Proceedings of the 9th International Conference on Educational Data Mining

T. Barnes, M. Chi, and M. Feng (Eds)
Preface


The EDM conference is the leading international forum for high-quality research that leverages educational data, learning analytics, and machine learning to answer research questions that shed light on the learning processes. Educational data may come from traces that students leave when they interact with learning management systems, interactive learning environments, intelligent tutoring systems, educational games or when they participate in other data-rich learning contexts. The types of data range from raw log files to data captured by eye-tracking devices or other kind of sensors. The methods used by EDM researchers include analytics, data science, data mining, machine learning, as well as social network analysis, graph mining, recommender systems, and model building.

This years conference features three invited talks by: Rakesh Agrawal, President and Founder of Data Insights Laboratories; Marcia C. Linn, Professor of the University of California at Berkeley; and Judy Kay, Professor of the University of Sydney. Judy Kay’s invited paper entitled “Enabling people to harness and control EDM for lifelong, life-wide learning” is also presented in the proceedings. Together with the Journal of Educational Data Mining (JEDM), the EDM 2016 conference supports a JEDM Track that provides researchers a venue to deliver more substantial mature work than is possible in a conference proceedings and to present their work to a live audience. The papers submitted to this track followed the JEDM peer review process; three papers have been accepted to the track and will be presented at the conference. The abstracts of the invited talks and accepted JEDM Track papers can be found in the proceedings. The main conference invited contributions to the Research Track and Industry Track. We received 161 submissions (109 full, 45 short, and 7 industry). We accepted 16 exemplary full papers (15% acceptance rate), 14 full papers (27.5% acceptance rate) and 51 short papers for oral presentation (52% acceptance rate ) and an additional 40 for poster presentation. Of the industry papers, 3 were selected for oral presentations and 2 for posters.

This year’s best paper and best student paper awards were generously sponsored by the Prof. Ram Kumar Memorial Foundation. The best paper was awarded to the paper entitled “How Deep is Knowledge Tracing?” while the best student paper was awarded to the paper entitled “Calibrated Self-Assessment.”

The EDM conference traditionally provides opportunities for young researchers, and particularly for PhD students, to present their research ideas and receive feedback from the peers and more senior researchers. This year, the Doctoral Consortium features 6 such presentations. In addition to the main program, the conference also includes 3 workshops: WS1: Computer-Supported Peer Review in Education (CSPRED-2016), WS2: Writing Analytics, Data Mining, and Writing Studies; WS3: Educational Data Analysis using LearnSphere, and 2 tutorials: T1: SAS Tools for Educational Data Mining, and T2: Massively Scalable EDM with Spark. This year we expand our electronic presence with Whova a social app for conference attendees that provided services for personal scheduling, social linking and personalized recommendations of papers.

We thank the sponsors of EDM 2016 for their generous support: Civitas Learning, Blackboard, MARi, SAS, Cengage Learning and the Prof. Ram Kumar Memorial Foundation. We thank North Carolina State University for their in-kind support, and the Sheraton Downtown Raleigh for their excellent conference services. We also thank the program committee members and reviewers, who with their enthusiastic contributions gave us invaluable support in putting this conference together. Last but not least we thank the organizing team: David Lindrum – Sponsorship Chair; Paul Inventado – Proceedings Chair & Webmaster; Collin Lynch – Posters Chair; Ed Gehringer – Student Volunteer Chair; Sidney D’Mello – DC Chair; Erica Snow and Jonathan Rowe – Workshop & Tutorials Chairs; and Piotr Mitros – Industry Track Chair.

Min Chi
North Carolina State University
Program Co-chair

Mingyu Feng
SRI
Program Co-chair

Tiffany Barnes
North Carolina State University
General Chair
Organization

Conference Chair

Tiffany Barnes  
North Carolina State University, USA

Program Chairs

Min Chi  
North Carolina State University, USA
Mingyu Feng  
SRI International

Workshop and Tutorials Chairs

Jonathan Rowe  
North Carolina State University, USA
Erica Snow  
SRI

Poster Chair and Local Arrangements Chair

Collin Lynch  
North Carolina State University, USA

Doctoral Consortium Chair

Sidney D’Mello  
University of Notre Dame, USA

Student Volunteers Chair

Ed Gehringer  
North Carolina State University, USA

Industry Track Chair

Piotr Mitros  
edX

Sponsors Chair

David Lindrum  
Soomo Learning

Proceedings Chair and Webmaster

Paul Salvador Inventado  
Carnegie Mellon University, USA
### Steering Committee / IEDMS Board of Directors

