IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 2, Ver. II (Mar – Apr. 2015), PP 43-49
www.iosrjournals.org

A Survey on Research work in Educational Data Mining

D.Fatima[1], Dr.Sameen Fatima[2], Dr.A.V.Krishna Prasad[3]
[1]*Associate Professor, MCA Department, MVSR Engineering College, Osmania University, Hyderabad, INDIA*
[2]*Professor & Dean, Department of Informatics, Osmania University, Hyderabad, INDIA*
[3] *Associate Professor, IT Department, MVSR Engineering College, Osmania University, Hyderabad, INDIA*

Abstract: *Educational Data Mining is an emerging discipline that focuses on applying Data Mining tools and techniques to educationally related data. The discipline focuses on analyzing educational data to develop models for improving learning experiences and institutional effectiveness. A literature review on educational data mining follows, which covers topics such as student retention and attrition, personal recommender systems with in education and how data mining can be used to analyze course management system data. Gaps in the current literature and opportunities for further research are presented.*

Keywords: *Data mining, Educational Data Mining, Student Modelling, Student Retention, Recommendation Systems, Learning Experience etc.*

I. Introduction

EDM is growing at a very fast pace. The main aim of EDM is to develop methods in order to explore the unique type of data that comes from educational institutes and to use those methods to better understand the students and their learning environments. All types of educational data independent of their source have multiple levels of meaningful hierarchy which is determined by properties in the data itself and not in advance. Other issues like time, sequence, and context also plays important roles in the study of educational data.

International Educational Data Mining Society has been formed with an aim to support collaboration and scientific development in this area. To realize its objectives EDM society organizes a series of conferences, bringing out a journal, development of community resources for sharing of data and techniques.

EDM deals with mining of large data sets of educational data to answer educational research questions. These data sets may come from learning management systems, interactive learning environments, intelligent tutoring systems, or any system used in a learning context. The types of data ranges from raw log files to eye – tracking devices and other sensor data. EDM is interdisciplinary research and may require adaptation of existing or development of new approaches that build upon techniques from a combination of areas like statistics, psychometrics, machine learning, information retrieval, recommender systems and scientific computing.

This survey features some of the innovative and fascinating basic and applied research centered on data mining, education and learning technologies. Survey includes diverse set of papers spanning the field of Machine Learning, Artificial Intelligence, Learning Technologies, Education, Linguistics and Psychology. These papers study application of data mining to analyze data generated by various information systems supporting learning or education. They also deal with EDM applications with an actual impact on the future of learning and teaching. Papers are contributed by researchers from computer science, machine learning and data mining, artificial intelligence in education, intelligent tutoring systems, education, learning sciences, psychometrics, statistics and cognitive psychology.

II. Literature Survey

Educational data mining is emerging as a research area with a suite of computational and psychological methods and research approaches for understanding how students learn

2.1 Student Modelling Research:

Student modelling is the major area of research in EDM, work done in student modelling ranges from automatic improvement of student model, unified discovery of student and cognitive model ,impact of individualizing student models on practice opportunities, technique for automated improvement of student model is presented which covers data sets from intelligent tutors to games. The improvements highlights flaws in original model which can lead to new insights into the learning process thereby improving the tutor design. The unified model is called as Dynamic Cognitive Tracing which expresses student learning in terms of skill mastery overtime by simultaneously building the student and cognitive models.

Limits to Accuracy: How well Can we do at Student modelling (predicting Student's next attempt): Here student modelling approach is used to predict whether student's next attempt will be correct. Many student modelling techniques are relatively close to ceiling performance, and there are probably not large gains in accuracy to be had. Knowledge tracing and performance factor analysis has very few differences between them.

Predicting Future Learning Better Using Quantitative Analysis of Moment-by-Moment Learning: Student models have been extended from predicting students future performance on the skills leaned in a tutor to predicting student's preparation for future learning. To predict PFL a combinations of features of student behavior from meta-cognition is used. An alternate method for predicting PFL is proposed which used quantitative aspects of moment by moment learning graph. Learning trajectories are analyzed very deeply.

