NBA Draft 2022

Final Project

Ethan Denenberg, Kenneth Garcia, Rayan Jaipuriyar
MSDS 460: Decision Analytics
June 8th, 2024

Table of Contents

Abstract

An Integer Programming optimization approach was used to predict the 2022 NBA 1st Round Draft. The aim was to make optimal picks for the Houston Rockets, who have the 3rd and the 17th first-round picks. To do this we designed a model that makes optimal picks for all teams up to the 17th pick. The model predicts that the Houston Rockets should choose Jalen Williams and David Roddy on the 3rd and 17th picks.

Introduction

The 2022 NBA Draft presented a crucial opportunity for the Houston Rockets to rebuild and strengthen their roster, particularly in the forward positions. With the 3rd and 17th picks in the first round, our objective is to optimize these selections by utilizing decision analytics to re-predict the draft outcomes. This project aims to leverage data modeling techniques to analyze player performance metrics and draft outcomes, thereby enabling us to make informed decisions on the optimal picks for the Houston Rockets.

Our primary focus is to select two highly talented forwards, assuming the right talent is available when it's our turn to draft. To achieve this, we have collected comprehensive data on player performances, draft projections, and team needs. By re-evaluating the draft with a data-driven approach, we can identify which players should have been selected by each team at their respective spots. This analysis will help us determine the best available options for the Rockets at both the 3rd and 17th picks, ensuring that our selections align with the team's strategic goals and long-term vision.

In this report, we will detail the methodology and analytical techniques used to model the draft data, the criteria for player evaluation, and the decision-making process behind our optimal draft choices. By the end of this analysis, we aim to provide a clear, evidence-based recommendation for the Houston Rockets' picks in the 2022 NBA Draft.

Literature Review

Predicting the seasonal draft for sports like basketball and football have been popular avenues of research for data modelers. The commercial and entertainment viability of such projects makes these predictive models a popular pet project. Machine learning has been used in the past to predict the NFL draft (Grassy, n.d.). While an interesting approach, we believe using a neural network algorithm for what is a simple maximization problem is not optimal. Instead, less computationally heavy approaches can be used to model the same problem. Closer in scope to the problem we are trying to solve, linear programming has been used to optimize picks for a fantasy football draft (Blagojevic 2018). Given that multiple picks can be made in a draft in this situation, a standard continuous variable optimization approach was utilized to select picks, subject to the positional and salary cap constraints. Our approach will be similar, but instead with an integer programming implementation.

Methodology

The success of our project hinges on the development and application of a robust player rating equation to evaluate the potential draft picks accurately. Our data source for this analysis is the official NBA website, from which we obtained an XLSX file titled "2022 Draft - Stats (*Appendix A)*." This file contains comprehensive college statistics for players drafted in 2022. After careful consideration, we determined that the most relevant metrics for assessing player quality were field goal percentage (FG%), points per game (PPG), rebounds per game (RPG), and assists per game (APG).

To quantify player performance, we created a "Player Rating" (PR) metric, which incorporates these key statistics from their most recent college season. The formula we devised is as follows:

Player Rating = (PPG*FG%) + RPG + APG

This equation is designed to provide a balanced assessment of a player's offensive efficiency (PPG *FG%), rebounding ability (RPG), and playmaking skills (APG). By multiplying points per game by field goal percentage, we emphasize scoring efficiency, ensuring that players who score frequently and effectively are rated higher. Adding rebounds and assists per game ensures that all-around performance is taken into account, not just scoring.

Once the Player Ratings were calculated, our team, well-versed in NBA player evaluations, validated the results. This cross-check with our extensive knowledge of the players confirmed that our ratings were both accurate and aligned with the expected outcomes.

In addition to the Player Rating, we introduced a variable called "Position Quant" to account for team-specific needs in our model. Each team was assigned a weight based on their positional demands: Center (C), Forward (F), or Guard (G). The weights were determined as follows:

- 2 for the most critical position a team needs to fill
- 1.5 for the secondary position of importance
- 1 for the least critical position

These weights were assigned based on each team's roster necessities, ensuring that our model not only identified the best available talent but also addressed team-specific strategic needs. We created a Google Sheet titled “2022_Draft_Weights” (*Appendix B)* to consolidate this information.