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ryan Baker</td>
<td>Teachers College, Columbia University, USA</td>
</tr>
<tr>
<td>Tiffany Barnes</td>
<td>University of North Carolina at Charlotte, USA</td>
</tr>
<tr>
<td>Michel Desmarais</td>
<td>Ecole Polytechnique de Montreal, Canada</td>
</tr>
<tr>
<td>Sidney D’Mello</td>
<td>University of Notre Dame, USA</td>
</tr>
<tr>
<td>Neile Heffernan III</td>
<td>Worcester Polytechnic Institute, USA</td>
</tr>
<tr>
<td>Agathe Merceron</td>
<td>Beuth University of Applied Sciences, Germany</td>
</tr>
<tr>
<td>Mykola Pechenizkiy</td>
<td>Eindhoven University of Technology, The Netherlands</td>
</tr>
<tr>
<td>John Stamper</td>
<td>Carnegie Mellon University, USA</td>
</tr>
<tr>
<td>Kalina Yacef</td>
<td>The University of Sydney, Australia</td>
</tr>
</tbody>
</table>

### Program Committee

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lalitha Agnihotri</td>
<td>McGraw Hill Education</td>
</tr>
<tr>
<td>Esma Aimeur</td>
<td>University of Montreal</td>
</tr>
<tr>
<td>Vincent Alevin</td>
<td>Human-Computer Interaction Institute, Carnegie Mellon University</td>
</tr>
<tr>
<td>Ivon Arroyo</td>
<td>Worcester Polytechnic Institute</td>
</tr>
<tr>
<td>Mirjam Augstein</td>
<td>Upper Austria University of Applied Sciences, Communication and Knowledge Media</td>
</tr>
<tr>
<td>Roger Azevedo</td>
<td>North Carolina State University</td>
</tr>
<tr>
<td>Costin Badica</td>
<td>University of Craiova, Software Engineering Department, Romania</td>
</tr>
<tr>
<td>Ryan Baker</td>
<td>Teachers College, Columbia University</td>
</tr>
<tr>
<td>Tiffany Barnes</td>
<td>North Carolina State University</td>
</tr>
<tr>
<td>Joseph Beck</td>
<td>Worcester Polytechnic Institute</td>
</tr>
<tr>
<td>Yoav Bergner</td>
<td>Educational Testing Service</td>
</tr>
<tr>
<td>Gautam Biswas</td>
<td>Vanderbilt University</td>
</tr>
<tr>
<td>Mary Jean Blink</td>
<td>TutorGen, Inc.</td>
</tr>
<tr>
<td>Jesus G. Boticario</td>
<td>UNED</td>
</tr>
<tr>
<td>François Bouchet</td>
<td>LIP6 - Universit Pierre et Marie Curie</td>
</tr>
<tr>
<td>Alex Bowers</td>
<td>Teachers College, Columbia University</td>
</tr>
<tr>
<td>Jared Boyce</td>
<td>SRI International</td>
</tr>
<tr>
<td>Kristy Elizabeth Boyer</td>
<td>University of Florida</td>
</tr>
<tr>
<td>Javier Bravo Agapito</td>
<td>Universidad Autonoma de Madrid</td>
</tr>
<tr>
<td>Keith Brawner</td>
<td>United States Army Research Laboratory</td>
</tr>
<tr>
<td>Emma Brunskill</td>
<td>Carnegie Mellon University</td>
</tr>
<tr>
<td>Renza Campagni</td>
<td>Universit degli Studi di Firenze</td>
</tr>
<tr>
<td>Alberto Cano</td>
<td>Department of Computer Sciences and Numerical Analysis</td>
</tr>
<tr>
<td>Min Chi</td>
<td>North Carolina State University</td>
</tr>
<tr>
<td>Mihaela Cocea</td>
<td>School of Computing, University of Portsmouth</td>
</tr>
<tr>
<td>Scott Crossley</td>
<td>Georgia State University</td>
</tr>
<tr>
<td>Cynthia D’Angelo</td>
<td>SRI International</td>
</tr>
<tr>
<td>Sidney D’Mello</td>
<td>University of Notre Dame</td>
</tr>
<tr>
<td>Michel Desmarais</td>
<td>Ecole Polytechnique de Montreal</td>
</tr>
<tr>
<td>Hendrik Drachsler</td>
<td>Open University of the Netherlands</td>
</tr>
<tr>
<td>Michael Eagle</td>
<td>North Carolina State University</td>
</tr>
<tr>
<td>Stephen Fancsali</td>
<td>Carnegie Learning, Inc.</td>
</tr>
<tr>
<td>Mingyu Feng</td>
<td>SRI International</td>
</tr>
<tr>
<td>Vladimir Fomichov</td>
<td>Faculty of Business Informatics,National Research University Higher School of Economics</td>
</tr>
</tbody>
</table>
Awards

