Predicting an Optimal NBA Fantasy Lineup

Eunice Park

Advised by Ethan Meyers, Yale Department of Statistics & Data Science

Introduction

This project will examine the applications of machine learning to DraftKings NBA Daily Fantasy Sports (DFS).

DraftKings is a dominant player in the DFS space. DraftKings NBA DFS allows users to draft a team lineup on a daily basis, picking players from any team playing on a particular game night. The team lineup comprises eight players, including a Point Guard, Shooting Guard, Small Forward, Power Forward, Center, general Guard (point or shooting), general Forward (small or power), and a Util (any of the ive positions). Each player in the lineup has an associated salary cost, typically ranging from $3,000 to $11,000 in DraftKings currency, and users have a $50,000 salary constraint to draft their team lineup. Players earn DraftKings fantasy points based on their performance in 9 various categories. These categories are as follows: points (PTS), 3 points made (3P), total rebounds (TRB), assists (AST), steals (STL), blocks (BLK), turnovers (TOV), double doubles (DD), and triple doubles (TD). Player i's total fantasy points (yi) is defined as follows:

$$yi = (PTSi) + 0.5(3Pi) + 1.25(TRBi) + 1.25(ASTi) + 2(STLi) + 2(BLKi) - 0.5(TOVi) + 1.5(DDi) + 3(TDi)$$

In fantasy sports betting, the crux of the problem is in trying to determine and predict fantasy scores for all players such that the most efficient lineup can be constructed in terms of fantasy value per salary unit. These defined limitations of a salary cap and player positions create a unique opportunity for a machine learning model to best predict which individual players can construct the most efficiently generating team. An optimal lineup will generate the highest number of fantasy points while staying within the salary and position constraints.

Dataset & Preprocessing

NBA Players Box-score Statistics

Kaggle, web scrapes player statistics for all NBA seasons, ranging from 1949-50 to 2022-23, from Basketball-Reference.com.

NBA Player Positions

Basketball-Reference.com, web scrapes player position data for all NBA seasons, player positions are important in constructing lineup

NBA Players Salary

Rotoguru, data includes player name, game date, various player statistics, and player's draft king salaries for each game date

Final Player Dataframe

Preprocessing steps yielded 4 clean dataframes: iltered_1_df (2022-2023 NBA season), iltered_2_df (2021-2022 NBA season), iltered_3_df (2020-2021 NBA season), and iltered_4_df (2022-2023 NBA season). Each of these dfs had all player statistics and total fantasy points for each player's historical game performance, with each row representing each game a NBA player has played within a particular season.

Implementation & Methods

LSTM
Projected Fantasy Points

'predict_next_game_scores', return the predicted fantasy pts for each player's next game or game after a cutoff point

Dynamic Programming
Optimal Lineup Construction

maximizes total projected fantasy score while adhering to constraints: 1) $50,000 salary and 2) lineup player positions

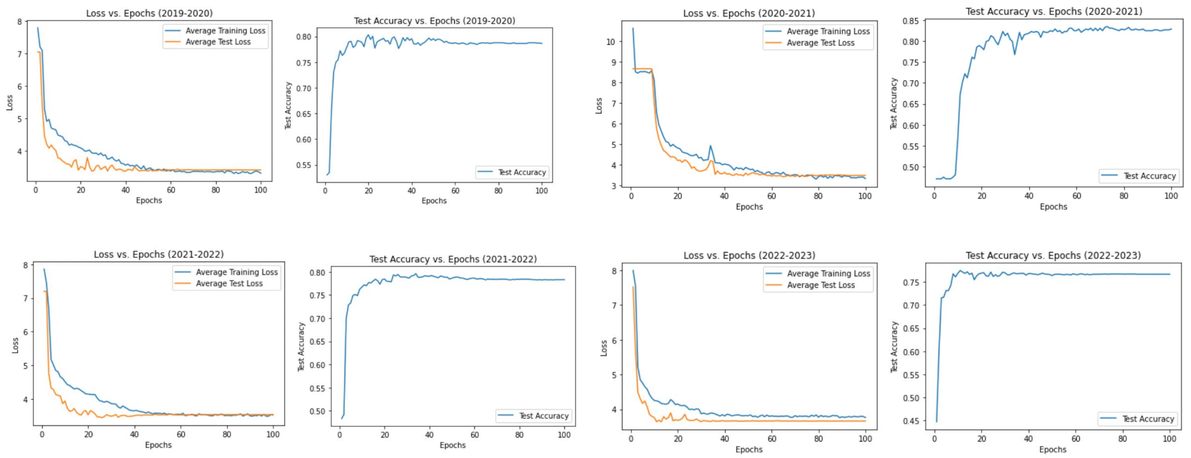
Results

```
Maximum Predicted Fantasy Score: 244.1683464050293
Optimal Lineup: {'PG': 'T.J. McConnell', 'SG': 'Josh Richardson', 'SF': 'Justin Holiday', 'PF': 'JaMychal Green',
'C': 'Rudy Gobert', 'G': 'Paul George', 'F': 'Justise Winslow', 'UTIL': 'Russell Westbrook'}
Total Cost of Optimal Lineup: 50000.0
```

Figure 1. NBA DFS Projected Optimal Lineup (based on predicted total fantasy points), 1/16/21

```
Maximum Predicted Fantasy Score: 277.5
Optimal Lineup: {'PG': 'Cole Anthony', 'SG': 'Frank Jackson', 'SF': 'Saddiq Bey', 'PF': 'Kenyon Martin Jr.', 'C': 'Ka
rl-Anthony Towns', 'G': 'Ricky Rubio', 'F': 'Cody Martin', 'UTIL': 'Russell Westbrook'}
Total Cost of Optimal Lineup: 50000.0
```

Figure 2. NBA DFS Best Optimal Lineup (based on historical actual total fantasy points)



Figures 3-7. Predicted Total Fantasy Points LSTM results for NBA Seasons (2019-20, 2020-21, 2021-22, 2022-23)

Conclusion

This project predicts an optimal fantasy lineup for DraftKings NBA DFS contest on any given game date. With historical game data, a LSTM model is constructed to predict a player's total fantasy points. With these predictions, an optimization algorithm using a dynamic programming-based approach generates a lineup of 8 players with the maximum number of predicted fantasy points.

Future work can be conducted through live testing on current DraftKings NBA DFS contests. Optimal lineups can further be compared with Fantasy Cruncher, which provides historical DFS contest results to see if generated optimal lineups are historically profitable. Furthermore, although this project focused on constructing one optimal fantasy lineup, future work can explore the benefit of submitting multiple fantasy lineups for DraftKings NBA DFS contest, which allows a user to submit up to 150 unique lineups. Similar to a portfolio optimization problem that balances risk to maximize reward, the construction of a fantasy lineup portfolio is an interesting future problem to explore.

The problem of constructing an optimal fantasy lineup is ripe for exploration. Utilizing the results of this project, a model to predict a player's total fantasy points and an algorithm that optimizes a valid lineup can be fine-tuned to increase accuracy.