Research Methodologies for Student Performance Evaluation Using Educational Analytics Tools and Approaches

Sri Laxmi Kuna, A.V. Krishna Prasad

ABSTRACT--- Educational Data Learning Analytics (EDLA) are about learning and is used to study and analyzing data from academic databases, Which is focuses on faculty teaching and student learning process, developing new tools and algorithms for discovering data patterns to improve student success. Recently Educational institutions are also using Learning analytics for the decision making to improve the student learning process. We can also develop new techniques from statistics for the analysis of educational data. It is also tests the learning theories and improves the teaching – learning process. One more important feature of EDLA is that it can focus on university level, institutional level, classroom level, session level and individual student answer level also. Educational Data Learning Analytics techniques are used for the benefit of all the stakeholders in the educational system such as student, faculty, engineering institutions and engineering education universities, which will help study, predict and improve a students' academic performance. In this paper we define Educational Data Learning Analytics and discussed about existing tools, how the stack holders can make use of it for student success.

Keywords—Educational Data Learning Analytics (EDLA), Classification, Clustering, Outlier analysis, Association Rule Mining, tools and stack holders.

I. INTRODUCTION

Everyone accepts that we are in data world. Generation and use of data is increasing day by day. Understanding the data is very important in these days in all the fields. One of the emerging fields is Education. Educational data Analytics is a learning science, and an emerging discipline, concerned with analyzing and studying data from academic databases. India has great demand for Education. Engineering education is one of the key factors in the Indian education system. Even though a large number of engineering students are completing engineering education from various government and private engineering institutions in INDIA, not even 7% of graduates are getting employment. According the survey conducted by the "Aspiring Minds National Employability Report Engineers,2016" more than 80% of engineering graduates remaining unemployed in INDIA due to lack of quality- oriented education in the institution and lack of skills acquired by the students. All over INDIA 3500 small or big companies have implementing the standardized assessment for hiring employees. Now a day the colleges are also showing interest

to improve the employability of their students by assessing their existing data. By using this analysis the institutes are designing their curriculum according to the industry needs and they are also changing their internal assessment of the student. The institutes are showing interest to implement example oriented teaching. The AICTE also suggesting to follow choice based credit system. Even though any method is followed by the institution the student should be in a position to adapt new learning method and selection of the subjects in which he can become perfect. Educational Data Learning Analytics (EDLA) is the technique which estimates the individual student strength and weaknesses in the subjects and suggests the selection of subjects. Educational analytics combines educational institution data, analysis by statistics, and predictive model on which students, faculty, and administrators can change teaching and learning behavior. For example if one student is not good in programming language and further if he selects the advanced programming language as elective then he may get less marks or sometimes he may fail in the specific subject. So based on the strength of the subject knowledge of the student, he has to select the subject. EDLA suggests the selection of subjects based on the strength of the individual student. So, that the student can choose the comfortable subjects according to their strength and improves his/ her knowledge in the subject. This professional skills are also reflects in getting employment after the completion of their education.

1.1 GOALS OF EDLA

Educational Data Learning Analytics main objective is to improve educational outcome of a student. That is, improving the learning and teaching performance for educational institutions. In particular,

1.1.1. For Students

i. To progress learning performance.
ii. Assess Learning effectiveness
iii. Understand social, cognitive and behavioral aspects**.**

    Sri Laxmi Kuna, research Scholar in KLU & Asst. Prof in Sreenidhi Institute of Science and Technology, India (E-mail: srilaxmiravi2016@gmail.com)
    Dr. A.V. Krishna Prasad, Prof in KLU & Prof in MVSR, India. (E-mail: kpvambati@gmail.com)

1.1.2. For Faculties

i. Get better teaching performance.
ii. Understand student learning experience and his / her learning preferences.
iii. Accommodate to each individual student.

1.1.3. For Educational Institutions

i. Flexible tutorial for the individual students needs.
ii. Assess teaching effectiveness which includes how teacher is responding to the individual student needs.
iii. Predict student academic performance like risk in failing a subject or dropping out of a subject.

1.1.4. For Universities

i. Curriculum amendments according to the industrial requirements.
ii. Respond to student and industry needs.
iii. Adopting the new and effective student learning methodology.

1.2 Objectives of EDLA

One of the main objectives of learning analytics is to minimizing the failure of a students and improving the learning standard of the student. It helps the students to become better learners and teacher to become better teacher. It gives the prior information about the students who are at risk. So, that the teacher can concentrate on a specific or group of students to improve their knowledge in the subject it automatically minimizes the failure.