EDM 2016 thanks the Prof. Ram Kumar Memorial Foundation for generously sponsoring the 2016 best paper and best student paper awards.

Best papers and exemplary paper selection

A total of 16 exemplary papers were selected by the program chairs as those that represent the best work submitted to EDM 2016. Candidates for exemplary papers were selected among those accepted as full papers using the following criteria: 1) the average ratings of all reviewers and 2) at least one reviewer indicated the paper should be considered for best paper. Program Chairs and the General Chair then performed meta-reviews for all of these papers to make the final exemplary paper selections.

Finally, the Best Paper Committee was formed to review the top papers in the conference to select best paper nominations. We used random selection to divide both the 16 exemplary papers and the 10 committee members into two groups. All members in the same Best Paper Sub-committee received the same 8 exemplary papers together with their reviews. Sub-committee members were asked to rank the three best papers from the 8 papers, and to provide a 1-2 sentence justification for each of the top 3 they chose. Based on these rankings, four Best Paper nominees were selected.

Best Paper Committee:

Koedinger, Kenneth  Pavlik Jr., Philip I.  Alevin, Vincent  Baker, Ryan  Galyardt, April
Goldin, Ilya  Heffernan, Neil  Ritter, Steven  Olney, Andrew  Pechenizkiy, Mykola

Award winners

<table>
<thead>
<tr>
<th>Category</th>
<th>Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best paper</td>
<td>How Deep is Knowledge Tracing?</td>
</tr>
<tr>
<td></td>
<td>Mohammad Khajah, Robert Lindsey and Michael Mozer</td>
</tr>
<tr>
<td>Best student paper</td>
<td>Calibrated Self-Assessment</td>
</tr>
<tr>
<td></td>
<td>Igor Lobutov and Christoph Studer</td>
</tr>
<tr>
<td>Best paper nominees</td>
<td>LIVELINET: A Multimodal Deep Recurrent Neural Network to Predict Liveliness in Educational Videos</td>
</tr>
<tr>
<td></td>
<td>Arjun Sharma, Arijit Biswas, Ankit Gandhi, Sonal Patil and Om Deshmukh</td>
</tr>
<tr>
<td></td>
<td>How to Model Implicit Knowledge? Similarity Learning Methods to Assess Perceptions of Visual Representations</td>
</tr>
<tr>
<td></td>
<td>Martina Rau, Blake Mason and Robert Nowak</td>
</tr>
<tr>
<td></td>
<td>Measuring Gameplay Affordances of User-Generated Content in an Educational Game</td>
</tr>
<tr>
<td></td>
<td>Andrew Hicks, Zhongxiu Liu and Tiffany Barnes</td>
</tr>
</tbody>
</table>
Table of Contents

Invited Talks (abstracts)

Data-Driven Education: Some opportunities and Challenges ............................................................ 2
  Rakesh Agrawal
WISE Ways to Strengthen Inquiry Science Learning ................................................................. 3
  Marcia Linn
Enabling people to harness and control EDM for lifelong, life-wide learning .......................... 4
  Judy Kay

JEDM Track Journal Papers (abstracts)