Discovering Student Models with a Clustering Algorithm Using Problem Content: Student model plays a crucial role in the instructional decisions of ITS. A good student model delivers good instruction on ITS. Traditional ways of making student models are time consuming. Automated methods can be used to make better student models, but requires some engineering effort and are hard to interpret.

Automated Student Model Improvement: Learning factor analysis algorithm is used. Improvements isolate flawed parts in student model. Focused investigation of flawed parts of model leads to new insights into the student learning process and suggests specific improvements of tutor design. Student models are directly improved by using data.

2.2 Improving educational software

Search variables and models to find out what is the mechanism of learning from multiple representations. Multiple representation increase error rate which inhibits learning. Designing multi-representational ITS to help students in reducing errors during practice and learning phase. This finding will benefit both educational psychology literature and ITS. Path Analysis and model search is being used here.

Identifying student learning behaviors especially those that either characterize or distinguish students, can be helpful in the design of adaptation and feedback mechanism in ITS. Differential Sequence Mining technique is used. Differentially frequent activity pattern is identified and interpreted in terms of student relevant learning behaviors.

Extension to the technique is done by contextualizing the sequence mining with information on the student's task performance and learning activities. Piecewise linear segmentation algorithm is used in conjunction with differential sequence mining and action transformation. This methodology is very effective in identifying and interpreting learning behavior patterns at multiple levels of details. Future work deals with more efficient and effective interpretation of learning behavior. Expand and revise the feedback triggering conditions and student modelling to improve learning behavior feedback.

Learner differences in hint processing Adaptation of ITS to differences in how students learn from help. Students may not be able to comprehend and use help of ITS in same way. Such individual differences can be measured by using logistic regression models - ProfHelp and ProfHelp-ID. These models extended the performance factor analysis with parameters that represent the effect of hints on performance on same step on which help was given. Models were implemented using multi-level Bayesian networks. Students differ in individual hint processing proficiency and these differences depend on hint levels.

Student Profiling from Tutoring System Log Data: When Do Multiple Graphical Representations matter: Log data generated by an experiment conducted with Fractions tutor an ITS is analyzed. Comparison of effectiveness of instruction with single and multiple representations is done. Error making and hint seeking behaviors of each student is extracted to characterize their learning strategy. Expectation maximization is used to cluster students by leaning strategy. Educational gains are more from instructions with multiple rather than single representation. This methodology can be implemented in an on-line tutoring system to dynamically tailor individualized instruction.

Investigating the solution space of an open ended educational game using conceptual feature extraction: As there are many different ways of using educational games, the interaction space is large. This large interaction space becomes a challenge for designers as well as researchers who strive to help students in achieving specific learning outcomes. Players are given total freedom to perform a complex game task, which makes it difficult to guess what they will do. To handle these situations designers need to ask some series of questions. In order to answer these questions designers needs methods that give the details of student play. Two dimensional context free grammar is used to automatically extract conceptual features from logs of student play sessions within an open educational game.

2.3 Automated Discovery of Speech act Categories in educational games:

Automated discovery of speech act categories in dialogue based multi-party educational games based on utterance clustering.

Predicting Player Moves in an Educational Game: A Hybrid Approach Modeling and Predicting learner performance in an open ended educational tools to assist the students and to refine the tool is very critical. The range of input in open ended educational tools is also very broad. Building the same type of models which are used to track and predict student behavior in ITS for educational games is very challenging. Classification methods cannot be used here as the range of inputs is very broad at the same time observed data is very sparse.

Sequences of Frustration and Confusion, and Learning: Sensor free affect detection and discovery with models is used to study the relationship between affect which occurs at different durations and learning outcomes among students using online tutors. The study indicates that frustration have stronger effect than confusion, the effect is strongest when both states are taken together. The role of frustration and confusion in online learning is the main topic of this paper. Work to understand and model these affective states in their full complexity will be an essential area of future research.