TABLE I
OBJECTIVES OF EDLA

Name of method	Aim of method	Beneficiaries
Learning Analytics	Student - Level: To achieve knowledge and its application to the real time problems. Course – Level: Improve the ability of problem analysis and applying the knowledge on problem solving. Program – Level: Improve the problem analysis and engineering knowledge. Ability to cross disciplinary lines to find solutions in different disciplines. Institutional - Level: Improve the Problem solving skills of the students who are studying in the particular institution. University - Level: Amendments of curriculum according to the industrial needs and adopting effective student learning methodology which meets industrial needs for placement.	Students Faculty

II. LITERATURE SURVEY

The authors explained a novel method to know about – 'how different data mining techniques are applied on student data set using educational data mining'. In this the authors analyzed student data set using data mining techniques such as association rule mining and classification. In this they found performance in the end semester examination. In classification they used decision tree algorithm used to predict the program outcome and course outcome. Association Rule mining algorithms is used to study the student understands capability and outcome of the student. This combination of algorithms was found very acceptable in predicting and analyzing the performance. Visual analytics was used in the framework to illustrate the analysis of the students' performance [1].

Omer Deperlioglu et al, analyzed the failure causes of high school students using data mining technique, the authors used clustering method for analysis. The collected data were separated into three clusters and each cluster was investigated to determine the factors that play a role together in students' failure, as well as the relations between them. Finally, the findings of the applied survey, taking into account the work aimed at reducing the causes of failure, are expressed [2].

Usamah Bin Mat, Norlida Buniyamin and Pauziah Mohd Arshad, explained about the importance of prediction of academic performance and the assessment of students at early in the beginning of the each semester of the course. They used data mining classification technique to categorize the students' based on their performance at the beginning of the each semester in the course. By this the teacher can predict the students' final attainment at the end of the semester. Neuro-Fuzzy classifier was used to study first semester data and predict the students' attainment at the final graduation [3]

Tripti Mishra et al, explained a classification method to predict the employability of the PG students. For this analysis they used Bayesian technique of data mining, Multilayer Perceptions and Sequential Minimal Optimization (SMO), Ensemble Methods and Decision Trees, to predict the employability of MCA students. The relative study among J48, Random Forest and SMO. By this study they found SMO is the best algorithm to prediction of the student employability in the PG students [4].

Brijesh Kumar Baradwaj et al, analyzed the student data to predict the enrolment of students in a particular course for higher education system in the university. The authors used the classification methods as decision tree. They considered the attributes such as Internal Exam, Seminar marks, Assignment marks and an Attendance, from the existing data set to predict the end semester result. By this the teachers can know the students who need special care to minimize the student fail percentage in the particular course [5].

Edin Osmanbegovic, Mirza Suljic described the prediction of students' success. They used the summer

semester Tuzla university data for the year 2010-11. The data set contains the first year students' data which is taken from the students at the time of enrollment. The success was evaluated with the passing grade at the exam. They considered the attributes like the percentage of end exam marks of higher school and the entrance exam score [6].

David J. Finch, Leah K. Hamilton, Riley Baldwin and Mark Zehner recognized the university graduates influencing factors of the employability. The authors used both Qualitative and Quantitative techniques. By this analysis they found 17 factors are influences the new graduate employability such as pre graduation experience, job specification function skills, academic reputation, leadership skills, creative thinking skills, critical thinking skills, interpersonal skills, verbal communi-cation skills, written communication skills, etc., they made 5 classes based on these 17 factors as soft skills, problem solving skills, pre graduation experience, functional skills and academic reputation. Out of these 17 factors two are more influencing factors such as problem solving skills and soft skills [7].

Bangsuk Jantawan, Cheng-Fa Tsai build the Graduates Employability Model using classification. They used various classification algorithms like AODE, BayesNet, HNB, NaviveBayes, WAODE algorithms to analyze the data using WEKA tool. This model is used to predicting whether a graduate was employed, unemployed, or in an undecided situation [8].

Lumbini P et al, attempted to detect students' failure to improve the academic and to preventing dropping out .differentiated fast learners with slow learners using classification and feature selection algorithm. In this analysis. He considered personal information, SSC, Intermediate and under graduation end semester marks and designed a model to predict student failure [9].

Dr. Emine Onder explained the failure cause from teacher and student perspective by using Descriptive research model it includes quantitative and qualitative research approach. He designed the data base with the 11 schools and 171 students of the school. Quantitative data was collected from students through the causes of failure determined and qualitative data was obtained from teachers through open-ended questions. By using chy-square test quantitative data was analyzed and by using content analysis the qualitative data was analyzed. The reasons are quite different for the student failure [10].

