

# 💰 NBA MONEYBALL 💰





# Machine Learning Contributors

- **Project Manager:**  
Chandan Yadav
- **MVP Predictor:**  
Brian McHorney
- **Team Predictor:**  
Bryce Pribyl



# Predictions and Market

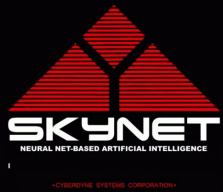
- **Vision - Wanted to spend time on something that is happening live and be able to train the model with new data as its available**
- **What's trending - NBA series had just started and there was a lot of buzz in the media and news about who will be the winners**
- **Goal - Create model that scores players performance and predict winning team based on regular season data**

## Our dataset consists of 25 records, each representing a player's performance in the NBA Finals

### The dataset contains the following features:

- Minutes Played (MIN)
- Field Goals Made (FGM)
- Field Goals Attempted (FGA)
- Field Goal Percentage (FG%)
- Three-Pointers Made (3PM)
- Three-Pointers Attempted (3PA)
- Three-Point Percentage (3P%)
- Free Throws Made (FTM)
- Free Throws Attempted (FTA)
- Free Throw Percentage (FT%)
- Offensive Rebounds (OREB)
- Defensive Rebounds (DREB)
- Total Rebounds (REB)
- Assists (AST)
- Turnovers (TOV)
- Steals (STL)
- Blocks (BLK)
- Personal Fouls (PF)
- Points (PTS)
- Plus/Minus (P/M)
- Versatility Index (VI)
- Prediction (Target variable indicating MVP status)\*\*

(Sample View)	MIN	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	...	REB	AST	TOV	STL	BLK	PF	PTS	P/M	VI	Prediction
PLAYER																					
Bam Adebayo	41.1	9.0	19.8	45.6	0.0	0.3	0.0	4.3	4.8	89.5	...	12.5	3.8	2.8	0.0	1.0	2.8	22.3	-6.8	9.4	1
Jimmy Butler	41.0	8.3	18.5	44.6	1.3	3.5	35.7	4.0	5.0	80.0	...	5.0	6.8	1.3	0.5	0.8	1.0	21.8	-8.5	10.6	1
Gabe Vincent	30.0	4.5	10.5	42.9	2.5	6.5	38.5	1.3	1.5	83.3	...	0.5	2.3	1.0	1.0	0.3	2.8	12.8	-3.8	5.7	1
Kyle Lowry	27.9	3.3	6.8	48.1	1.8	4.3	41.2	2.3	2.3	100.0	...	2.8	5.0	2.0	0.8	0.0	2.3	10.5	-3.5	8.0	1
Duncan Robinson	19.4	3.3	6.0	54.2	2.0	4.5	44.4	0.0	0.3	0.0	...	0.8	1.3	0.5	0.3	0.0	2.5	8.5	-5.0	6.5	1



# PREDICTING MVP WITH MACHINE LEARNING

VI	PLUS/MINUS	Prediction	Cluster
0.900238	-1.268401	1	1
1.221752	-1.587945	1	1
-0.091096	-0.704500	1	2
0.525139	-0.648110	1	2
0.123247	-0.930060	1	2

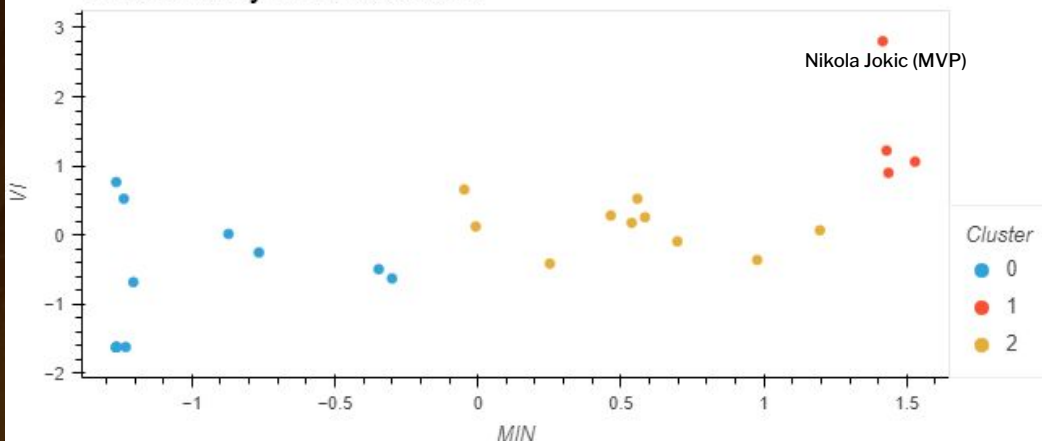
## LinearRegression() model

Based on the in game stats the most valuable player is: Name: Nikola Jokic

```
#Accuracy score for the model  
y_pred = model.predict(X_test)  
accuracy = accuracy_score(y_test, y_pred)  
print("Accuracy:", accuracy)
```

