# PREDICTION OF THE MOST VALUABLE PLAYER (MVP) ON NBA USING MACHINE LEARNING ALGORITHMS

## DAVUTCAN KÖSEMEN CEYHUN AY

MARMARA UNIVERSITY FACULTY OF TECHNOLOGY

## Abstract

This research develops a machine learning approach to predict NBA MVP awards using 1997-2022 player statistics, incorporating unique clutch time metrics and comparing multiple algorithms for optimal accuracy.

## Keywords

Machine Learning, Artificial Intelligence, MVP, NBA, Data Analytics

## Introduction

Team sports are followed with great interest around the world and hold significant social and economic importance. One of the most prestigious individual honors in the NBA is the Most Valuable Player (MVP) award. The prediction of the NBA MVP award is a challenging research topic that lies at the intersection of sports analytics and artificial intelligence.To achieve this, the performances of various algorithms—including Random Forest, Gradient Boosting, ANN, CNN, RNN, GNN were comparatively analyzed

## Metodology

The dataset used in this study consists of basic and advanced performance statistics of NBA players from 1997 to 2022, obtained via Selenium. In addition to standard and advanced player statistics (e.g., points, assists, rebounds as basic metrics, along with PER, WS, and VORP as advanced ones), two innovative features rarely addressed in the literature were incorporated into the dataset: Clutch Time Score and Team Standing
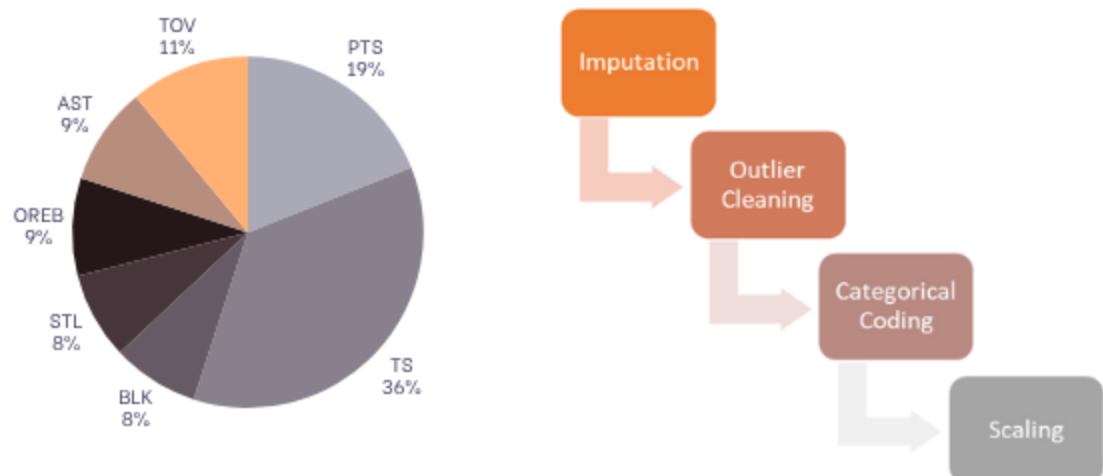
For deep learning algorithms the performance impact of activation functions, optimal number of layers and neurons,
In tree-based algorithms, hyperparameters such as tree depth, minimum number of leaf nodes, and the number of epochs.
For the linear classification method SVM, the C and kernel parameters were tested to analys the best result under different parameters.

## Literature

Machine learning techniques in NBA MVP prediction have evolved through various studies, with previous research achieving accuracies ranging from 25% to 88.7% using models like Random Forest, XGBoost, and Neural Networks. Unlike most previous work focusing solely on overall season statistics, this study introduces distinctive contributions by incorporating clutch-time performance metrics, using team ranking instead of win percentage, applying feature extraction to varius machine learning and deep learning models.

## Results and Analyses

- XGBoost achieved highest recall (80%) among traditional ML models
- CNN model demonstrated superior real-world performance with 80% accuracy
- Deep learning models (ANN, CNN, GNN) showed varied performance on imbalanced dataset
- CNN model correctly predicted 4/5 actual MVPs (80% accuracy)
- GNN model achieved 60% top-three accuracy
- Both models consistently identified real MVP candidates within top predictions
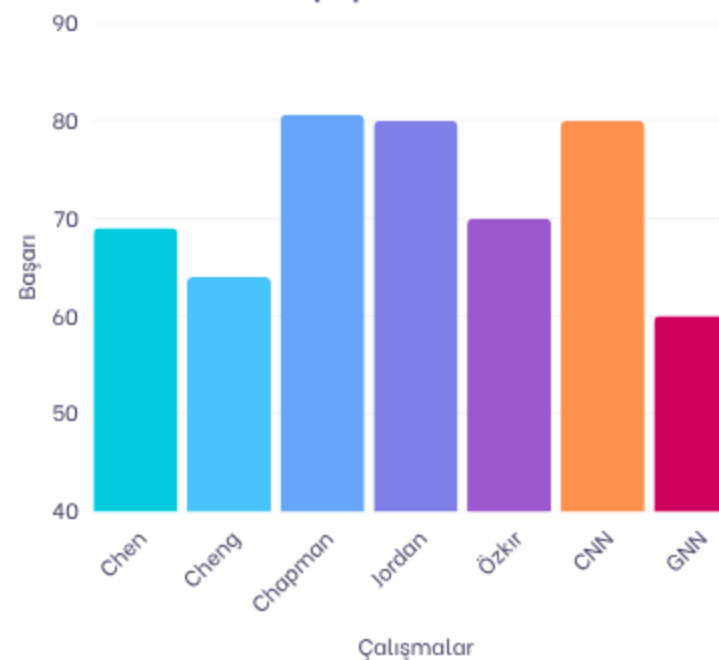
| Season | Real MVP | Predict (1) | Predict (2) | Predict (3) |
|---|---|---|---|---|
| 2018 | James Harden | Lebron James | Demarcus Cousins | Kevin Durant |
| 2019 | Giannis Antetokounmpo | Giannis Antetokounmpo | Karl-Anthony towns | James Harden |
| 2020 | Giannis Antetokounmpo | Giannis Antetokounmpo | Anthony davis | James Harden |
| 2021 | Nikola Jokić | Nikola Jokić | Giannis Antetokounmpo | Joel Embiid |
| 2022 | Nikola Jokić | Nikola Jokić | Giannis Antetokounmpo | Karl-Anthony towns |

## Conclusion

The proposed model achieves high MVP prediction accuracy, with innovative clutch time performance variables significantly contributing to success, demonstrating potential applications in sports analytics, player evaluation, and team strategy development.