THE 10th INTERNATIONAL CONFERENCE
ON EDUCATIONAL DATA MINING
WUHAN, CHINA
JUNE 25 - 28, 2017

# Proceedings of the 10th International Conference on Educational Data Mining

X. Hu, T. Barnes, A. Hershkovitz, L. Paquette (Eds)

*International Conference on Educational Data Mining (EDM) 2017*
Proceedings of the 10th International Conference on Educational Data Mining
Xiangen Hu, Tiffany Barnes, Arnon Hershkovitz, Luc Paquette(eds)
Wuhan, China, June 25-28, 2017

# Preface

The 10th International Conference on Educational Data Mining (EDM 2017) is held under the auspices of the International Educational Data Mining Society at the Optics Velley Kingdom Plaza Hotel, Wuhan, Hubei Province, in China. The conference, held June 25th - June 28th, 2017, follows the nine previous editions (Raleigh 2016, Madrid 2015, London 2014, Memphis 2013, Chania 2012, Eindhoven 2011, Pittsburgh 2010, Cordoba, 2009 and Montréal 2008).

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 two invited talks by: Dr. Jie Tang, Associate Professor with the Department of Computer Science and Technology at Tsinghua University; and Dr. Ron Cole, President of Boulder Learning Inc. Together with the Journal of Educational Data Mining (JEDM), the EDM 2017 conference supports a JEDM Track that provides researchers with 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; five papers have been accepted to the track and will be presented at the conference. The abstract for 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 122 submissions (71 full, 47 short, 4 industry). We accepted 18 full papers (25% acceptance rate) and 32 short papers for oral presentation (42% acceptance rate) and an additional 39 for poster presentations, 3 demonstrations. The industry track includes all 4 submitted industry papers and 1 paper initially submitted as a full paper.

The EDM conference provides opportunities for young researchers, and particularly Ph.D. 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 includes 3 workshops: Graph-based Educational Data Mining (G-EDM 2017); Sharing and Reusing Data & Analytics Methods with LearnSphere; Deep Learning with Educational Data, and 2 tutorials: Why Data Standards are Critical for EDM and AIED; and Principal Stratification for EDM Experiments.

We thank the sponsors of EDM 2017 for their generous support: 17Zuoye, Coursera, Learnta, and the Prof. Ram Kumar Memorial Foundation. 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.

| Xiangen Hu | Tiffany Barnes | Arnon Hershkovitz | Luc Paquette |
|---|---|---|---|
| University of Memphis | North Carolina State | University of Tel Aviv | University of Illinois at |
| Central China Normal University | University | Program Chair | Urbana-Champaign |
| Conference Chair | Conference Chair | | Program Chair |

# Organization

## Conference Chairs

| | |
|---|---|
| Xiangen Hu | University of Memphis, USA,<br>Central China Normal University, China |
| Tiffany Barnes | North Carolina State University, USA |

## Program Chairs

| | |
|---|---|
| Arnon Hershkovitz | Tel Aviv University, Israel |
| Luc Paquette | The University of Illinois at Urbana-Champaign, USA |

## Workshop and Tutorials Chairs

| | |
|---|---|
| Ran Liu | Carnegie Mellon University, USA |
| Michael Eagle | Carnegie Mellon University, USA |

## Poster Chairs

| | |
|---|---|
| Mariluz Guenaga | Deusto University, Spain |
| Pedro J. Muñoz-Merino | Universidad Carlos III de Madrid, Spain |
| Rinat Rosenberg-Kima | Tel Aviv University, Israel |

## Industry Track Chairs

| | |
|---|---|
| Giora Alexandron | Center for Educational Technology, Israel |
| Qian Zhou | Tsinghua University, China |

## Doctoral Consortium Chair

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

## JEDM Track Chair

| | |
|---|---|
| Radek Pelanek | Masaryk University, Czech Republic |
| Agathe Merceron | Beuth University of Applied Sciences Berlin, Germany |

## Student Volunteers Chair

| | |
|---|---|
| Ying Fang | University of Memphis, USA |

## Sponsorship Chair

| | |
|---|---|
| Steve Ritter | Carnegie Learning, USA |

## Web Chairs

| | |
|---|---|
| Alexandra Andres | Ateneo de Manila University, Philippines |
| Yun Tang | Central China Normal University, China |

