International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 6, Issue 2, March - April 2017 ISSN 2278-6856

EDM: A field Of Data Analysis To Improve Education

Rishi Kumar Dubey[1], Umesh Kumar Pandey[2]

[1] Kushabhau Thakre Patrakarita Avam Jansanchar Vishwavidyalaya
Sundernagar, Raipur Chhattisgarh India

[2]MATS School of IT, MATS University,
Raipur Chhattishgarh India

Abstract: ***Data mining proved its worth to business world. Due to its capability of complex data study it got prominent place among the researcher of 21st century. After business domain other domains also focused on data mining. Education domain is one of them. Elements of education like student, faculty pedagogy, course content etc. generate huge size and complex data. This attracts researcher to focus on educational data with data mining tools and a new field emerge named educational data mining (EDM). This paper review the research organized in recent years, need of educational data mining and challenges to educational data mining.***

Keywords: Educational Data Mining, paradigm of education, Why EDM, challenges for EDM, EDM methods

1. INTRODUCTION

21st century looks everything digitally. This digital age has technology and gadgets which satisfies and helps in achieving human expectation. Now humans try to control everything using these gadgets and technology.

Gadgets are only medium of collecting data. These data are source of information, knowledge and wisdom. But the question is "How collector will transform this data into other useful form". The answer is hidden in data mining and analytics. Data mining offer various tools of analytics for data which in turn return valuable information.

Increasing accountability toward student created pressure on educational institution to offer most recent updates, information; increasing skill; customizing student on requirement, offering best method of teaching learning process etc. To cope up with pressure researcher applies data mining on educational data sets to understand hidden information inside obtained data.

Power of data analysis is first time realized by business industry. And they got benefit. To receive benefits, realized by business industry, a new research area has been developed in the field of computer who studies education data and known as Educational Data Mining.

In this paper different educational data mining paper reviewed to reveal the journey of EDM and who are the beneficiary and how they can use EDM.

2. EDM IS PARADIGM SHIFT IN EDUCATION

That was the time when novitiate learn everything by observing, Knowledge transferred by heredity. This causes heredity based professions. Thus there were no organized schools. Later a structured architecture of schooling came into existence and it has been realized that novitiate may learn and take expertise as per there interest field. This was a win-win situation in every aspect of society.

After privatization of schooling all income generation responsibility comes on the institution compare to early situation where government findings provided. Now a day's novitiate's learning capacity is not only measured but various methods are used to enhance their learning and skill. Along with curricular activities, other than curricular activities are organized for their experience.

After paradigm shift in education from informal to formal, educational data study transforms process of learning. This data analysis based paradigm shift provides educational institution to take firm decision and offer logical learning path.

This paradigm shift will bring a charismatic change in the field of education when applied with data and tools of data mining.

According to Kay J. [7] learning is lifelong activity and it needs technological support to increase its efficiency. They proposed personal data mining. They argued for evidence based EDM design guidelines and start discussion on building infrastructure for EDM which control self data and associated EDM process.

3. LITERATURE REVIEW OF EDM

Romero C & Venutra [8] suggested that in future EDM will focus on e-learning system, useful for educator and non educator and customizing algorithm for the purpose of educational context.

Baker RSJ and Inventado PS[4] in their chapter discussed and provide a timeline of development of educational data study (see figure 1). They also discussed four major areas i.e. Prediction model; structure discovery; relationship modeling and discovery with models; in which EDM community is engaged and trying to find solution.

Jindal R and Borah M D [5] studied EDM objective, EDM component, EDM Software available with its key features and challenges faced by EDM. According to them educational objective can be divided into two groups academic and administrative. Academic objective fulfills the student and faculty person oriented study; study department development; study any specific domain data to help them. They [5] also describe components of EDM, i.e. Stakeholders: (primary group-Student, Faculties);

(secondary group- Parents and alumni); Hybrid group-(Employers, Administrator/educational planner and experts); EDM Environment: Formal, Informal, Computer Supported environment-Collaborative learning, Adaptive Educational System, Learning Management System, Cognitive learning, recommender system, user modeling; Educational data: Offline data(traditional and modern classroom), Online data(Web logs, Email, Spreadsheet, text data, in detail.

