

A379_2017AbstractHoblinTim

Abstract

Since 1940 the National Collegiate Athletic Association (NCAA) has held an annual competition pitting the best college basketball teams against each other. This single-elimination tournament has grown to include sixty-eight teams vying to be crowned national champion.

The teams are not the only ones competing for glory during the tournament. From its humble beginnings, the tournament that is aptly nicknamed March Madness has grown to include tens of millions of people betting billions of dollars on who they think will win each of the sixty-three games that make up the tournament. For decades, people ranging from die-hard sports fanatics to people who have never watched a game of basketball have attempted the difficult task of predicting the outcome of this tournament.

Many have debated whether there is a true statistical method for predicting outcomes in this tournament, so we put that to the test. We attempted to predict the 2017 NCAA College Basketball Tournament by applying generalized linear models and random forests, predictive modeling tools widely used in statistics. Based on our various predictive models, we submitted 21 brackets to ESPN's Tournament Challenge and tracked their success against the 18.8 million other entries submitted by the general population. We analyzed our findings based on the overall rankings of our entries on ESPN to determine if our predictive models held a statistical advantage over the population.

Honors College
Ball State University
Muncie, IN 47306