## Local Arrangement Chair

| | |
|---|---|
| Qinglin Cao | Central China Normal University, China |

## Social Media Chair

| | |
|---|---|
| Sharon Hsiao | Arizona State University, USA |

## Proceedings Chair

| | |
|---|---|
| Paul Salvador Inventado | Carnegie Mellon University, USA |

## Steering Committee / IEDMS Board of Directors

| | |
|---|---|
| Mykola Pechenizkiy | Eindhoven University of Technology, Netherlands |
| Mingyu Feng | SRI International, USA |
| Rakesh Agrawal | Data Insights Laboratories, USA |
| Ryan Baker | Teachers College, Columbia University, USA |
| Tiffany Barnes | University of North Carolina at Charlotte, USA |
| Michel Desmarais | Ecole Polytechnique de Montreal, Canada |
| Sidney D'Mello | University of Notre Dame, USA |
| Neil Heffernan III | Worcester Polytechnic Institute, USA |
| Kalina Yacef | University of Sydney, Australia |

## Senior Program Committee

| | |
|---|---|
| Vincent Aleven | Human-Computer Interaction Institute, Carnegie Mellon University, USA |
| Roger Azevedo | North Carolina State University, USA |
| Joseph Beck | Worcester Polytechnic Institute, USA |
| Matthew Berland | University of Wisconsin Madison, USA |
| Sidney D'Mello | University of Notre Dame, USA |
| Dragan Gasevic | University of Edinburgh, UK |
| Art Graesser | University of Memphis, USA |
| Kenneth Koedinger | Carnegie Mellon University, USA |
| Agathe Merceron | Beuth University of Applied Sciences Berlin, Germany |
| Ma. Mercedes T. Rodrigo | Department of Information Systems and Computer Science, Ateneo de Manila University, Philippines |
| Cristobal Romero | Department of Computer Sciences and Numerical Analysis, University of Cordoba, Spain |
| Carolyn Rose | Carnegie Mellon University, USA |
| George Siemens | UT Arlington, USA |
| John Stamper | Carnegie Mellon University, USA |

## Program Committee

| | |
|---|---|
| Lalitha Agnihotri | McGraw Hill Education, USA |
| Esma Aimeur | University of Montreal, Canada |
| Carlos Alario-Hoyos | Universidad Carlos III de Madrid, Spain |
| Giora Alexandron | Massachusetts Institute of Technology, USA |
| Ma. Victoria Almeda | Teachers College Columbia University, USA |
| Aitor Almeida | DeustoTech Deusto Institute of Technology, Spain |
| Ivon Arroyo | Worcester Polytechnic Institute, USA |
| Juan I. Asensio-Prez | Collaborative and Intelligent Systems Group, University of Valladolid, Spain |
| Mirjam Augstein | Upper Austria University of Applied Sciences, Communication and Knowledge Media, Austria |
| Costin Badica | University of Craiova, Software Engineering Department, Romania |
| Gautam Biswas | Vanderbilt University, USA |
| Mary Jean Blink | TutorGen, Inc., USA |
| Nigel Bosch | University of Illinois Urbana-Champaign, USA |
| Miguel L. Bote-Lorenzo | Universidad de Valladolid, Spain |
| Jesus G. Boticario | UNED, Spain |

| | |
|---|---|
| Franois Bouchet | LIP6 - Universit Pierre et Marie Curie, France |
| Alex Bowers | Teachers College, Columbia University, USA |
| Kristy Elizabeth Boyer | University of Florida, USA |
| Javier Bravo Agapito | Universidad Autonoma de Madrid, Spain |
| Keith Brawner | United States Army Research Laboratory, USA |
| Manuel Caeiro Rodrguez | University of Vigo, Spain |
| Renza Campagni | Universit degli Studi di Firenze, Italy |
| Alberto Cano | Virginia Commonwealth University, USA |
| Zhongzhou Chen | MIT physics, USA |
| Min Chi | North Carolina State University, USA |
| Mihaela Cocea | School of Computing, University of Portsmouth, UK |
| Miguel ngel Conde | University of Len, Spain |
| Scott Crossley | Georgia State University, USA |
| Juan Cruz-Benito | GRIAL Research Group. University of Salamanca, Spain |
| Cynthia D'Angelo | SRI International, USA |
| Rosanna De Rosa | University of Naples Federico II, Italy |
| Hendrik Drachsler | Open University, The Netherlands |
| Bruno Emond | National Research Council Canada, Canada |
| Maka Eradze | Tallinn University, Estonia |
| Stephen Fancsali | Carnegie Learning, Inc., USA |
| Vladimir A. Fomichov | School of Business Informatics,National Research University Higher School of Economics, Russia |
| Davide Fossati | Emory University, ISA |
| Kobi Gal | Ben Gurion University, Israel |
| April Galyardt | University of Georgia, USA |
| Carlos Garca-Martnez | Computing and Numerical Analysis Dept. Univ. of Crdoba, Spain |
| Eva Gibaja | Department of Computer Science and Numerical Analysis, University of Cordoba, Spain |
| Daniela Godoy | ISISTAN Research Institute, Argentina |
| Ilya Goldin | 2U, Inc., USA |
| Eduardo Gmez-Snchez | University of Valladolid, Spain |
| Yue Gong | CS department, Worcester Polytechnic Institute, USA |
| Jos Gonzlez-Brenes | Chegg, USA |
| Joseph Grafsgaard | North Carolina State University, USA |
| Mariluz Guenaga | Deusto Institute of Technology - University of Deusto, Spain |
| Philip Guo | UC San Diego, USA |
| Jiangang Hao | Educational Testing Service, USA |
| ngel Hernndez-Garca | Universidad Politcnica de Madrid, Spain |
| Davinia Hernandez-Leo | Universitat Pompeu Fabra, Barcelona, Spain |
| Andrew Hicks | North Carolina State University, USA |
| Sharon Hsiao | Arizona State University, USA |
| Stephen Hutt | University Of Notre Dame, USA |
| Seiji Isotani | University of Sao Paulo, Brazil |
| Vladimir Ivanevi | University of Novi Sad, Faculty of Technical Sciences, Serbia |
| Yang Jiang | Teachers College, Columbia University, USA |
| Jelena Jovanovic | University of Belgrade, Serbia |
| Mike Joy | University of Warwick, UK |
| Ralf Klamma | RWTH Aachen University, Germany |
| Irena Koprinska | The University of Sydney, Australia |
| Sotiris Kotsiantis | Univerisity of Patras, Greece |
| Sbastien Lall | University of British Columbia, Canada |
| Charles Lang | Teachers College, Columbia University, USA |