**Accuracy: 0.8015**

Scatter Plot by VI with 3 clusters

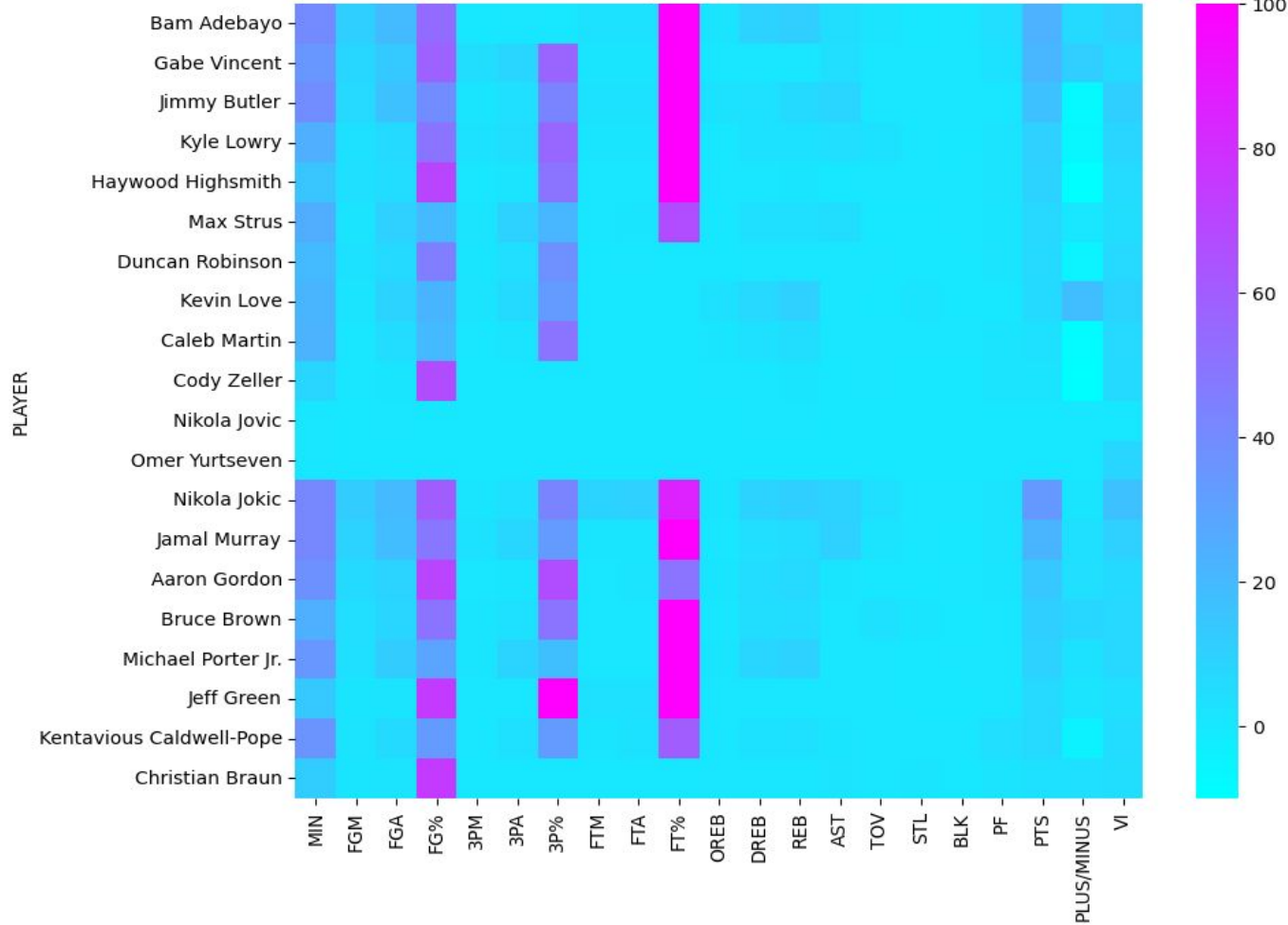


Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 9)	27
dense_1 (Dense)	(None, 6)	60
dense_2 (Dense)	(None, 2)	14
Total params: 101		
Trainable params: 101		
Non-trainable params: 0		



