

A Hybrid Machine Learning and MCDM Approach for Optimal Basketball Team Selection: A Decision Support System Framework

Burcu Ozcan Turkkan, Ejder Aycin, Sibel Fettahoglu*

Department of Business Administration, Faculty of Business, Kocaeli University, Kocaeli, Turkiye. *Corresponding Author's Email: stopdemir@kocaeli.edu.tr

Abstract

The goal of this study was to create a decision-support system for establishing the best team roster for basketball using machine learning and MCDM techniques. The research was performed using the statistical analysis of data pertaining to the fifty best collegians who played in the 2024-2025 NBA regular season, as reported by ESPN. The players were assessed on ten separate performance metrics such as: points per game, field goal percentage, three-point shooting percentage, free-throw shooting percentage, total rebounds - total and defensive rebounds, assists, turnovers, steals and blocks. The first phase used the Two-Step clustering algorithm to categorize the players into five groups based on performance criteria, or traditional basketball positions. The second phase employed RAWEC for player rankings in each cluster based on performance metrics to identify the five highest ranked players to start for that team. Ultimately an analytical, data-driven decision-support system model was created, providing an objective means to identify players for starting positions, which will assist in the sports industry with player analysis, roster design and prediction of players' future performance. Overall, this study provides evidence that the use of AI based models will significantly increase the efficiency of decision making in professional sports.

Keywords: Basketball Team Selection, Multi-Criteria Decision-Making (MCDM), RAWEC, Two-step Clustering, Machine Learning.

Introduction

Basketball is one of the most popular team games in the world. To gain a competitive advantage, team owners make significant investments in players, coaches, staff, facilities and amenities. During the 2022-2023 NBA season, clubs spent an average of US\$156 million on player salaries (1). Sports analytics means making decisions about game strategy, player health and financial management with data science and machine learning (ML) tools. For example, unplanned absences due to injuries can lead to significant financial losses for basketball teams (2). In today's sports industry, performance measurement and team-building processes have evolved from relying solely on technical knowledge or intuitive decisions to effectively utilizing data science and AI-powered analytical methods. Individual player performance can be evaluated using multidimensional criteria and these criteria can be combined to facilitate team optimization efforts (3, 4). In traditional methods, these selections are largely based on coaches' experiences,

observations and intuitive judgments. However, this approach can lead to inconsistent results due to factors such as subjective assessments, human errors and lack of information. Therefore, in recent years, studies on the use of decision support systems (DSS) and ML-based models in sports have increased (5). A review of studies on player performance evaluation and selection reveals that, in addition to clustering and classification algorithms, multi-objective decision-making techniques are frequently employed. Against this background, decision-making processes in sports have long relied on intuitive and experience-based approaches. However, recent advances in data collection technologies have enabled the digital analysis of player performance data. This development has paved the way for the more effective use of data science, artificial intelligence and MCDM methods in sports management. In a statistically rich sport like basketball, evaluating player performance with analytical methods provides a scientific basis for team strategies.

This is an Open Access article distributed under the terms of the Creative Commons Attribution CC BY license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution and reproduction in any medium, provided the original work is properly cited.

(Received 15th March 2026; Accepted 25th August 2026; Published 02nd October 2026)

Some researchers stated that machine learning methods can give good results in predicting the match results of teams that have not played each other before and developed a multi-objective decision model that includes DEMATEL and ANP methods to encourage elderly individuals to participate in sports (6, 7). Some researchers made predictions by modeling many NBA games with deep learning and artificial neural networks and using artificial neural networks by including the aging factor into the analysis (8). Machine learning was integrated with some algorithms for prediction of professional basketball games and identifying the variations within performance indicators that are critical in determining the outcomes of the games and created the optimal team roster based on the clustering, team model, rates and player salaries in their study (9-13).

In team sports like basketball, the roster building process presents a complex optimization problem due to diverse player profiles, multi-dimensional performance indicators and position-based constraints. This study proposes a unique hybrid model integrating the clustering algorithm with the Rank Analysis of Weighted Evaluations of Criteria (RAWEC) method. Unlike commonly used approaches in the literature that rely on only one of the clustering or ranking methods, this study first divides players into homogeneous clusters based on their similar characteristics using unsupervised machine learning and then performs a multi-criteria ranking analysis for each cluster.

This study uses an integrated Two-Stage clustering algorithm and RAWEC methods to create an ideal roster of the best players for the 2024-2025 NBA season. The data set used in the study includes ten performance criteria for the top fifty NBA players as determined by ESPN rankings. The Two-Stage clustering algorithm was used to automatically group players with similar characteristics. Clusters representing five positions were created using the Two-Stage method; then, players within each cluster were ranked according to their

performance using the RAWEC method. The RAWEC method allows for ranking alternatives by normalizing the decision matrix and comparing ideal-anti-ideal solution values. Players with the highest performance scores in each position were included in the top five, while the others were included in the reserve roster.

Three main purposes of the research can be summarized under the headings:

- (a) It presents a unique decision support architecture that integrates machine learning and multi-criteria decision-making methods for the team roster building process,
- (b) This study demonstrates the usability of the RAWEC method, which has not yet found sufficient application in the field of sports analytics, in this context. The study encourages the usage of a more widespread and systematic use of AI-powered decision support systems in the sports industry,
- (c) It offers an objective and repeatable analytical framework that minimizes subjective evaluations in player selection processes.

The proposed model was tested on performance data from the top fifty NBA players during the 2024-2025 regular season. The findings reveal that hybrid analytics approaches have the potential to significantly improve roster optimization strategies and offer actionable implications for sports management professionals.

Methodology

The RAWEC method is employed to prioritize alternatives within the solution process of decision-making problems involving multiple criteria (14). RAWEC is distinguished by its integration of two conventional procedures: the creation of the weighted normalized decision matrix and the calculation of deviations from ideal and anti-ideal values. The computational procedure of the RAWEC method comprises the four steps delineated in Table 1. These steps comprise of Equations [1-6].

