Teaching Case

Bracketology:
Predicting Winners from Music March Madness

Kevin Mentzer
kmentzer@bryant.edu

Zachary Galante
zachgalante62@gmail.com

Information Systems and Analytics Department
Bryant University
Smithfield, RI 02917, USA

Mark Frydenberg
mfrydenberg@bentley.edu
Computer Information Systems Department
Bentley University
Waltham, MA 02452, USA

Abstract

Organizations are keenly interested in data gathering from websites where discussions of products and brands occur. This increasingly means that programmers need an understanding of how to work with website application programming interfaces (APIs) for data acquisition. In this hands-on lab activity, students will learn how to gather data from several prominent websites using APIs and then build predictive models using that data. Unlike popular challenges on competition sites such as Kaggle where challenges often supply the data, this project emphasizes the data acquisition step of the analytics lifecycle. Working with data from Spotify, YouTube, and Twitter, students will fill out a music based March Madness bracket to predict the winner of the annual Locura De Marzo, a popular middle and high school Spanish competition. By becoming familiar with the data available from each site, through the analysis of the JSON formatted data returned by the APIs, students will be able to explore which features of a song might lend themselves to higher voting by high school students in order to build better prediction models.

Keywords: Bracketology, March Madness, Python, Data Science, Data Analytics, APIs, Hit Song Science

A full and updated manuscript may be found at https://isedj.org