Fatima D. et.al [2] research work focused on different data mining task and used technique for data analysis in the perspective of education and its future work. Paper identifies some field of study in educational data mining i.e. "predicting next item correctness", "inferring student knowledge", "Mastery learning assessment", Improving student model", "Speech Act classification", "Multiple Graphical Representation", "Dialogue act modeling". At the end they recommended "to create and continue strong collaboration across research, commercial and educational sectors.

Figure 1 Development of educational data analytics [4]

Saranya et al[11] organized study using Naïve Bayesian classification method for student (job placement prognosis, academic performance, co curricular activity, extracurricular activity) as well as management (job placement analysis, student intake prognosis) personal detail to provide review for their holistic performance improvement.

Osmanbegovic E, Sulijic M,[12] made comparative study of three classification algorithm i.e. Naïve Bayesian, Multi Layer Perceptron and J48, on different student data attribute (gender, family, distance, highschool, GPA, entrance exam, scholarship, time, Material-book, internet, grade importance, earnings) for the purpose of classification. In this paper they concluded that data mining prediction success is good. In the context of decision tree preparation naïve Bayesian classifier is very good.

Mills C. et. al. [13] research focuses on eye movement during film viewing to study mind wandering for developing next generation educational interface "Mind wandering(MW) reflects a shift in attention from task-related to task unrelated thoughts".

Hutt S. et al. [16] built machine learning model using eye-gaze data collected from ITS GuruTutor. Bayesian network, logistic regression, multilayer perceptrons, random forest and support vector machine classifier method is used for this purpose. Finding of this research shows that "eye gaze can be powerful signal of attention regardless of the learning context".

Chen et. al. [14] developed a "COMMAND (Combine student Modeling and prerequisite Discovery)" algorithm, using expectation maximization and Bayesian network method, that "simultaneously infers a prerequisite graph and a student data model with less human intervention."

Doroudi S et. al. [15] research focused on impact of pedagogical activities on student learning. They introduced SCOVA (Sequencing Constraint Violation Analysis) method for this purpose. Proposed analysis method is useful for researchers and practitioners for "refining hypothesis and determining question" and "improving design of tutor problem".

Ma Y et al. [17] studied relationship between question difficulty level and time taken to predict performance. They concluded that better performer student take much time for difficult question and those student take more time for easy question are poor performer.

Bydžovská H [18] research focuses to "provide recommendation of selective and optional courses with respect to knowledge, interest and free time slots in their timetables". Bydžovská H [18] proposed two algorithms first for semester selection and another for finding path in template.

Bydžovská H [19] research, presents two different approaches for predicting success or failure and final grades of students. The first approach is based on classification and regression algorithms that search for patterns in study-related data and also data about students' social behavior. The second approach is based on collaborative filtering techniques. The research, predict the final grades based on previous achievements of similar students.

Clement B, Oudeyer P, Lopes M [20] research compare two main approaches - Partially Observed Markov Decision Process (POMDP) is a markovian decision process and Zone of Proximal Development and Empirical Success (ZPDES) that can find an optimal long-term path, and Multi-armed bandits that optimize policies locally and greedily but that are computationally more efficient while requiring a simpler learner model for online planning of exact teaching sequences has the potential to offer a sincerely customized coaching enjoy with a huge effect on the motivation and gaining knowledge of students.

Hao J, Liu L, Davier A, Kyllonen P and Kitchen C [21] research presents the aid of educational data mining and statistical analysis. Hao J, Liu L, Davier A, Kyllonen P and Kitchen C [21] investigate the relationship between collaboration outcomes and collaborative problem solving (CPS) skills exhibited during the collaboration process.

Mostafavi B, Barnes T [22] adding data-driven methods into the Deep Thought logic tutor for the purpose of creating a fully data-driven intelligent tutoring system. Mostafavi B, Barnes T [22] investigate how the addition of data-driven methods affects students' demonstrative knowledge of logic proof solving using their post-tutor examination scores. They determine which methods are most beneficial to students who demonstrate higher or lower knowledge of the subject matter.

Klingler S, Käser T, Solenthaler B, Gross M[23] used regression analysis and analyze the influence of different sampling parameters on the performance of the models and study their robustness under different model assumption violations. Original Bayesian Knowledge Tracing (BKT) model have been proposed, among them two novel models that unify BKT and Item Response Theory (IRT). Latent Factor Knowledge Tracing (LFKT) and Feature Aware Student knowledge Tracing (FAST) exhibit state of the art prediction accuracy. Klingler S, Käser T, Solenthaler B, Gross M [23] evaluate and compare properties of the models using synthetic data sets.

