IST707 – Applied Machine Learning

Predicting NBA Free Agent Salaries for the 2022-23 Season

Jack Dolitsky

jdolitsk@syr.edu

Abstract

Every offseason in the National Basketball Association (NBA), a pool of free agents looks to sign a new contract with a team. Every team has a salary cap, and has a given amount of money that they can spend amongst their players. It can be difficult to figure out how much a player is worth, and teams may overpay or lose player's due to low ball offers. This project aims to create a model to capture the monetary value of every single player, and that can be used for teams to see how much a player is worth on the open market.

Predicting NBA Free Agent Salaries

Free agency is critical for every NBA team, and will usually have a major impact on the direction and future of a franchise. Overpaying a single player, can have major negative implications for a team for the several years following. Understanding how to value players can be very difficult, because there are so many different factors that come into play. The Miami Heat signed Chris Bosh and LeBron James to massive contracts and ended up winning two championships. On the other hand, the New York Knicks signed Joakim Noah in 2016, and never had any playoff success with him on the roster, and still had to pay him through the 2022 season after he retired. I am using a random forest model to look at every single NBA contract from the 2021-2022 season and use those player's stats to predict what they are really worth on the open market. I will be running two Random Forests models, one with integers and one with classification. I will also run an SVM Model. I will then predict what the free agents in 2022 are worth and see who is currently overpaid and underpaid. The main question that is being answered is what salaries are the upcoming free agent class expected to get? This will help team front offices make an educated decision before signing a free agent, and potentially setting themselves up for years of being confined to unmovable contracts.

Method

In order to get all of the data, I needed to get three separate tables from basketball-reference.com. I gathered a table for all player salaries, all players' advanced stats, and all players' general stats. I then joined all the tables on the player name and dropped the unwanted columns that were irrelevant to the study I am trying to conduct. For the training and testing split, I divided all players who were not going to be free agents into a training data frame, and all upcoming free agents into a testing data frame. I was able to create the free agents table by finding those who had null values for the 2022-23 season, meaning they did not have a salary for that year. I also dropped all unwanted columns that were not relevant to the study.

For the general stats I decided to keep, they were the stats such as PPG (points per game), TRB (total rebounds per game), MPG (minutes per game), and some other stats that are typically shown in a basic box score. This is important because these are the stats the casual watcher looks at, and the first thing that is noticed about a player among everybody, so these stats definitely contribute to a player's worth. The next set of stats I used were advanced stats. These stats are similar but adjusted for playing time and efficiency. This helps the model take into account role. Some players may have high general stats, because they play for a bad team and have a major role, but it does not contribute to winning. On the other hand, there are players who do not have eye-popping stats due to playing as a lesser option on a team but can contribute a lot more with a larger role. I also included age, so the model can take into account that players who have more experience, have a higher max and minimum salary that they can receive. This created a training data set with 317 observations, and a testing data set with 125 observations. The final steps were some extra cleaning with changing some column names and dropping duplicate values.

The next step was to turn this into a classification project, so I created salary buckets, as opposed to a set number. This is important for a variety of reasons. The first reason is because teams value players differently. One team may be looking for experienced hustle players to support the star player, while another team may be looking to rebuild with young scorers, and neither team would pay the other player the same amount, respectively. The salary buckets create a range, so it allows the model to show what a player is worth, while also giving leeway for teams to pay a bit more or less depending on their need. I split the buckets into 4 million dollars, which creates

12 categories of players, from $0 to $48 million. I chose 4 million dollars because I found that it created the best balance of being able to have a large enough salary range for front offices to bargain with, but also not too large where the model would not return any meaningful insight.

Results

One model ran was a Support Vector Machine model. This model returned an accuracy of predicting a player's salary of 49.6%. This is very accurate, as if the model were guessing randomly, we'd expect an accuracy of 8.33%. After tuning the model with center and scale, I could not get the accuracy higher. I also ran a Random Forest Model. This originally came out to 51% accuracy, but after tuning with mtry, I got the model to be 61.6% accurate. As you can see (see Appendix A), most of the correct predictions came in the 0-4-million-dollar range. That is because this salary range has the most NBA players. The players that make in this range usually do not have as much impact, but this is still an important bucket to consider because this could be where you get the most value for low risk. Since this free agency class does not have as many superstars and the previous few years, there are not too many players who were making such high salaries, so for this year especially, it will be important for teams to try and find the low-risk value players.

