Any any sports ranking is going to be contentious. Players, coaches, and teams are often judged subjectively, not so much by how many games they win so much as how inspiring their story is. When we were given the challenge of deciding the greatest college coaches of all time, we knew that we'd have to be as objective as possible. Rather than choose any one way of picking the greatest coaches, we explored several metrics and compared them to existing opinion polls and awards.

1 Common Sense Doesn't Work

One way of ranking coaches that everyone's probably seen is by win/loss percentage. Often this is the easiest way to compare two coaches when all you have are their stat totals. However, when you're looking at the stat totals of every coach in NCAA Division I history, you'll find this metric breaks down quickly. The ratio metric will tell you that a coach with a 15-1 career record is better than one with a 700-200 record. This is plain wrong. Another common metric is wins minus losses. This works well enough when all the coaches play about the same number of games. That, of course, isn't true, and you could end up favoring coaches who have longer careers.

2 Statistics to the Rescue

We solve this problem with the use of a confidence interval. That means, given a coach's wins and losses, we come up with an upper and lower bound on their "innate" winning potential with 95% confidence. In layman's terms, we can say confidently the range where a coach's average win ratio while being more robust against flukes like an oddly good or bad season. The more games a coach has played, the higher we can raise the lower bound on their true potential.

Using this, we took career statistics from every NCAA Division I coach we could find, ran them through some filters ("has the coach ever been in the tournament, or won a bowl game?") to get rid of outliers, and ran our model on the list of coaches we determined to be contenders for "greatest ever". We ended up sorting through about fifty coaches for each of the three sports.

3 Connecting the Dots

One problem with this model is that it doesn't consider the connections between coaches. Surely career stats don't matter when one coach consistently beats another. We explored the interactions between coaches with what is known as a graph model. We looked at what games we could, paired teams to coaches, and compiled statistics about the interactions between pairs of coaches.

For Basketball, we collected score data from every NCAA tournament game, and for Football we collected the results of every postseason bowl game. The original plan was to

collect data on every regular season game ever as well, but that would have violated the data use policies of our sources. We were also unable to find any more data that tournament final scores for Baseball (the data exists but not in a usable format), and only considering two coaches every year would not have been enough data.

With the data that we could get, we did some math relating to the score differences and game importance (a first-round game is not as important as a Final Four game). We used this data to say whether one coach beats another over the course of their games together. With this, we used a variant of Google's Pagerank algorithm (originally designed to rank website results) to rank coaches. We call this system CoachRank.

The idea is like this: I get points from every coach that I beat, and I evenly distribute those points to coaches who beat me. We give every coach the same amount of points to start, and see where they stabilize. This adds transitivity to our model: if I beat another coach and he beats 100 other coaches, then I get a good share of all those points. This also has the effect of eliminating small outliers. A coach with a 15-0 record gets a lower score than one with a 200-30 record simply because the better coach has a better inflow of points.

4   Results

Results shit

5   Commentary

Akilesh's ML shit