Table 1: The Procedure and Equations of RAWEC Method

Step	Description	Equation No.
Step 1: Constructing Decision Matrix	$X = \begin{bmatrix} C_1 & C_2 & \dots & C_n & A_1 & A_2 \\ \vdots & A_m & [x_{11} & x_{12} & \dots & x_{1n} & x_{21} & \vdots & x_{m1} & x_{22} & \dots & x_{2n} & \vdots \\ \vdots & x_{m2} & \dots & x_{mn} & \end{bmatrix}$	[1]

Step 2: Normalization	$n_{ij} = \frac{x_{ij}}{x_{jmax}} \text{ and } n'_{ij} = \frac{x_{jmin}}{x_{ij}} \text{ for maximization criteria}$	[2]
	$n_{ij} = \frac{x_{jmin}}{x_{ij}} \text{ and } n'_{ij} = \frac{x_{ij}}{x_{jmax}} \text{ for minimization criteria}$	[3]
Step 3: Calculating the deviation from the criterion weight	$v_{ij} = \sum_{i=1}^n w_j(1 - n_{ij})$	[4]
	$v'_{ij} = \sum_{i=1}^n w_j(1 - n'_{ij})$	[5]
Step 4: Calculating the performance value	$Q_i = \frac{v'_{ij} - v_{ij}}{v'_{ij} + v_{ij}}$	[6]

Note: w_j represents the weight coefficient of the crit, Q_i value can range between -1 and 1. The highest Q_i value is more desirable.

The RAWEC method, which is used in many different decision-making areas such as facility location selection (15), banking sustainability (16) and energy source (17) decisions, has not been found in any research where it has been used in sports-related decisions. The study was conducted in the field of basketball. It aimed to improve personal efficiency and accelerate decision-making. The Two-Step method was used as a machine learning technique and five clusters were identified for the five positions. The Two-stage clustering algorithm was chosen because of its suitability for medium-sized datasets and its ability to automatically determine the optimal

number of clusters (18) through statistical criteria such as the Bayesian Information Criterion (BIC). Furthermore, the method integrates hierarchical and partitioning approaches, increasing its robustness in the face of non-normally distributed data and variable scale differences. Given the heterogeneous distributions of basketball performance statistics, the Two-Stage method offers methodological flexibility and improved clustering reliability. Five different RAWEC models were developed for each cluster and the top players were chosen for each group. The game schedule was planned based on these rankings. The model shown in Figure 1 is detailed.

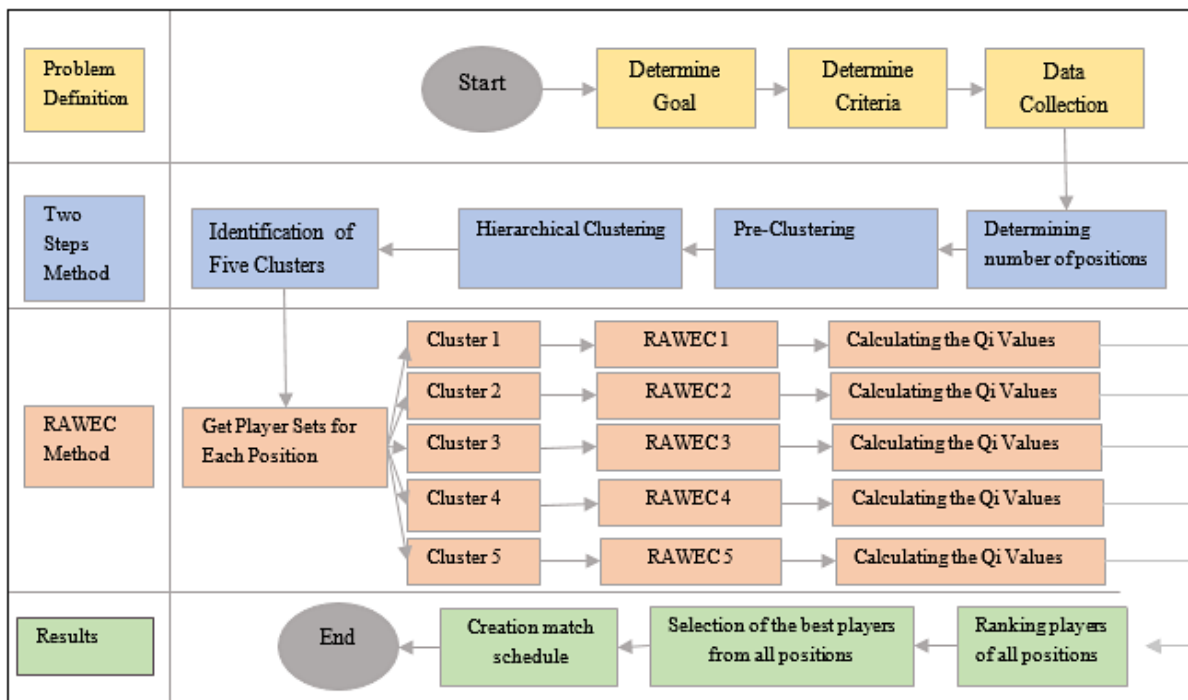


Figure 1: Generated Model

Compared with prior work, the proposed Two-Step Clustering-RAWEC framework differs from existing sports-analytics approaches in two key ways: First, studies that rely solely on clustering (19, 20) or apply MCDM methods directly to the full player pool (6, 7, 21-24) either stop short of producing a ranked roster or rank players across fundamentally different positions on a single scale, risking unfair comparisons (e.g., a center rebounding load against a guard assist rate). Second, ML prediction models (8, 25-30) and studies combining clustering with salary or rate data (9-13) do not apply a formal ideal-anti-ideal ranking within each group. The present framework addresses both issues by first clustering players into position-homogeneous groups, so that ranking is only performed among comparable players and then applying RAWEC, whose ideal-anti-ideal calculation weighs beneficial and non-beneficial criteria together in one transparent

score. This sequential design, together with the first documented use of RAWEC in sports team selection, distinguishes the framework from prior approaches.