Paquette L, Rowe J, Baker R, Mott B, Lester J, DeFalco J, Brawner K, Sottilare R, Georgoulas V [24] describe work using the Generalized Intelligent Framework for Tutoring (GIFT) to build multi-channel affect detection models for a serious game on tactical combat casualty care. Comparing the creation and predictive performance of models developed for two different data modalities: 1) software logs of learner interactions with the serious game, and 2) posture data from a Microsoft Kinect sensor and find that interaction-based detectors outperform posture-based detectors for population.

Sabourin J, Kosturko L, FitzGerald C, McQuiggan S [25] present three primary areas of concern related to student privacy in practice: policy, corporate social responsibility, and public opinion, based on experience as academic researchers transitioning into industry.Sabourin J, Kosturko L, FitzGerald C, McQuiggan S [25] discussion will describe the key challenges faced within these categories, strategies for overcoming them, and ways in which the academic Educational Data Mining (EDM)community can support the adoption of innovative technologies in large-scale production.

Rollinson J, Brunskill E[26] compare the predictive similarity policy to the mastery threshold policy and see if using different student asinput to the predictive similarity policy yields quantitatively different policies.

González-Brenes J P, Huang Y[27] presented the Learner Effort-Outcomes Paradigm (Leopard), a new framework to evaluate adaptive tutoring. González-Brenes J P, Huang Y [27] introducesTheoretical Evaluation of Adaptive Learning Systems(Teal) and Whole Intelligent Tutoring System Evaluation (White), novel automatic metrics that apply Leopard and quantify the amount of effort required to achieve a learning outcome. Teal focuses on models of the Knowledge Tracing Family a very popular set of student models.Knowledge Tracing uses a Hidden Markov Model (HMM) per skill to model the student's knowledge as latent variables.

Olsen J K, Aleven V, Rummel N[28]adjust the Additive Factors Model (AFM), a standard logistic regression model for modeling individual learning, often used in conjunction with knowledge component models and tutor log data. The extended model predicts performance of students solving problems collaboratively with an ITS.

Luo L, Koprinska I, Liu W[29] presented discrimination-aware classification for mining of educational data, with a case study in predicting student exam performance based on enrolment information and assessment marks during the semester, in the context of a computer programming course. Luo L, Koprinska I, Liu W [29] applied discrimination-aware method DAAR, which is based on association rules, and also DADT, a discrimination-aware decision tree method, and compared DAAR and DADT with their non-discrimination-aware alternatives.

Feild J [30] providing students with continuous and personalized feedback on their performance are an important part of encouraging self-regulated learning. As part of higher education platform, Feild Jbuilds a set of data visualizations to provide feedback to students on their assignment performance. These visualizations give students information about how they are doing compared to the rest of the class, and allow them to compare the time they spent on assignments across their courses. Included in the feedback are `nudges' which provide guidance on how students might improve theirperformance by adjusting when they start or submit assignments.

Bravo J, Romero S J, Luna M, Pamplona S [31] evaluated differences according to age in digital competence, usages, and attitude towards ICT. To fulfill this goal, hypothesis testing, correlation analysis, and data mining techniques were performed on the basis of a 72-item survey. Results showed no strong differences between extreme groups of age. Besides, some interesting correlations between variables and additional information through association rules were found.

Pelánek R [32] studied two flexible approaches to skill estimation: time decay functions and the Elo rating system. Results of experiments in several different settings show that these simple approaches provide good and consistent performance.

Chi M, Schwartz D L, Chin D B [33] research mined students' sequential behaviors from an instructional game for color mixing called Lightlet. Students playing the game have two broad strategies. They can either test candidate color combinations in an experiment room without risking an incorrect answer. Or they can choose colors from a faux shopping Catalog containing several different mixing charts. Chi M, Schwartz D L, Chin D B [33] identified the crucial choice pattern(s) in students' game play that would contribute to their learning or subsequent performance.

Käser T, Koedinger K R, Gross M [34] explores a wider set of modeling techniques and by using a data set with additional observations of longer term retention that provide a check on whether judged mastery is maintained in case of non-random factors drive. The data set at hand contains math learning data from children with and without developmental dyscalculia. Käser T, Koedinger K R, Gross M [34] test variations on logistic regression, including the

Additive Factors Model and others explicitly designed to adjust for mastery-based data, as well as Bayesian Knowledge Tracing.

