BREAK OUT SESSION

Learning Analytics & Educational Data Mining

Co-leaders

Aditya Johri, George Mason Univ

Gautam Biswas, Vanderbilt Univ

BIG DATA PI Meeting 2016

Overarching Themes

- How can learning analytics support innovations in education?
  - This includes individual and collaborative learning environments, intelligent tutors, open-ended learning environments, adaptive hypermedia, recommender systems, flipped classrooms, Maker spaces, informal/blended learning, Learning @ scale, visual analytics
  - Learning analytics as a change agent for educational innovation
- Focus on Personalized Education
  - What kind of data to collect to support personalized education?
  - Learn about human cognition
- Infrastructure needs to advance learning analytics and educational data mining
  - Access to data/share datasets
  - Access to new/novel techniques for analysis

Overarching Themes

- What exactly are the Data Sciences?
  - How do we train a new generation of data scientists?
  - Use the vast amounts of data collected by Flickr, Amazon, etc.
- Professional development across all levels
  - Teachers, Researchers
- Undergraduate education
  - Develop analytics and mining techniques to support students through their degree curricula
  - What kind of support do we provide?

Recent Success

- Data Collection/Online Analysis/Assessment Reports for Students and Teachers
  - Assisstments: web based intelligent tutor design by Neil Heffernan at WPI – widely used
- Analyzing the success of MOOCs
  - Large dropout rates
  - Participation of students and instructors
  - MOOCdb – Common interface to edX & Coursera data
- Data Sharing for Analytics
  - MIT Scratch online Community
  - CMU's LEARNLAB – microanalysis of student learning with intelligent tutors
- Algorithmic advances
  - From HMMs to Recurrent neural networks to Deep Learning of student models
  - Mining and Analytics to understand students learning behaviors – cognitive, metacognitive, affect, and self-regulation processes
- Dashboards for student progress
- The very fact that we are even talking about LA/EDM

Major obstacles

- Educational institutions/administrators have no idea of how to use data collected in any significant way
- Access to data/lack of shared datasets
  - How do we integrate data from multiple sources
  - Lack of standardized formats
  - Need instrumented environments to collect data
- Infrastructure to keep up with novel methods/techniques
- Still haven't been able to use collected data to develop comprehensive models of "how people learn"
  - What kind of data do we need to support comprehensive personalized learning?
  - Collaborative learning
- Privacy issues
  - Educational institutions are risk averse when it comes to sharing data
- Need Interdisciplinary/Multidisciplinary approaches
  - Domain experts + data science experts

Areas of Neglect

- Interpretive understanding of context
  - Who defines students success (graduating on time/not going to prison?)
  - How do we define learning?
- Professional development
  - Education researchers still largely do not use analytics/data science
  - Practitioners have little knowledge of how to access analytics to support/improve their learning
- Analytics focused on micro-interactions
  - Lack of macro or cross-level understanding
- Policy decisions to support data sharing while preserving privacy and security

Strategic Priorities to Advance Innovation

- Infrastructure
  - Shared data and repositories
  - Shared tools for analysis
  - Building tool chains that support end to end analysis
- Bridging perspectives
  - Need domain experts and end-users involved from the start
  - May require more cross-disciplinary courses and curricula
- Professional development
- Support to Scale-up innovations