Data

This study aims to identify the best NBA team lineup for the 2024-2025 regular season. It assumes that the specific positions played by a randomly selected group of players are unknown. Five clusters will be created corresponding to the five basketball positions. The team will be formed by choosing five starters and seven substitutes from these clusters. MCDM methods will be employed to streamline the selection process. Considering these factors, the Two-Step algorithm for clustering players and the RAWEC methods were selected for player selection. Ten criteria were identified for this process, as shown in Table 2.

Table 2: Criteria Used in Multidimensional Performance Evaluation

Criteria	Definition (Per Game)	Criteria Type
C1	Total Points	Max
C2	Field Goal Percentage	Max
C3	Three-Point Percentage	Max
C4	Free Throw Percentage	Max
C5	Offensive Rebound	Max
C6	Defensive Rebound	Max
C7	Assist	Max
C8	Turnover	Min
C9	Steal	Max
C10	Block	Max

ESPN is a premier data source for player evaluation, with two different methods. Below are the standard NBA Rank Top 100 Players list (as voted on by analysts prior to the season) and the

live ESPN Season Rating (which factors in statistical performance throughout the season). The study is based on the top 50 players (by the 2024-2025 ESPN Season Rating).

Table 3: Data of the Study

Player	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Shai Gilgeous-Alexander	32.70	51.90	37.50	89.80	0.90	4.10	6.40	2.40	1.70	1.00
Giannis Antetokounmpo	30.40	60.10	22.20	61.70	2.20	9.70	6.50	3.10	0.90	1.20
Nikola Jokić	29.60	57.60	41.70	80.00	2.90	9.90	10.20	3.30	1.80	0.60
Luka Dončić	28.20	45.00	36.80	78.20	0.80	7.40	7.70	3.60	1.80	0.40
Anthony Edwards	27.60	44.70	39.50	83.70	0.80	4.90	4.50	3.20	1.20	0.60
Jayson Tatum	26.80	45.20	34.30	81.40	0.70	8.00	6.00	2.90	1.10	0.50
Kevin Durant	26.60	52.70	43.00	83.90	0.40	5.70	4.20	3.10	0.80	1.20
Tyrese Maxey	26.30	43.70	33.70	87.90	0.30	3.10	6.10	2.40	1.80	0.40
Cade Cunningham	26.10	46.90	35.60	84.60	0.80	5.30	9.10	4.40	1.00	0.80
Jalen Brunson	26.00	48.80	38.30	82.10	0.40	2.50	7.30	2.50	0.90	0.10
Paolo Banchero	25.90	45.20	32.00	72.70	1.10	6.40	4.80	3.00	0.80	0.60
Devin Booker	25.60	46.10	33.20	89.40	1.00	3.10	7.10	2.90	0.90	0.20
LaMelo Ball	25.20	40.50	33.90	84.30	1.00	3.90	7.40	3.60	1.10	0.30
Damian Lillard	24.90	44.80	37.60	92.10	0.50	4.20	7.10	2.80	1.20	0.20

Anthony Davis	24.70	51.60	28.20	77.50	2.60	8.90	3.50	2.20	1.20	2.20
Kyrie Irving	24.70	47.30	40.10	91.60	1.20	3.60	4.60	2.20	1.30	0.50
Zion Williamson	24.60	56.70	23.10	65.60	2.50	4.70	5.30	3.00	1.20	0.90
Stephen Curry	24.50	44.80	39.70	93.30	0.60	3.90	6.00	2.90	1.10	0.40
Karl-Anthony Towns	24.40	52.60	42.00	82.90	2.90	9.80	3.10	2.70	1.00	0.70
LeBron James	24.40	51.30	37.60	78.20	1.00	6.80	8.20	3.70	1.00	0.60
Victor Wembanyama	24.30	47.60	35.20	83.60	1.80	9.20	3.70	3.20	1.10	3.80
Trae Young	24.20	41.10	34.00	87.50	0.50	2.60	11.60	4.70	1.20	0.20
Franz Wagner	24.20	46.30	29.50	87.10	0.90	4.80	4.70	2.30	1.30	0.40
Donovan Mitchell	24.00	44.30	36.80	82.30	0.80	3.70	5.00	2.10	1.30	0.20
Tyler Herro	23.90	47.20	37.50	87.80	0.50	4.70	5.50	2.60	0.90	0.20
Joel Embiid	23.80	44.40	29.90	88.20	1.90	6.30	4.50	3.30	0.70	0.90
De'Aaron Fox	23.50	46.30	31.00	82.70	0.90	3.90	6.30	2.80	1.50	0.40
Zach LaVine	23.30	51.10	44.60	82.50	0.30	4.00	4.20	2.80	0.80	0.20
Ja Morant	23.20	45.40	30.90	82.40	0.70	3.40	7.30	3.70	1.20	0.20
James Harden	22.80	41.00	35.20	87.40	0.70	5.10	8.70	4.30	1.50	0.70
Brandon Ingram	22.20	46.50	37.40	85.50	0.90	4.60	5.20	3.80	0.90	0.60
Jaylen Brown	22.20	46.30	32.40	76.40	1.30	4.50	4.50	2.60	1.20	0.30
Jaren Jackson Jr.	22.20	48.80	37.50	78.10	1.20	4.40	2.00	2.10	1.20	1.50
Jalen Williams	21.60	48.40	36.50	78.90	0.90	4.50	5.10	2.20	1.60	0.70
Kawhi Leonard	21.50	49.80	41.10	81.00	0.90	5.00	3.10	1.90	1.60	0.50
Jamal Murray	21.40	47.40	39.30	88.60	0.70	3.20	6.00	2.10	1.40	0.50
RJ Barrett	21.10	46.80	35.00	63.00	1.10	5.30	5.40	2.90	0.80	0.30
Darius Garland	20.60	47.20	40.10	87.80	0.60	2.20	6.70	2.50	1.20	0.10
Pascal Siakam	20.20	51.90	38.90	73.40	1.70	5.20	3.40	1.40	0.90	0.50
Austin Reaves	20.20	46.00	37.70	87.70	0.80	3.70	5.80	2.40	1.10	0.30
Domantas Sabonis	19.10	59.00	41.70	75.40	3.80	10.10	6.00	2.90	0.70	0.40
Jalen Johnson	18.90	50.00	31.20	74.60	1.70	8.30	5.00	2.90	1.60	1.00
Tyrese Haliburton	18.60	47.30	38.80	85.10	0.60	3.00	9.20	1.60	1.40	0.70
Alperen Sengun	19.10	49.60	23.30	69.20	3.40	6.90	4.90	2.60	1.10	0.80
Scottie Barnes	19.30	44.60	27.10	75.50	1.70	6.00	5.80	2.80	1.40	1.00
Desmond Bane	19.20	48.40	39.20	89.40	0.90	5.20	5.30	2.40	1.20	0.40
Nicola Vucevic	18.50	53.00	40.20	80.50	2.40	7.60	3.50	1.60	0.80	0.70
Evan Mobley	18.50	55.70	37.00	72.50	2.30	7.00	3.20	2.00	0.90	1.60
Bam Adebayo	18.10	48.50	35.70	76.50	2.40	7.20	4.30	2.10	1.30	0.70
Ivica Zubac	16.80	62.80	0.01	66.10	3.80	8.90	2.70	1.60	0.70	1.10

