**NBA & WNBA Player Evaluation & Roster Optimization**

This project leverages advanced supervised learning techniques and optimization algorithms to create a useful metric for evaluating player performance and optimizing basketball team rosters.

**Abstract**

We developed a performance metric aimed at maximizing wins, which was subsequently used to optimize team rosters through the CVXPY optimization library. This approach provides a systematic way to evaluate players and ensure optimal team compositions based on various constraints such as budget and positional needs.

**Introduction**

The project addresses the need for reliable and comprehensive player evaluation metrics, especially for expansion teams constructing competitive rosters. Traditional methods often rely on subjective assessments and basic statistics, leading to inconsistent and suboptimal decisions. By combining supervised learning techniques with optimization algorithms, our metric provides a holistic view of player performance, enhancing decision-making for free agency, drafts, and trades.

**Methods**

**Data Collection and Preparation**

* **Datasets**: NBA box scores (2003-2024), RAPM data, and player statistics (2010-2022).
* **Data Cleaning**: Separated 'TeamSeason' into 'team' and 'season', merged home and away datasets, and aggregated statistics.
* **Modeling**: Addressed autocorrelation with a fixed effects model and included two lags of total wins. Assigned seasons before 2020 to the training set and tested on 2020-2022 seasons.
* **Regression Models**: Applied Ridge, LASSO, and ElasticNet regression models with five-fold cross-validation. LASSO performed best with an MSE of 15.25 and an R² of 88.04%.

**Player Coefficient Mapping**

* **Normalization**: Standardized player statistics to match model data.
* **Coefficient Calculation**: Computed coefficients by taking the dot product of model coefficients with player statistics, standardizing to an average of 100 for interpretability.

**Optimization**

* **Roster Construction**: Used CVXPY to maximize player coefficients under constraints (e.g., positional requirements and salary caps).
* **Salary and Season Constraints**: Constructed optimal rosters for 2023-24 season based on 2022-2023 statistics and salaries.

**Results**

* **Model Performance**: LASSO model produced the best results, favoring taller, more physical teams.
* **Optimization Outcomes**: Optimized rosters included many top players, validating the metric's effectiveness. Introduced salary constraints yielded practical NBA 10-man rosters.
* **Future Work**: Plans to integrate more granular financial data and identify free agents or trade targets for tangible team recommendations.

**Discussion**

This project demonstrates the power of combining supervised learning and optimization techniques to evaluate and select basketball players, constructing competitive rosters. Potential enhancements include:

* Integrating injury data for risk evaluation.
* Adding detailed contract information for nuanced analysis.
* Including market availability data for real-time optimization.

The utility of this model extends beyond basketball, providing valuable insights for any team sport where player performance and team composition are critical.

**Conclusion**

The integration of supervised learning and optimization techniques represents a significant advancement in sports management. By refining and expanding this model, its applications across different sports and contexts are vast. Leveraging sophisticated models for data-driven decisions will lead to better performance and greater success on the field or court.