Toward Data-Driven Design of Educational Courses: A Feasibility Study .......................... 6
  Rakesh Agrawal, Behzad Golshan and Evangelos Papalexakis
Next-Term Student Performance Prediction: A Recommender Systems Approach .............. 7
  Mack Sweeney, Jaime Lester, Huzefa Rangwala and Aditya Johri
Exploring the Effect of Student Confusion in Massive Open Online Courses ...................... 8
  Diyi Yang, Robert Kraut, and Carolyn Rosé

Invited Paper

Enabling people to harness and control EDM for lifelong, life-wide learning ...................... 10
  Judy Kay

Full Papers

{ENTER}ing the Time Series {SPACE}: Uncovering the Writing Process through Keystroke Analyses .... 22
  Laura Allen, Matthew Jacovina, Mihai Dascalu, Rod Roscoe, Kevin Kent, Aaron Likens and Danielle McNamara
Automatic Gaze-Based Detection of Mind Wandering during Film Viewing ............................ 30
  Caitlin Mills, Robert Bixler, Xingyi Wang and Sidney D’Mello
Riding an emotional roller-coaster: A multimodal study of young child’s math problem solving activities ...... 38
  Lujie Chen, Xin Li, Zhuyun Xia, Zhanmei Song, Louis-Philippe Morency and Artur Dubrawski
Joint Discovery of Skill Prerequisite Graphs and Student Models ........................................... 46
  Yetian Chen, José González-Brenes and Jin Tian
Gauging MOOC Learners’ Adherence to the Designed Learning Path .................................... 54
  Daniel Davis, Guanliang Chen, Claudia Hauff and Geert-Jan Houben
Dynamics of Peer Grading: An Empirical Study ................................................................. 62
  Luca de Alfaro and Michael Shavlovsky
Sequence Matters, But How Exactly? A Method for Evaluating Activity Sequences from Data .......... 70
  Shayan Doroudi, Kenneth Holstein, Vincent Aleven and Emma Brunskill
Measuring Gameplay Affordances of User-Generated Content in an Educational Game ............ 78
  Andrew Hicks, Zhongxru Liu, Michael Eagle and Tiffany Barnes
The Eyes Have It: Gaze-based Detection of Mind Wandering during Learning with an Intelligent Tutoring System
Stephen Hutt, Caitlin Mills, Shelby White, Patrick J. Donnelly and Sidney K. D’Mello

How Deep is Knowledge Tracing?
Mohammad Khajah, Robert Lindsey and Michael Mozer

Temporally Coherent Clustering of Student Data
Severin Klingler, Tanja Käser, Barbara Solenthaler and Markus Gross

Web as a textbook: Curating Targeted Learning Paths through the Heterogeneous Learning Resources on the Web
Igor Labutov and Hod Lipson

Calibrated Self-Assessment
Igor Labutov and Christoph Studer

MOOC Learner Behaviors by Country and Culture; an Exploratory Analysis
Zhongxiu Liu, Rebecca Brown, Collin Lynch, Tiffany Barnes, Ryan Baker, Yoav Bergner and Danielle McNamara

Effect of student ability and question difficulty on duration
Yijun Ma, Lalitha Agnihotri, Ryan Baker and Shirin Mojarad

Modeling the Influence of Format and Depth during Effortful Retrieval Practice
Jaclyn K. Maass and Philip I. Pavlik Jr.

The Apprentice Learner architecture: Closing the loop between learning theory and educational data
Christopher Maclellan, Erik Harpstead, Rony Patel and Kenneth Koedinger

Mining behaviors of students in autograding submission system logs
Jessica McBroom, Bryn Jeffries, Irena Koprinska and Kalina Yacef

Modelling the way: Using action sequence archetypes to differentiate learning pathways from learning outcomes
Kelvin H. R. Ng, Kevin Hartman, Kai Liu and Andy W H Khong

A Coupled User Clustering Algorithm for Web-based Learning Systems
Ke Niu, Zhendong Niu, Xiangyu Zhao, Can Wang, Kai Kang and Min Ye

Execution Traces as a Powerful Data Representation for Intelligent Tutoring Systems for Programming
Benjamin Paaßen, Joris Jensen and Barbara Hammer

Generating Data-driven Hints for Open-ended Programming
Thomas Price, Yihuan Dong and Tiffany Barnes

