Understanding Student Learning in Online Courses with Data Mining

Philip Chan (Computer Sciences) and Joshua Pritchard (Behavior Analysis)

1 A list of project goals and objectives.

Recently, MOOCs (Massive Open Online Courses) have gained significant adoption by major universities [3]. Many students have registered for MOOCs, however completion rates are low [1]. This phenomenon raises many questions in how to understand and measure the effectiveness of online courses [2]. Our goal is to advance understanding of how students learn in online courses and use the insights to foster deeper learning. We will use data mining algorithms to analyze data from the Applied Behavior Analysis online courses (an approximation of MOOCs) in the past 2 years.

Our objectives are to investigate:

1. How is accessing course materials (frequency and duration) correlated with student performance?
2. How are starting and submitting assignments (date and time) correlated with performance?
3. How are discussion activities (frequency and length) correlated with performance?
4. How is the level of discussion (with respect to Bloom’s Taxonomy) correlated with performance?
5. How is student motivation (want to learn, want the certificate/credential, want to start a career in the field, want a salary raise or promotion) for the course correlated with the different activities? Can we use activities to predict motivation?
6. How to use student motivation to identify student subpopulations and answer Questions 1 through 4 within each student subpopulation?
7. How are activities of students with stronger performance different from those of students with weaker performance?
8. How are different activities of a student correlated with different performance on different assessments?
9. Which intervention techniques can be applied to the different student subpopulations to foster deeper learning.
A description of how the project is distinct from the PI's current area(s) of research.

Chan’s research has been in studying data mining (machine learning) algorithms. He and his students have designed data mining algorithms in the areas of health, web personalization, device monitoring, and computer security. Applying data mining algorithms to educational data to understand student learning is new. Pritchard’s research has been the development of online versions of teaching technologies (Interteach and embedded Active Student Responding exercises). This is an interdisciplinary branching of that interest into the use of data mining to identify how/why some of these technologies work. The integration of data mining and behavioral science is new.

A description of the use of information technology in the proposed work.

We will use data mining algorithms to study the relationship among student learning factors. Data mining algorithms can identify combinations of variables that are correlated with the interested target variable.

A description of the type(s) of external funding that could be pursued should the proposed research under the minigrant program prove successful.

Possible sources of funding include the educational research programs at National Science Foundation (such as Research on Education and Learning) and Institute of Education Sciences (such as Postsecondary and Adult Education). Additionally, the MOOC Research Initiative (moocresearch.com) funded by the Gates Foundation is another possibility.

References

