NBA 2024-25 season MVP Predictions

Group 4: Kang An, Kunbo Chen, Huiyu Li, Shengbo Qian, Zheng Jiang

Our Goals:

- Identify which player should be considered a candidate for the NBA Most Valuable Player (MVP) .

Our methods:

Three machine learning algorithms to predict:

- Support Vector Machine (SVM)
- Decision Tree Classification
- K-Nearest Neighbors Classifier (KNN)

Compare accuracy of the three methods and apply the best model to predict the result.

Our Data:

Our raw data from the official NBA web: https://www.basketball-reference.com/leagues/NBA_2025_per_game.html

2024-25 NBA Season | Standings | Schedule and Results | Leaders | Coaches | Player Stats ▾ | Other ▾ | Back to top ▴

Totals | Per Game | Per 36 Min | Per 100 Poss | Advanced | Play-by-Play | Shooting | Adjusted Shooting

Per Game

✨Upgraded✨ Share & Export ▾ ☑ When table is sorted, hide non-qualifiers for rate stats Glossary

Regular Season

Rk	Player	Age	Team	Pos	G	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	Awards
1	Giannis Antetokounmpo	30	MIL	PF	17	17	35.2	13.2	21.6	.609	0.2	0.8	.214	13.0	20.8	.624	.613	6.4	10.5	.612	2.1	9.8	11.9	6.6	0.5	1.4	3.3	2.7	**32.9**	
2	LaMelo Ball	23	CHO	PG	18	18	34.1	10.7	24.9	.430	4.7	13.1	.356	6.1	11.8	.512	.523	4.9	5.8	.848	0.9	4.4	5.4	6.9	1.1	0.2	4.5	4.1	31.1	
3	Shai Gilgeous-Alexander	26	OKC	SG	20	20	34.6	10.4	20.7	.504	2.1	6.1	.344	8.3	14.6	.570	.554	7.1	8.3	.855	1.0	4.5	5.5	6.5	1.7	1.1	2.8	1.9	30.0	
4	Nikola Jokić	29	DEN	C	15	15	37.5	11.1	19.8	.562	2.2	4.3	.508	8.9	15.5	.578	.618	5.1	6.3	.819	4.1	9.1	**13.2**	10.7	1.5	0.9	3.7	1.9	29.6	
5	Paolo Banchero	22	ORL	PF	5	5	36.4	9.6	19.4	.495	2.2	6.4	.344	7.4	13.0	.569	.552	7.6	11.8	.644	2.4	6.4	8.8	5.6	0.6	0.8	2.2	2.6	29.0	
6	Luka Dončić	25	DAL	PG	15	15	36.6	10.2	22.7	.449	3.3	9.9	.329	6.9	12.8	.542	.521	4.9	6.4	.771	0.5	7.1	7.6	8.0	1.7	0.3	3.3	2.7	28.6	
7	Jayson Tatum	26	BOS	PF	21	21	36.4	9.1	20.1	.454	4.0	10.6	.374	5.2	9.6	.542	.552	6.2	7.7	.802	0.5	8.1	8.6	5.6	1.3	0.6	3.0	2.4	28.4	
8	Anthony Davis	31	LAL	C	20	20	35.4	10.0	18.4	.542	0.8	2.1	.357	9.2	16.3	.566	.563	7.2	9.1	.786	2.6	9.0	11.5	3.3	1.3	2.0	2.0	2.0	27.8	
9	De'Aaron Fox	27	SAC	PG	21	21	37.8	10.0	20.3	.493	2.1	6.3	.341	7.9	14.0	.561	.546	5.3	6.6	.806	1.0	4.0	5.0	5.9	1.6	0.1	3.6	2.6	27.5	
10	Kevin Durant	36	PHO	PF	12	12	37.3	9.6	18.0	.532	2.8	6.3	.434	6.8	11.7	.586	.609	4.9	6.1	.808	0.3	6.7	6.9	3.2	0.6	1.6	3.3	1.8	26.8	