To establish the best team, the top fifty NBA players, ranked based on the ESPN Ratings for the 2024-2025 Regular Season, were analyzed. The statistics of these players, computed in accordance with the criteria detailed in Table 2, comprise the dataset (31) for this study and are documented in Table 3.

Results

The SPSS software was used for data clustering. A Two-Step clustering algorithm was applied, dividing the data into five clusters. These five clusters represent different positions, as shown in Table 4.

Table 4: NBA Player Clusters Based on Positional Roles

Cluster 1 (Guard)	Cluster 2 (Guard)	Cluster 3 (Forward)	Cluster 4 (Forward)	Cluster 5 (Center)
Gilgeous Alexander	Jalen Brunson	Kevin Durant	Anthony Davis	Nikola Jokic
Luka Dončić	Devin Booker	Paolo Banchero	Karl-Anthony Towns	Giannis Antetokounmpo
Anthony Edwards	Damian Lillard	Franz Wagner	Victor Wembanyama	Zion Williamson
Tyrese Maxey	Kyrie Irving	Jayson Tatum	Pascal Siakam	Alperen Sengun
Cade Cunningham	Stephen Curry	Brandon Ingram	Domantas Sabonis	Ivica Zubac
LaMelo Ball	Donovan Mitchell	Jaylen Brown	Nicola Vucevic	Joel Embiid
LeBron James	Tyler Herro	Jaren Jackson Jr.	Evan Mobley	
Trae Young	Zach LaVine	Jalen Williams	Bam Adebayo	
De'Aaron Fox	Jamal Murray	Kawhi Leonard		

Ja Morant	Darius Garland	RJ Barrett
James Harden	Austin Reaves	Jalen Johnson
	Tyrese Haliburton	Scottie Barnes
	Desmond Bane	

As indicated in Table 4, the playmakers within Cluster 1 show notable differences in terms of playmaking skills and scoring ability (with averages of 7.6 assists and 25.9 points per game). In addition, their outside shooting is very effective (35% from three-point range, 84% from the free-throw line), but defensively they have limited contributions (0.5 blocks) and/or offensive rebounds [0.75]. Therefore, it appears that this would indicate that the majority of the individuals in Cluster 1 are point guards. When examining players within Cluster 2, we can conclude that their three-point shooting percentage (38.7%) and free-throw shooting percentage (87.7%) are the highest of any cluster, indicating that they can be classified as shooting guards. In addition, the average number of turnovers [2.4] for players in this cluster is lower than that of players in other clusters and therefore these players are proven to be successful shooters while also being extremely ball-dominating and ball-controlling players. Players from Cluster 3 add to the contributions of outside shooters (34.8% three-point shooting) and utilize a defensive style of play, as evidenced by their averages of 5.6 defensive rebounds and 0.72 blocked shots per game. Their average points [22.7] and turnovers [2.7] are consistent with the average statistics of all other clusters in their league.

While players from Cluster 4 are focused on defensive type positions, they are also among the most successful in terms of three-point shooting

(37.4%). Furthermore, their free-throw shooting percentage (77.8%) is decent. Additionally, when accounting for the number of blocks [1.3] and rebounds (8.1 defensive rebounds per game), they contribute well defensively as well as offensively, averaging 20.9 points per game and 3.8 assists per game. There are five players in Cluster 5 and they have the most noted characteristic of being extremely efficient with field goals (48% field-goal percentage).

The RAWEC method uses not only the distances between the optimal and worst values but also incorporates both the optimal and worst distances when evaluating athletes for inclusion in a team. The method considers how well-rounded each player is compared to their teammates. Therefore, it provides a much more equalized average weight for players being considered for the roster than other methods that evaluate only players rated as top players. The RAWEC method is advantageous when evaluating players using multiple criteria such as performance metrics and situational factors, including their ability to contribute to the overall performance of their team.

The calculations were carried out individually for the five clusters, following the mathematical procedures of the RAWEC method. The detailed application steps for Cluster-5 are described below.

Initially, the normalized decision matrix was created using Equation [2] for benefit criteria and Equation [3] for cost criteria, as detailed in Table 5.