| | |
|---|---|
| Mikel Larraaga | University of the Basque Country, Spain |
| Sunbok Lee | MIT, USA |
| Young-Jin Lee | University of Kansas, USA |
| James Lester | North Carolina State University, USA |
| Innar Liiv | Tallinn University of Technology, Estonia |
| Ran Liu | Carnegie Mellon University, USA |
| Martn Llamas-Nistal | Universidad de Vigo, Spain |
| Vanda Luengo | Laboratoire d'informatique de Paris, LIP6, Universit Pierre et Marie Curie, France |
| Ivan Lukovi | University of Novi Sad, Faculty of Technical Sciences, Serbia |
| J. M. Luna | Dept. of Computer Science and Numerical Analysis, University of Cordoba, Spain |
| Mihai Lupu | Vienna University of Technology, Austria |
| Maria Luque | University of Cordoba, Spain |
| Collin Lynch | North Carolina State University, USA |
| Christopher Maclellan | Carnegie Mellon University, USA |
| Juan Jose Martins | University of Deusto, Spain |
| Noboru Matsuda | Texas A&M University, USA |
| Victor Menendez | Universidad Autnoma de Yucatn, Mexico |
| Donatella Merlini | Universit di Firenze, Italy |
| Cristian Mihaescu | University of Craiova, Romania |
| Piotr Mitros | edX/MIT, USA |
| Shirin Mojarad | McGraw Hill Education, USA |
| Carlos Monroy | Rice University, USA |
| Behrooz Mostafavi | North Carolina State University, USA |
| Bradford Mott | North Carolina State University, USA |
| Mario Muoz-Organero | Carlos III University of Madrid, Spain |
| Tristan Nixon | University of Memphis, USA |
| Roger Nkambou | Universit du Qubec  Montral (UQAM), Canada |
| Cristian Olivares-Rodrguez | Universidad Andres Bello, Chile |
| Andrew Olney | University of Memphis, USA |
| Shai Olsher | University of Haifa, Israel |
| Abelardo Pardo | The University of Sydney, Australia |
| Zach Pardos | UC Berkeley, USA |
| Philip I. Pavlik Jr. | University of Memphis, USA |
| Radek Pelnek | Masaryk University Brno, Czech Republic |
| Niels Pinkwart | Humboldt-Universitt zu Berlin, Germany |
| Paul Stefan Popescu | University of Craiova, Faculty of Automation, Conputers an Electronics, Craiova, Romania |
| Thomas Price | North Carolina State University, USA |
| David Pritchard | MIT, USA |
| Arti Ramesh | University of Maryland, College Park, USA |
| Martina Rau | University of Wisconsin - Madison, Department of Educational Psychology, USA |
| Steven Ritter | Carnegie Learning, Inc., USA |
| Jos Ral Romero | University of Cordoba, Spain |
| Susana Romero | Universidad de Deusto, Spain |
| Salvador Ros | UNED, Spain |
| Yigal Rosen | Harvard University, USA |
| Adolfo Ruiz Calleja | Tallinn University, Estonia |
| Vasile Rus | The University of Memphis, USA |
| Shaghayegh Sahebi | University at Albany - SUNY, USA |