11	Anthony Edwards	23	MIN	SG	20	20	37.0	9.3	20.8	.446	4.5	10.7	.423	4.8	10.1	.470	.554	3.7	4.7	.796	0.7	4.8	5.5	3.7	1.3	0.6	3.0	2.0	26.7	
12	Damian Lillard	34	MIL	PG	16	16	36.2	7.9	18.0	.441	3.3	9.2	.361	4.6	8.8	.525	.533	6.8	7.3	.923	0.6	3.9	4.5	7.8	1.0	0.3	2.8	1.9	25.9	
13	Tyrese Maxey	24	PHI	PG	12	12	36.2	9.3	22.2	.417	3.3	9.8	.333	6.0	12.4	.483	.491	4.2	4.9	.847	0.2	3.0	3.2	4.5	2.0	0.3	1.8	2.7	25.9	
14	Jalen Brunson	28	NYK	PG	20	20	34.5	8.7	17.9	.486	2.6	6.2	.411	6.2	11.7	.526	.557	5.5	6.5	.845	0.5	2.5	3.0	7.9	0.7	0.1	2.3	2.5	25.4	
15	Karl-Anthony Towns	29	NYK	C	19	19	33.4	9.1	17.0	.533	2.4	5.2	.455	6.7	11.8	.567	.602	4.7	5.5	.857	2.9	10.1	13.0	3.2	0.8	0.8	2.2	3.3	25.2	
16	Jaylen Brown	28	BOS	SF	16	16	36.1	8.4	19.7	.429	2.7	7.9	.341	5.8	11.8	.487	.497	5.4	7.0	.777	1.6	4.8	6.4	4.3	1.1	0.4	2.4	2.8	25.0	
17	Kyrie Irving	32	DAL	SG	19	19	35.4	9.3	18.4	.504	3.2	6.9	.466	6.1	11.5	.528	.592	3.1	3.5	.879	1.2	3.7	4.8	5.6	1.2	0.2	2.1	2.3	24.8	
18	Devin Booker	28	PHO	SG	19	19	36.8	8.0	18.1	.443	2.6	7.6	.345	5.4	10.4	.515	.516	6.1	6.8	.892	1.2	2.3	3.5	6.6	1.0	0.4	2.1	2.2	24.7	
19	Cam Thomas	23	BRK	SG	17	17	33.4	8.1	17.5	.461	2.9	7.4	.389	5.2	10.1	.515	.544	5.7	6.6	.866	0.4	2.9	3.2	3.4	0.7	0.1	2.8	2.0	24.7	
20	Donovan Mitchell	28	CLE	SG	20	20	31.9	8.8	19.3	.457	3.7	9.3	.398	5.1	10.0	.513	.553	3.3	4.1	.805	0.6	4.3	4.9	4.0	1.5	0.3	2.1	2.3	24.6	
21	Victor Wembanyama	21	SAS	C	17	17	32.4	8.8	18.2	.484	3.2	9.2	.350	5.6	9.0	.621	.573	3.1	3.6	.855	2.1	8.2	10.2	3.6	1.3	**3.5**	3.6	2.1	24.0	
22	RJ Barrett	24	TOR	SG	18	18	34.5	8.7	18.9	.457	2.1	5.8	.356	6.6	13.2	.502	.512	4.3	6.2	.694	0.9	5.4	6.3	6.2	0.8	0.3	3.4	2.5	23.7	
23	Tyler Herro	25	MIA	SG	19	19	35.2	8.1	17.6	.461	4.1	9.9	.407	4.1	7.6	.531	.576	3.5	4.0	.868	0.3	4.9	5.2	4.9	0.7	0.1	2.4	1.7	23.7	
24	Norman Powell	31	LAC	SG	16	16	32.9	8.1	16.5	.492	3.9	7.9	.496	4.2	8.6	.489	.612	3.4	4.1	.831	0.3	2.6	2.9	2.3	1.1	0.2	2.1	1.8	23.6	
25	Cade Cunningham	23	DET	PG	18	18	36.2	8.8	19.9	.443	2.3	6.2	.375	6.5	13.7	.474	.501	3.5	4.3	.808	1.1	6.1	7.2	9.0	0.9	0.8	4.7	2.6	23.5	
26	Franz Wagner	23	ORL	SF	**22**	22	32.5	8.5	18.4	.464	2.1	6.2	.338	6.5	12.2	.528	.521	4.2	5.0	.853	1.0	4.7	5.7	5.9	1.8	0.5	2.3	2.4	23.4	
27	DeMar DeRozan	35	SAC	SF	16	16	36.4	8.6	17.0	.507	0.6	2.2	.286	8.0	14.8	.540	.526	5.1	6.1	.827	0.6	3.6	4.2	4.4	1.6	0.6	1.8	1.4	22.9	