How to Model Implicit Knowledge? Similarity Learning Methods to Assess Perceptions of Visual Representations
Martina Rau, Blake Mason and Robert Nowak

Student Usage Predicts Treatment Effect Heterogeneity in the Cognitive Tutor Algebra I Program
Adam Sales, Asa Wilks and John Pane

LIVELINET: A Multimodal Deep Recurrent Neural Network to Predict Liveliness in Educational Videos
Arjun Sharma, Arijit Biswas, Ankit Gandhi, Sonal Patil and Om Deshmukh

Semantic Features of Math Problems: Relationships to Student Learning and Engagement
Stefan Slater, Jaclyn Ocumpaugh, Ryan Baker, Peter Scupelli, Paul Salvador Inventado and Neil Heffernan

An Ensemble Method to Predict Student Performance in an Online Math Learning Environment
Martin Stapel, Zhilin Zheng and Niels Pinkwart
Predicting Post-Test Performance from Student Behavior: A High School MOOC Case Study .......................... 239
   Sabina Tomkins, Arti Ramesh and Lise Getoor

The Affective Impact of Tutor Questions: Predicting Frustration and Engagement .............................. 247
   Alexandria Vail, Joseph Wiggins, Joseph Grafsgaard, Kristy Boyer, Eric Wiebe and James Lester

Unnatural Feature Engineering: Evolving Augmented Graph Grammars for Argument Diagrams ............. 255
   Linting Xue, Collin Lynch and Min Chi

Short Papers

Investigating Swarm Intelligence for Performance Prediction ................................................................. 264
   Mohammad Majid Al-Rifaie, Matthew Yee-King and Mark d’Inverno

Predicting Student Progress from Peer-Assessment Data ................................................................. 270
   Michael Mogessie Ashenafi, Marco Ronchetti and Giuseppe Riccardi

Topic-wise Classification of MOOC Discussions: A Visual Analytics Approach ................................... 276
   Thushari Atapattu, Katrina Falkner and Hamid Tarmazdi

Document Segmentation for Labeling with Academic Learning Objectives .......................................... 282
   Divyanshu Bhartiya, Danish Contractor, Sovan Biswas, Bikram Sengupta and Mukes Mohania

Semi-Automatic Detection of Teacher Questions from Human-Transcripts of Audio in Live Classrooms .... 288
   Nathaniel Blanchard, Patrick Donnelly, Andrew Olney, Borhan Samei, Sean Kelly, Xiaoyi Sun, Brooke
   Ward, Martin Nystrand and Sidney D’Mello

Modeling Interactions Across Skills: A Method to Construct and Compare Models Predicting the Existence
   of Skill Relationships ......................................................................................................................... 292
   Anthony F. Botelho, Seth Adjei and Neil Heffernan

Robust Predictive Models on MOOCs : Transferring Knowledge across Courses .................................. 298
   Sebastien Boyer and Kalyan Veeramachaneni

A Comparative Analysis of Techniques for Predicting Student Performance ............................................ 306
   Hana Bydžovská

Course Enrollment Recommender System .............................................................................................. 312
   Hana Bydžovská

Data-driven Automated Induction of Prerequisite Structure Graphs ...................................................... 318
   Devendra Singh Chaplot, Yiming Yang, Jaime Carbonell and Kenneth R. Koedinger

Exploring Learning Management System Interaction Data: Combining Data-driven and Theory-driven
   Approaches ........................................................................................................................................... 324
   Hongkyu Choi, Ji Eun Lee, Won-Joon Hong, Kyumin Lee, Mimi Recker and Andy Walker

A Comparison of Automatic Teaching Strategies for Heterogeneous Student Populations .................... 330
   Benjamin Clement, Pierre-Yves Oudeyer and Manuel Lopes

Automatic Assessment of Constructed Response Data in a Chemistry Tutor ........................................ 336
   Scott Crossley, Kris Kyle, Jodi Davenport and Danielle McNamara

Choosing versus Receiving Feedback: The Impact of Feedback Valence on Learning in an Assessment Game... 341
   Maria Cutumisu and Daniel L. Schwartz

