

### Basketball Game Predictor

This would be a model where you can input two NBA teams and their stats, record, etc... and it will predict who will win a matchup between them. I'm sure this has been done a thousand times over especially with sports betting being such a big industry, but I think it would be fun to build this on my own and have it give a predicted victor to future games. I'm not sure how exactly would be the best way to set this model up, but I'm sure through trial and error of a bunch of different models I could come up with a fairly good predictor.

There are a few different ways of gathering data for a project like this. You could either collect the given team's overall stats from the previous season in several major categories or you could take their stats heading into each game or a third option would be to break it up by players playing in each game. I'd have to think more into each option as to what makes more sense, but those are a few ideas. As far as collecting the data, it would be very easy to collect overall team data from the previous season so maybe instead of doing that I could go through and try and figure out how the team's exact stats before each game which would give a good predictor of that given game. This would take much more time and require a lot more effort to get a large enough sample size. An example dataset might look something like the following:

Team	Games Played	Win %	Average pts/game	F%	Rebounds	Assists	Turnovers	Steals	Won?
Bucks	82	.707	116.9	47.3	48.6	25.8	14.6	6.4	Yes

\*The statline for the opposing team would be incorporated in this as well, either in addition to these stats or maybe I take the difference between the two and have either positive or negative values in each feature. Not sure which makes more sense and will work better. I'd have to think more about it.

There are a lot of data on NBA stats online through various websites such as the NBA's website, ESPN, etc... It might be difficult to find the teams given stats the day of each game so that might require some digging through different websites to piece that together. As far as selecting which features to use, I think I will just pick out the main stats that seem to have the biggest influence as to whether a team wins or not. The NBA tracks a ton of stats so there might be some trial and error as to which ones should be included or not. I think for the most part the data will be real numbers other than maybe the team names and the target.