Table 5: Normalized Decision Matrix

Player/Criteria	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
<i>n_{ij}</i> Values										
Giannis Antetokounmpo	1.000	0.957	0.532	0.700	0.579	0.980	0.637	0.516	0.500	1.000
Nikola Jokić	0.974	0.917	1.000	0.907	0.763	1.000	1.000	0.485	1.000	0.500
Zion Williamson	0.809	0.903	0.554	0.744	0.658	0.475	0.520	0.533	0.667	0.750
Joel Embiid	0.783	0.707	0.717	1.000	0.500	0.636	0.441	0.485	0.389	0.750
Alperen Sengun	0.628	0.790	0.559	0.785	0.895	0.697	0.480	0.615	0.611	0.667
Ivica Zubac	0.553	1.000	0.000	0.749	1.000	0.899	0.265	1.000	0.389	0.917
<i>n'_{ij}</i> Values										
Giannis Antetokounmpo	1.000	0.957	0.532	0.700	0.579	0.980	0.637	0.516	0.500	1.000
Nikola Jokić	0.974	0.917	1.000	0.907	0.763	1.000	1.000	0.485	1.000	0.500
Zion Williamson	0.809	0.903	0.554	0.744	0.658	0.475	0.520	0.533	0.667	0.750
Joel Embiid	0.783	0.707	0.717	1.000	0.500	0.636	0.441	0.485	0.389	0.750
Alperen Sengun	0.628	0.790	0.559	0.785	0.895	0.697	0.480	0.615	0.611	0.667
Ivica Zubac	0.553	1.000	0.000	0.749	1.000	0.899	0.265	1.000	0.389	0.917

Table 6: Deviations from the criterion weight

Player/Criteria	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	v_{ij} Values
Giannis Antetokounmpo	0.000	0.004	0.047	0.030	0.042	0.002	0.036	0.048	0.050	0.000	0.260
Nikola Jokić	0.003	0.008	0.000	0.009	0.024	0.000	0.000	0.052	0.000	0.050	0.145
Zion Williamson	0.019	0.010	0.045	0.026	0.034	0.053	0.048	0.047	0.033	0.025	0.339
Joel Embiid	0.022	0.029	0.028	0.000	0.050	0.036	0.056	0.052	0.061	0.025	0.359
Alperen Sengun	0.037	0.021	0.044	0.022	0.011	0.030	0.052	0.038	0.039	0.033	0.327
Ivica Zubac	0.045	0.000	0.100	0.025	0.000	0.010	0.074	0.000	0.061	0.008	0.323
v'_{ij} Values											
Giannis Antetokounmpo	0.045	0.026	0.100	0.000	0.014	0.052	0.058	0.006	0.022	0.050	0.373
Nikola Jokić	0.043	0.023	0.100	0.023	0.034	0.053	0.074	0.000	0.061	0.000	0.411
Zion Williamson	0.032	0.022	0.100	0.006	0.024	0.000	0.049	0.009	0.042	0.033	0.316
Joel Embiid	0.029	0.000	0.100	0.030	0.000	0.025	0.040	0.000	0.000	0.033	0.258
Alperen Sengun	0.012	0.010	0.100	0.011	0.044	0.032	0.045	0.021	0.036	0.025	0.337
Ivica Zubac	0.000	0.029	0.000	0.007	0.050	0.047	0.000	0.052	0.000	0.045	0.230

Next, deviations from the criterion weights were determined using Equations [4] and [5]. At this point, all criteria were assigned equal importance weights ($w_j=0.10$). The results are presented in Table 6.

Finally, the performance scores for Cluster 5 were computed using Equation [6] and are presented in Table 7. The solution was applied to all clusters and summarized in Table 8.

Table 7: Results of the RAWEC Method for Cluster 5

Player/Criteria	Qi Value	Rank
Giannis Antetokounmpo	0.178	2
Nikola Jokić	0.477	1
Zion Williamson	-0.034	4
Joel Embiid	-0.164	5
Alperen Sengun	0.014	3
Ivica Zubac	-0.168	6

Table 8: Results of the RAWEC Method (All Clusters)

Cluster 1 (Guard)	Qi	Rank	Cluster 2 (Guard)	Qi	Rank	Cluster 3 (Forward)	Qi	Rank
Shai Gilgeous-Alexander	0.5562	1	Jalen Brunson	-0.3562	13	Jayson Tatum	0.151	4
Luka Dončić	0.2629	2	Devin Booker	-0.0633	9	Kevin Durant	0.045	7
Anthony Edwards	0.0571	6	Damian Lillard	0.0033	8	Paolo Banchero	-0.006	9
Tyrese Maxey	-0.1207	9	Kyrie Irving	0.3165	2	Franz Wagner	0.022	8
Cade Cunningham	0.1296	4	Stephen Curry	0.0596	5	Brandon Ingram	-0.085	10
LaMelo Ball	-0.1085	8	Donovan Mitchell	0.0135	6	Jaylen Brown	-0.090	11
LeBron James	0.2131	3	Tyler Herro	-0.1079	10	Jaren Jackson Jr.	0.134	6
Trae Young	-0.4030	11	Zach LaVine	-0.3546	12	Jalen Williams	0.191	3
De'Aaron Fox	0.0211	7	Jamal Murray	0.1963	3	Kawhi Leonard	0.149	5
Ja Morant	-0.3230	10	Darius Garland	-0.2887	11	RJ Barrett	-0.226	12
James Harden	0.0935	5	Austin Reaves	0.0129	7	Jalen Johnson	0.390	1
			Tyrese Haliburton	0.3733	1	Scottie Barnes	0.234	2
			Desmond Bane	0.1959	4			
Cluster 4 (Forward)	Qi	Rank	Cluster 5 (Center)	Qi	Rank			
Anthony Davis	0.117	1	Giannis Antetokounmpo	0.178	2			
Karl-Anthony Towns	0.034	3	Nikola Jokić	0.477	1			
Victor Wembanyama	0.090	2	Zion Williamson	-0.034	4			
Pascal Siakam	-0.341	8	Joel Embiid	-0.164	5			

