

CSI 4142 – DATA SCIENCE

Project Phase 4. Part B.2

Table 1: NBA MVP Prediction - Metrics

	Model	RMSE	R ²	Construction Time
0	SVM	0.019	0.871	66.417
1	Random Forest	0.013	0.938	26.330
2	Gradient Boosting	0.018	0.878	115.411

Table 2: NBA MVP Prediction - Results

	Player	MVP Share Real	MVP Rank SVM	MVP Share SVM	Season	MVP Rank Random Forest	MVP Share Random Forest	MVP Rank Gradient Boosting	MVP Share Gradient Boosting
3	Nikola Jokić	0.875	Nikola Jokić	0.775	2021-22	Nikola Jokić	0.681	Nikola Jokić	0.231
4	Joel Embiid	0.706	Giannis Antetokounmpo	0.634	2021-22	Giannis Antetokounmpo	0.520	Joel Embiid	0.187
5	Giannis Antetokounmpo	0.595	Joel Embiid	0.606	2021-22	Joel Embiid	0.514	Giannis Antetokounmpo	0.158

Summary

In our NBA MVP prediction project, we used three machine learning algorithms: Support Vector Machine (SVM), Random Forest, and Gradient Boosting, to forecast MVP vote shares for NBA players in the 2021-22 season.

We looked at the results and the Random Forest algorithm outperformed the other algorithms in predictive accuracy, it has a lower Root Mean Squared Error (RMSE) and higher R-square (R²) values.

If we look at the construction time metric, Random Forest was also the most efficient algorithm, taking significantly less time than Gradient Boosting and less than SVM. When we look at the MVP rankings generated by each algorithm, Nikola Jokić was ranked the highest across all models, there were variations for the second highest ranked. This shows us the importance of considering multiple models to gain a comprehensive understanding of the data analysis and avoid relying on only one model's prediction.

Our analysis highlights the importance of not only evaluating performance but also consider efficiency and robustness of models. By using different models, we were able to gain a comprehensive idea of who is the most likely NBA MVP candidate by leveraging machine learning techniques and using player performance to make strategic decisions.