Redefining NBA positions

and classifying those for incoming prospects

Adam Bushman

December 2024 | Machine Learning

Executive Summary

The National Basketball League (NBA) has grown more dynamic and versatile with its maturity. Traditional positions fail to roles and play styles, erring more toward legacy physical profiles. Despite playing Center for the *Denver Nuggets*, Nikola Jokic operates far more as the team's offensive orchestrator, not unlike Hall of Fame Point Guard John Stockton of the *Utah Jazz*.

Nikola Jokic, C | Denver

Are there natural NBA player groupings that better reflect playing style and roles than traditional positions? **Yes!**

Through clustering and predictive modeling techniques, we've derived <u>10 new player position classifications</u> for current league players and incoming prospects. Incorporation of these more nuanced position labels will improve scouting, roster management, and prospect evaluation development.

John Stockton, PG | Utah

Scope of Analysis

Addressing this opportunity is a two-step process: 1) deriving natural groupings from recent player activity data and 2) applying new classifications to future players. Each step leverages its own data sets and analysis techniques, as described below:

Step #1

Data sets:

- Player-level performance metrics
- Top 500 NBA players (2020-21 to present) by minutes played
- Sourced from PBPstats.com

Analysis techniques:

- Dimensionality reduction (via PCA)
- Clustering (via Kmeans & Hierarchical)
- Top feature study by cluster (via Elastic Net Logistic Regression)

Step #2

Data sets:

- Pre-NBA collegiate data
- 395 matches to data from step 1
- Sourced from Basketball-Reference.com

Analysis techniques:

- Predictive models (via Neural Network & XGBoost)

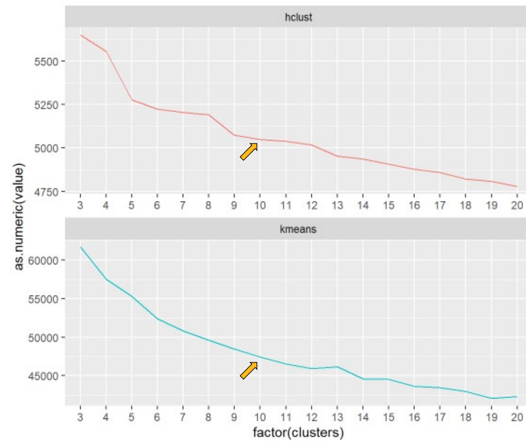
Step #1

Deriving new positions required clustering NBA players across nearly 200 attributes.

Ten (10) was determined to be an ideal breakpoint balancing too few (very diverse groupings) clusters and too many (far too similar groupings). "Within sum of squares" was the chosen metric to find the right balance (plotted on y-axis).

KMeans was leveraged for assigning each player to one of the 10 clusters. Giving clusters meaning was next.

"Within Sum of Squares" by number of clusters



New Positions

Versatile Anchor

Cluster 1 labeled as *Versatile Anchor* due to high activity in transition defense, on-off rating, and propensity for turnovers. Ex:

 ***Draymond Green***
 ***Russell Westbrook***

Perimeter Anchor

Cluster 6 labeled as *Perimeter Anchor* due to high defensive activity and scoring stability via assisted shots and free throw attempts. Ex:

 ***Fred VanVleet***
 ***Mike Conley***

Versatile Finisher

Cluster 2 labeled as *Versatile Finisher* due to shot orientation from high value areas: rim and 3P, particularly in second chance settings. Ex:

 ***Nikola Vucevic***
 ***Lauri Markkanen***

Versatile Connector

Cluster 7 labeled as *Versatile Connector* due to free throw activity, efficient scoring, and second change opportunity generation. Ex:

 ***P.J. Washington***
 ***Terance Mann***

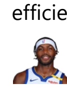
Versatile Engine

Cluster 3 labeled as *Versatile Engine* due to generate unassisted shots and capture rebounds for second change opportunities. Ex:

 ***Mikal Bridges***
 ***RJ Barrett***

Perimeter Finisher

Cluster 8 labeled as *Perimeter Finisher* due to high volume of perimeter, assisted shots volume and efficiency. Ex:

 ***Buddy Hield***
 ***Grayson Allen***

Perimeter Engine

Cluster 4 labeled as *Perimeter Engine* due to self-creation from the perimeter, including 3P shots and trips to the free throw line. Ex:

 ***Jayson Tatum***
 ***Donovan Mitchell***

Interior Anchor

Cluster 9 labeled as *Interior Anchor* due to the high activity in the paint on both ends: rebounding, shot blocking, and 2nd chance points. Ex:

 ***Jalen Duren***
 ***Clint Capela***

Interior Connector

Cluster 5 labeled as *Interior Connector* due to inside-outside nature of scoring, assisting, and rebounding. Ex:

 ***Wendell Carter Jr.***
 ***Isaiah Hartenstein***

Interior Engine

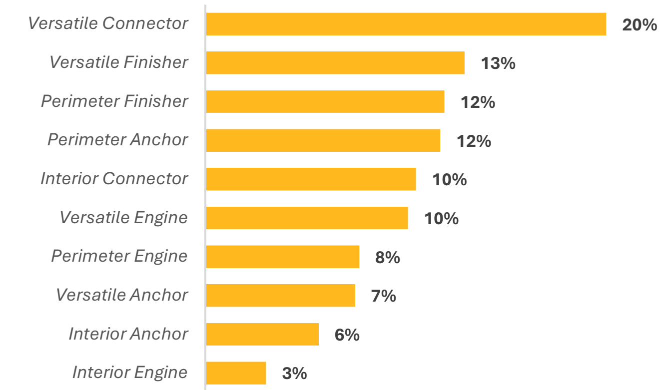
Cluster 10 labeled as *Interior Engine* due to high assist rates, fouls drawn, and propensity to score inside to outside. Ex:

 ***Nikola Jokic***
 ***Giannis Antetokounmpo***

These ten (10) positions aren't equally distributed across the league or teams (which is expected). To the right is a visualization of which positions see the most and least concentration of players (i.e. "Interior Engines" are hard to come by, just 3% of all NBA players).

Many interesting analyses and studies could derive from these classifications. One of interest now is what positions incoming NBA prospects are likely to assume.

Distribution of players by new position label



Step #2

Classifying incoming NBA prospects requires modeling the relationship between pre-NBA performance of existing players for which we know their classified positions. A tuned Neural Network model was employed capture these patterns. It boast a fairly high performance metric, *AUC* of 0.85. XGBoost validated results.

Neural Network

- 3 tuned hyperparameters
- Cross-validated
- **AUC: 0.85**

XGBoost

- 4 tuned hyperparameters
- Cross-validated
- **AUC: 0.82**

Classified positions

Perimeter Finisher

Tre Johnson ***Texas***
Liam McNeeley ***UConn***
Kon Knueppel ***Duke***
Alex Karaban ***UConn***
Will Riley ***Illinois***
Hunter Sallis ***Wake Forest***
Ian Jackson ***UNC***
Jalil Bethea ***Miami***

Versatile Connector

Cooper Flagg ***Duke***
Ace Bailey ***Rutgers***
Nique Clifford ***Colorado St.***
KJ Lewis ***Arizona***
Drake Powell ***UNC***
Kanon Catchings ***BYU***
Jalil Bethea ***Arizona***
Mackenzie Mgbako ***Indiana***

Interior Connector

Khaman Maluach ***Duke***
Derik Queen ***Maryland***
Collin Murray-Boyles ***South Carolina***

Perimeter Anchor

Egor Demin ***BYU***
Boogie Fland ***Arkansas***
Kasparas Jakucionis ***Illinois***

Versatile Anchor

VJ Edgecomb ***Baylor***
Labaron Philon ***Alabama***

Perimeter Engine

Dylan Harper ***Rutgers***

Versatile Finisher

Asa Newell ***Georgia***

Recommendations for NBA teams & staff

I recommend implementing the new position classifications across NBA front offices to improve player evaluation, scouting, and team-building strategies. Using the predicted roles for incoming prospects can enhance draft decision-making and scouting processes. Additionally, incorporating these classifications into discussions about team construction can help align player acquisitions with strategic goals. Follow-up studies on positional value, such as their impact on team success or salary structures, could provide further insights to optimize resource allocation and player development.

Limitations of findings

The conclusions are limited by the reliance on publicly available data, which lacks the granularity of premium sources like Synergy, and by combining offensive and defensive metrics, which may have led to overly generic clusters. Additionally, excluding non-NCAA players narrows applicability, leaving gaps for those entering the NBA via high school or international routes. Addressing these issues would improve the accuracy and utility of the findings.

Future analyses

Future analyses could explore the relative value of the newly defined positions by examining their correlation with team success metrics, such as win shares or playoff performance, and their impact on salary structures to assess market inefficiencies. Separating offensive and defensive clusters could provide role-specific insights, enabling a deeper understanding of how players contribute in different contexts. Expanding pre-NBA data to include high school and international leagues would make predictions more comprehensive and inclusive. Additionally, integrating temporal trends could reveal how player roles evolve over time, offering insights into strategic shifts in team construction and player development. These directions could validate and expand the utility of the findings for broader applications in the NBA.