Course Content Analysis: An Initiative Step toward Learning Object Recommendation Systems for MOOC
   Learners ............................................................................................................................................... 347
   Yiling Dai, Yasuhito Asano and Masatoshi Yoshikawa
Exploring the Impact of Data-driven Tutoring Methods on Students’ Demonstrative Knowledge in Logic Problem Solving ............................. 460
   Behrooz Mostafavi and Tiffany Barnes

Properties and Applications of Wrong Answers in Online Educational Systems .................. 466
   Radek Pelánek and Jiří Řihák

Using Inverse Planning for Personalized Feedback .................................................. 472
   Anna Rafferty, Rachel Jansen and Thomas Griffiths

Pattern mining uncovers social prompts of conceptual learning with physical and virtual representations .......... 478
   Martina Rau

Predicting Performance on MOOC Assessments using Multi-Regression Models .................. 484
   Zhiyun Ren, Huzefa Rangwala and Aditya Johri

Validating Game-based Measures of Implicit Science Learning ......................... 490
   Elizabeth Rowe, Jodi Asbell-Clarke, Michael Eagle, Andrew Hicks, Tiffany Barnes, Rebecca Brown and Teon Edwards

Assessing Student-Generated Design Justifications in Virtual Engineering Internships .................. 496
   Vasile Rus, Dipesh Gautam, Zach Swiecki, David Shaffer and Art Graesser

Tensor Factorization for Student Modeling and Performance Prediction in Unstructured Domain ........ 502
   Shaghayegh Sahebi, Yu-Ru Lin and Peter Brusilovsky

Aim Low: Correlation-based Feature Selection for Model-based Reinforcement Learning .......... 507
   Shitian Shen and Min Chi

Personalization of Learning Paths in Online Communities of Creators ...................... 513
   Mingquan Sun and Seungwon Yang

Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models ........ 517
   Mike Tissenbaum, Matthew Berland and Vishesh Kumar

Learning Curves for Problems with Multiple Knowledge Components ....................... 523
   Brett van de Sande

A Nonlinear State Space Model for Identifying At-Risk Students in Open Online Courses .......... 527
   Feng Wang and Li Chen

Transactivity as a Predictor of Future Collaborative Knowledge Integration in Team-Based Learning in
Online Courses .................................................. 533
   Miaomiao Wen, Keith Maki, Xu Wang, Steven Dow, James Herbsleb and Carolyn Rose

Back to the basics: Bayesian extensions of IRT outperform neural networks for proficiency estimation .......... 539
   Kevin Wilson, Yan Karklin, Bojian Han and Chaitanya Ekanadham

Going Deeper with Deep Knowledge Tracing .................................................. 545
   Xiaolu Xiong, Siyuan Zhao, Eric Vaninwegen and Joseph Beck

Boosted Decision Tree for Q-matrix Refinement .................................................. 551
   Peng Xu and Michel Desmarais

Individualizing Bayesian Knowledge Tracing. Are Skill Parameters More Important Than Student Parameters? 556
   Michael Yudelson

Deep Learning + Student Modeling + Clustering: a Recipe for Effective Automatic Short Answer Grading .... 562
   Yuan Zhang, Rajat Shah and Min Chi
Posters and Demo

Redefining “What” in Analyses of Who Does What in MOOCs ................................................................. 569
    Alok Baikadi, Carrie Demmans Epp, Yanjin Long and Christian Schunn

Text Classification of Student Self-Explanations in College Physics Questions ........................................... 571
    Sameer Bhatnagar, Michel Desmarais, Nathaniel Lasry and Elizabeth Charles

Automated Feedback on the Quality of Collaborative Processes: An Experience Report ............................ 573
    Marcela Borge and Carolyn Rosé

Mining Sequences of Gameplay for Embedded Assessment in Collaborative Learning ............................... 575
    Philip Buffum, Megan Frankosky, Kristy Boyer, Eric Wiebe, Bradford Mott and James Lester

Can Word Probabilities from LDA be Simply Added up to Represent Documents? ..................................... 577
    Zhiquiang Cai, Haiying Li, Xiangen Hu and Art Graesser

Examining the necessity of problem diagrams using MOOC AB experiments ............................................ 579
    Zhongzhou Chen, Neset Demirci and David Pritchard