2.4 Mining assessment data

Optimal and Worst-Case performance of Mastery Learning Assessment with BKT: By implementing mastery learning, ITS aim to present students with exactly the amount of instruction they need to master a concept. Determination of mastery is imperfect. A standard method is to set a threshold for mastery representing a level of certainty that the student has attained mastery. Mastery threshold can be viewed as a parameter that controls the relative frequency of false positives and false negatives. Here a framework has been provided to understand the role of the mastery threshold in BKT. The effects of setting different thresholds under different best and worst case skill modelling assumptions have been studied.

Predicting drop out from social behavior of students: Social behavior data describes social dependencies as described by emails and discussion board's conversation. A new method is suggested to extract features from both student data as well as behavior data which are in the form graph. Novel method is used to learn a classifier for student failure prediction that uses cost sensitive learning to reduce the number of incorrectly classifieds unsuccessful students. Use of social behavior data improves prediction accuracy. DM and SNA methods were used. Structured data obtained by means of linked based data analysis increased the classification accuracy. For future work incorporate faculty data, use more information from social networkw. Building heterogeneous networks and use learning methods like multi-label classification.

2.5 Generic frameworks, Methods and Approaches for EDM

A Spectral learning approach to knowledge tracing: EM was traditionally used in BKT. Here spectral learning is used to learn PSR that represents BKT. A heuristic is then used to extract BKT parameters from PSR using basic matrix operations.

Extending the assistance model: Analyzing the use of assistance over time: There are multiple ways for predicting student performance. Bayesian networks with KT or logistic regression with PFA. Another approach uses raw data which uses Assistance Model which takes into account the number of attempts and hints required to answer previous question correctly. This work is extended by introducing a general framework for predicting student performance with raw data and a new way of predictions within this framework called Assistance Progress model. APM makes predictions on the basis of relationship between the assistance used on previous two problems. The importance of reporting multiple accuracy measures when evaluating student models is also discussed.

2.6 Mining Meaningful Patterns from Students Handwritten Coursework: A key challenge in educational data mining is capturing student work in form suitable for computational analysis. ITS accomplishes this task efficiently. A method to capture student handwriting in digital form is investigated. Data mining techniques are applied to digital copies of handwritten work to understand the cognitive process used by students in an ordinary work environment. Pen stroke data is transformed into a sequence of discrete actions.

InVis: An Interactive Visualization Tool for Exploring Interaction Networks: inVis is a novel visualization technique and tool for exploring, navigating and understanding user interaction data. InVis built an interaction network from student interaction data extracted from large number of students using educational systems and helps instructors to make new insights and discoveries about student learning. This is the first step in creating domain independent visualization tool for understanding student behavior in software tutors and the initial results are promising for the future development of InVis.

2.7 Tag-Aware Ordinal Sparse Factor Analysis for Learning and Content Analytics: Machine learning provides novel ways and means to design personalized learning systems, where each student's educational experiences are customized in real time depending on their background, learning goals, and performance to date. SPARFA is a new framework for machine learning based learning analytics which estimates a learner's knowledge of concepts underlying a domain and content analytics which estimates the relationship between a collection of questions and those concepts. SPARFA jointly learns the associations among the questions and the concepts, learner concept knowledge profiles, and the underlying question difficulties, solely on the basis of correct/incorrect graded responses of a population of students to collection of questions. SPARFA framework is extended to enable it also helps instructors to discover new question-concept associations underlying their learning material.

Assisting instructional assessment of Undergraduate collaborative Wiki and SVN Activities: Assessing the collaborative performance of students who work on shared project. Team Analytics tool is implemented. Document content is processed using machine learning techniques. Summaries of students contribution to coding activities was used to evaluate and coordinate team projects. Future works involves tracing how manager uses the extracted information in team coordination and assisting students. Analyzing errors in NLP to Propositional logic translation using edit distance, so that it facilitates the development of tools and infrastructure to solve problems that these errors represent. The ultimate goal is to produce evidence based pedagogy in this area.

2.8 Emotion, affect, and choice

Sensor free affects detection from students' interaction with a cognitive tutor for algebra. These detectors are developed from students' semantic actions with the interface so that they can be used for driving intervention and labelling log files in the PSLC data shop facilitating future discovery with models analyses at scale. Generalizing detectors is the future work.

