The Grand Challenge: Using Analytics to Predict Student Success

John P. Campbell

2007 EDUCAUSE Midwest Regional Conference
Tuesday, March 13, 2007
2:30-3:15 PM

Copyright John P. Campbell (2007). This work is the intellectual property of the author. Permission is granted for this material to be shared for non-commercial, educational purposes, provided that this copyright statement appears on the reproduced materials and notice is given that the copying is by permission of the author. To disseminate otherwise or to republish requires written permission from the author.

Analytics

- Gartner: "Analytics leverage data in a particular functional process (or application) to enable context-specific insight that is actionable."

Academic Analytics

- ECAR Publication (2005): Academic Analytics: The Use of Management Information and Technology in Higher Education – Phil Goldstein
- Umbrella term – reporting, analysis, and decision support within higher education
Academic Analytics

- Mining data from systems that support teaching and learning to provide customization, tutoring, or intervention within the learning environment
- “actionable intelligence”

Grand Challenge

- What are the potential sources of data?
- How would you make an impact at your institution?
- What would you do if you could harness the information from learning systems to improve learning?

Sources of Data

- Audience response systems (clickers)
- Course management systems
- Electronic grade books
- ePortfolios
- Library
- Media servers (video streaming, podcasting, etc.)
### Potential Applications

- Retention activities
- Customization
  - Learning styles
  - Path through content
- Analytics will change the role of academic IT units on campus

### Critical Aspects of Analytics

- Build upon “normal” faculty and student activities within the course
- “Real time” information
- Provide institutional view
- Scale to large institutions

### Initial Focus: CMS and Academic Success

- Adopted at most campuses
- Wide adoption by faculty members
- Large amount of data tracked automatically
Premise: Academic Success

Initial Project

- Use of analytics to predict which students need help
- Selection of CMS Users
  - 27,276 unique students
  - 597 courses – over 3,000 sections
  - 75 departments and 9 colleges
- One semester of data

Step One: Standardizing the Data

- Student comparisons are made to their peers within the course
- Process allows to “outsider” to make assumptions about the course
Variables

- 20 Course management variables (time, tool usage, etc.)
- 15 Student demographic variables (standardized test scores, high school performance, etc.)
- 10 Course variables

Reducing the Data: Key Factors

- Usage Composite (logins, time, content)
- Assessment composite
- Assignment composite
- Calendar composite
- Communications composite (discussions, email)

Simplified Model

- Standardized test score
- Grade point average
- Usage composite
- Assessment composite
- Assignment composite
- Calendar composite
- Freshman
- Science course
Correct classification – Main Model

- Students needing help: 65.7%
- Students doing well: 87.4%
- Overall: 79.3%

Improving Prediction

![Correct Classification of Students Needing Help](image)

Freshman Model

- Standardized test score
- Grade point average
- Usage composite
- Assessment composite
- Assignment composite
- Calendar composite
- Science course
Correct classification – Freshman Model

- Students needing help: 78.9%
- Students doing well: 86.0%
- Overall: 82.8%

Comparing Models

Correct Classification - Freshmen
Testing the Model

- Models tested with another sample from the CMS
- Results were not significantly different – the model remains valid

Model Effective Despite Limitations

- Course management measures do not necessarily record “active” usage
- Grade data not included
- Limited knowledge of instructor use of the CMS
- Caution: study based on one semester and one institution

Where do we go from here?

- Focus: Correctly classifying the students needing help
- Determining what interventions should be implemented (and when)
**How do we respond?**

- Week 1 - email to student, “my progress” display
- Week 2 - faculty contact with student, reinforcement of available programs
- Week 4 - advisor contact with student

**Ethical Dilemma?**

- What is the student/faculty/institutional obligations to do “anything” based on this information?
- What are the implications of having this data?

**Numerous Policy Decisions**

- How much information is shared (and to whom)
- What information is shared
- What interventions are more appropriate at what times
Collaborations

- Interested in extending the work beyond one institution
- Contact me after this session if you are interested

Questions?

Contact Information

John P. Campbell
Associate Vice President for Teaching and Learning Technologies
Information Technology at Purdue (ITaP)
Purdue University
Young Hall - 5th Floor
302 Wood Street
West Lafayette, IN 47907-2108
(765) 494-1289
jpcampbe@purdue.edu