| | |
|---|---|
| Maria Ofelia San Pedro | Teachers College, Columbia University, USA |
| Olga C. Santos | aDeNu Research Group (UNED), Spain |
| Erica Snow | Arizona State University, USA |
| Angela Stewart | University of Notre Dame, USA |
| Jun-Ming Su | Department of Information and Learning Technology, National University of Tainan, Taiwan |
| Ling Tan | Australian Council for Educational Research, Australia |
| Mike Tissenbaum | Massachusetts Institute of Technology, USA |
| Stefan Trausan-Matu | University Politehnica of Bucharest, Romania |
| Peter Van Rosmalen | Open University, The Netherlands |
| Sebastin Ventura | Department of Computer Sciences and Numerical Analysis, University of Cordoba, Spain |
| Katrien Verbert | KU Leuven, Belgium |
| Lucian Vintan | "Lucian Blaga" University of Sibiu, Romania |
| Feng-Hsu Wang | Ming Chuan University, Taiwan |
| Stephan Weibelzahl | Private University of Applied Sciences Gttingen, Germany |
| Fridolin Wild | Oxford Brookes University, UK |
| Michael Yudelson | Carnegie Mellon University, USA |
| Amelia Zafra Gmez | Department of Computer Sciences and Numerical Analysis, University of Cordoba, Spain |
| Alfredo Zapata Gonzlez | Universidad Autonoma de Yucatan, Mexico |
| Diego Zapata-Rivera | Educational Testing Service, USA |
| Marta Zorrilla | University of Cantabria, Spain |

## Sponsors

Prof. Ram Kumar Memorial Foundation

一起作业

coursera

learnta

CENTRAL CHINA NORMAL UNIVERSITY

# Awards

## Best papers and exemplary paper selection

The two program chairs selected 5 best paper nominees based on the reviews and meta-reviews for each of those paper. The nominees were then sent to the members of the best paper awards committee. Each committee member read and ranked each one of the nominees. Ranking was compiled and the best paper award was attributed to the most highly ranked paper. The best student paper award was attributed to the most highly ranked paper that was also authored by a student. The winner of the best paper award was not eligible to also win the best student paper award.

## Best paper/best student papers committee:

| | | |
|---|---|---|
| Ryan Baker | Michel Desmarais | Zach Pardos |
| Cristobal Romero | Danielle McNamara | Didith Rodrigo |

## Award winners

| | |
|---|---|
| Best paper | Efficient Feature Embeddings for Student Classification with Variational Auto-encoders<br>*Severin Klingler, Rafael Wampfler, Tanja Kser, Barbara Solenthaler and Markus Gross* |
| Best student paper | Generalizability of Face-Based Mind Wandering Detection Across Task Contexts<br>*Angela Stewart, Nigel Bosch and Sidney DMello* |
| Best paper nominees | Zone out no more: Mitigating mind wandering during computerized reading<br>*Sidney D'Mello, Caitlin Mills, Robert Bixler and Nigel Bosch* |
| | Efficient Feature Embeddings for Student Classification with Variational Auto-encoders<br>*Severin Klingler, Rafael Wampfler, Tanja Kser, Barbara Solenthaler and Markus Gross* |
| | Generalizability of Face-Based Mind Wandering Detection Across Task Contexts<br>*Angela Stewart, Nigel Bosch and Sidney D'Mello* |
| | Towards Closing the Loop: Bridging Machine-induced Pedagogical Policies to Learning Theories<br>*Guojing Zhou, Jianxun Wang, Collin Lynch and Min Chi* |
| | The Misidentified Identifiability Problem in Bayesian Knowledge Tracing<br>*Shayan Doroudi and Emma Brunskill* |
| Best Poster Award | Identifying the relationships Between Students' Questions Type and Their Behavior<br>*Fatima Harrak, Franois Bouchet and Vanda Luengo* |

# Table of Contents

**Invited Talks (abstracts)**

Can AI help MOOCs? . . . . . 1
*Jie Tang*

The evolution of virtual tutors, clinician, and companions: A 20-year perspective on conversational agents in real-world applications . . . . . 2
*Ronald Cole*

**JEDM Track Journal Papers (abstracts)**

Identifiability of the Bayesian Knowledge Tracing Model . . . . . 3
*Junchen Feng*

RiPLE: Recommendation in Peer-Learning Environments Based on Knowledge Gaps and Interests . . . . . 4
*Hassan Khosravi, Kendra Cooper and Kirsty Kitto*

Modeling Wheel-spinning and Productive Persistence in Skill Builders . . . . . 5
*Shimin Kai, Ma. Victoria Almeda, Ryan Baker, Nicole Shechtman, Cristina Heffernan and Neil Heffernan*

Modeling MOOC Student Behavior With Two-Layer Hidden Markov Models . . . . . 6
*Chase Geigle and Chengxiang Zhai*

Closing the loop: Automated data-driven cognitive model discoveries lead to improved instruction and learning . . . . . 7
*Ran Liu and Kenneth Koedinger*

