How effectively can a team’s March Madness success be predicted based off the success of historical teams with comparable stats?

Beginning during last year’s tournament and over the last few weeks, I have developed a k-nearest neighbors machine learning algorithm to predict the “tier” of each team in the 2023 tournament based off the success of the k-most statistically comparable teams from tournaments ranging from 2008 to 2022. Tiers are used to predict a bracket by choosing the higher tier team to move to the next round (random if equal). That means not all predicted Final Four teams will be “F4” tier, depending on their competition.

After using cross-validation training on this model, the optimal k-value was 8, meaning that the tier of a team is predicted based on the most common round that it’s 8 statistically closest neighbors reached. The model also uses distance-based weighting to weight closer neighbors more than those farther away.

For 2023, no teams were predicted as “Champions” or “2nd” tier and there was only a single “F4” tier team. Does this mean that this year’s tournament will hold even more madness without a clear top team and nearly 50% of the field as mid-tier, Round of 32 teams?

One of my top rated teams is Purdue. Take a look at the teams this model says are most statistically comparable and their tournament results:

Pittsburgh (2009 E8), North Carolina (2016 2ND), North Carolina (2009 Champions), Kentucky (2019 E8), Pittsburgh (2011 R32), Kentucky (2022 R64), Duke (2009 S16), Baylor (2017 S16)

Will this method of modeling March Madness reveal some insight into the madness?

My elite eight ---

historical March Madness teams with similar stats help predict the success of comparable teams in this year’s tournament?

During last year's tournament and over the last few weeks, I have examined how successful a prediction method based off the success of the k-most comparable teams from previous years can be on predicting the bracket of a given year.

For this project, I use a k-nearest neighbors machine learning algorithm with cross validation training on tournament results from 2008-2022 to compare the this year’s teams to the k-most comparable teams, statistically speaking, from previous years and average the results (Elite 8, Final 4, etc) of the historical teams. This gives a “tier” for the this year’s team that is used to predict their success.

Example: Predicting the success of Duke in the 2022 bracket

k-most similar teams: Kentucky (2017 E8), Texas (2008 E8), Notre Dame (2011 R32), Gonzaga (2018 S16), Oregon (2016 E8), Tennessee (2019 S16), BYU (2011 S16), Duke (2017 R32), Michigan St. (2016 R64)

Tier prediction: Elite 8

Actual: Final 4

Analysis: Although Duke was predicted as an E8 caliber team, the 2022 field had less “high tier” teams so the algorithm would have pushed them all the way to the F4 based on matchups.

How well will it work for 2023? Check out my bracket and the algorithm’s tiers for this year’s tournament field!

Arkansas,E8,S16

Arkansas: Texas Tech (2018 E8), Marquette (2012 S16), Tennessee (2018 R32), Texas A&M (2016 S16), VCU (2014 R64), Connecticut (2014 Champions), Louisville (2011 R64), Tennessee (2010 E8), Purdue (2008 R32)

Duke,E8,F4

Duke: Kentucky (2017 E8), Texas (2008 E8), Notre Dame (2011 R32), Gonzaga (2018 S16), Oregon (2016 E8), Tennessee (2019 S16), BYU (2011 S16), Duke (2017 R32), Michigan St. (2016 R64)

Headline: Excited to announce the completion of our latest project!

Image: (Include a high-quality image or video that showcases the project)

Introduction: I am thrilled to announce the completion of our latest project! We have been working tirelessly over the past few months to bring this project to fruition, and we are proud to share the results with you.

Project Description: Our latest project is a web-based platform that allows users to easily schedule and manage appointments with healthcare professionals. It is designed to streamline the scheduling process and improve the overall patient experience. With features like automated appointment reminders and secure messaging, our platform is changing the way people interact with healthcare providers.

Key Features:

* Simple, user-friendly interface
* Automated appointment reminders
* Secure messaging for patients and providers
* Customizable appointment types and availability
* Advanced reporting and analytics

Call-to-Action: Visit our website to learn more about our platform and how it can benefit your healthcare practice or organization.

Hashtags: #healthcaretechnology #healthtech #appointmentscheduling #patientexperience

We are thrilled to have completed this project and can't wait to see the positive impact it will have on the healthcare industry. Thank you for your continued support!