Discussion

After running both the SVM and Random Forest model, I found the Random Forest was more accurate and decided to focus on that to make my predictions. We saw it most accurately predict the players who fell into the 0-4-million-dollar category, and that is the prime spot to look to make high reward investment.

When looking at the table of players (see Appendix B), you can see even the players that the model guessed wrong, it was usually only one salary bucket off from what they made in 2021-22. There are some noteworthy predictions. Blake Griffin was making in the 32-36 range this past season, but the model predicts him to make under 4 million dollars next season. Griffin, a former starting all-star and MVP candidate, suffered an injury a couple of years back, and has saw his role diminish ever since. It is important for GMs to note this, because name brand can influence overpaying. This was by far the largest discrepancy in the model. All three of the currently highest paid upcoming free agents are expected to make considerably less than they are making right now.

Another notable player is DeAndre Ayton who is currently making \$12-16 million. He is expected to make \$16-20 million next year, which is a jump from his current salary, but there are talks about him receiving a max contract, which, according to Hoops Rumors, means he can make up to almost \$30 million dollars in the upcoming season. This is important for the Phoenix Suns, his current team, to take into account. Since Ayton is a restricted free agent, the Suns are able to match any offer another team gives him, even if he signs their offer. If another team does offer him the max, this model would suggest that the Suns should not match and find a different option at center. Teams who are on a budget constraint should look towards players such as Dennis Schroder Carmelo Anthony, and Victor Oladipo. All of these players are making \$0 to \$8 million, and the model suggests that they are worth at least \$12 million. Due to their old age, they make garner as much interest from teams as some younger players will, so they can be great value signings.

The model has some limitations. The first limitation is that this model does account for the potential of a player. Potential is not something that can necessarily be quantitative, so it can be arbitrary. Some players are expected to make big jumps in the future due to a variety of factors, and that can raise their value on the open market. Since there is no “potential” statistic, it would have been very difficult for this model to have been able to account for that. Another limitation is that is only based of one season of data. A player could have suffered an injury, or just had an off-year, and the model would not take that into account. Additionally, the model does not account for salary minimums and maximums. Some players are currently making less than they are worth because they are still on their rookie contract or they do not have enough years in the league to be making more.

References

"Basketball Statistics and History." *Basketball*, https://www.basketball-reference.com/.

NBA Maximum Salaries For 2020/21. (2020, November 10). Hoops Rumors. https://www.hoopsrumors.com/2020/11/nba-maximum-salaries-for-202021.html

Appendix

A.

```
Confusion Matrix and Statistics

              Reference
Prediction    0-4mil 4mil-8mil 8mil-12mil 12mil-16mil 16mil-20mil 20mil-24mil 24mil-28mil 28mil-32mil 32mil-36mil 36mil-40mil 40mil-44mil
  0-4mil          75        13          8           2           0           0           0           0           1           0           0
  4mil-8mil       10         2          2           0           0           2           0           0           0           0           0
  8mil-12mil       0         1          0           2           0           0           0           0           0           0           0
  12mil-16mil      1         1          0           0           1           0           0           0           0           0           0
  16mil-20mil      1         0          1           1           0           0           0           0           0           0           0
  20mil-24mil      0         0          0           0           0           0           0           0           0           0           0
  24mil-28mil      0         0          0           0           0           0           0           0           0           0           0
  28mil-32mil      0         0          0           0           1           0           0           0           0           0           0
  32mil-36mil      0         0          0           0           0           0           0           0           0           0           0
  36mil-40mil      0         0          0           0           0           0           0           0           0           0           0
  40mil-44mil      0         0          0           0           0           0           0           0           0           0           0
  44mil-48mil      0         0          0           0           0           0           0           0           0           0           0
              Reference
Prediction    44mil-48mil
  0-4mil                0
  4mil-8mil             0
  8mil-12mil            0
  12mil-16mil           0
  16mil-20mil           0
  20mil-24mil           0
  24mil-28mil           0
  28mil-32mil           0
  32mil-36mil           0
  36mil-40mil           0
  40mil-44mil           0
  44mil-48mil           0

Overall Statistics

               Accuracy : 0.616
                 95% CI : (0.5248, 0.7016)
    No Information Rate : 0.696
    P-Value [Acc > NIR] : 0.9777

                  Kappa : 0.1026

 Mcnemar's Test P-Value : NA
```

B.