Domantas Sabonis	0.004	4	Alperen Sengun	0.014	3
Nicola Vucevic	-0.097	6	Ivica Zubac	-0.168	6
Evan Mobley	-0.117	7			
Bam Adebayo	-0.079	5			

The RAWEC method was applied, under the assumption that all criteria had equal weighting. Since the athletes were grouped into clusters based on similar performance characteristics in accordance with the basic design of the study using the two-step cluster analysis methodology, the RAWEC method was utilized within a homogeneous group for each athlete by position or group within the respective clusters. Because of this homogeneity, the criteria in each of the respective clusters provided an unbiased basis for comparing the athletes in each cluster by position; as such, the impact of the weighting of the different classification criteria was limited in terms of effecting the final selection of the best team. The objective of the study was to minimize the incidental bias that is found in many of the evaluation methods used in traditional athlete evaluations; thus, if variable weighing was applied to individual criteria, it would not have been possible to develop accurate conclusions.

The 2024-2025 season's best team will consist of five guards, five forwards and two centers based on the application of the Two-Step and RAWEC methods. The player with the highest score at each position will be included in the starting five, while all other players will be classified as reserves. After conducting the clustering process independently for each cluster, the RAWEC Method was applied and all of the players were ranked by multi-criteria performance scores. Nikola Jokić received the highest performance score ($Q_i = 0.477$) within Cluster 5, followed by Giannis Antetokounmpo (Q_i

$= 0.178$); Anthony Davis received the highest performance score ($Q_i = 0.117$) within Cluster 4, followed by Victor Wembanyama ($Q_i = 0.090$); Jalen Johnson received the highest performance score ($Q_i = 0.390$) within Cluster 3, with Scottie Barnes in a close second ($Q_i = 0.234$); Shai Gilgeous-Alexander received the highest performance score ($Q_i = 0.556$) within Cluster 1, with Luka Dončić in the second position ($Q_i = 0.263$); and finally Tyrese Haliburton received the highest performance score ($Q_i = 0.373$) within Cluster 2, followed closely by Kyrie Irving ($Q_i = 0.317$). Based on this information, each player's highest performance score will be considered and the top five players from each cluster will be designated as the starting roster, with the remainder of the highest-ranked players being assigned to the reserve roster, as shown in Table 9. The analysis will provide the best team will consist of these twelve players.

The proposed decision support system is a tool for team managers, analysts and sports administrators in activities such as roster organization, player assessment and talent identification. The existing framework reduces the ambiguities associated with subjective assessment, thereby providing a more structured, transparent and accountable evaluation process. Further, the integration of cluster analysis and multi-criteria decision-making methods enables teams to develop a thorough squad composition that corresponds to positional demands and performance indicators.

Table 9: Team Roster

Players	Position	Role
Shai Gilgeous-Alexander	Guard	Starting
Tyrese Haliburton	Guard	Starting
Jalen Johnson	Forward	Starting
Anthony Davis	Forward	Starting
Nikola Jokić	Center	Starting
Luka Doncic	Guard	Substitute
Kyrie Irving	Guard	Substitute
LeBron James	Guard	Substitute
Scottie Barnes	Forward	Substitute
Jalen Williams	Forward	Substitute
Victor Wembanyama	Forward	Substitute
Giannis Antetokounmpo	Center	Substitute

Discussion

To make the process of constructing the NBA roster objective and data-driven, the present research proposes and validates a hybrid decision support framework combining the Two-Stage Clustering algorithm with the RAWEC multi-criteria decision-making approach. The results demonstrate that the framework can effectively cluster players into position-oriented groups and rank them based on a comprehensive assessment of their statistical inputs. Results are consistent with professional assessments in a number of key areas and demonstrate the value of a multi-faceted analytical approach.

The model's validity is strongly supported by its highest-ranked players. Nikola Jokić, who received the top performance score ($Q_i = 0.477$) in the Center cluster (Cluster 5), is widely regarded as one of the most versatile and impactful players in the league. His high ranking is a direct result of his outstanding play, scoring, rebounding and playmaking, indicative of the kind of multi-faceted success the RAWEC methodology is designed to recognize. Similarly, Shai Gilgeous-Alexander's top ranking ($Q_i = 0.556$) in the primary guard cluster (Cluster 1) reflects his well-rounded offensive skill set, combining high-volume scoring with effective distribution. The model's ability to identify consensus top-tier talent confirms its alignment with established standards of player evaluation. Perhaps more revealing are the instances where the model's rankings diverge from traditional superstar hierarchies. For example, Jalen Johnson's emergence as the top-ranked player in Cluster 3, ahead of established scorers like Jayson Tatum and Kevin Durant, highlights the framework's emphasis on all-around efficiency over raw scoring output. Johnson's profile—characterized by strong defensive contributions (1.6 steals, 1.0 blocks), excellent rebounding for his position and efficient scoring with low turnover rates—makes him an ideal candidate under the RAWEC method's balanced evaluation. The model effectively quantifies the significant on-court impact of such players.

This result is consistent with several studies (29) that have successfully applied clustering analysis to NBA data and identified metrics such as shooting efficiency and offensive/defensive workloads as key differentiators among player profiles. The framework first classifies players into

uniform groups (e.g., high-assist guards and high-rebound centers) such that players can be ranked among similar players in the later MCDM step. This systematic approach directly overcomes a key limitation of research applying MCDM techniques on the whole population of players that can lead to unfair comparisons between essentially different roles (21). Furthermore, it represents an advancement over clustering-only approaches (19-20) as it goes beyond simple grouping to produce a final, viable ranking for roster selection. The goal of roster building aligns our work with some studies (10) that combine clustering with a rating system. However, there is a significant methodological difference that explains our unique results. The framework of some studies (10) includes player salaries and uses a more complex, data-driven weighting scheme that optimizes financial efficiency. In contrast, the model deliberately used equal weighting and excluded salary data to create a transparent framework focusing solely on multi-criteria on-field performance. This distinction may explain why a balanced and statistically efficient player like Jokić ranked highest in our model; whereas a salary-driven model might prioritize players who offer more money-value. The use of the RAWEC method, a new contribution to sports analytics, is central to this result. Its ideal/anti-ideal calculation provides a single, transparent score that holistically assesses a player's strengths and weaknesses, offering a more complete picture than methods that focus on a limited set of performance indicators.

