Data Mining ePortfolios for Engagement Analytics
Finding Your Vocation in STEM: Development of a STEM Learning Community Network at the University of Notre Dame (NSF-DUE 1161222)
Everaldo Aguiar  Nitesh V. Chawla  Jay Brockman  G. Alex Ambrose  Victoria Goodrich  Leo McWilliams
College of Engineering
University of Notre Dame

1 Goals
In general, we may characterize incoming college students' preparation for success in STEM fields along two axes, their interest and their proficiency, giving rise to the four quadrants shown in the figure below.

2 Datasets
Our data analysis process utilized a variety of datasets, each describing a different facet of the students in focus. Below we cover the three main categories of data that were used.

3 ePortfolios
Electronic portfolios consist of a collection of electronic evidence assembled and managed by the students themselves. This set of evidences can be provided in a rich variety of formats (e.g., text, images, multimedia files, blog entries and hyperlinks)

4 Example
The goal of the proposed set of programs is to support incoming students in following a trajectory toward the high interest/high proficiency quadrant.

5 Preliminaries
While we have identified a potential means for tracking STEM students' interest levels and shown how it can be used to predict retention, it remains as future work to map the impact of each proposed program on how students move across the four quadrants described in (1). Lastly, we also intend to analyse the content of ePortfolios as they may provide insights as to what best describes successful STEM students.