

Predicting NCAA March Madness 2019 with SpringRank

Brady Metherall
University of Oxford
(Dated: 16 March, 2020)

Each spring the National Collegiate Athletic Association (NCAA) hold a single-elimination tournament across the United States known as March Madness. The tournament features 68 of the top US college men's basketball teams in Division I, with the winner being crowned the national champion. The champions of each of the 32 Division I conferences automatically qualify, and the remaining 36 spots are determined by a selection committee. Since this process is done by a committee the teams given the final spots can be quite controversial. For this reason, the final eight teams selected play for four spots in what is called the First Four, which takes the field from 68 teams down to 64. We use the network ranking algorithm SpringRank [1] to predict the outcome of March Madness 2019 games based on the scores from the regular season.

To construct our network we use the box scores of 5910 NCAA games from 2019 before March Madness began [2]. This includes games played in all three divisions of NCAA men's basketball—a total of 649 teams. From this, we construct a directed graph and the associated adjacency matrix, where each team is represented by a node and each game is represented by an edge directed from the winner to the loser. We then compute the SpringRank from this matrix, and predict the team with the higher SpringRank to win the game.

To compute the accuracy of this prediction we compare to the actual seedings and outcomes from March Madness 2019 [3]. The graph for the top 64 teams by SpringRank is depicted in Figure 1, along with how far through the tournament they advanced. Eight of the top ten SpringRank teams had a top ten seed in the tournament. SpringRank performs well at predicting the teams that advanced to the round of 16, however, the performance decreases further down the ranking. Several teams that did not qualify for the tournament are predicted to be in the top 64 by SpringRank—this is due to the qualification process. Any team that wins their conference automatically qualifies for March Madness, but, the relative strength of each conference varies significantly. Weaker conferences must have a team represented, even though that team may be considerably weaker than a team that did not qualify from a much stronger conference. Since SpringRank only considers the relative strength of each team, many teams that were eliminated in the round of 64 are ranked below teams that were unable to qualify. Additionally, Temple and St. John's, both of whom lost in the First Four, were ranked very highly out of the teams that did not make it to the top 64, as one would expect.

Using SpringRank to predict the winner of each game yields accuracies of 23/32, 15/16, 4/8, 2/4, 1/2,

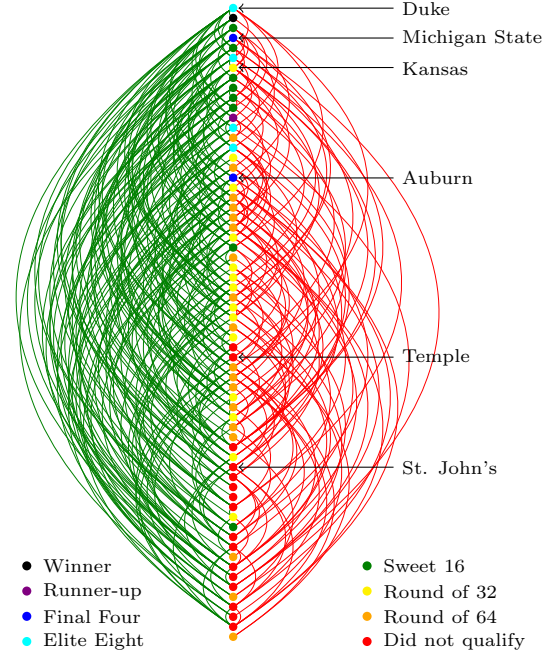


FIG. 1. Top 64 teams from SpringRank and their progress in March Madness 2019. A green (red) edge denotes a game in which the higher (lower) ranked team won.

1/1 for each round, respectively, and a total accuracy of $46/63 \approx 73.0\%$. The only game SpringRank incorrectly predicted in the round of 32 was when Auburn (18) upset Kansas (7) by a score of 89–75; Auburn continued on to the Final Four. Another point of interest is that the number one ranked team in the country, Duke, only made it to the round of 8. Duke (1) faced the very strong Michigan State (4), narrowly losing by a score of 67–68.

SpringRank correctly predicted the winner of 73.0% of the games during March Madness 2019—a large improvement over simply guessing the winner, and a 5% increase from assuming the team with the higher seed would win. Moreover, SpringRank generally predicted how well the top teams would fare, correctly predicting 14 of the teams to make the Sweet 16.

- [1] C. D. Bacco, D. B. Larremore, and C. Moore, Science Advances 4 (2018), 10.1126/sciadv.aar8260.
- [2] K. Wirth, GitHub Repository (Last commit Jan 2020), github.com/kurtawirth/ncaahoopsscraaper.
- [3] D. Vanderkam, GitHub Repository (Last commit May 2019), github.com/danvk/march-madness-data.