Bazaldua D A L, Baker R S, Pedro M O Z S[35] examined metrics by distilling association rules from real educational data relevant to established research questions in the areas of affect and disengagement. Then ask three domain experts to rate the interestingness of the resultant rules and finally analyze the data to determine which metric(s) best agree with expert judgments of interestingness.

Khajah M M, Wing R M, Lindsey R V, Mozer M C[36] propesed a principled synthesis of the two approaches Latent-factor models and Knowledge-tracing models in a hierarchical Bayesian model that predicts student performance by integrating a theory of the temporal dynamics of learning with a theory of individual differences among students and problems. They find significant predictive value in considering the difficulty of specific problems, a source of information that has rarely been exploited.

Stefanescu D, Rus V, Graesser A C [37] paper describes a study which is part of a project whose goal is to detect students' prior knowledge levels with respect to a target domain based solely on characteristics of the natural language interaction between students and a state-of-the-art conversational Intelligent Tutoring System (ITS). They collected dialogues from two versions of the intelligent tutoring system DeepTutor: a micro-adaptive-only version and a fully-adaptive (micro and macro-adaptive) version.

Kokkodis M, Kannan A,Kenthapadi K[38] research focuses to rigorous formulation of the video assignment problem and presented an algorithm for assigning each video to the optimum subset of logical units. Kokkodis M, Kannan A,Kenthapadi K [38] experimental evaluation using a diverse collection of educational videos relevant to multiple chapters in a textbook demonstrates the efficacy of the proposed techniques for inferring the granularity at which a relevant video should be assigned.

Agnihotri L, Ott A [39]research created a firstyear at risk model using educational data mining and to apply that model at New York Institute of Technology (NYIT). Building the model creates new challenges: (1)the model must be welcomed by counseling staff and the outputs need to be user friendly, and (2)the model needs to work automatically from data collection to processing and prediction in order to eliminate the bottleneck of a human operator which can slow down the process. They chose to build four different initial models: Neural Networks, Naïve Bayes, Decision Tree, and Logistic Regression.

Srimani P K, Patil M M[40] developed a linear regression model for Edu-Mining using the statistical approach. The statistical results obtained help the management to predict the semester results and also helps in proper decision making processes in TES.

Jackson D, Chapman E [41assessed the typical performance levels of Australian business graduates against a comprehensive framework of 20 skills and 45 associated workplace behaviors. Ratings were examined within and across the two samples and variations analyzed by work area, business activity and business discipline.

Potgieter I, Coetzee M [42] determined the relationship between employees' employability attributes and their personality. The authors conducted a quantitative survey. It involved a non-probability sample of 304 early career adults enrolled for an Honor's degree in business management in an open distance learning higher education institution. They used correlational statistics and multiple regression analyses to analyze the data.

Thakar P, Mehta A, Manisha[43]compared various classification algorithms on two datasets of MCA (Masters in Computer Applications) students collected from various affiliated colleges of a reputed state university in India. One dataset includes only primary attributes, whereas other dataset is feeded with secondary psychometric attributes in it. The study analyzes and stresses the role of secondary psychometric attributes for better prediction accuracy and analysis of students' performance. Timely prediction and analysis of students' performance can help Management, Teachers and Students to work on their gray areas for better results and employment opportunities.

Arora R K, Badal D[44] research focuses to identify those set of students that are likely to face difficulty in getting the placements. The analysis using decision tree is being done with the help of WEKA tool.

Arora R K, Badal D [45] described a system that analyze the performance of students using association analysis algorithm.. Arora R K, Badal D [45] research assist the academic planners in identification of students that need more attention such that the extra efforts can be employed on these set of students to improve the results.

Srimani P K, Patil M M [46] mainly discusses on the application of data mining algorithms and techniques on academic data to potentially increase some of the aspects of education system by developing a method called Edu-mining which is a novel approach.

Srimani P K, Patil M M [47] compares the data mining (DM) and Massive Data Mining(MDM) techniques. In the case of MDM, Massive (M) Online (O) Analysis (A) framework is used to generate the data streams and Naive Bayes classifier is used for the analysis purpose.

Pumpuang P, Srivihok A, Praneetpolgrang P[48] research proposes the classifier algorithm for building Course Registration Planning Model (CRPM) from historical dataset. The algorithm is selected by comparing performances of four classifiers include Bayesian Network, C4.5, Decision Forest and NBTree.

