

# Machine Learning Optimization Model to Predict Fantasy Basketball Teams

Author 1<sup>1</sup> and Author 2<sup>2</sup>

**Abstract**—The task of constructing optimal fantasy basketball teams presents a significant challenge. To address this challenge, a machine learning model leveraging Gurobi-Py Optimization was implemented. This approach provides an effective means of assembling top-performing teams while adhering to specific constraints, such as a \$50,000 salary cap and restrictions on the number of players in various positions.

The model was trained using an initial dataset sourced from DraftKings, supplemented by an additional dataset from Action Network. These datasets contain granular player statistics, including performance data from the last 5 days, ensuring alignment with recent player performance trends.

The linear search capabilities of the optimization model proved highly efficient in identifying the most promising players within specified positions, while respecting the imposed salary cap. Future investigations will explore advanced techniques, including vectorization and neural networks, to predict fantasy scores by segmenting player data samples. This approach promises to further enhance the efficacy of the optimization technique, optimizing the selection of fantasy basketball rosters.

## I. INTRODUCTION

In recent years, the computational analysis of statistics has emerged as a pivotal tool within the National Basketball Association (NBA) for a multitude of purposes. Professional organizations have harnessed player statistics to discern optimal team compositions. Notably, the realm of Fantasy Basketball has seen a surge in popularity, with platforms like FanDuel and DraftKings providing an enticing arena for users to draft fantasy teams and compete against others.

These platforms introduce a compelling incentive for users to craft the most competitive team

possible, as a team's performance directly correlates with potential monetary gains. Users who assemble teams that outperform their competitors receive higher payouts relative to their initial investments. As such, the pursuit of constructing an exemplary fantasy team has taken center stage.

To address this challenge and elevate team selection to a further level, a machine learning optimization model emerges as a valuable asset. This model systematically identifies and assembles the most advantageous players for a fantasy team while navigating significant constraints, such as salary limits and player positions. Through this innovative approach, the model aims to optimize team composition, maximizing the team's potential to accrue points and outperform competitors on platforms like FanDuel and DraftKings.

## II. GOAL

The optimization model under consideration is designed with a focus on the DraftKings platform, although its applicability extends to any platform furnished with adequate data. DraftKings' fantasy basketball format involves the selection of an 8-player team, each with designated positions that encompass a Point Guard, Shooting Guard, Small Forward, Power Forward, Center, any Guard (Point or Shooting), any Forward (Small or Power), and a Utility player, chosen from any of the five primary positions.

Accompanying these positional constraints are associated salary caps, with premier players commanding prices of approximately \$12,000. Users are allocated a budget of \$50,000 to curate their team, with the collective cost of all selected players not surpassing this threshold.

<sup>1</sup> Author 2, Author 1 Email, Author 1 Affiliation

<sup>1</sup> Author 2, Author 2 Email, Author 2 Affiliation

The overarching objective for users is to strategically choose eight NBA players who can yield the highest fantasy point totals while adhering to the stipulated constraints, encompassing both positions and salaries. Fantasy points are accrued based on fundamental box-score statistics, including points, rebounds, blocks, and steals.

In pursuit of this goal, the model leverages machine learning techniques to optimize the composition of a fantasy team, a task for which existing optimization models primarily reside within fantasy sites such as DraftKings and FanDuel, accessible only through paywalls. This endeavor aspires to contribute to the community by creating an open-source, freely accessible fantasy basketball optimization model, thereby broadening accessibility to this valuable tool.

### III. DATA COLLECTION & PREPROCESSING

In this study, we conducted a comprehensive data preprocessing procedure to obtain a refined dataset for player selection in daily fantasy sports platforms, specifically focusing on DraftKings and FanDuel. The initial player dataset was extracted from DraftKings and comprised player information including salary cap restrictions, player positions, and player season averages for various statistics such as points, rebounds, assists, etc.

Recognizing the significance of considering player injuries in the player selection process, we employed web scraping techniques to retrieve injury data from Action Network ([actionnetwork.com](https://www.actionnetwork.com)). This additional dataset provided us with the names of injured players along with their corresponding injury status. To consolidate the data, we merged the original DraftKings dataset with the injury data, enabling us to associate each player with their injury status. Consequently, we removed players who were labeled as "out" or "questionable" due to injuries from the dataset. This meticulous curation process played a pivotal role in ensuring that the

players selected for fantasy teams would actively participate in forthcoming games. Consequently, it mitigated the potential risk associated with the selection of inactive or underperforming players owing to injury-related absences.