**Full Papers**

Zone out no more: Mitigating mind wandering during computerized reading . . . . . 8
*Sidney D'Mello, Caitlin Mills, Robert Bixler and Nigel Bosch*

Measuring Similarity of Educational Items Using Data on Learners' Performance . . . . . 16
*Jiří Řihák and Radek Pelánek*

Adaptive Sequential Recommendation for Discussion Forums on MOOCs using Context Trees . . . . . 24
*Fei Mi and Boi Faltings*

Analysis of problem-solving behavior in open-ended scientific-discovery game challenges . . . 32
*Aaron Bauer, Jeff Flatten and Zoran Popović*

The Antecedents of and Associations with Elective Replay in An Educational Game: Is Replay Worth It? . . . . . 40
*Zhongxiu Liu, Christa Cody, Tiffany Barnes, Collin Lynch and Teomara Rutherford*

Grade Prediction with Temporal Course-wise Influence . . . . . 48
*Zhiyun Ren, Xia Ning and Huzefa Rangwala*

Toward the Automatic Labeling of Course Questions for Ensuring their Alignment with Learning Outcomes ..... 56
*S. Supraja, Kevin Hartman, Sivanagaraja Tatinati and Andy Khong*

Behavior-Based Latent Variable Model for Learner Engagement ..... 64
*Andrew Lan, Christopher Brinton, Tsung-Yen Yang and Mung Chiang*

Efficient Feature Embeddings for Student Classification with Variational Auto-encoders ... 72
*Severin Klingler, Rafael Wampfler, Tanja Käser, Barbara Solenthaler and Markus Gross*

Predicting Short- and Long-Term Vocabulary Learning via Semantic Features of Partial Word Knowledge ..... 80
*Sungjin Nam, Gwen Frishkoff and Kevyn Collins-Thompson*

Generalizability of Face-Based Mind Wandering Detection Across Task Contexts ..... 88
*Angela Stewart, Nigel Bosch and Sidney D'Mello*

Addressing Student Behavior and Affect with Empathy and Growth Mindset ..... 96
*Shamya Karumbaiah, Rafael Lizarralde, Danielle Allessio, Beverly Woolf and Ivon Arroyo*

Epistemic Network Analysis and Topic Modeling for Chat Data from Collaborative Learning Environment ..... 104
*Zhiqiang Cai, Brendan Eagan, Nia Dowell, James Pennebaker, Arthur Graesser and David Shaffer*

Towards Closing the Loop: Bridging Machine-induced Pedagogical Policies to Learning Theories ..... 112
*Guojing Zhou, Jianxun Wang, Collin Lynch and Min Chi*

On the Influence on Learning of Student Compliance with Prompts Fostering Self-Regulated Learning ..... 120
*Sébastien Lallé, Cristina Conati, Roger Azevedo, Michelle Taub and Nicholas Mudrick*

Assessing Computer Literacy of Adults with Low Literacy Skills ..... 128
*Andrew Olney, Dariush Bakhtiari, Daphne Greenberg and Arthur Graesser*

Towards reliable and valid measurement of individualized student parameters ..... 135
*Ran Liu and Kenneth Koedinger*

The Misidentified Identifiability Problem of Bayesian Knowledge Tracing ..... 143
*Shayan Doroudi and Emma Brunskill*

**Short Papers**

An Effective Framework for Automatically Generating and Ranking Topics in MOOC Videos ..... 150
*Jile Zhu, Xiang Li, Zhuo Wang and Ming Zhang*

Grouping Students for Maximizing Learning from Peers ..... 156
*Rakesh Agrawal, Sharad Nandanwar and Narasimha Murty Musti*

Assessing the Dialogic Properties of Classroom Discourse: Proportion Models for Imbalanced Classes ..... 162
*Andrew Olney, Borhan Samei, Patrick Donnelly and Sidney D'Mello*

When and who at risk? Call back at these critical points ..... 168
*Yuntao Li, Chengzhen Fu and Yan Zhang*

Characterizing Collaboration in the Pair Program Tracing and Debugging Eye-Tracking Experiment: A Preliminary Analysis ..... 174
*Maureen Villamor and Ma. Mercedes Rodrigo*

Linking Language to Math Success in a Blended Course ..... 180
*Scott Crossley, Tiffany Barnes, Collin Lynch and Danielle McNamara*

Task and Timing: Separating Procedural and Tactical Knowledge in Student Models ..... 186
*Joshua Cook, Collin Lynch, Andrew Hicks and Behrooz Mostafavi*

Evaluation of a Data-driven Feedback Algorithm for Open-ended Programming ..... 192
*Thomas Price, Rui Zhi and Tiffany Barnes*

