48 receptions
- Tight Ends: minimum of 300 receiving yards and at least 40 receptions

However, as a result, the player pool became too small, making it impossible to fill rosters for larger leagues (8 or 10 teams). We needed to find the right balance for the thresholds to find the middle ground between players excluded and players with minimal data. We adjusted these thresholds while manually checking the prediction of our model until we found the right values. In the end, we decided on the following values:

- Quarterback: minimum of 1500 passing yards
- Running back: minimum of 490 rushing yards
- Wide Receiver: minimum of 450 receiving yards and at least 40 receptions
- Tight Ends: minimum of 300 receiving yards and at least 40 receptions

This approach allowed us to maintain good predictions while ensuring that there were enough players for draft simulations of larger team sizes.

VI. FUTURE WORK

In terms of future work, one of the primary metrics the team would like to incorporate is the players' teams. Currently, our model does not consider any of the 32 NFL teams. Instead, our model considers player performance independent of their corresponding team. Our model is decent using only player metrics, but could be greatly improved by incorporating teams. This is because players on high-performing teams have a higher ceiling for the number of points they can score. For example, a high-performing wide receiver on a top team is likely to score more points than that same wide receiver on a bad team. The reason for this lies in the fact that better teams combine for more total offense and more total points scored than worse teams. Football is a team sport, so player production is not completely independent from team performance. An NFL team could have the best quarterback in the league, but without a good offensive line that is capable of keeping the pocket clear for an adequate amount of time, it will be hard to get passes off. In addition to this, if the quarterback's wide receivers are on the lower end of performance capabilities, the number of dropped passes is likely to increase. This is to say that it is difficult for individual star players to shine without good players around them. So, players on good teams should be weighted and ranked higher than those on lower-end teams.

Another potentially very beneficial tool that could be integrated into our system is player removal once that player has been picked by one of the draftees. Currently, our system just lists the players and performs full mock drafts rather than being integrated into a live draft system. This means that a user would have to manually keep track of what players have already been selected in our current system. It would be much more beneficial for the user if players were marked off the ordered list as draftees selected them. This feature would significantly improve the usability of our draft assistant. While this would be a much bigger project, integrating this system as a browser extension could allow users to utilize our assistant on any fantasy football platform of their choice.

VII. MEMBER RESPONSIBILITIES

While both team members collaborated throughout the project, each member was responsible for leading specific areas in which they had more knowledge.

Chris led the preprocessing, data cleaning, and filtering. As previously mentioned, the original dataset had significantly more columns than necessary, which meant that we needed to figure out which columns would actually aid us in creating a successful machine learning model. Since we used two datasets, we had to do this for both of them, reducing the offensive player dataset from 660 columns to 17 relevant features and reducing the defensive dataset from 45 to 7. This required us to go through each column manually to figure out which metrics would actually be helpful in predicting player performance.

Braeden led the machine learning model development and defined the metrics used to rank the players. As this was Chris's first time working with Fantasy Football, Braeden helped to choose which statistics and columns mattered the most for determining the performance of players. Braeden was also responsible for implementing the snake draft simulation algorithm, determining the ideal composition, and performing the random search to find the best hyperparameters for the model.

Both team members came to the conclusion that z-scores would be the best approach for calculating the overall rating of each player. Both members also worked together on setting and adjusting the thresholds that filtered out players. As previously mentioned, this required running the model multiple times, checking the player pools for each position, and making sure enough players remained. Since Braeden was knowledgeable about current players and their performance, he helped spot the best middle ground for these thresholds.

To evaluate the model, both team members ran mock drafts where teams were built using the model and then compared against teams that Braeden's current league created. This process involved going to the ESPN website and manually grabbing statistics for each player on both the model's draft team and the team we were comparing against. While the team members split up to each do one team, the process was still time-consuming as there was no automated way to get this data.

Lastly, both teammates worked on creating the visualization for the data. This included creating a Flask backend that would call the main Python function and use the returned prediction data in the frontend to display top players by position and provide a button to run the simulated draft with the desired number of teams.

Below is a table of the project's timeline:

Table 1: Timeline

Weeks	Tasks
Week: 1	Finish gathering data
Week: 2-3	Clean/Filter data to include only offensive positions and DEF team
Week: 4-5	Add scoring metrics to each position
Week: 6-7	Implement a Random Forest ML algorithm to draft players based on the player draft value metric
Week: 8-9	Finalize the algorithm for drafting and compare it to Braeden's team. Create the website for visualization.

REFERENCES

[1] Hyde, Philip. NFL Stats 2012-2024. Kaggle, 2025. https://www.kaggle.com/datasets/philiphyde1/nfl-stats-2012-2024