Heatmap of Combined Data



**MVP  
Player  
Heatmap**

# MVP PREDICTOR

- Binary classification model to predict the NBA Finals MVP
- Linear Regression machine learning
- Exclusively based on NBA Finals Player Data (per game)
- **Correctly predicted the NBA Finals MVP: Nikola Jokic!!**

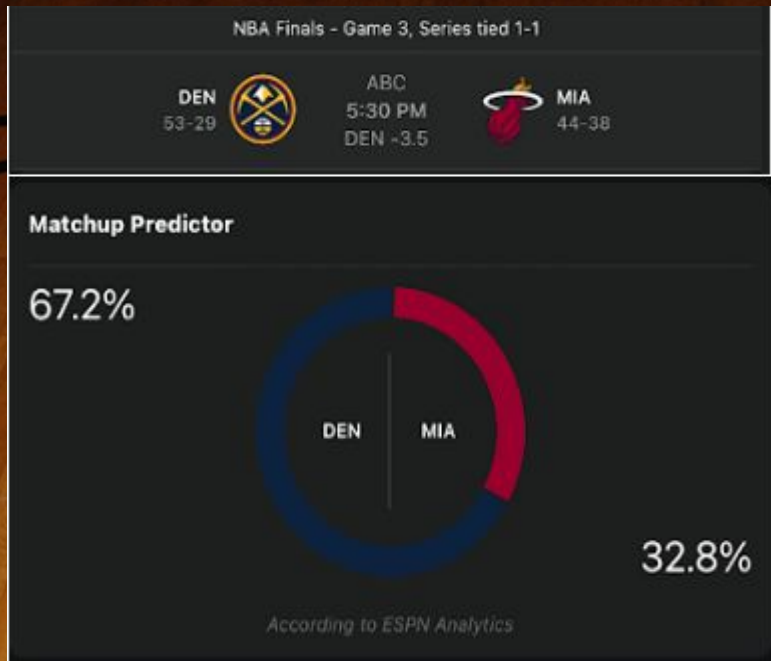
# ML Team Model

- RandomForestRegressor Machine Learning
- Inception: post-game 2 & prior to the start of game 3
- 2022-2023 Regular season & NBA Finals Data
  - CSV data from NBA.com
  - ESPN Matchup Predictor
- Oddspegia API integration with live betting odds: FanDuel, DraftKings
- NBA Moneyball Metrics: Points, Rebounds, Assists
- Surprise in store... wait for it.....



# 2023 NBA Finals Winning Game Predictor: Game 3 sample

## ESPN Matchup Predictor



## BCB Model Prediction

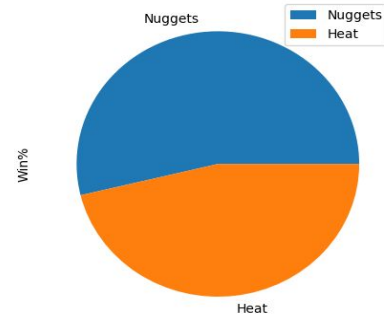
```
[19]: # Make predictions for game 3
game3_predictions = model.predict(df2)

# Print the predicted winner for game 3
print("Probability of Winning Game 3 (Nuggets):", round(game3_predictions[0],2), '%')
print("Probability of Winning Game 3 (Heat):", round(game3_predictions[1],2), '%')
# Evaluate the model accuracy using the testing set
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print("Model mean squared error:", mse)

Probability of Winning Game 3 (Nuggets): 53.97 %
Probability of Winning Game 3 (Heat): 46.38 %
Model mean squared error: 1.0773733854168586

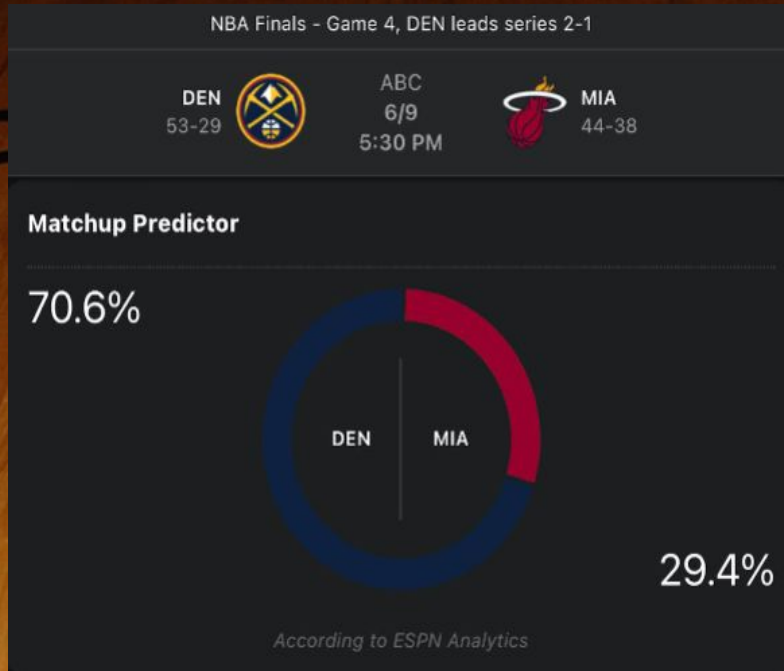
[20]: %matplotlib inline
pd.DataFrame(game3_predictions, index = ['Nuggets', 'Heat'], columns = ['win%']).plot.pie(y = 'win%')

[20]: <AxesSubplot: ylabel='win%'>
```