Making the Grade: How Learner Engagement Changes After Passing a Course ..... 198
*David Lang, Ben Domingue, Alex Kindel and Andreas Paepcke*

Using a Single Model Trained across Multiple Experiments to Improve the Detection of Treatment Effects ..... 202
*Thanaporn Patikorn, Douglas Selent, Neil Heffernan, Joseph Beck and Jian Zou*

Data-Mining Textual Responses to Uncover Misconception Patterns ..... 208
*Joshua Michalenko, Andrew Lan, Andrew Waters, Phillip Grimaldi and Richard Baraniuk*

Automated Assessment for Scientific Explanations in On-line Science Inquiry ..... 214
*Haiying Li, Janice Gobert and Rachel Dickler*

Can Typical Behaviors Identified in MOOCs be Discovered in Other Courses? ..... 220
*Truong-Sinh An, Christopher Krauss and Agathe Merceron*

Gaze-based Detection of Mind Wandering during Lecture Viewing ..... 226
*Stephen Hutt, Jessica Hardey, Robert Bixler, Angela Stewart, Evan Risko and Sidney D'Mello*

Sequence Modelling For Analysing Student Interaction with Educational Systems ..... 232
*Christian Hansen, Casper Hansen, Niklas Hjuler, Stephen Alstrup and Christina Lioma*

Predicting Prospective Peer Helpers to Provide Just-In-Time Help to Users in Question and Answer Forums ..... 238
*Oluwabukola Ishola and Gordon McCalla*

Combining Machine Learning and Natural Language Processing Approach to Assess Literary Text Comprehension ..... 244
*Renu Balyan, Kathryn McCarthy and Danielle McNamara*

Predicting Student Retention from Behavior in an Online Orientation Course ..... 250
*Shimin Kai, Juan Miguel Andres, Luc Paquette, Ryan Baker, Kati Molnar, Harriet Watkins and Michael Moore*

Inferring Frequently Asked Questions from Student Question Answering Forums ..... 256
*Renuka Sindhgatta, Smit Marvaniya, Tejas Dhamecha and Bikram Sengupta*

On the Prevalence of Multiple-Account Cheating in Massive Open Online Learning ..... 262
*Yingying Bao, Guanliang Chen and Claudia Hauff*

Clustering Student Sequential Trajectories Using Dynamic Time Wrapping ..... 266
*Shitian Shen and Min Chi*

Learner Affect Through the Looking Glass: Characterization and Detection of Confusion in Online Courses ..... 272
*Ziheng Zeng, Snigdha Chaturvedi and Suma Bhat*

Modeling Classifiers for Virtual Internships Without Participant Data ..... 278
*Dipesh Gautam, Zachari Swiecki, David Shaffer, Vasile Rus and Arthur Graesser*

Convolutional Neural Network for Automatic Detection of Sociomoral Reasoning Level ... 284
*Ange Adrienne Nyamen Tato, Roger Nkambou and Aude Dufresne*

A Latent Factor Model For Instructor Content Preference Analysis ..... 290
*Jack Wang, Andrew Lan, Phillip Grimaldi and Richard Baraniuk*

Mining Innovative Augmented Graph Grammars for Argument Diagrams through Novelty Selection ..... 296
*Linting Xue, Collin Lynch and Min Chi*

An Extended Learner Modeling Method to Assess Students' Learning Behaviors ..... 302
*Yi Dong and Gautam Biswas*

Estimating Individual Treatment Effect from Educational Studies with Residual Counterfactual Networks ..... 306
*Siyuan Zhao and Neil Heffernan*

Online Learning Persistence and Academic Achievement ..... 312
*Ying Fang, Benjamin Nye, Philip Pavlik Jr., Yonghong Xu, Arthur Graesser and Xiangen Hu*

Using Temporal Association Rule Mining to Predict Dyadic Rapport in Peer Tutoring .... 318
*Michael Madaio, Rae Lasko, Justine Cassell and Amy Ogan*

Learning to Represent Student Knowledge on Programming Exercises Using Deep Learning ..... 324
*Lisa Wang, Angela Sy, Larry Liu and Chris Piech*

Development of a Trajectory Model for Visualizing Teacher ICT Usage Based on Event Segmentation Data ..... 330
*Longwei Zheng, Rui Shi, Xiaoqing Gu, Bingcong Wu and Yuanyuan Feng*

**Posters**

Modeling Network Dynamics of MOOC Discussion Interactions at Scale ..... 336
*Jingjing Zhang and Maxim Skryabin*

Studying MOOC Completion at Scale Using the MOOC Replication Framework ..... 338
*Juan Miguel Andres, Ryan Baker, George Siemens, Dragan Gašević, Catherine Spann and Scott Crossley*