Conclusion

While creating players and teams across all sports disciplines, selecting athletes is a vital process. During these selections, data such as intuition, perception, attention, technique, quickness, stamina, motor skills, athletes' attitudes and behaviors, their performance in previous training and competitions, past injuries and behavioral risks are used. When viewed broadly, this study combines two different application methods by clustering the medium-sized data group and performing sorting and selection operations separately for each cluster.

Initially, a thorough review of the relevant literature was conducted and the methods used in

similar applications were analyzed. As a result of brainstorming at the end of this review, the topic and objectives were defined. The goal is to select the top players among NBA players for the 2024-2025 regular season. From this perspective, the top 50 NBA players were selected, their performances evaluated based on ten criteria and players were grouped into five categories corresponding to the five basketball positions. Clustering was performed using the SPSS program, producing highly realistic results. Subsequently, the sorting and selection process was carried out using the RAWEC method for these five groups. The results showed that this study yielded satisfactory outcomes. It is crucial for machine learning to become more widespread in the sports industry.

This study contributes to the literature by presenting a fully integrated analytical process that first structures the decision environment using unsupervised learning and then optimizes alternatives through multi-criteria evaluation. The findings demonstrate that AI-powered hybrid decision support systems can significantly improve roster optimization strategies and reduce subjectivity in sports management.

This study also has several limitations that suggest directions for future research. The analysis relies on a single season of aggregated, per-game statistics for the top fifty ESPN-ranked players, so the framework's stability across multiple seasons or different player pools has not yet been examined. The five clusters were deliberately fixed to the traditional basketball positions, which may not fully capture increasingly positionless playing styles in the modern NBA. Similarly, the ten criteria used, while covering the core box-score dimensions, do not include advanced tracking-based metrics (e.g., defensive matchup difficulty, on/off-court impact) or contextual factors such as injury history and team chemistry. Future studies could extend the framework to multiple seasons and leagues, incorporate advanced tracking data and alternative weighting schemes and as noted above, benchmark its rankings against the judgment of coaches, scouts, or analysts.

The present framework was intentionally designed around objective, criteria-based rankings rather than subjective expert judgment, which ensures full transparency and reproducibility. As a complementary direction, future studies could

benchmark these rankings against the assessments of basketball coaches, scouts, or analysts, for example through structured interviews or expert surveys. Such a comparison would not only offer external validation but could also help surface additional qualitative factors -- such as team chemistry or matchup-specific tactics -- that could be incorporated as further criteria in future extensions of the model.

Abbreviations

DSS: Decision Support Systems, MCDM: Multi-Criteria Decision-Making, ML: Machine Learning, RAWEC: Rank Analysis of Weighted Evaluations of Criteria.

Acknowledgment

This work was done without any support from any supporting agency, facility, or public agency.

Author Contributions

Burcu Ozcan Turkkan: Conceptualization, Project administration, Writing – Review and Editing, Supervision, Ejder Ayçin: Data curation, Validation, Methodology, Software, Formal analysis, Writing – Original Draft, Sibel Fettahoglu: Investigation, Resources, Visualization, Writing – Review and Editing.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability

The primary data supporting the findings of this study are openly available on the ESPN NBA Statistics page at

<https://www.espn.com/nba/seasonleaders>. The aggregated dataset used for clustering and ranking is fully presented in Table 3 of this manuscript.

Declaration of Artificial Intelligence (AI) Assistance

During the preparation of this work, the authors used generative AI tools for language polishing, grammar correction and phrasing improvements. The authors take full responsibility for the content's originality, interpretation and accuracy.

Ethics Approval

No ethical clearance certificate is applicable for present study. The authors of the submitted paper did not receive support from any organization. This study utilized publicly available statistical

performance data from the 2024-2025 NBA regular season (ESPN). No human subjects, animals, or sensitive personal data were involved. Therefore, ethics committee approval was not required.

Funding

No funding was received for this study.