Furthermore, our pursuit of an enhanced player selection process prompted us to integrate more contemporary player statistics. To realize this objective, we procured player performance metrics encompassing crucial facets such as points scored, three-point shooting percentages, steals, and other pertinent indicators, which were sourced from basketball-reference.

Notably, we chose to focus on a more recent data window, specifically spanning the last ten days, which encompassed approximately five games. This selection was made purposefully, as it accounts for the dynamic nature of the NBA season. The performance trajectories of players can exhibit substantial fluctuations over relatively brief intervals, and the early phases of the season may not offer accurate insights into their long-term capabilities.

By embracing these updated statistics, we were equipped to glean invaluable insights into players' recent on-court exploits. Furthermore, we harnessed this data to compute each player's average fantasy points per game over their most recent five-game stretch. This calculation adhered to DraftKings' established point calculation system, enabling a more robust assessment of their current form and suitability for selection.

Through the outlined data preprocessing pipeline, we obtained a refined dataset that accounted for player positions, recent performance trends, injury status, and comprehensive player statistics. This enhanced dataset served as a solid foundation for subsequent player selection and optimization strategies in daily fantasy sports competitions.

### IV. OPTIMIZATION

As previously stated, the overarching objective of this project is to select the most exceptional

players to optimize point accumulation within the constraints of a predefined salary cap and the available player positions. A pivotal facet of this endeavor is the development of a machine learning model, capable of discerning nuanced steps and progressively refining its accuracy in fulfilling this mission.

To achieve this aim, optimization methodologies are used as a means to enhance the performance, efficiency, and overall efficacy of our machine learning algorithms. This optimization framework is thoughtfully framed through the rigorous definition of pertinent variables and constraints, driving the systematic selection of the optimal player lineup as seen below:

Binary variable  $(x_{ij})$ : Takes a value of 1 if player  $i$  is selected for position  $j$  in the fantasy team; 0 otherwise.

Parameters:

$(P_i)$ : Average fantasy points per game for player  $i$ .

$(C_i)$ : Salary cap cost for player  $i$ .

$(S_j)$ : The number of players required for position  $j$  (e.g.,  $(S_j) = 3$  for guards,  $(S_j) = 1$  for center, etc.).

Budget: The total budget available to select the team (e.g., \$50,000 for DraftKings). The objective function aims to maximize the total fantasy points scored by the selected team:

Maximize:  $\sum (P_i * x_{ij})$

The specialized tool employed for this purpose is Gurobi-Py, strategically leveraged to streamline the player selection process. Diverging from conventional iterative for-loops, Gurobi-Py harnesses its robust linear search capabilities to traverse expansive solution spaces with remarkable celerity. This advanced capability facilitates the rapid identification of the most advantageous team composition, aligned with predefined criteria.

This pioneering approach yields substantial acceleration in the optimization process, rendering Gurobi-Py the preeminent choice for resolving

intricate challenges within the realm of fantasy sports. Its capacity to efficiently navigate complex problem domains underscores its indispensability in the pursuit of superior player selection strategies.

In the initial phase, our model establishes a comprehensive set of variables that encapsulate player identities and their associated statistical attributes, prominently including metrics such as the average points scored per game. These variables constitute the fundamental building blocks upon which our fantasy team construction process is firmly grounded, as they facilitate the seamless integration of player-specific data directly into the model's computational framework.

During the subsequent phase, we craft the central objective of our model, which is the maximization of the projected points per game for the assembled team. This pivotal objective is strategically engineered to ensure that the resulting team configuration attains the highest attainable fantasy score, thereby affording the user an unparalleled advantage in the competitive landscape of fantasy sports leagues.

