Educational Data Analytics

Credit: 3-0-0

Course Description:

Education is increasingly taking place on online platforms. As a result, these platforms are generating humongous amount of interaction and behavioral data that can be used to design better models for learning, teaching and assessment. A major goal of this course is to teach when and how to use methods for mining patterns from educational data. In particular, the students will learn about the methods that are being developed in learning analytics, students' behavior modeling, personalized learning material recommendation etc. The course will be covered both at the theoretical level as well as the practical level where software tools (such as Vowpal Wabbit for fast learning algorithms and Weka) will be used to analyze the data.

Course Objective:

Upon completion of the course the students will learn the following

Major problem areas in educational data mining
Data mining methods for educational data analytics
How to use methods to answer important educational questions
Application in learning analytics
Use of software tools (such as Vowpal Wabbit, Weka) for analytics
Validity and generalizability of the conclusions drawn from the models

Course Content:

- **Prediction Modeling (4)**: Regressors, Classifiers, and their role in Educational Data Analytics
- **Behavior Detection (5):** Data Synchronization, Feature Engineering, Feature Generation and Feature Selection for behavior detection
- **Latent Knowledge Estimation (7)**: Bayesian Knowledge Tracing, Performance Factors Analysis, Item Response Theory
- **Relationship Mining (4)**: Correlation Mining, Causal Mining, Sequential Pattern Mining, Students' Interaction Network Analysis
- **Visualization (6)**: Educational Visualization and Learning Curves, Moment-by-Moment Learning Graphs, Heat Maps, Parameter Space Maps, State-space Network
- **Structure Discovery (5)**: Applications of Clustering in EDA, Factor Analysis, Knowledge Inference (Q-matrix and Learning Factor Analysis)
- **Model Goodness (5)**: Detector Confidence, Diagnostic Metrics, Cross-validation and Over-fitting
- **Personalized Recommendation (3)**: Topic-based Content Recommendation, Course Recommendation

Books:

1. Handbook of Educational Data Mining, Cristobal Romero, Sebastian Ventura, Mykola Pechenizkiy, Ryan Baker, CRC Press
2. Big Data and Education, R. S. Baker, 2nd Edition. Teachers College, Columbia University, NY
3. Data Mining: Practical Machine Learning Tools and Techniques, Eibe Frank, Ian H. Witten, Morgan Kaufmann

References:

1. Toward integrating feature selection algorithms for classification and clustering, H. Liu and L. Yu, In IEEE Transactions on Knowledge and Data Engineering, 2005.
2. Performance Factors Analysis: A New Alternative to Knowledge Tracing, P.I. Pavlik, H. Cen, K.R. Koedinger, In Proceedings of the International Conference on Artificial Intelligence and Education, 2009.
3. Clustering and Sequential Pattern Mining of Online Collaborative Learning Data, D. Perera et al., In IEEE Transactions on Knowledge and Data Engineering, 2009
4. Predicting robust learning with the visual form of the moment-by-moment learning curve, Baker et al., In Journal of the Learning Sciences, 2013.
5. Discovery with Models: A Case Study on Carelessness in Computer-based Science Inquiry, Hershkovitz et al., American Behavioral Scientist, 2013.
6. Searching for Variables and Models to Investigate Mediators of Learning from Multiple Representations, M. A. Rau, & R. Scheines, In International Conference on Educational Data Mining, 2013.
7. The Q-matrix Method: Mining Student Response Data for Knowledge, T. Barnes, Proceedings of the Workshop on Educational Data Mining at the Annual Meeting of the American Association for Artificial Intelligence, 2005.