2.9 Mining browsing or interaction data

Data Mining in the Classroom: Discovering Groups' Strategies at a Multi-tabletop Environment The data generated when students interact with computer based learning systems can be analyzed to find patterns or train models that help students tutoring systems or teachers to provide better support.

2.10 Comparison of methods to trace multiple sub skills: A long standing challenge to knowledge tracing is how to update estimates of multiple sub skills that underlie a single observable step. Various approaches to this problem are characterized by how they model knowledge tracing, fit its parameters, predict performance and update sub skill estimates. Previous methods allocated blame and credit among sub skills in ways based on relation to observe performance. LR-DBN relaxes this assumption.LR-DBN is very useful in predicting performance there is dramatic improvement when it is jointly used to estimate sub skills. Future work is to use LR-DBN to improve other DBN

2.11 Co Clustering by Bipartite Spectral graph partitioning for Out of tutor prediction: Learning from a distributed representation of input feature space boosts the performance of predictor to achieve this data is portioned into homogenous groups by clustering so that separate model can be trained on each cluster. The drawback is students are clustered but not features. Co Clustering measures the degree of homogeneity in students as well as features thereby achieving clustering and dimensionality reduction simultaneously. Students and features are modelled as bipartite graphs and simultaneous clustering could be shown as bipartite graph portioning problem. Effective bagging strategy is integrated with clustering and is used for prediction of out-of-tutor performance of students. For future work use this technique on co-occurrence table.

III. Summary Of Research Work

Area/Technique Used/DM Task	Problem Statement	Field	Future work
Student Modeling, Performance Factor Analysis, Bayesian Knowledge Tracing Classification	Explore potential reasons behind the inability to create highly accurate models.	Predicting next item correctness	Construct student models that could detect student behaviors like boredom, frustration and discouragement, retention instead of just using it for deterring whether a student is learning or not.
Student Modeling Predictive State Representation, Spectral algorithm, classification	Using spectral learning to learn a Predictive State Representation that represents the BKT HMM. We then use a heuristic to extract the BKT parameters from the learned PSR using basic matrix operations.	Inferring student knowledge	Learning complex latent variable models (Variations of BKT) directly from student performance data.
Student Modeling Bayesian Knowledge Tracing Classification	Understanding the role of the mastery threshold in Bayesian Knowledge.	Mastery Learning Assessment	Beyond considering relatively limited best-case and worst-case scenarios, we should investigate a greater range of average-case possibilities. Future work should also address a broader, more exhaustive range of BKT parameter quadruples.
Student Modelling/Learning Factor Analysis /Discovery with models	Accelerate the process of improving student models.	Improving student models	Applying the idea broadly
Improving educational software/K-Means, Expectation Maximization/Clustering	A fully automated method to speech act discovery.	Speech Act Classification	Use lexico-semantic distance to represent dialogue utterances.
Grouping Students Expectation Maximization clustering approach	Classify students into four strategic profiles based on their Error-rates and hint-seeking behaviors which reveals interesting differences in student learning strategies.	Multiple Graphical Representation	Investigating additional features will better characterize student's behaviors and help in clustering students accurately, construct more informative features from log data.
Natural language Processing Query-likelihood, Clustering	A novel unsupervised Frame work, query-likelihood clustering, for classifying student dialogue acts.	Dialog Act Modeling	Developing research techniques for evaluating unsupervised dialogue act classification. Modeling higher-level dialogue structure and discourse structure.

IV. Conclusion

This paper presents the research work carried out on Educational Data Mining by several research scholars and professional experts. There are a wide variety of applications of EDM discussed in this paper i.e. Improving Student Models, Discovering or improving models of the knowledge structure of the domain, studying the pedagogical support provided by learning software, Scientific discovery about learning and learners. Discovery with models being the key method EDM have lot of scope to the Researchers and software developers. A final recommendation is to create and continue strong collaboration across research, commercial, and educational sectors. Commercial companies operate on fast development cycles and can produce data useful for research.