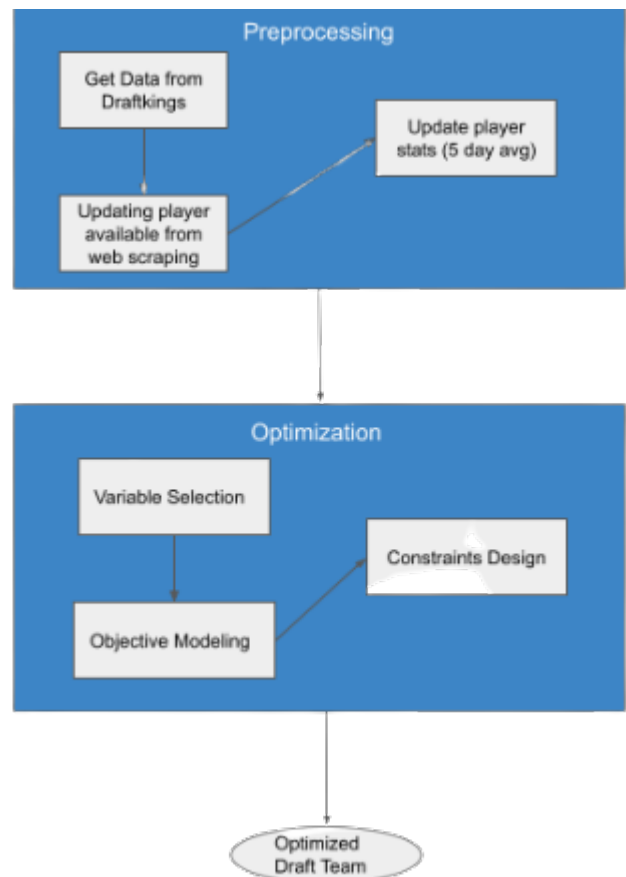
In the final phase, our model systematically introduces a set of pivotal constraints to orchestrate the assembly of a well-balanced and highly competitive fantasy basketball team, all while rigorously adhering to the stipulated regulations of the fantasy sports platform. The foremost among these constraints revolves around the platform's stringent salary cap, which imposes an upper bound on the aggregate cost of the players' salaries, typically set at a predefined threshold (e.g., below \$50,000 for DraftKings). This paramount constraint serves a dual purpose: it ensures the selected team remains prudent, abiding by budgetary restrictions, while simultaneously fostering strategic decision-making by necessitating the thoughtful selection of high-value players within these financial confines.

Furthermore, our model rigorously imposes position-specific constraints to meticulously curate a team composition characterized by diversity and

adherence to the platform's stipulations. These constraints meticulously delineate the requisite number of players designated for each specific position, commonly stipulating the inclusion of three guards, three forwards, one center, and one versatile utility player. This consideration of distinct positions effectively shields against the inadvertent generation of team configurations that run afoul of the stringent rules and regulations governing the NBA fantasy platform.

While the model ensures compliance with the foundational requisites for player positions, we recognize the significance of discerning between sub-positions, such as point guards and shooting guards. This differentiation is instrumental to avoid forming teams overloaded with one category of guard at the expense of others. The nuanced consideration of these sub-positions decisively contributes to the cultivation of team rosters that are inherently more balanced and competitive in nature.

As a result of our comprehensive optimization process, our model consistently delivers optimal fantasy basketball teams each game day, taking into consideration account salary and positional restraints. By providing a free and open-source solution, we aim to expand accessibility and enable users to make informed decisions in managing fantasy basketball draft teams.



Flowchart Caption Goes Here...

## V. FUTURE WORK

By persistently advancing our efforts to optimize player selection for NBA fantasy, we lay the foundation for further refinement and expansion of our methodologies. These ongoing refinements are poised to propel the model beyond its current scope, transitioning from an analysis reliant on 5-day averages of player scores for optimization to a paradigm underpinned by sophisticated techniques such as vectorization and neural networks. This strategic methodology represents a pivotal step toward the creation of fantasy scores that are engineered for superior performance and precision.

This process can be enhanced with the usage of each player datapoint as a sample, dividing into vector/usage rates of the following:

1. Playing time
2. Shots taken
3. Turnovers
4. Assists/blocks

This partitioning approach opens up the avenue for the implementation of advanced techniques, namely vectorization and deep learning neural network models. These methodologies can be leveraged to analyze player performance in the realm of fantasy sports with remarkable precision, surpassing the efficacy of conventional strategies that rely solely on the use of 5-day averages for optimization purposes.

Furthermore, it's noteworthy that this model's versatility extends beyond the domain of basketball, encompassing various other fantasy and scoring platforms across different sports. One such arena where its applicability shines is within the context of the National Football League (NFL). Football enthusiasts, much like their basketball counterparts, frequently engage with fantasy and gambling platforms, including FanDuel and DraftKings, to curate strategic teams geared toward accruing the highest possible point totals.

The optimization techniques harnessed through Gurobi, which have demonstrated their effectiveness in the realm of basketball, are poised to make a transformative impact within the NFL fantasy sports landscape as well. This model, adept at crafting optimal teams while considering salary and player position constraints, holds the potential to revolutionize team selection strategies in the pursuit of maximizing fantasy sports success.