Dimokas D, Mittas N, Nanopoulos A, Angelis L [49] presents the design and development of the proposed data warehouse solution, which facilitates better and more thorough analysis of department's data. The proposed system constitutes an integrated platform for a thorough analysis of department's past data. Analysis of data could be achieved with OLAP operations. Also proposes a thorough statistical analysis with an array of data mining techniques that are appropriate for the examined tasks.

Cha S [50] describes distance or similarity measures are essential to solve many pattern recognition problems such as classification, clustering, and retrieval problems. Various distance/similarity measures that are applicable to compare two probability density functions, pdf in short, are reviewed and categorized in both syntactic and semantic relationships. A correlation coefficient and a hierarchical clustering technique are adopted to reveal similarities among numerous distance/similarity measures.

Tsien T B K &Tsui M [51] presents a student centered, collaborative and participative practice teaching model for social work. The model is learning and teaching method undertaken 'with' rather than 'for' students, and its process and outcomes are owned by the team. The relationship between students and teachers develops during a mutual learning process and there is less power disparity between the two parties.

4. WHY EDM

In an education process three groups play a vital role i.e. Academician, Administration and learner. Every groups wish a unlax learning environment with enhance learning skill. Today education sector become industry beside charity work. Still its objective is not to make profit openly. But its working and behavior is almost channelized in private player's educational institution as other industry viz. manufacturing, service sector etc. These sectors got immense growth after applying various data study.

Educational data mining with artificial intelligence, statistics, database management data analytics provide intrinsic knowledge of teaching and learning process so an effective education become true.

Academician utilize EDM in direct participation domain, Administrators utilize EDM for measuring the effort and success of resources provided by the institution for improved learning skill. Last but not least learners also get benefit by knowing their weak and strong point. Knowing oneself is key way to success. Various saying in different language and region of world classify knows and knows not[1] as depicted in following chart:

Table 1: Classification of knows and knows not [1]

Who He	Knows not	Knows
Knows Not	Fool/Shun or avoid him	Asleep/Awake
Knows	Student/Teach him	Wise/Follow him

EDM tries to bring knows-knows stage so that component of learning process i.e. academician, administrator and learner act like wise. Thus overall learning will improve and effective.

5. EDM METHODS

- Prediction: Prediction is the method to find new value from known value. In EDM prediction is used to predict the student performance, building future learning model, role of Meta cognition on learning, predicting motivational factor for learner.
- Classification: Sometimes prediction and classification seems similar but Han & Kamber [9] mentioned that predicting class labels known as classification and predicting values is known as prediction. Classification is used to label class name. Each class represents a specific characteristic. Classification is used for performance analysis and prediction of class. Kaur P, Singh M & Josan G S[10] studied "152 student to predict and analyze student performance also find slow learner among them".
- Regression: A regression evaluation generates an equation to explain the statistical dating among one or more predictors and the reaction variable and watching for new observations.
- Latent Knowledge Estimation: Latent Knowledge Estimation, a student's information of particular skills and concepts is assessed via the usage of their kinds of correctness on those talents.
- Structure Discovery-characterizes the content to be learned and optimal instructional sequences.
- Clustering: Clustering is process for finding clusters of data objects that are similar in some sense to one another.
- Factor Analysis: a process in which the values of observed data are expressed as functions of a number of possible causes in order to find which are the most important.
- Student Characterization- purpose summarization of features in a target group/class
- Relationship mining: Is it possible to pick out and define mechanism to generate beneficial causal inter-relationships across information collections.
- Outlier analysis:-Data objects, that square measure grossly completely different from remaining set of information, square measure referred to as outliers. The outliers are also of explicit interest, like within the case of fraud detection, wherever outliers could indicate deceitful activity. Thus, outlier detection and analysis is a noteworthy data processing task, stated as outlier analysis.
- Association Rule Mining: Association rules area unit created by analyzing information for frequent patterns and mistreatment the standards support and confidence to spot the foremost necessary relationships.
- Pattern Mining: pattern mining could be a data processing technique involved with finding statistically relevant patterns between information.
- Correlation Mining: Indicates the strength and direction of a linear relationship between random variables.
- Causal Data Mining: Casual data processing, makes an attempt to search out whether or not one event was the explanation for another event, either by analyzing the variance of two events
- Discovery with Models: A model of a development is developed via prediction, bunch or data engineering. This model is then used as a element in another analysis, like relationship mining or prediction
- Student modeling: A model provides elaborate info a few student's characteristics or states, like information motivation and perspective.