Data processing:

[404]:

	Player	Age	Team	Pos	G	GS	MP	FG	FGA	FG%	...	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	Awards
0	Giannis Antetokounmpo	30.0	MIL	PF	9.0	9.0	34.8	12.9	21.2	0.607	...	2.3	10.4	12.8	5.2	0.4	0.9	2.7	3.6	31.6	NaN
1	Anthony Davis	31.0	LAL	PF	9.0	9.0	35.1	10.8	18.7	0.577	...	2.1	8.3	10.4	2.8	1.3	2.0	2.2	1.2	31.2	NaN
2	Jayson Tatum	26.0	BOS	SF	11.0	11.0	36.0	9.5	20.5	0.465	...	0.5	7.2	7.6	5.0	1.6	0.5	2.9	2.5	30.5	NaN
3	Nikola Jokić	29.0	DEN	C	10.0	10.0	38.1	10.8	19.2	0.563	...	4.5	9.2	13.7	11.7	1.7	1.0	4.1	2.0	29.7	NaN
4	LaMelo Ball	23.0	CHO	PG	10.0	10.0	33.4	10.2	23.0	0.443	...	1.0	3.9	4.9	6.2	1.5	0.3	4.7	4.1	29.4	NaN

Only Keep Useful Data:

	Player	Team	Pos	G	GS	MP	FG	FGA	FG%	3P	...	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS
0	Giannis Antetokounmpo	MIL	PF	9.0	9.0	34.8	12.9	21.2	0.607	0.1	...	0.554	2.3	10.4	12.8	5.2	0.4	0.9	2.7	3.6	31.6
1	Anthony Davis	LAL	PF	9.0	9.0	35.1	10.8	18.7	0.577	0.7	...	0.794	2.1	8.3	10.4	2.8	1.3	2.0	2.2	1.2	31.2
2	Jayson Tatum	BOS	SF	11.0	11.0	36.0	9.5	20.5	0.465	4.3	...	0.798	0.5	7.2	7.6	5.0	1.6	0.5	2.9	2.5	30.5
3	Nikola Jokić	DEN	C	10.0	10.0	38.1	10.8	19.2	0.563	2.2	...	0.843	4.5	9.2	13.7	11.7	1.7	1.0	4.1	2.0	29.7
4	LaMelo Ball	CHO	PG	10.0	10.0	33.4	10.2	23.0	0.443	4.8	...	0.840	1.0	3.9	4.9	6.2	1.5	0.3	4.7	4.1	29.4

Adding Variables:

Add variable ***efficiency (EFF)*** to measure a player's goal efficiency:

$$\text{Efficiency} = \frac{\text{PTS} + \text{TRB} + \text{AST} + \text{STL} + \text{BLK} - (\text{FGA} - \text{FG}) - (\text{FTA} - \text{FT}) - \text{TOV}}{MP}$$

	Player	Team	Pos	G	GS	MP	FG	FGA	FG%	3P	...	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	EFF
0	Giannis Antetokounmpo	MIL	PF	9.0	9.0	34.8	12.9	21.2	0.607	0.1	...	10.4	12.8	5.2	0.4	0.9	2.7	3.6	31.6	1.017241
1	Anthony Davis	LAL	PF	9.0	9.0	35.1	10.8	18.7	0.577	0.7	...	8.3	10.4	2.8	1.3	2.0	2.2	1.2	31.2	1.005698
2	Jayson Tatum	BOS	SF	11.0	11.0	36.0	9.5	20.5	0.465	4.3	...	7.2	7.6	5.0	1.6	0.5	2.9	2.5	30.5	0.819444
3	Nikola Jokić	DEN	C	10.0	10.0	38.1	10.8	19.2	0.563	2.2	...	9.2	13.7	11.7	1.7	1.0	4.1	2.0	29.7	1.160105
4	LaMelo Ball	CHO	PG	10.0	10.0	33.4	10.2	23.0	0.443	4.8	...	3.9	4.9	6.2	1.5	0.3	4.7	4.1	29.4	0.718563

SVM Features and Variable:

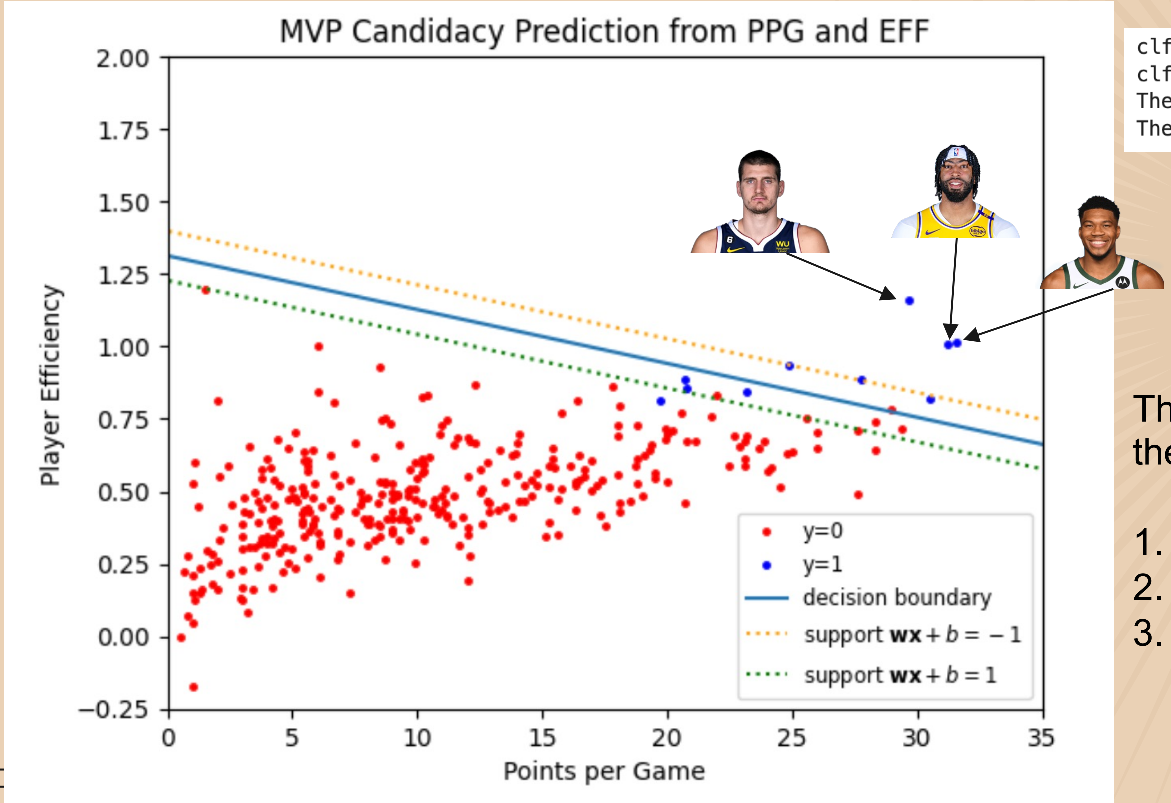
Training Feature X:

- Points Scored (PTS)
- Efficiency (EFF)

Binary Variable y:

- Y = 1 when G>5, MP>30, and EFF>0.8
- Y = 0 otherwise

Support Vector Machine (SVM)

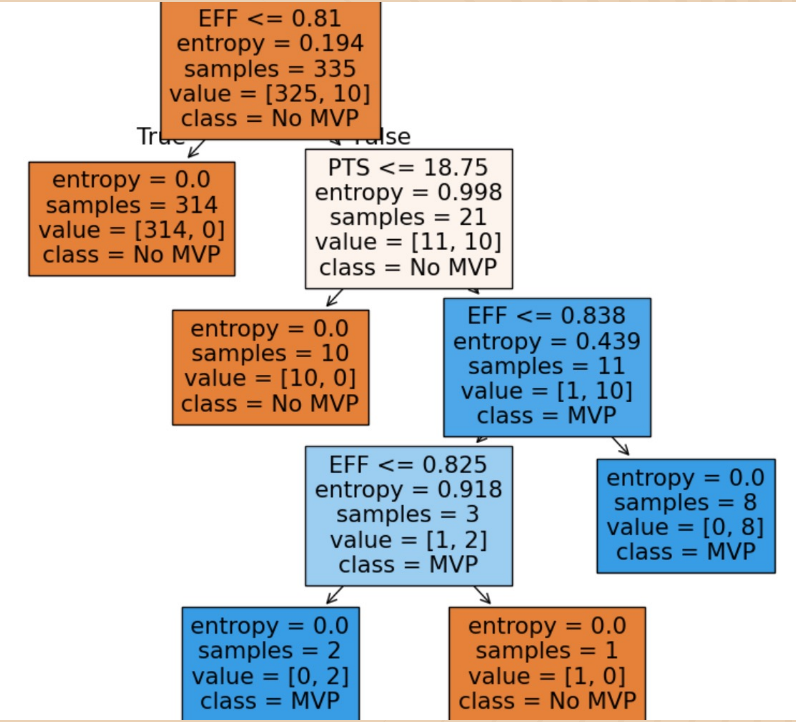



```
clf.coef_=[[ 0.21774983 11.71920381]]
clf.intercept_=[-15.38975733]
The decision boundary is 0.218 * PTS + 11.7 * EFF = -15.4
The training accuracy is 0.985.
```