References

[1]. C. Romero, S. Ventura, Educational data mining: A survey from 1995 to 2005

[2]. Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on management of data, Washington, DC (pp. 207–216).

[3]. Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In Eleventh international conference on data engineering (pp. 3–14). Taipei, Taiwan: IEEE Computer Society Press.

[4]. Arroyo, I., Murray, T., Woolf, B., & Beal, C. (2004). Inferring unobservable learning variables from students' help seeking behaviour. In Intelligent tutoring systems (pp. 782–784).

[5]. Baker, R., Corbett, A., & Koedinger, K. (2004). Detecting student misuse of intelligent tutoring systems. In Intelligent tutoring systems (pp. 531– 540).

[6]. Beck, J., & Woolf, B. (2000). High-level student modelling with machine learning. In Intelligent tutoring systems (pp. 584–593).

[7]. Becker, K., Ghedini, C., & Terra, E. (2000). Using kdd to analyze the impact of curriculum revisions in a Brazilian university. In Eleventh international conference on data engineering. Proceedings of the SPIE 14th annual international conference on aerospace/defense, sensing, simulation and controls, Orlando (pp. 412–419).

[8]. Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. International Journal of Artificial Intelligence in Education, 13, 156–169.

[9]. Chen, G., Liu, C., Ou, K., & Liu, B. (2000). Discovering decision knowledge from web log portfolio for managing classroom processes by applying decision tree and data cube technology. Journal of Educational Computing Research, 23(3), 305–332.

[10]. Chen, J., Li, Q., Wang, L., & Jia, W. (2004). Automatically generating an e-textbook on the web. In International conference on advances in web based learning (pp. 35–42).

[11]. Farzan, R. (2004). Adaptive socio-recommender system for open-corpus e-learning. In Doctoral consortium of the third international conference on adaptive hypermedia and adaptive web-based systems.

[12]. Feng, M., Heffernan, N., & Koedinger, K. (2005). Looking for sources of error in predicting student's knowledge. In Proceedings of AAAI'05 workshop on educational data mining.

[13]. Freyberger, J., Heffernan, N., & Ruiz, C. (2004). Using association rules to guide a search for best fitting transfer models of student learning. In Workshop on analyzing student–tutor interactions logs to improve educational outcomes at ITS conference.

[14]. Grob, H., Bensberg, F., & Kaderali, F. (2004). Controlling open source intermediaries – a web log mining approach. In Proceedings of the 26th international conference on information technology interfaces (pp. 233–242).

[15]. Hanna, M. (2004). Data mining in the e-learning domain. Computers & Education Journal, 42(3), 267–287.

[16]. Heiner, C., Beck, J., & Mostow, J. (2004). Lessons on using its data to answer educational research questions. In Proceedings of the ITS2004 workshop on analyzing student–tutor interaction logs to improve educational outcomes (pp. 1–9).

[17]. Hwang, W., Chang, C., & Chen, G. (2004). The relationship of learning traits, motivation and performance-learning response dynamics. Computers & Education Journal, 42(3), 267–287.

[18]. Iksal, S., & Choquet, C. (2005). Usage analysis driven by models in a pedagogical context. Ingram, A. (1999). Using web server logs in evaluating instructional web sites. Journal of Educational Technology Systems, 28(2), 137–157.

[19]. Johnson, S., Arago, S., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. Journal of Interactive Learning Research, 11(1), 29–49.

[20]. Klosgen, W., & Zytkow, J. (2002). Handbook of data mining and knowledge discovery. New York: Oxford University Press.

[21]. Koutri, M., Avouris, N., & Daskalaki, S. (2004). Ch. A survey on web usage mining techniques for web-based adaptive hypermedia systems.

[22]. Lu, J. (2004). Personalized e-learning material recommender system. In International conference on information technology for application (pp. 374–379).

[23]. Luan, J. (2002). Data mining, knowledge management in higher education, potential applications. In Workshop associate of institutional research international conference, Toronto (pp. 1–18).