The combined use of Player Rating and Position Quant allows us to model draft scenarios effectively, identifying the optimal picks for the Houston Rockets at their respective spots. This methodology ensures that our draft recommendations are data-driven, addressing both player quality and team needs comprehensively.

Using the described weights, we have chosen to implement an integer programming approach to solve a knapsack problem in order to ascertain the optimal picks for each team. We created an objective function for each team in the draft. This function is a summation of the products of the two weights and the integer variable x_i, for each player:

$$\sum_i Player\ Rating_i \cdot Position\ Quant_i \cdot x_i \qquad \text{For players in the draft } i$$

The LP was maximized for each team, subject to the constraint that the sum of all integer variables x_i cannot exceed 1, thereby ensuring that only one player is chosen by a team during an individual pick. Additionally, once a pick was made, the list of available players was updated to reflect the change in available players for the next team, thereby ensuring that the model optimized for only the available players in the current pick.

Results

The player rating formula we applied produced different outcomes compared to the real 2022 NBA draft. Our mock draft showed a difference of five players within the first 17 picks compared to the real draft selections. The model's predictions are as follows for the mock draft:

Pick	Team	Player	Position	Player Rating
1	ORL	Paolo Banchero	F	20.9146
2	OKC	Jalen Duren	C	19.0934
3	HOU	Jalen Williams	F	16.878
4	SAC	Jabari Smith Jr.	F	14.719
5	DET	Jaden Ivey	G	14.8676
6	IND	Chet Holmgren	C	19.045
7	POR	Mark Williams	C	14.9842
8	NOP	Walker Kessler	C	14.903
9	SAS	Jeremy Sochan	F	13.9172
10	WAS	Shaedon Sharpe	G	10.362
11	NYK	Bennedict Mathur:	F	12.6362
12	OKC	Keegan Murray	F	12.6198
13	CHO	Tari Eason	F	11.5488
14	CLE	Nikola Jović	F	8.4453
15	CHO	Malaki Branham	G	8.5292
16	ATL	Christian Braun	F	7.1853
17	HOU	David Roddy	F	6.7456

As for the real draft results for comparison:

Pick	Team	Player	Position	Player Rating
1	ORL	Paolo Banchero	F	20.9146
2	OKC	Chet Holmgren	C	19.045
3	HOU	Jabari Smith Jr.	F	14.719
4	SAC	Keegan Murray	F	12.6198
5	DET	Jaden Ivey	G	14.8676
6	IND	Bennedict Mathurin	F	12.6362
7	POR	Shaedon Sharpe	G	10.362
8	NOP	Dyson Daniels	G	8.088
9	SAS	Jeremy Sochan	F	13.9172
10	WAS	Johnny Davis	G	4.176
11	NYK	Ousmane Dieng	F	5.2945
12	OKC	Jalen Williams	F	16.878
13	CHO	Jalen Duren	C	19.0934
14	CLE	Ochai Agbaji	F	6.4073
15	CHO	Mark Williams	C	14.9842
16	ATL	AJ Griffin	F	6.0525
17	HOU	Tari Eason	F	11.5488

The Rockets didn't have a better overall draft in this mock draft compared to the real draft. The Rockets mock-drafted Jalen Williams with the 3rd pick which was 2.159 PR points over Jabari Smith Jr. who they picked up in the real draft. Unfortunately, the forward Tari Eason was picked up earlier by CHO at pick 13, so the Rockets couldn't pick him up at pick 17. The Rockets then downgraded to David Roddy to fill the same forward position with a player rating, PR, of 6.7456 compared to Tari Eason's rating of 11.5488, which is a loss of 4.8032. The Rockets are down overall in this mock draft by -2.6442 for both picks when compared to their real draft results.

