Advanced NBA Analytics for Fantasy Basketball

Minsub Kim, Siavash Tahamtan, Claudio Fofiu, Kangin Choi, Anh Hoang, Jaeyong Kim

Summary

With the rise of NBA teams utilizing data science to search for the best players to build a championship team, advanced data analytics focused in fantasy basketball provide the same benefits to fantasy players by predicting NBA player's performance and visualizing the data for fantasy players to build the best fantasy team possible.

What is Fantasy Sports?

Fantasy sports are online prediction games where a participant puts together a virtual team of real sports players and earn points based on real life statistics that are converted into fantasy points. The participant receives more fantasy points when a player performs better in real life.

Problems

• Current advanced NBA analytics are not created for fantasy sports and are hidden away from the public.

• Lack of services providing NBA players predictions using advanced data analytics.

Example: Current providers only give current NBA season rankings and not future predictions.

Objectives

• Provide NBA players predictions using advanced data analytics for fantasy players.

• Provide the users with a dashboard that allow customization based on user's team needs.

NBA Data

Collected NBA player stats with third-party Python package (pypi.org/project/nba-api/)

• NBA player stats from seasons 2010 to 2020

• # of unique players: 507

• # of properties per player: 69

• After calculation: 2812 rows

Experiments and Evaluation

Used Python's SciKits package

Initial Experiment

• Linear Regression model only

Started with only one model and fed in 12 features: **GP, W, L, W, PCT, MIN, PLUS MINUS, DD2, TD3, HEIGHT, WEIGHT, SEASON, EXP,** and **POSITION.**

Using one model was not sufficient to produce reliable predictions due to lack of necessary features to predict 2021 season for NBA fantasy use.

Final Experiment

• **Linear Regression and Ridge Regression models**

Adjusted to 2 models in series with 1 model feeding into another.

Features split into two separate groups: **pre-features** and **features**.

• **Pre-Features (Fig. 1):** Most constant features in the dataset. Used to predict **Features (Fig. 1)**.

• **Features (Fig. 1):** Used as inputs to the final model to determine **Output (Fig. 1)**.

• **Output (Fig. 1):** Used to calculate Fantasy Points.

Pre-Features	Features	Output
PLAYER_AGE	GP = Games Played	FGM = Field Goals Made
POSITION	W = Wins	FGA = Field Goals Attempted
HEIGHT	L = Losses	FG3M = Field Goal Threes Made
WEIGHT	W_PCT = Win Percentage	FTM = Free Throws Made
	PLUS_MINUS	FTA = Free Throws Attempted
	MIN = Minutes	REB = Rebounds
	DD2 = Double Doubles	AST = Assists
	TD3 = Triple Doubles	TOV = Turnovers
		STL = Steals
		BLK = Blocks
		PTS = Points

Fig. 1

Dashboard/Result

• Players ranking based on Fantasy Points **(Fig. 2)**

• Allows customization based on category **(Fig. 3)**

Fig. 2

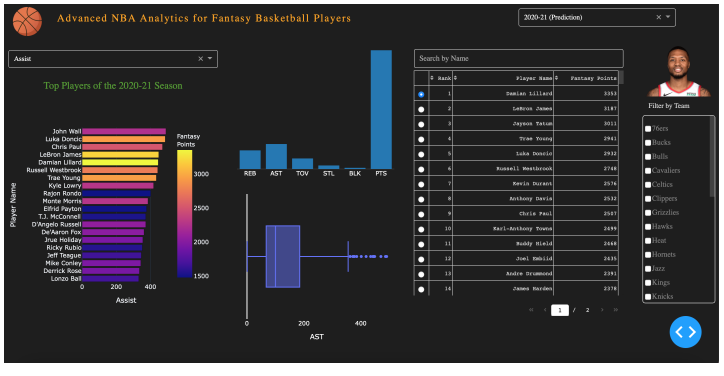

Fig. 3