References

- Wang H, Sarker A, Hosoi A. The effect of basketball analytics investment on national basketball association (NBA) team performance. *Journal of Sports Economies*. 2025;26(6):668-688. <https://doi.org/10.1177/15270025251328264>
- Sarlis V, Papageorgiou G, Tjortjis C. Leveraging sports analytics and association rule mining to uncover recovery and economic impacts in NBA basketball. *Data*. 2024;9(7):83. <https://doi.org/10.3390/data9070083>
- Li J. Machine learning-based analysis of defensive strategies in basketball using player movement data. *Scientific Reports*. 2025;15:13887. <https://doi.org/10.1038/s41598-025-98877-1>
- Mănescu DC. Big data analytics framework for decision-making in sports performance optimization. *Data*. 2025;10(7):116. <https://doi.org/10.3390/data10070116>
- Madrigal-Cerezo R, Domínguez-Sanz N, Martín-Rodríguez A. Wearable biosensing and machine learning for data-driven training and coaching support. *Biosensors*. 2026;16(2):97. <https://doi.org/10.3390/bios16020097>
- Min B, Kim J, Choe C, *et al.* A compound framework for sports results prediction: A football case study. *Knowledge-Based Systems*. 2008;21(7):551-562. <https://doi.org/10.1016/j.knosys.2008.03.016>
- Chen T, Sun KS. Exploring the strategy to improve senior citizens' participations on recreational sports. *Knowledge-Based Systems*. 2012;26:86-92. <https://doi.org/10.1016/j.knosys.2011.07.008>
- Young WA, Weckman G.R. Evaluating the effects of aging for professional football players in combine events using performance-aging curves. *International Journal of Sports Science and Engineering*. 2008;2(3):131-143. https://www.academia.edu/58816945/Evaluating_the_effects_of_aging_for_professional_football_players_in_combine_events_using_performance_aging_curves
- Ouyang Y, Li X, Zhou W, *et al.* Integration of machine learning XGBoost and SHAP models for NBA game outcome prediction and quantitative analysis methodology. *Plos One*. 2024;19(7):e0307478. <https://doi.org/10.1371/journal.pone.0307478>
- Ke Y, Bian R, Chandra R. A Unified machine learning framework for basketball team roster construction: NBA and WNBA. *Applied Soft Computing*. 2024;153:111298. <https://doi.org/10.1016/j.asoc.2024.111298>
- Alves JM, Barbosa RS. Machine learning for basketball game outcomes: NBA and WNBA leagues. *Computation*. 2025;13(10):230. <https://doi.org/10.3390/computation13100230>
- Cavan E, Cao J, Swartz TB. NHL aging curves using functional principal component analysis. *Journal of Quantitative Analysis in Sports*. 2025;21(3):177-189. <https://doi.org/10.1515/jqas-2024-0083>
- Herold M, Goes F, Nopp S, *et al.* Machine learning in men's professional football: Current applications and future directions for improving attacking play. *International Journal of Sports Science & Coaching*. 2019;14(6):798-817. <https://doi.org/10.1177/1747954119879350>
- Puška A, Štilić A, Pamučar D, *et al.* Introducing a novel multi-criteria ranking of alternatives with weights of criterion (RAWEC) model. *MethodsX*. 2024;12:102628. <https://doi.org/10.1016/j.mex.2024.102628>
- Dundar S, Karadağ İ. Selection of a facility location for a cosmetics company by integrated f-lbwa and i-rawec methods. *International Journal of Fuzzy Systems*. 2025;28:424-443. <https://doi.org/10.1007/s40815-024-01960-4>
- Akbulut OY, Aydın Y. A hybrid multidimensional performance measurement model using the msd-mpsi-rawec model for Turkish banks. *Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*. 2024;11(3):1157-1183. <https://doi.org/10.30798/makuiibf.1464469>
- Demir G. Selection of biomass resources with fuzzy multi-criteria decision-making methods: alternatives for sustainable energy production. *Journal of Operations Intelligence*. 2025;3(1):161-179. <https://doi.org/10.31181/jopi31202537>
- Wang M, Zhang YY, Min F, *et al.* A two-stage density clustering algorithm. *Soft Computing*. 2020;24:17797-17819. <https://doi.org/10.1007/s00500-020-05028-x>
- An N. Sports competition data analysis and strategy optimization using K-means clustering algorithm. *Journal of Computational Methods in Sciences and Engineering*. 2025;25(1):888-902. <https://doi.org/10.1177/14727978251321959>
- Kılıç Depren S. The effectiveness of different machine learning algorithms on basketball players' shooting performance. *Spor ve Performans Araştırmaları Dergisi*. 2019;10(3):256-269. <https://doi.org/10.17155/omuspd.507797>
- Qader M., Zaidan BB, Zaidan AA, *et al.* A methodology for football players selection problem based on multi-measurements criteria analysis. *Measurement*. 2017;111:38-50. <https://doi.org/10.1016/j.measurement.2017.07.024>
- Zuccolotto P, Manisera M, Sandri M. Big data analytics for modeling scoring probability in basketball: The effect of shooting under high-pressure conditions. *International Journal of Sports Science & Coaching*. 2017;13(4):569-589. <https://doi.org/10.1177/1747954117737492>
- Kizielewicz B, Dobryakova L. MCDA based approach to sports players' evaluation under incomplete knowledge. *Procedia Computer Science*. 2020;176:3524-3535. <https://doi.org/10.1016/j.procs.2020.09.034>

24. Gazi KH, Raisa N, Biswas A, *et al.* Finding the most important criteria in women's empowerment for sports sector by pentagonal fuzzy dematel methodology. *Spectrum of Decision Making and Applications*. 2025;2(1):28-52.
<https://doi.org/10.31181/sdmap21202510>
25. Ozkan IA. A novel basketball result prediction model using a concurrent neuro-fuzzy system. *Applied Artificial Intelligence*. 2020;34(13):1038–1054.
<https://doi.org/10.1080/08839514.2020.1804229>
26. Li L, Hao A. Research on the sports training effect based on GABP neural network and artificial intelligence. *Scientific Reports*. 2025;15:39214.
<https://doi.org/10.1038/s41598-025-20426-7>
27. Lampis T, Ioannis N, Vasilios V, *et al.* Predictions of european basketball match results with machine learning algorithms. *Journal of Sports Analytics*. 2023;9(2):171-190.
<https://doi.org/10.3233/JSA-220639>
28. Huang ML, Lin YJ. Regression tree model for predicting game scores for the golden state warriors in the national basketball association. *Symmetry*. 2020;12(5):835.
<https://doi.org/10.3390/sym12050835>
29. Dehesa R, Vaquera A, Gomez-Ruano MA, *et al.* Key performance indicators in NBA players' performance profiles. *Kinesiology*. 2019;51(1):92-101.
<https://ojs.srce.hr/kinesiology/article/view/5456>
30. Taber CB, Sharma S, Raval MS, *et al.* Holistic approach to performance prediction in collegiate athletics: player, team and conference perspectives. *Scientific Reports*. 2024;14:1162.
<https://doi.org/10.1038/s41598-024-51658-8>
31. 2024-2025 NBA Season Leaders 2025. ESPN.
<https://www.espn.com/nba/seasonleaders>

How to Cite: Turkkan BO, Aycin E, Fettahoglu S. A Hybrid Machine Learning and MCDM Approach for Optimal Basketball Team Selection: A Decision Support System Framework. *Int Res J Multidiscip Scope*. 2026;7(4):84-95. DOI: 10.47857/irjms.2026.v07i04.011194