Even though the mock draft had fewer forward players drafted before pick 17, nine forwards, compared to 10 forwards in the real draft; the Rockets still lost out because all the teams before them had better picks using this player rating model. This is highlighted by Tari being drafted earlier and when three forwards with lower player ratings were drafted before Tari in the real draft: NYK Ousmane Dieng, forward PR 5.2945, CLE Ochai Agbaji PR 6.4073, and ATL AJ Griffin PR 6.0525. Instead, NYK, CLE, and ATL mock-drafted higher-rated forward players Bennedict Mathurin PR 12.6362, Nikola Jović PR 8.4453, and Christian Braun PR 7.1853, respectively.

Other notable changes were that five players in the top 17 picks from the real draft that were not selected in the top 17 in the mock, these players fell short in our model's player rating formula: Dyson Daniels, G, PR 8.088; Johnny Davis, G, PR 4.176; Ousmane Dieng, F, PR 5.2945; Ochai Agbaji, F, PR 6.4073; and AJ Griffin, F, PR 6.0525. Conversely, five other players were elevated to the top 17 picks who fell below that in the real draft: Walker Kessler, C, PR 14.903; Nikola Jović, F, PR 8.4453; Christian Braun, F, PR 7.1853; David Roddy, F, PR 6.7456; and Malaki Branham, G, PR 8.5292.

The weights relative to the players' rating made an impact on five key picks in the first 17 picks of the mock draft when five teams changed what positions they drafted in the mock over their position of need because a better player was available. IND with the 6th pick, mock drafted Chet Holmgren, center PR 19.045, over their need for a forward. For IND Chet's PR had a weight of 1.5 so their overall score was 28.5675 while the next best forward was Jeremy Sochan, PR 13.9172, with a weight of 2 had an overall score of 27.8344. Despite needing a guard, POR used their 7th pick to select Mark Williams, center PR 14.9842. NOP needed a guard and got Walker Kessler, center PR 14.903, with the 8th pick. CHO changed both their 13th and 15th mock draft picks from two centers to Tari Eason, forward PR 11.5488, and Malaki Branham, guard PR 8.5292.

Taking the difference between the mock draft player's PR subtracted by the real draft player's PR:

Pick	Team	Player Rating Difference
1	ORL	0
2	OKC	0.0484
3	HOU	2.159
4	SAC	2.0992
5	DET	0
6	IND	6.4088
7	POR	4.6222
8	NOP	6.815
9	SAS	0
10	WAS	6.186
11	NYK	7.3417
12	OKC	-4.2582
13	CHO	-7.5446
14	CLE	2.038
15	CHO	-6.455
16	ATL	1.1328
17	HOU	-4.8032

The player rating difference shows that the majority of teams got the same or better results in the mock draft when using our rating system. Alternatively, four teams had worse results: OKC, CHO, and HOU. CHO had it the worst with both of their picks falling because their real drafted players were selected much earlier in the draft in the mock.

Conclusion

Of course, the real draft had different results as there are more variables, some quantifiable and others not so much, that are valued by certain teams more than others. Some teams could just value certain positions over others because of strategies and play styles the coaches favor, biased toward some metrics, will always take the position of need over the best player available, etc. Conversely, not everything is about numbers, some teams value what they see with their eyes over stats, gut feelings, intuition, and the like. In short, there is no one model or formula that all teams in the draft would use when there are so many variables and biases around, the real draft just won't be as uniform as our mock draft.

Initially, it seemed that the weight differences of 1, 1.5, and 2 would be too great and that only the most critical positions with a weight of 2 would be drafted but it did make an impact on five teams. It would be fun to play with different weights to see how the mock draft plays out, lowering the weights could potentially cause a lot of chaos in the draft. And while the Rockets didn't have a better draft, 10 of the other teams did end up with a better rating and the Rockets didn't have a vastly worse draft.

Surprisingly, the results across the 17 mock-draft picks were overall better than the real draft. Maybe this means that looking at fewer variables, basketball footage, and ignoring biases for a more simplistic approach has some real value. Maybe NBA teams should be more like Occam's razor, that the simplest solution is often the best.