- Effect of pedagogical support/educational support: Pedagogical coaching had a bearing on scales measure abstract change/student-focused approach and self-efficacy beliefs.
- User modeling : The important goal of consumer modeling is customization and edition of structures to the person's precise needs
- User knowledge modeling: Knowledge modeling is a process of creating a computer interpretable model of knowledge.
- User behavior modeling: Presence is a service that allows a user to be informed about the reachability, availability, and willingness of communication of another user.
- User profiling-performance identification, Tutoring system log: User identification could be a type of dynamic user behavior analysis.
- Domain modeling: a domain model may be a abstract model of the domain that includes each behavior and information
- Trend analysis: Trend analysis is that the rampant apply of aggregation data and making an attempt to identify a pattern, or trend, within the data
- Deviation analysis- purpose is to search out the explanation of the deviations from the anticipated price.

6. Challenges of EDM

In this decade study of educational data mining is taking pace. Worldwide researchers are focusing on this area to handle the data and mine the new knowledge hidden inside these data but they fail or unable to get fruitful result from this data. Following are the limitation while studying the data:

- Lack of data warehouse to store the data and apply data mining technique to cover larger geographical region and large number of respondents.
- Lack of interoperability between heterogeneous data source.
- Change of Respondent (Student/Teacher/Course Content)
- Uncertainty of respondent behavior in different region
- Researcher’s expertise also big issues because EDM researchers often belongs to computer science field whereas developing and understanding behavior and interpreting students need is educationist subject. Thus researcher must have good knowledge of data analysis and educationist approach.
- Scarcity of gadgets specific for educational data collection

7. Conclusion

This paper discusses the development of educational data mining and its recent trends. In this decade researcher around the world are focusing on data mining technique to help education field. But still few challenges are in front of educational data mining. Addressing these challenges will lead successful implementation of educational data mining.

References

[1] http://www.xenodochy.org/ex/quotes/knowsnot.html
[2] Fatima D., Fatima S. Prasad V K, “ A Survey on Research work in Educational Data Mining”, IOSR Journal of Computer Engineering (IOSR-JCE) Vol 17 Ver II (Mar-Apr 2015) PP 43-49 e-ISSN: 2278-0661, p-ISSN 2278-8727
[3] Huebner Richard A., “A survey of educational data mining”
[4] Baker R.S.J.d, Inventado P.S, “Educational Data Mining and Learning Analytics”, In J.A. Larusson, B White (Eds.) Learning Analytics: From Research to Practice. Berlin, Germany: Springer.
[5] Jinal Rajani and Borah Malaya Dutta, “A Survey on Educational Data Mining Research Trends”, International Journal of Database Management System (IJDMS), Vol. 5 NO.3 June 2013.
[6] Prabha S Lakshmi and Shanavas A.R Mohammad, “Educational Data Mining Applications”, Operation Research and Applications: An International Journal (ORAJ) Vol. 1, No. 1 August 2014
[7] Kay J., “Enabling people to harness and control EDM for lifelong, life-Wide learning”, Proceedings of the 9th International conference on educational data mining.
[8] Romero, C. & Ventura, S.(2010),Educational data mining: A review of the state of the art, IEEE Transactions on systems man and Cybernetics Part C. Applications and review, 40(6),601-618.
[9] Han and Kamber
[10] Kaur P, Singh M & Josan G S, “Classification and prediction based data mining algorithms to predict slow learners in education sector”, 3rd International conference on Recent trends in computing 2015, 500-508
[11] Saranya S., Ayappan R. and Kumar N, “Student progress analysis and educational institutional growth prognosis using data mining”, International Journal of Engineering Science and Research Technology, Vol 3(4) April 2014, pg no. 1982-87, ISSN 2277-9655
[12] Osmanbegovic E, Sulijic M, “Data Mining Approach for Predicting Student Performance”, Economic Review- Journal of Economics and Business, Vol. X Issue 1, May 2012.
[13] Mills C. Bixler R, Wang X and Sidney K D’mello, “ Automatic Gaze based Detection of Min Wandering during Narrative film comprehension” Proceedings of 9th International conference on Educational Data Mining 2016 pg 30-37.
[14] Chen Y, Gonzalez-Brenes J P and Tian J, “Joint Discovery of Skill Prerequisite Graph and Student Model”, Proceedings of 9th International conference on Educational Data Mining 2016 pg 46-53.
[15] Doroudi S, Holstein K, Aleven V and Brunskill E, “ Sequence Matters, But How Exactly? A method for

Evaluating Activity Sequences from Data, Proceedings of 9th International conference on Educational Data Mining 2016 pg 70-77.