# 2023 NBA Finals Winning Game Predictor: Game 4 sample

## ESPN Matchup Predictor



## BCB Model Prediction

```
[4]: # Make predictions for the winner of the next NBA finals game
next_predictions = model.predict(df2)

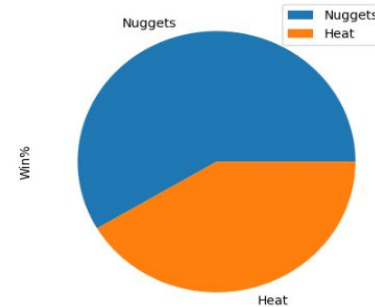
# Print the predicted winner for game 3
print("Probability of Winning Game 4 (Nuggets):", round(next_predictions[0],2), '%')
print("Probability of Winning Game 4 (Heat):", round(next_predictions[1],2), '%')

# Evaluate the model accuracy using the testing set
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print("Model mean squared error:", mse)

Probability of Winning Game 4 (Nuggets): 57.19 %
Probability of Winning Game 4 (Heat): 40.54 %
Model mean squared error: 1.0773733854168586

[9]: %matplotlib inline
pd.DataFrame(next_predictions, index = ['Nuggets', 'Heat'], columns = ['Win%']).plot.pie(y = 'Win%')

[9]: <AxesSubplot: ylabel='Win%'>
```





## Successful (potentially lucrative) Results

- Accurately predicted the winner of games 3-5 (Denver Nuggets)!
- Correctly predicted the NBA Finals MVP: Nikola Jokic!!



# We are **HIRING!**



Predict winners of each game for all the series by taking regular season data of current year

Share model to betting websites like oddspedia for all users to use the model

Take more historical data of all MVPs for the last 5 years to give better prediction results

Analyze Google trends or twitter for fan sentiment analysis

End game betting app to get real time odds.



# References

<b>Player Data</b>	<a href="https://www.nbastuffer.com/2022-2023-nba-player-stats/">https://www.nbastuffer.com/2022-2023-nba-player-stats/</a>
<b>US Sports Betting Revenue</b>	<a href="https://www.legalsportsreport.com/sports-betting/revenue/">https://www.legalsportsreport.com/sports-betting/revenue/</a>
<b>NBA Moneyball</b>	<a href="https://www.nbastuffer.com/analytics-101/nba-moneyball/">https://www.nbastuffer.com/analytics-101/nba-moneyball/</a>
<b>Model</b>	<a href="https://www.dataart.com/blog/5-use-cases-for-machine-learning-in-sports-betting">https://www.dataart.com/blog/5-use-cases-for-machine-learning-in-sports-betting</a> <a href="https://www.dataart.com/clients/case-studies/applications-suite-for-a-sports-analytics-company">https://www.dataart.com/clients/case-studies/applications-suite-for-a-sports-analytics-company</a>
<b>Odds API</b>	<a href="https://the-odds-api.com/#get-access">https://the-odds-api.com/#get-access</a>