(POSTER) Freshman Engineering Student Engagement through Sport’s Data Analytics

Adam Tennant
atennant@usi.edu

Follow this and additional works at: https://docs.lib.purdue.edu/aseeil-insectionconference

Part of the Engineering Education Commons, and the Sports Studies Commons

Tennant, Adam, "(POSTER) Freshman Engineering Student Engagement through Sport’s Data Analytics" (2019). ASEE IL-IN Section Conference. 12.
https://docs.lib.purdue.edu/aseeil-insectionconference/2019/posters/12

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Freshman Engineering Student Engagement through Sport’s Data Analytics

Adam Tennant, Ph.D., P.E. Assistant Professor of Engineering
University of Southern Indiana Department of Engineering

Introduction

Instructors of engineering for the last quarter of a century have attempted to engage undergraduates through project based learning. Often these topics can fall short due to a lack of interest by the students. Faculty led student research experiences addressing pop-culture topics of interest can often stimulate student engagement and have the possibility to positively impact student retention rates.

This presentation reports on the preliminary findings of a sport’s data analytics approach for faculty-led student investigations and increased engagement. Freshman engineering students will rank teams or individuals through the PageRank algorithm creating a model of the complexity of sports.

Methodology

The specifics of the research will be briefly discussed in this presentation along with a general feel of engagement before and after the project.

This presentation includes two semesters of findings and observations from the fundamentals of engineering course (ENGR 107) that all freshman-engineering students are required to take at the University of Southern Indiana. The course focuses on engineering problem-solving methods with much of the course dedicated to formulating problems for computer solutions using MATLAB or Excel.

Students often come in with little coding experience and often become frustrated and can be verbally heard questioning if they chose the right major while working. Utilizing an interesting topic such as sport’s data analytics is an attempt to engage students in the learning process and motivate them through difficulties.

Spring 2018: In NCAA (National Collegiate Athletic Association) basketball, Division-1 (D1) teams compete in regular season games to be given a chance at playing in the final tournament of the season, the NCAA Men’s Division 1 Basketball Tournament (i.e., March Madness, Big Dance). Each year, thousands of people fill out a bracket in an attempt to predict the outcome of the tournament.

Most individuals simply guess based on a very limited knowledge of the regular season performances of each team within the tournament. With an average of 5832 games per season between the regular season performances of each team within the tournament, it is impossible for any person to watch every game every season.

Instructors Impression: In the spring of 2018 the students were very excited to start the project and continued with enthusiasm throughout the way to completion. They were further challenged to add weights to the links so that all wins were not treated equally. For instance a win during a road game was worth more than a win at home and thus affected the PageRank scores. Additionally the students were assigned to code one improvement to the class project and thus affected the PageRank scores. Additionally the students were assigned to code one improvement to the class project.

During this spring semester, two students were identified to carry on the project in an independent research class. The two along with the instructor performed more of a longitudinal study of multiple seasons and have recently submitted an article to a peer reviewed journal.

Spring 2019: No other sport captures the imagination and is as fundamentally understood around the world as boxing. Although the science of boxing for the combatants is brutally simple, to hit and not get hit, the sport additionally offers a rich source of data for complexity science. Boxing, with its long history, has a depth to its data that can be explored with complexity science to yield insights that analysis through standard statistics has left in the dark. The data that can be explored with complexity science to yield insights that can have detailed directions of steps involved to perform the analytics. This allows them to work at their own speed and to review later for clarification.

The intent again is to motivate the students to learn a little about how Google’s PageRank algorithm is sufficient for two purposes. The first purpose is to provide an unbiased alternative method for deciding which teams will participate in the tournament. The second will be to find a more accurate way to predict outcomes in the NCAA tournament.

Methodology

Software Requirements

Microsoft Excel: Used for preprocessing data set to put in a source target arrangement to later build a directed graph. After the PageRank scores were calculated excel was used to sort the results.

MATLAB: Used for creating the directed graph from the data set and performing the PageRank algorithm.

Gephi: Planned to be used to create high quality images of the directed graph incorporating the PageRank scores.

Data Sets

Boxing: BoxRec is an online database that is the go to source for boxing of the complexity of sports.

Gephi: BoxRec is a JSON database that is used for importing data and creating graphs.

Basketball: cardboard is a platform that contains a collection of data sets from a variety of sports both on the collegiate and professional levels. The data set is built for the sports analytics enthusiast and offers links to download data and tools to analyze. The NCAA Division One basketball data set for each season is publicly made available prior to the start of the tournament. The data set includes results, venues, points scored, and dates; all in a heavily formatted Excel Spreadsheet that is user friendly.

NCAA Men’s Basketball Game Results

Teaching Tools

VoiceThread: Used to create instructional videos so that students can have detailed directions of steps involved to perform the analytics. This allows them to work at their own speed and to review later for clarification.

Blackboard: Instructional and informational videos are posted here along with data sets that cannot be downloaded. The project is also broken down assignments posted on Blackboard including grading rubrics.

Spring 2019: The project introduced to boxing by taking a historical perspective emphasizing the cultural significance of the sport. The students then explored the BoxRec website.

2. The instructor acquainted students with the PageRank Algorithm by the story of Google’s founding and impact on society.

3. The instructor taught the students the idea of a directed graph by a simple example of links on webpages. Then expanded the idea of a directed graph to boxing by a visual.

4. The instructor then gave the students the data set and explored a few entries of the bout history with them. Students gained clear understanding of how the data was not setup for the creation of a directed graph within MATLAB.

5. The instructor then began with the students to preprocess the data in Excel. Exploring the idea of filtering data, cell referencing, data management, and vlookup to create a source and target arrangement .csv for the creation later of a directed graph.

6. As a class the students worked through a small example of the creation of a directed graph in MATLAB.

7. Students then were assigned separate data sets from BoxRec to create directed graphs of boxers.

8. Explored the centrality function within MATLAB and obtained PageRank scores by using the function on the newly created directed graph.

9. Exported data set to Excel and began to process results.

10. Students prepared visualizations and began writing process for a shared journal article documenting the results.

References and Acknowledgements


They would like to acknowledge the support of the BoxRec.com and especially Martin Reechter within that organization.