[16] Hutt S, Mills C, White S, Donnelly P J & D'Mello S K, " The Eyes Have It: Gaze-based Detection of Mind Wandering during Learning with an Intelligent tutoring System", Proceedings of 9th International conference on Educational Data Mining 2016 pg 86-93.

[17] Ma Y, Agnihotri L, Baker R and Mojard S, Ëffect of student ability and question difficulty on duration", Proceedings of 9th International conference on Educational Data Mining 2016 pg 86-93.

[18] Bydžovská H, "Course Enrollment Recommender System", Proceedings of 9th International conference on Educational Data Mining 2016 pg 312-17.

[19] Bydžovská H, "A Comparative Analysis of Techniques for Predicting Student Performance", Proceedings of 9th International conference on Educational Data Mining 2016 pg 306-12.

[20] Clement B, Oudeyer P, Lopes M, "A Comparison of Automatic Teaching Strategies for Heterogeneous Student Populations", Proceedings of 9th International conference on Educational Data Mining 2016 pg 330-35.

[21] Hao J, Liu L, Davier A, Kyllonen P and Kitchen C, "Collaborative Problem Solving Skills versus Collaboration Outcomes: Findings from Statistical Analysis and Data Mining", Proceedings of 9th International conference on Educational Data Mining 2016 pg 382-87.

[22] Mostafavi B, Barnes T, "Exploring the Impact of Data-driven Tutoring Methods on Students' Demonstrative Knowledge in Logic Problem Solving", Proceedings of 9th International conference on Educational Data Mining 2016 pg 460-65.

[23] Klingler S, Käser T, Solenthaler B, Gross M, "On the Performance Characteristics of Latent-Factor and Knowledge Tracing Models", Proceedings of 8th International conference on Educational Data Mining 2015 pg 37-44.

[24] Paquette L, Rowe J, Baker R, Mott B, Lester J, DeFalco J, Brawner K, Sottilare R, Georgoulas V, "Sensor-Free or Sensor-Full: A Comparison of Data Modalities in Multi-Channel Affect Detection", Proceedings of 8th International conference on Educational Data Mining 2015 pg 93-100.

[25] Sabourin J, Kosturko L, FitzGerald C, McQuiggan S, "Student Privacy and Educational Data Mining: Perspectives from Industry", Proceedings of 8th International conference on Educational Data Mining 2015 pg 164-70.

[26] Rollinson J, Brunskill E, "From Predictive Models to Instructional Policies", Proceedings of 8th International conference on Educational Data Mining 2015 pg 179-86.

[27] González-Brenes J P, Huang Y , "Your model is predictive— but is it useful? Theoretical and Empirical Considerations of a New Paradigm for Adaptive Tutoring Evaluation", Proceedings of 8th International conference on Educational Data Mining 2015 pg 187-94.

[28] Olsen J K, Aleven V, Rummel N, "Predicting Student Performance In a Collaborative Learning Environment"", Proceedings of 8th International conference on Educational Data Mining 2015 pg 211-17.

[29] Luo L, Koprinska I, Liu W,"Discrimination-Aware Classifiers for Student Performance Prediction", Proceedings of 8th International conference on Educational Data Mining 2015 pg 383-87.

[30] Feild J, "Improving Student Performance Using Nudge Analytics", Proceedings of 8th International conference on Educational Data Mining 2015 pg 464-67.

[31] Bravo J, Romero S J, Luna M, Pamplona S, "Exploring the influence of ICT in online students through data mining tools", Proceedings of 8th International conference on Educational Data Mining 2015 pg 540-543.

[32] Pelánek R, "Application of Time Decay Functions and the Elo System in Student Modeling", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 21-27

[33] Chi M, Schwartz D L, Chin D B,"Choice-based Assessment: Can Choices Made in Digital Games Predict 6th-Grade Students' Math Test Scores?" , Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 36-43

[34] Käser T, Koedinger K R, Gross M,"Different parameters - same prediction: An analysis of learning curves", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 52-59

[35] Bazaldua D A L, Baker R S, Pedro M O Z S, "Comparing Expert and Metric-Based Assessments of Association Rule Interestingness", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 44-51.