The player with the best shot at winning the MVP for the 24-25 season is:

1. Giannis Antetokounmpo
2. Anthony Davis
3. Nikola Jokić

Decision Tree Classification



Key Features:

- The model uses Efficiency (EFF) and Points Per Game (PTS) to classify MVP candidates.
- **EFF ≤ 0.81** is the most significant split, dividing players into likely MVP and non-MVP groups.

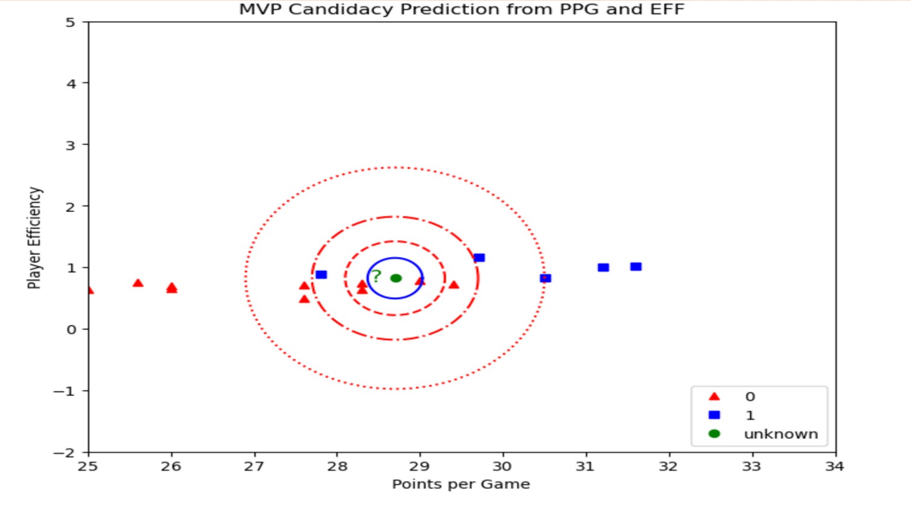
Results:

- Achieved 100% accuracy on the training data, perfectly classifying all samples.
- This suggests potential overfitting, as the model may not generalize well to new data.

Insight:

- Players with high EFF (> 0.81) but moderate PTS can still be classified as MVP candidates, showing the importance of efficiency in performance evaluation.

k-Nearest Neighbors Classifier




```
For k=1, predict green is 0.
For k=3, predict green is 0.
For k=5, predict green is 0.
For k=7, predict green is 0.
For k=1, predict green is 0.0.
For k=3, predict green is 0.0.
For k=5, predict green is 0.2.
For k=7, predict green is 0.286.
```

- X = PPG Y= player efficiency
- Green dot for NBA player averaging 28.7 PTS on 0.82 EFF
- the classifier would determine most player to not be a MVP candidate.

Conclusion:

- We have found that using a support vector machine is the best way to predict if a player is an MVP candidate. This method correctly identifies 'MVP-caliber' players based on their individual achievements, with very few mistakes.
- The decision tree classifier, while it seems strong at first, tends to fit the data too closely. This overfitting means it may not predict new data accurately.
- The k-nearest neighbors method did not work as well with our data. This is because most players in the data are not MVP candidates, which affects how this method performs.
- It is important to remember that deciding if a player is an MVP candidate is somewhat subjective. The MVP award is given based on votes from sports media people, who might consider factors like a player's story or "narrative." These factors are hard to measure with data. So, while our models are helpful, they might not capture all the reasons why a player becomes an MVP.