Predicted Salary	Current Salary	Name
0-4mil	32mil-36mil	Blake Griffin\griffbl01
4mil-8mil	20mil-24mil	Gary Harris\harriga01
4mil-8mil	20mil-24mil	Goran Dragi?\dragigo01
28mil-32mil	16mil-20mil	Zach LaVine\lavinza01
12mil-16mil	16mil-20mil	Ricky Rubio\rubiori01
16mil-20mil	12mil-16mil	Deandre Ayton\aytonde01
8mil-12mil	12mil-16mil	Joe Ingles\inglejo01
8mil-12mil	12mil-16mil	Robert Covington\covinro01
0-4mil	12mil-16mil	Taurean Prince\princta02
0-4mil	12mil-16mil	Thaddeus Young\youngth01
16mil-20mil	8mil-12mil	Jusuf Nurki?\nurkiju01
4mil-8mil	8mil-12mil	Marvin Bagley III\baglema01
4mil-8mil	8mil-12mil	Serge Ibaka\ibakase01
0-4mil	8mil-12mil	Jeremy Lamb\lambje01
0-4mil	8mil-12mil	Tomas Satoransky\satorto01
0-4mil	8mil-12mil	Kyle Anderson\anderky01
0-4mil	8mil-12mil	Tristan Thompson\thomptr01
0-4mil	8mil-12mil	Montrezl Harrell\harremo01
0-4mil	8mil-12mil	Derrick Jones Jr.\jonesde02
0-4mil	8mil-12mil	Thomas Bryant\bryanth01
0-4mil	8mil-12mil	Delon Wright\wrighde01
12mil-16mil	4mil-8mil	Dennis Schroder\schrode01
8mil-12mil	4mil-8mil	Miles Bridges\bridgmi02
4mil-8mil	4mil-8mil	Mo Bamba\bambamo01
4mil-8mil	4mil-8mil	Collin Sexton\sextoco01
0-4mil	4mil-8mil	Tyus Jones\jonesty01
0-4mil	4mil-8mil	Chris Boucher\bouchch01
0-4mil	4mil-8mil	Kevin Knox\knoxke01
0-4mil	4mil-8mil	Kevon Looney\looneke01
0-4mil	4mil-8mil	Troy Brown Jr.\browntr01
0-4mil	4mil-8mil	Josh Jackson\jacksjo02
0-4mil	4mil-8mil	JaVale McGee\mcgeeja01
0-4mil	4mil-8mil	Robin Lopez\lopezro01
0-4mil	4mil-8mil	Lou Williams\willilo02
0-4mil	4mil-8mil	Bruce Brown\brownbr01
0-4mil	4mil-8mil	Donte DiVincenzo\divindo01
0-4mil	4mil-8mil	Lonnie Walker IV\walkelo01
0-4mil	4mil-8mil	Josh Okogie\okogijo01
16mil-20mil	0-4mil	Victor Oladipo\oladivi01
12mil-16mil	0-4mil	Carmelo Anthony\anthoca01
4mil-8mil	0-4mil	Rajon Rondo\rondora01
4mil-8mil	0-4mil	Trevor Ariza\arizatr01
4mil-8mil	0-4mil	Dwight Howard\howardw01
4mil-8mil	0-4mil	Enes Freedom\kanteen01
4mil-8mil	0-4mil	Andre Iguodala\iguodan01
4mil-8mil	0-4mil	James Johnson\johnsja01