Identifying relevant user behavior and predicting learning and persistence in an ITS-based after-school program 581
    Scotty Craig, Xudong Huang, Jun Xie, Ying Fang and Xiangen Hu

Extracting Measures of Active Learning and Student Self-Regulated Learning Strategies from MOOC Data . . . 583
    Nicholas Diana, Michael Eagle, John Stamper and Kenneth Koedinger

Exploring Social Influence on the Usage of Resources in an Online Learning Community ........................... 585
    Oghenevbo Dibie, Tamara Sumner, Keith Maull and David Quigley

Time Series Cross Section method for monitoring students’ page views of course materials and improving classroom teaching ................................................................. 587
    Konomu Dobashi

Predicting STEM Achievement with Learning Management System Data: Prediction Modeling and a Test of an Early Warning System ................................................................. 589
    Michelle Dominguez, Matthew Bernacki and Merlin Uesbeck

Comparison of Selection Criteria for Multi-Feature Hierarchical Activity Mining in Open Ended Learning Environments ........................................................................................................ 591
    Yi Dong, John S. Kinnebrew and Gautam Biswas

A Data-Driven Framework of Modeling Skill Combinations for Deeper Knowledge Tracing .......................... 593
    Yun Huang, Julio Guerra and Peter Brusilovsky

Generating Semantic Concept Map for MOOCs ......................................................................................... 595
    Zhuoquan Jiang, Peng Li, Yan Zhang and Xiaoming Li

How to Judge Learning on Online Learning: Minimum Learning Judgment System ................................. 597
    Jaechoon Jo and Heuiseok Lim

Guiding Students Towards Frequent High-Utility Paths in an Ill-Defined Domain ..................................... 599
    Igor Jugo, Božidar Kovacić and Vanja Slavuj

Portrait of an Indexer - Computing Pointers Into Instructional Videos ...................................................... 601
    Andrew Lamb, Jose Hernandez, Jeffrey Ullman and Andreas Paepcke

Hierarchical Cluster Analysis Heatmaps and Pattern Analysis: An Approach for Visualizing Learning Management System Interaction Data ............................................................................... 603
    Ji Eun Lee, Mimi Recker, Alex Bowers and Min Yuan
Understanding Engagement in MOOCs
Qiujie Li and Rachel Baker

How quickly can wheel spinning be detected?
Noboru Matsuda, Sanjay Chandrasekaran and John Stamper

Exploring and Following Students’ Strategies When Completing Their Weekly Tasks
Jessica McBroom, Bryn Jeffries, Irena Koprinska and Kalina Yacef

Identifying Student Behaviors Early in the Term for Improving Online Course Performance
Makoto Mori and Philip Chan

Time Series Analysis of VLE Activity Data
Ewa Mlynarska, Pádraig Cunningham and Derek Greene

Massively Scalable EDM with Spark
Tristan Nixon

Study on Automatic Scoring of Descriptive Type Tests using Text Similarity Calculations
Izuru Nogaito, Keiji Yasuda and Hiroaki Kimura

Equity of Learning Opportunities in the Chicago City of Learning Program
David Quigley, Ogheneovo Dibie, Arafat Sultan, Katie Van Horne, William R. Pennebaker, Tamara Sumner, Ugochi Acholonu and Nichole Pinkard

Novel features for capturing cooccurrence behavior in dyadic collaborative problem solving tasks
Vikram Ramanarayanan and Saad Khan

Adding eye-tracking AOI data to models of representation skills does not improve prediction accuracy
Martina Rau and Zach Pardos

MATHia X: The Next Generation Cognitive Tutor
Steven Ritter and Stephen Fancsali

Towards Integrating Human and Automated Tutoring Systems
Steven Ritter, Michael Yudelson, Stephen Fancsali and Susan Berman

Toward Revision-Sensitive Feedback in Automated Writing Evaluation
Rod Roscoe, Matthew Jacovina, Laura Allen, Adam Johnson and Danielle McNamara

Preliminary Results On Dialogue Act and Subact Classification in Chat-based Online Tutorial Dialogues
Vasile Rus, Rajendra Banjade, Nabin Maharjan, Donald Morrison, Steve Ritter and Michael Yudelson