Appendix A

Draft Player Statistics (Basketball Reference 2024).

Pick	Team	Player	Position (G, F, C)	Position Quant	FG%	PTS	PT + FG%	TRB	AST	Player Rating
1	ORL	Paolo Banchero	F	2	0.442	21.3	9.4146	6.9	4.6	20.9146
2	OKC	Chet Holmgren	C	1.5	0.53	16.5	8.745	7.9	2.4	19.045
3	HOU	Jabari Smith Jr.	F	2	0.43	13.3	5.719	7.6	1.4	14.719
4	SAC	Keegan Murray	F	2	0.454	13.7	6.2198	5	1.4	12.6198
5	DET	Jaden Ivey	G	1	0.422	15.8	6.6676	3.7	4.5	14.8676
6	IND	Bennedict Mathurin	F	2	0.439	15.8	6.9362	4	1.7	12.6362
7	POR	Shaedon Sharpe	G	1	0.445	11.6	5.162	3.5	1.7	10.362
8	NOP	Dyson Daniels	G	1	0.435	4.8	2.088	3.5	2.5	8.088
9	SAS	Jeremy Sochan	F	2	0.444	11.3	5.0172	5.9	3	13.9172
10	WAS	Johnny Davis	G	1	0.394	4	1.576	1.8	0.8	4.176
11	NYK	Ousmane Dieng	F	2	0.421	4.5	1.8945	2.2	1.2	5.2945
12	OKC	Jalen Williams	F	2	0.532	16.5	8.778	4.2	3.9	16.878
13	CHO	Jalen Duren	C	1.5	0.631	11.4	7.1934	10.2	1.7	19.0934
14	CLE	Ochai Agbaji	F	2	0.419	6.7	2.8073	2.5	1.1	6.4073
15	CHO	Mark Williams	C	1.5	0.642	10.1	6.4842	7.9	0.6	14.9842
16	ATL	AJ Griffin	F	2	0.447	7.5	3.3525	1.9	0.8	6.0525
17	HOU	Tari Eason	F	2	0.452	9.4	4.2488	6.2	1.1	11.5488
18	CHI	Dalen Terry	G	1	0.441	2.7	1.1907	1.5	1.1	3.7907
19	MIN	Jake LaRavia	F	2	0.389	6.9	2.6841	2.7	1.1	6.4841
20	SAS	Malaki Branham	G	1	0.436	9.7	4.2292	2.3	2	8.5292
21	DEN	Christian Braun	F	2	0.473	6.1	2.8853	3.1	1.2	7.1853
22	MEM	Walker Kessler	C	1.5	0.69	8.7	6.003	8	0.9	14.903
23	PHI	David Roddy	F	2	0.416	6.6	2.7456	3	1	6.7456
24	MIL	MarJon Beauchamp	G	1	0.433	4.8	2.0784	2.1	0.7	4.8784
25	SAS	Blake Wesley	G	1	0.398	4.6	1.8308	1.8	2.7	6.3308
26	DAL	Wendell Moore Jr.	F	2	0.443	1.1	0.4873	0.6	0.4	1.4873
27	MIA	Nikola Jović	F	2	0.443	7.1	3.1453	3.7	1.6	8.4453
28	GSW	Patrick Baldwin Jr.	F	2	0.386	4.2	1.6212	2.4	0.6	4.6212
29	MEM	TyTy Washington Jr.	G	1	0.356	3.8	1.3528	1.2	1.2	3.7528
30	OKC	Peyton Watson	F	2	0.468	5.9	2.7612	2.9	0.9	6.5612

2022 Draft - Stats.xlsx

Appendix B

Position Demands per Team (Basketball Reference 2024).