4mil-8mil	0-4mil	Austin Rivers\riverau01
4mil-8mil	0-4mil	Wesley Matthews\matthwe02
4mil-8mil	0-4mil	Jalen Brunson\brunsja01
4mil-8mil	0-4mil	Malik Monk\monkma01
0-4mil	0-4mil	Gorgui Dieng\dienggo01
0-4mil	0-4mil	Aaron Holiday\holidaa01
0-4mil	0-4mil	Jake Layman\laymaja01
0-4mil	0-4mil	Anfernee Simons\simonan01
0-4mil	0-4mil	Facundo Campazzo\campafa01
0-4mil	0-4mil	Paul Millsap\millspa01
0-4mil	0-4mil	Udonis Haslem\hasleud01
0-4mil	0-4mil	Markieff Morris\morrima02
0-4mil	0-4mil	Wayne Ellington\ellinwa01
0-4mil	0-4mil	LaMarcus Aldridge\aldrila01
0-4mil	0-4mil	Avery Bradley\bradlav01
0-4mil	0-4mil	Kent Bazemore\bazemke01
0-4mil	0-4mil	Andre Drummond\drumman01
0-4mil	0-4mil	Cody Zeller\zelleco01
0-4mil	0-4mil	Ben McLemore\mclembe01
0-4mil	0-4mil	Tony Snell\snellto01
0-4mil	0-4mil	Otto Porter Jr.\porteot01
0-4mil	0-4mil	Elfrid Payton\paytoel01
0-4mil	0-4mil	Rodney Hood\hoodro01
0-4mil	0-4mil	Bol Bol\bolbo01
0-4mil	0-4mil	Raul Neto\netora01
0-4mil	0-4mil	Frank Kaminsky\kaminfr01
0-4mil	0-4mil	Nemanja Bjelica\bjeline01
0-4mil	0-4mil	Abdel Nader\naderab01
0-4mil	0-4mil	Damian Jones\jonesda03
0-4mil	0-4mil	Rodney McGruder\mcgruro01
0-4mil	0-4mil	DeAndre' Bembry\bembrde01
0-4mil	0-4mil	Gary Payton II\paytoga02
0-4mil	0-4mil	Timothe Luwawu-Cabarrot\luwawti01
0-4mil	0-4mil	PJ Dozier\doziepj01
0-4mil	0-4mil	Damion Lee\leeda03
0-4mil	0-4mil	Dennis Smith Jr.\smithde03
0-4mil	0-4mil	Semi Ojeleye\ojelese01
0-4mil	0-4mil	Eric Paschall\pascher01
0-4mil	0-4mil	KZ Okpala\okpalkz01
0-4mil	0-4mil	Bruno Fernando\fernabr01
0-4mil	0-4mil	Cody Martin\martico01

0-4mil	0-4mil	Vlatko Cancar\cancavl01
0-4mil	0-4mil	Drew Eubanks\eubandr01
0-4mil	0-4mil	Yuta Watanabe\watanyu01
0-4mil	0-4mil	Isaac Bonga\bongais01
0-4mil	0-4mil	Isaiah Hartenstein\harteis01
0-4mil	0-4mil	Juan Toscano-Anderson\toscaju01
0-4mil	0-4mil	Kevin Pangos\pangoke01
0-4mil	0-4mil	Matt Thomas\thomama02
0-4mil	0-4mil	Robert Woodard II\woodaro01
0-4mil	0-4mil	Jahmi'us Ramsey\ramseja01
0-4mil	0-4mil	CJ Elleby\ellebcj01
0-4mil	0-4mil	Elijah Hughes\hugheel01
0-4mil	0-4mil	Jordan Nwora\nworajo01
0-4mil	0-4mil	Anthony Gill\gillan01
0-4mil	0-4mil	Danuel House Jr.\houseda01
0-4mil	0-4mil	Alfonzo McKinnie\mckinal01
0-4mil	0-4mil	Jabari Parker\parkeja01
0-4mil	0-4mil	Austin Reaves\reaveau01
0-4mil	0-4mil	Luka Garza\garzalu01
0-4mil	0-4mil	Lance Stephenson\stephla01
0-4mil	0-4mil	Miye Oni\onimi01
0-4mil	0-4mil	Wayne Selden\seldewa01
0-4mil	0-4mil	Brad Wanamaker\wanambr01
0-4mil	0-4mil	Alize Johnson\johnsal02
0-4mil	0-4mil	DeMarcus Cousins\couside01
0-4mil	0-4mil	Sam Merrill\merrisa01
0-4mil	0-4mil	Luke Kornet\kornelu01
0-4mil	0-4mil	Keifer Sykes\sykeske01
0-4mil	0-4mil	Caleb Martin\martica02
0-4mil	0-4mil	Georgios Kalaitzakis\kalaige01
0-4mil	0-4mil	Gary Clark\clarkga01
0-4mil	0-4mil	Isaiah Thomas\thomais02
0-4mil	0-4mil	Ish Wainright\wainris01
0-4mil	0-4mil	Joe Wieskamp\wieskjo01
0-4mil	0-4mil	Amir Coffey\coffeam01
0-4mil	0-4mil	RJ Nembhard Jr.\nembhrj01
0-4mil	0-4mil	Skylar Mays\mayssk01
0-4mil	0-4mil	Trent Forrest\forretr01