[24]. Ma, Y., Liu, B., Wong, C., Yu, P., & Lee, S. (2000). Targeting the right students using data mining. In KDD '00: Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 457–464).

[25]. Markham, S., Ceddia, J., Sheard, J., Burvill, C., Weir, J., Field, B., et al. (2003). Applying agent technology to evaluation tasks in e-learning environments. In Proceedings of the exploring educational technologies conference.

[26]. Mazza, R., & Milani, C. (2005). Exploring usage analysis in learning systems: Gaining insights from visualisations. In Workshop on usage analysis in learning systems at 12th international conference on artificial intelligence in education.

[27]. Merceron, A., & Yacef, K. (2003). A web-based tutoring tool with mining facilities to improve learning and teaching. In Proceedings of 11th international conference on artificial intelligence in education (pp. 201–208).

[28]. Merceron, A., & Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. Journal of Interactive Learning Research, 15(4), 319–346.

[29]. Merceron, A., & Yacef, K. (2005). Tada-ed for educational data mining. Interactive Multimedia Electronic Journal of Computer-Enhanced Learning, 7(1), 267–287.

[30]. Minaei-Bidgoli, B., & Punch, W. (2003). Using genetic algorithms for data mining optimization in an educational web-based system. In GECCO (pp. 2252–2263).

[31]. Mor, E., & Minguillon, J. (2004). E-learning personalization based on itineraries and long-term navigational behaviour. In Proceedings of the 13th international World Wide Web conference (pp. 264–265).

[32]. Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., & Heiner, C. (2005). An educational data mining tool to browse tutor–student interactions: Time will tell! In Proceedings of the workshop on educational data mining (pp. 15–22).

[33]. Nilakant, K., & Mitrovic, A. (2005). Application of data mining in constraint-based intelligent tutoring systems. In Proceedings of the artificial intelligence in education, AIED (pp. 896–898).

[34]. Peled, A., & Rashty, D. (1999). Logging for success: Advancing the use of www logs to improve computer mediated distance learning. Journal of Educational Computing Research, 21(4), 413–431.

[35]. Rahkila, M., & Karjalainen, M. (1999). Evaluation of learning in computer based education using log systems. In ASEE/IEEE frontiers in education conference, San Juan, Puerto Rico (pp. 16–21).

[36]. Romero, C., Ventura, S., & Bra, P. D. (2004). Knowledge discovery with genetic programming for providing feedback to courseware author. User Modeling and User-Adapted Interaction: The Journal of Personalization Research, 14(5), 425–464.

[37]. Sanjeev, P., & Zytkow, J. M. (1995). Discovering enrolment knowledge in university databases. In KDD (pp. 246–251).

[38]. Shen, R., Yang, F., & Han, P. (2002). Data analysis center based on e-learning platform. In Proceedings of the 5th international workshop on the internet challenge: Technology and applications (pp. 19–28).

[39]. Silva, D., & Vieira, M. (2002). Using data warehouse and data mining resources for ongoing assessment in distance learning. In IEEE international conference on advanced learning technologies, Kazan, Russia (pp. 40–45).

[40]. Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. (2000). Web usage mining: Discovery and applications of usage patterns from web data. SIGKDD Explorations, 1(2), 12–23.

[41]. Talavera, L., & Gaudioso, E. (2004). Mining student data to characterize similar behaviour groups in unstructured collaboration spaces. In Workshop on artificial intelligence in CSCL. 16th European conference on artificial intelligence (pp. 17–23).

[42]. Tane, J., Schmitz, C., & Stumme, G. (2004). Semantic resource management for the web: An e-learning application. In Proceedings of the WWW conference, New York, USA (pp. 1–10).

[43]. Tang, C., Yin, H., Li, T., Lau, R., Li, Q., & Kilis, D. (2000). Personalized courseware construction based on web data mining. In Proceedings of the first international conference on web information systems engineering, Washington, DC, USA (pp. 204–211).

[44]. Tang, T., & McCalla, G. (2002). Student modelling for a web-based learning environment: A data mining approach. In Eighteenth national conference on artificial intelligence, Menlo Park, CA, USA (pp. 967– 968).