Team	G Weight	F Weight	C Weight
ORL	1	2	1.5
OKC	1	1.5	2
HOU	1.5	2	1
SAC	1.5	2	1
DET	2	1.5	1
IND	1	2	1.5
POR	2	1	1.5
NOP	2	1	1.5
SAS	1.5	2	1
WAS	2	1	1.5
NYK	1.5	2	1
OKC	1.5	2	1
CHO	1	1.5	2
CLE	1.5	2	1
CHO	1.5	1	2
ATL	1	2	1.5
HOU	1.5	2	1
CHI	2	1	1.5
MIN	1.5	2	1
SAS	2	1	1.5
DEN	1.5	2	1
MEM	1	1.5	2
PHI	1.5	2	1
MIL	2	1.5	1
SAS	2	1	1.5
DAL	1	2	1.5
MIA	1.5	2	1
GSW	1	2	1.5
MEM	2	1.5	1
OKC	1.5	2	1

2022_Draft_Weights

Appendix C

The Python code used to implement the model.

```python
# The final model that makes picks for all teams

teams = df.columns[7:24]
chosen=[]
players = list(df.index)
playernames = list(df['Player'])
playerrating = list(df['Player Rating'])

for team in teams:
    prob = LpProblem(sense=LpMaximize)
    obfunc=[]
    var=[]

    for player in players:
        c = df[team][player]
        q = df['Player Rating'][player]
        x = LpVariable(f'x{player}', 0, 1, cat='Integer')
        var.append(x)
        obfunc.append(c*q*x)
    prob += sum(obfunc)
    prob += sum(var) <=1
    prob.solve()
    #prob.writeLP(f"{team}.lp")
    for i in enumerate(var):
        if i[1].value() == 1:
            chosen.append([team, playernames[i[0]], playerrating[i[0]]])
            del players[i[0]]
            del playernames[i[0]]
            del playerrating[i[0]]


print('\nTeam picks')
chosen
print('\nplayers not chosen')
#players
playernames
```

The Output:

```
Team picks

[['ORL', 'Paolo Banchero', 20.9146],
 ['OKC', 'Jalen Duren', 19.0934],
 ['T_HOU', 'Jalen Williams', 16.878],
 ['SAC', 'Jabari Smith Jr.', 14.719],
 ['DET', 'Jaden Ivey', 14.8676],
 ['IND', 'Chet Holmgren', 19.045],
 ['POR', 'Mark Williams', 14.9842],
 ['NOP', 'Walker Kessler', 14.903],
 ['SAS', 'Jeremy Sochan', 13.9172],
 ['WAS', 'Shaedon Sharpe', 10.362],
 ['NYK', 'Bennedict Mathurin', 12.6362],
 ['OKC2', 'Keegan Murray', 12.6198],
 ['CHO', 'Tari Eason', 11.5488],
 ['CLE', 'Nikola Jović', 8.4453],
 ['CHO2', 'Malaki Branham', 8.5292],
 ['ATL', 'Christian Braun', 7.1853],
 ['T_HOU2', 'David Roddy', 6.7456]]


players not chosen

['Dyson Daniels',
 'Johnny Davis',
 'Ousmane Dieng',
 'Ochai Agbaji',
 'AJ Griffin',
 'Dalen Terry',
 'Jake LaRavia',
 'MarJon Beauchamp',
 'Blake Wesley',
 'Wendell Moore Jr.',
 'Patrick Baldwin Jr.',
 'TyTy Washington Jr.',
 'Peyton Watson']
```

References

Blagojevic, Branko. 2018. “Using Python and Linear Programming to Optimize Fantasy Football Picks.” Medium. https://medium.com/ml-everything/using-python-and-linear-programming-to-optimize-fantasy-football-picks-dc9d1229db81.

Grassy, Brock. n.d. “Applying Machine Learning to Predict the NFL Draft.” BG. Accessed June 4, 2024. https://www.brockgrassy.com/blog/applying-machine-learning-to-predict-the-nfl-draft/.

Basketball Reference. 2022. "2022 NBA Draft." *Basketball Reference*. Accessed April 26, 2024. https://www.basketball-reference.com/draft/NBA_2022.html.