SAS Tools for Educational Data Mining
Jennifer Sabourin, Scott Mcquiggan and André De Waal

Applicability of Educational Data Mining in Afghanistan: Opportunities and Challenges
Abdul Rahman Sherzad

Browsing-Pattern Mining from e-Book Logs with Non-negative Matrix Factorization
Atsushi Shimada, Fumiya Okubo and Hiroaki Ogata

How employment constrains participation in MOOCs?
Mina Shirvani Boroujeni, Lukasz Kidziński and Pierre Dillenbourg

Quantifying How Students Use an Online Learning System: A Focus on Transitions and Performance
Erica Snow, Andrew Krumm, Timothy Podkul, Mingyu Feng and Alex Bowers

A Platform for Integrating and Analyzing Data to Evaluate the Impacts of Educational Technologies
Daniel Stanhope and Karl Rectanus
Educational Technology: What 49 Schools Discovered about Usage when the Data were Uncovered ............ 644

Daniel Stanhope and Karl Rectanus

Learning curves versus problem difficulty: an analysis of the Knowledge Component picture for a given context 646

Brett van de Sande

Validating Automated Triggers and Notifications @ Scale in Blackboard Learn ................................. 648

John Whitmer, Aleksander Dietrichson and Bryan O’Haver

Discovering 'Tough Love' Interventions Despite Dropout ................................................................. 650

Joseph Jay Williams, Anthony Botelho, Adam Sales, Neil Heffernan and Charles Lang

Stimulating collaborative activity in online social learning environments with Markov decision processes ....652

Matthew Yee-King and Mark D’Inverno

Predicting student grades from online, collaborative social learning metrics using K-NN .................... 654

Matthew Yee-King, Andreu Grimalt-Reynes and Mark D’Inverno

Meta-learning for predicting the best vote aggregation method: Case study in collaborative searching of LOs...656

Alfredo Zapata González, Victor Menendez, Cristobal Romero and Manuel E. Prieto Méndez

Soft Clustering of Physics Misconceptions Using a Mixed Membership Model ................................. 658

Guoguo Zheng, Seohyun Kim, Yanyan Tan and April Galyardt

Perfect Scores Indicate Good Students !? The Case of One Hundred Percenters in a Math Learning System .660

Zhilin Zheng, Martin Stapel and Niels Pinkwart

Doctoral Consortium

Towards the Understanding of Gestures and Vocalization Coordination in Teaching Context .............. 663

Roghayeh Barmaki and Charles E. Hughes

Towards Modeling Chunks in a Knowledge Tracing Framework for Students’ Deep Learning ............... 666

Yun Huang and Peter Brusilovsky

Using Case-Based Reasoning to Automatically Generate High-Quality Feedback for Programming Exercises ...669

Angelo Kyrilov

Predicting Off-task Behaviors for Adaptive Vocabulary Learning System ................................. 672

Swangin Nam

Estimation of prerequisite skills model from large scale assessment data using semantic data mining ....675

Bruno Penteado

Designing Interactive and Personalized Concept Mapping Learning Environments ....................... 678

Shang Wang

Industry Track - Short Papers

Analysing and Refining Pilot Training ........................................ 682

Bruno Emond, Scott Buffett, Cyril Goutte and Jaff Guo

A Scalable Learning Analytics Platform for Automated Writing Feedback .................................. 688

Jacqueline Feild, Nicolas Lewkow, Neil Zimmerman, Mark Riedesel and Alfred Essa

An Automated Test of Motor Skills for Job Selection and Feedback ................................... 694

Bhanu Pratap Singh Rawat and Varun Aggarwal

Proceedings of the 9th International Conference on Educational Data Mining xvi
Industry Track - Posters

Studying Assignment Size and Student Performance Using Propensity Score Matching ....................... 701
  Shirin Mojarad

Toward Automated Support for Teacher-Facilitated Formative Feedback on Student Writing ................. 703
  Jennifer Sabourin, Lucy Kosturko, Kristin Hoffmann and Scott Mcquiggan

TutorSpace: Content-centric Platform for Enabling Blended Learning in Developing Countries ........... 705
  Kuldeep Yadav, Kundan Shrivastava, Ranjeet Kumar, Saurabh Srivastava and Om Deshmukh