[45]. Tang, T., & McCalla, G. (2005). Smart recommendation for an evolving e-learning system. International Journal on E-Learning, 4(1), 105– 129.

[46]. Ueno, M. (2004b). Online outlier detection system for learning time data in e-learning and its evaluation. In International conference on computers and advanced technology in education (pp. 248–253).

[47]. Urbancic, T., Skrjanc, M., & Flach, P. (2002). Web-based analysis of data mining and decision support education. AI Communications, 15, 199–204.

[48]. Wang, F. (2002). On using data-mining technology for browsing log file analysis in asynchronous learning environment. In Conference on educational multimedia, hypermedia and telecommunications (pp. 2005–2006).

[49]. Wang, W., Weng, J., Su, J., & Tseng, S. (2004). Learning portfolio analysis and mining in SCORM compliant environment. In ASEE/IEEE frontiers in education conference (pp. 17–24).
[50]. Zaıane, O. (2002). Building a recommender agent for e-learning systems. In ICCE (pp. 55–59).
[51]. Zaıane, O., & Luo, J. (2001). Web usage mining for a better web-based learning environment. In Proceedings of conference on advanced technology for education, Banff, Alberta (pp. 60–64).
[52]. Zaıane, O., Xin, M., & Han, J. (1998). Discovering web access patterns and trends by applying OLAP and data mining technology on web logs. In Advances in digital libraries (pp. 19–29).
[53]. Kenneth R. Koedinger, Elizabeth A. McLaughlin and John C. Stamper, "Automated Student Model Improvement ",In Fifth international conference on Educational Data Mining- 2012, (PP:17-24).
[54]. Vasile Rus, Arthur Graesser, Cristian Moldovan and Nobal Niraula, "Automatic Discovery of Speech Act Categories in Educational Games", In Fifth international conference on Educational Data Mining-2012,(PP: 25-32).
[55]. Shubhendu Trivedi, Zachary Pardos, Gábor Sárközy and Neil Heffernan, "Co-Clustering by Bipartite Spectral Graph Partitioning for Out-of-Tutor Prediction", In Fifth international conference on Educational Data Mining-2012, (PP: 33-40).
[56]. Yanbo Xu and Jack Mostow,"Comparison of methods to trace multiple subskills: Is LR-DBN best? ", In Fifth international conference on Educational Data Mining-2012, (PP: 41-48).
[57]. Jose Gonzalez-Brenes and Jack Mostow, "Dynamic Cognitive Tracing: Towards Unified Discovery of Student and Cognitive Models", In Fifth international conference on Educational Data Mining-2012, (PP: 49-56).
[58]. John Kinnebrew and Gautam Biswas, "Identifying Learning Behaviours by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution", In Fifth international conference on Educational Data Mining-2012, (PP: 57 – 64).
[59]. François Bouchet, John Kinnebrew, Gautam Biswas and Roger Azevedo, " Identifying Student's Characteristic Learning Behaviours in an Intelligent Tutoring System Fostering Self-Regulated Learning", In Fifth international conference on Educational Data Mining-2012, (PP: 65-72).
[60]. Ilya Goldin, Kenneth Koedinger and Vincent Aleven, "Learner Differences in Hint Processing", In Fifth international conference on Educational Data Mining-2012, (PP: 73-80).
[61]. Behzad Beheshti, Michel Desmarais and Rhouma Naceur, "Methods to find the number of latent skills", In Fifth international conference on Educational Data Mining-2012, (PP: 81-86).
[62]. Terry Peckham and Gordon McCalla, "Mining Student Behavior Patterns in Reading Comprehension Tasks", In Fifth international conference on Educational Data Mining-2012, (PP: 87- 94).
[63]. Yoav Bergner, Stefan Droschler, Gerd Kortemeyer, Saif Rayyan, Daniel Seaton and David Pritchard, " Model-Based Collaborative Filtering Analysis of Student Response Data: Machine-Learning Item Response Theory", In Fifth international conference on Educational Data Mining-2012, (PP: 95 -102).