[36] Khajah M M, Wing R M, Lindsey R V, Mozer M C, "Integrating Latent-Factor and Knowledge-Tracing Models to Predict Individual Differences in Learning", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 99-106.

[37] Stefanescu D, Rus V, Graesser A C, "Towards Assessing Students' Prior Knowledge From Tutorial Dialogues", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)pg 197-200.

[38] Kokkodis M, Kannan A,Kenthapadi K, "Assigning Educational Videos at Appropriate Locations in Textbooks", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 201-04.

[39] Agnihotri L, Ott A,"Building a Student At-Risk Model: An End-to-End Perspective", Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014) pg 209-12.

[40] Srimani P K, Patil M M,“Linear Regression Model for Edu-mining in TES”, International Journal of Conceptions on Electrical and Electronics Engineering Vol. 1, Issue 1, Oct 2013,pg 45-49.
[41] Jackson D, Chapman E, "Non-technical skill gaps in Australian business graduates", ECU Publications 2012, pg 95-113
[42] Potgieter I, Coetzee M , “Employability attributes and personality preferences of postgraduate business management students”, SA Journal of Industrial Psychology/SA TydskrifvirBedryfsielkunde, 39(1), Art. #1064, 10 pages.
[43] Thakar P, Mehta A, Manisha , “Role of Secondary Attributes to Boost the Prediction Accuracy of Students’ Employability Via Data Mining”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 11, 2015, pg 84-90.
[44] Arora R K, Badal D, "Placement Prediction through Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 4, Issue 7, July 2014, pg 447-51
[45] Arora R K, Badal D,"Mining Association Rules to Improve Academic Performance", International Journal of Computer Science and Mobile Computing(IJCSMC), Vol. 3, Issue. 1, January 2014, pg.428 – 33
[46] Srimani P K, Patil M M, “A Classification Model for Edu-Mining”, International Conference on Intelligent Computational Systems (ICICS'2012) Jan. 7-8, 2012 Dubai, pg. 35-40
[47] Srimani P K, Patil M M, “A comparative study of data mining (DM) and massive data mining (MDM)" International Journal of Conceptions on Computing and Information Technology Vol. 1, Issue. 1, November 2013, pg.42-46
[48] Pumpuang P, Srivihok A, Praneetpolgrang P, "Comparisons of Classifier Algorithms: Bayesian Network, C4.5, Decision Forest and NBTree for Course Registration Planning Model of Undergraduate Students", IEEE International Conference on Systems, Man and Cybernetics (SMC 2008), pg.3647-51.
[49] Dimokas D, Mittas N, Nanopoulos A, Angelis L, "A Prototype System for Educational Data Warehousing and Mining", InInformatics, 2008. PCI'08. Panhellenic Conference on 2008 Aug 28, IEEE pg. 199-203.
[50] Cha S, “Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions”, International Journal Of Mathematical Models And Methods In Applied Sciences, Issue 4, Volume 1, 2007, pg.300-307.
[51] Tsien T B K &Tsui M, "A Participative Learning and Teaching Model: The Partnership of Students and Teachers in Practice Teaching", Social Work Education: The International Journal, Volume 26, No.4, pg. 348-358.

AUTHOR

Rishi Kumar Dubey is Computer Programmer at Kushabhau Thakre Patrakarita Avam Jansanchar Vishwavidyalay Raipur Chhattisgarh. He obtained his MCA degree from MP Bhoj (Open) University, Bhopal MP in 2007. He is also research scholar of MSIT dep’t of MATS University Raipur. His one paper is presented at ICTIS 2017, International conference, at Ahemdabad and same paper is accepted for publication in Springer journal SIST series.

Umesh Kumar Pandey is Assistant Professor in MSIT Dep’t of MATS University Raipur CG. He obtained his PhD in Year 2013 and MCA in 2003. He has more than 10 year of teaching experience and continuously engaged in research area of data mining. He published 6 international and 1 national publications. He is also editorial board member of 3 international journals and reviewer of more than 8 journals. He has professional membership of ISCA, IAENG and IACSIT. Currently 3 research scholars are pursuing their PhD.