Clustering Students in ASSISTments: Exploring System- and School-Level Traits to Advance Personalization ..... 340
*Seth Adjei, Korinn Ostrow, Erik Erickson and Neil Heffernan*

Application of the Dynamic Time Warping Distance for the Student Drop-out Prediction on Time Series Data ..... 342
*Alexander Askinadze and Stefan Conrad*

Student Use of Scaffolded Inquiry Simulations in Middle School Science ..... 344
*Elizabeth McBride, Jonathan Vitale and Marcia Linn*

Modeling Dormitory Occupancy Using Markov Chains ..... 346
*David Pokrajac, Kimberley Sudler, Diana Yankovich and Teresa Hardee*

Improving Models of Peer Grading in SPOC ..... 348
*Yong Han, Wenjun Wu and Xuan Zhou*

Personalized Feedback for Open-Response Mathematical Questions using Long Short-Term Memory Networks ..... 350
*Joshua Michalenko, Andrew Lan and Richard Baraniuk*

Intelligent Composition of Test Papers based on MOOC Learning Data ..... 352
*Lin Ma and Yuchun Ma*

Toward Replicable Predictive Model Evaluation in MOOCs ..... 354
*Josh Gardner and Christopher Brooks*

Modeling the Zone of Proximal Development with a Computational Approach ..... 356
*Irene-Angelica Chounta, Bruce Mclaren, Patricia Albacete, Pamela Jordan and Sandra Katz*

A Prediction and Early Alert Model Using Learning Management System Data and Grounded in Learning Science Theory ..... 358
*Wonjoon Hong and Matthew Bernacki*

Cluster Analysis of Real Time Location Data - An Application of Gaussian Mixture Models ..... 360
*Alvaro Ortiz-Vazquez, Xiang Liu, Ching-Fu Lan, Hui Soo Chae and Gary Natriello*

A Topic Model and Social Network Analysis of a School Blogging Platform ..... 362
*Xiaoting Kuang, Hui Soo Chae, Brian Hughes and Gary Natriello*

Supporting the Encouragement of Forum Participation ..... 364
*Aashna Garg and Andreas Paepcke*

Untangling The Program Name Versus The Curriculum: An Investigation of Titles and Curriculum Content ..... 366
*R. Wes Crues*

Emerging Patterns in Student's Learning Attributes through Text Mining ..... 368
*Kejkaew Thanasuan, Warasinee Chaisangmongkon and Chanikarn Wongviriyawong*

A Neural Network Approach to Estimate Student Skill Mastery in Cognitive Diagnostic Assessments ..... 370
*Qi Guo, Maria Cutumisu and Ying Cui*

Automatic Peer Tutor Matching: Data-Driven Methods to Enable New Opportunities for Help . . . 372
*Nicholas Diana, Michael Eagle, John Stamper, Shuchi Grover, Marie Bienkowski and Satabdi Basu*

Short-Answer Responses to STEM Exercises: Measuring Response Validity and Its Impact on Learning . . . 374
*Andrew Waters, Phillip Grimaldi, Andrew Lan and Richard Baraniuk*

Using an Additive Factor Model and Performance Factor Analysis to Assess Learning Gains in a Tutoring System to Help Adults with Reading Difficulties . . . 376
*Genghu Shi, Philip Pavlik Jr. and Arthur Graesser*

Identifying student communities in blended courses . . . 378
*Niki Gitinabard, Collin Lynch, Sarah Heckman and Tiffany Barnes*

Automatic Scoring Method for Descriptive Test Using Recurrent Neural Network . . . 380
*Keiji Yasuda, Izuru Nogaito, Hiroyuki Kawashima, Hiroaki Kimura and Masayuki Hashimoto*

Using Graph-based Modelling to explore changes in students' affective states during exploratory learning tasks . . . 382
*Beate Grawemeyer, Alex Wollenschlaeger, Sergio Gutierrez-Santos, Wayne Holmes, Manolis Mavrikis and Alex Poulovassilis*

Predicting Performance in a Small Private Online Course . . . 384
*Han Wan, Jun Ding, Xiaopeng Gao, Qiaoye Yu and Kangxu Liu*

Social work in the classroom? A tool to evaluate topical relevance in student writing . . . 386
*Heeryung Choi, Zijian Wang, Christopher Brooks, Kevyn Collins-Thompson, Beth Glover Reed and Dale Fitch*

Causal Forest vs. Naive Causal Forest in Detecting Personalization: An Empirical Study in ASSISTments . . . 388
*Biao Yin, Anthony F. Botelho, Thanaporn Patikorn, Neil Heffernan and Jian Zou*

An Offline Evaluation Method for Individual Treatment Rules and How to Find Heterogeneous Treatment Effect . . . 390
*Thanaporn Patikorn, Neil Heffernan and Jian Zou*