[64]. Tomas Obsivac, Lubos Popelinsky, Jaroslav Bayer, Jan Geryk and Hana Bydzovska, "Predicting drop-out from social behaviour of students", In Fifth international conference on Educational Data Mining-2012, (PP: 103 – 109).
[65]. Martina Rau and Richard Scheines, "Searching for Variables and Models to Investigate Mediators of Learning from Multiple Representations", In Fifth international conference on Educational Data Mining-2012, (PP:110-117).
[66]. Jung In Lee and Emma Brunskill, "The Impact on Individualizing Student Models on Necessary Practice Opportunities ",In Fifth international conference on Educational Data Mining-2012, (PP:118-125).
[67]. Leigh Ann Sudol, Kelly Rivers and Thomas K. Harris, "Calculating Probabilistic Distance to Solution in a Complex Problem Solving Domain", In Fifth international conference on Educational Data Mining-2012, (PP:144-147).
[68]. Manuel Ignacio Lopez, Cristobal Romero, Sebastián Ventura and J.M. Luna, "Classification via clustering for predicting final marks starting from the student participation in Forums", In Fifth international conference on Educational Data Mining-2012, (PP:148-151).
[69]. Ma. Mercedes Rodrigo, Ryan S. J. D. Baker, Bruce McLaren, Alejandra Jayme and Thomas Dy, "Development of a Workbench to Address the Educational Data Mining Bottleneck", In Fifth international conference on Educational Data Mining-2012, (PP: 152-155).
[70]. Jennifer Sabourin, Bradford Mott and James Lester, "Early Prediction of Student Self-Regulation Strategies by Combining Multiple Models", In Fifth international conference on Educational Data Mining-2012, (PP:156-159).
[71]. Judi Mccuaig and Julia Baldwin, "Identifying Successful Learners from Interaction Behaviour", In Fifth international conference on Educational Data Mining-2012, (PP:160-163).
[72]. Michael Eagle, Matthew Johnson and Tiffany Barnes, "Interaction Networks: Generating High Level Hints Based on Network Community Clustering", In Fifth international conference on Educational Data Mining-2012, (PP: 164-167).
[73]. Martina Rau and Zachary Pardos, "Interleaved Practice with Multiple Representations: Analyses with Knowledge Tracing Based Techniques", In Fifth international conference on Educational Data Mining-2012, (PP: 168-171).
[74]. Carol Forsyth, Philip Pavlik Jr, Arthur Graesser, Zhiqiang Cai, Mae-Lynn Germany, Keith Millis, Heather Butler, Diane Halpern and Robert Dolan, " Learning Gains for Core Concepts in a Serious Game on Scientific Reasoning", In Fifth international conference on Educational Data Mining-2012, (PP: 172-175).
[75]. Yutao Wang and Neil Heffernan, "Leveraging First Response Time into the Knowledge Tracing Model", In Fifth international conference on Educational Data Mining-2012, (PP: 176-179).
[76]. Jin Soung Yoo and Moon-Heum Cho, "Mining Concept Maps to Understand University Students' Learning", In Fifth international conference on Educational Data Mining-2012, (PP:184-187).
[77]. Michael Yudelson and Emma Brunskill, "Policy Building – An Extension to User Modeling", In Fifth international conference on Educational Data Mining-2012, (PP: 188- 191).
[78]. Zachary Pardos, Qing Yang Wang and Shubhendu Trivedi, "The real world significance of performance prediction", In Fifth international conference on Educational Data Mining-2012, (PP: 192-195).
[79]. John Stamper, Derek Lomas, Dixie Ching, Steven Ritter, Kenneth Koedinger and Jonathan Steinhart, " The Rise of the Super Experiment", In Fifth international conference on Educational Data Mining-2012, (PP: 196-199).
[80]. Yutao Wang and Joseph Beck, "Using Student Modeling to Estimate Student Knowledge Retention", In Fifth international conference on Educational Data Mining-2012, (PP:200-203).