MyCOS Intelligent Teaching Assistant . . . 392
*Jiao Guo, Xinhua Huang and Boqing Wang*

Towards Automatic Classification of Learning Objects: Reducing the Number of Used Features . . . 394
*Cristobal Romero, Pedro Gonzalez Espejo, Eva Gibaja, Alfredo Zapata González and Victor Menendez*

The Reading Ability of College Freshmen . . . 396
*Andrew Olney, Breya Walker, Raven Davis and Arthur Graesser*

Discovering skill prerequisite structure through Bayesian estimation and nested model comparison . . . 398
*Soo-Yun Han, Jiyoung Yoon and Yun Joo Yoo*

Text analysis with LIWC and Coh-Metrix: Portraying MOOCs Instructors ..... 400
*Junyi Li, Yun Tang, Lijun Sun and Xiangen Hu*

Identifying relationships between students' questions type and their behavior ..... 402
*Fatima Harrak, François Bouchet and Vanda Luengo*

Metacognitive Prompt Overdose: Positive and Negative Effects of Prompts in iSTART ... 404
*Kathryn McCarthy, Amy Johnson, Aaron Likens, Zachary Martin and Danielle McNamara*

Tracking Online Reading of College Students ..... 406
*Andrew Olney, Eric Hosman, Arthur Graesser and Sidney D'Mello*

Dropout Prediction in MOOCs using Learners' Study Habits Features ..... 408
*Wan Han, Jun Ding, Xiaopeng Gao and David Pritchard*

Exploring the Relationship Between Student Pre-knowledge and Engagement in MOOC Class Using Polytomous IRT ..... 410
*Jingxuan Liu and Hongli Li*

An Analysis of Students' Questions in MOOCs Forums ..... 412
*Meng Cao, Yun Tang and Xiangen Hu*

**Tutorials**

Real-time programming exercise feedback in MOOCs ..... 414
*Zhenghao Chen, Andy Nguyen, Amory Schlender and Jiquan Ngiam*

Why data standards are critical for EDM and AIED ..... 416
*Xiangen Hu, Robby Robson and Avron Barr*

Tutorial: Principal Stratification for EDM Experiments ..... 418
*Adam Sales*

Whitebox: A Device To Assist Group Work Evaluation ..... 420
*Daisuke Yukita*

Understanding Student's Reviewing and Reflection Behaviors Using Web-based Programming Grading Assistant ..... 422
*Yancy Vance Paredes, Po-Kai Huang and Sharon Hsiao*

**Doctoral Consortium**

A Framework for the Estimation of Students' Programming Abilities ..... 424
*Ella Albrecht*

Student Use of Inquiry Simulations in Middle School Science ..... 427
*Elizabeth McBride*

Developing Chinese Automated Essay Scoring Model to Assess College Students' Essay Quality ..... 430
*Yu-Ju Lu, Bor-Chen Kuo and Kai-Chih Pai*

Teaching Informal Logical Fallacy Identification with a Cognitive Tutor ..... 433
*Nicholas Diana, John Stamper and Kenneth Koedinger*

Automated Extraction of Results from Full Text Journal Articles . . . . . 436
*R. Wes Crues*

Intelligent Argument Grading System forStudent-produced Argument Diagrams . . . . . 439
*Linting Xue*

**Industry Track**

Dropout Prediction in Home Care Training . . . . . 442
*Wenjun Zeng, Si-Chi Chin, Brenda Zeimet, Rui Kuang and Chih-Lin Chi*

Few hundred parameters outperform few hundred thousand? . . . . . 448
*Amar Lalwani and Sweety Agrawal*

Tell Me More: Digital Eyes to the Physical World for Early Childhood Learning . . . . . 454
*Vijay Ekambaram, Ruhi Sharma Mittal, Prasenjit Dey, Ravindranath Kokku, Aditya K Sinha and Satya V Nitta*

Student Learning Strategies to Predict Success in an Online Adaptive Mathematics Tutoring System . . . . . 460
*Jun Xie, Shirin Mojarad, Keith Shubeck, Alfred Essa, Ryan Baker and Xiangen Hu*

Adaptive Assessment Experiment in a HarvardX MOOC . . . . . 466
*Ilia Rushkin, Yigal Rosen, Andrew Ang, Colin Fredericks, Dustin Tingley, Mary Jean Blink and Glenn Lopez*

**Workshops**

Graph-based Educational Data Mining . . . . . 472
*Collin Lynch, Tiffany Barnes, Linting Xue and Niki Gitinabard*

Workshop on deep learning with educational data . . . . . 474
*Joseph Beck, Min Chi and Ryan Baker*

Sharing and Reusing Data and Analytic Methods with LearnSphere . . . . . 476
*Ran Liu, Kenneth Koedinger, John Stamper and Philip Pavlik Jr.*