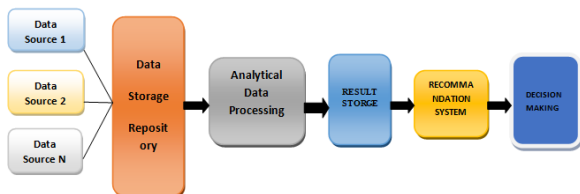### III. ARCHITECTURE FOR EDLA PROCESS

Fig 1: Educational Analytics Traditional Architecture

Data collection: Data is collected from the students which are belonging to different regions and branches.
This step include the identifying the various data generated sources and collection.

Data Storage Server: Data is stored in a single system called Repository which is collected from various sources
This step includes the storing the various formatted data into single repository.

Processing: Applying the analytics on the vast data collected from the different sources for the insight knowledge.

In this step first we need to identify the related attributes and delete non related attributes using attribute reduction process and use the relevant attributes for the analysis to get the result. If the result is not satisfactory repeat the step 1 to step3 with the different data sources and with different analysis processing techniques until we find the insight knowledge from the data set.

Result Storage: Various Analysis techniques give various results and all various results are stored in one place. So that we can easily compare the result.

Recommendation Systems: Recommendations are given based on the result generated in the above step. To understand the result very effectively we can generate the graphs using the results and suggest the recommendations for the future purpose.

Decision making: Based on the recommendations given by the analytics we can make decisions accordingly.

IV. AVAILABLE TOOLS IN THE FIELD OF EDLA

Learning Management System (LMS) by Brigham Young University (BYU): The LMS is the extension of blackboard LMS in 2004-2005 at Brigham Young University. In this they used entire institutional data which is belongs to variety of courses for the analysis. They implemented same like blackboard tool which is analyzed the data and how it finds the patterns. By using this tool they tested with different colleges data and also found the similar patterns among the colleges [11].

CLASS Management System (CMS) used by the 120 teachers of Israel region. This was used to identify the weak points in the educational system and to improve the educational outcomes [12].

University of South Florida used the CMS (Content Management System) data logs for analysis of graduate course with 67 students and found the some common among affective attributes among the student with the sample data logs. CMS is used to analyze course data using data mining method [13].

Course Management System is also developed as a tool called Moodle by the Cordoba University in Spain. By using classification technique they developed a tool called KEEL which allows designers of course for an analysis on single course or more courses. This tool used rule induction, neural networks, and statistical inference and decision trees for the analysis. By using of this tool every teacher can analyses their own course. This tool can also use by the students to analyses their own data usage [14].

Purdue Early Warning System (PAWS) Signals Project: In Prudue institution, these people started exams from paper work, then bubble sheets and then migrated to web. The

main aim of this is to identify the students who were in at high risk. For this institution administrator taken the term-1 result, based on this data, the institution engaged a student with necessary tutorial classes to support his/her academic improvement [15].

Limitations: event thought PAWS succeeded in improvement of students' performance, it has some limitations like the assessment of a student requires so much time so that they could not assisted in a proper way in the early in the semester. More over which ever the procedure they were following that are general in nature but not according to the course.

Institutional Research, Planning, and Effectiveness (IRP&E) implemented the EDLA at Colorado State University. In this university the full information about the students who are at risk is shared with the faculties. They concentrated on them and adapted the new methods to guide them. So, that the students improved a lot in course.

Middle Tennessee State University (MTSU): In this university the first year students' retention increased by 3%. By implementing in the successive years they increased high retention rate in the university.

In Grinnell College, they combined the learning analytics with the human intelligent networks to improve students' course completion as well as student retention in the course. In this the researchers found the various students psychological and social factors influence the student success in the course. This data is linked with the academic data to find the similarity of the student success rate.

Students Success Network: It is a student dash board which integrates LMS, financial aid, educational planning information and early alert. It is accessed by the teacher and student. By using this students improved a lot in their studies by getting feedback from the tool. This tool is also useful for the faculty and they adapted the new methods to improve the students' performance, so that they increased the result in the midterm from 75% to 90%.

TABLE II
COMPARISON OF TOOLS EXISTING

Sl. No	Learning Analytic Tool	Improvement by using the tool	Institution	Purpose of tool	Limitations
1	Class Management System(CMS)	Administrator work made easy.	Brigham Young University (BYU)[16]	To make admin work quick.	No student interaction with this tool
2	CLASS Management System	Analysis made easy.	University of Granada	Find the similarities of attribute values.	This is an analysis tool. There is no prediction of student performance
3	Course Management System(Sakai and Moodle)	Analysis made easy	Cordoba University	To find weak points of educational system.	This is an analysis tool. There is no prediction of student performance
4	Signals Project	Improved students' academic performance.	Purdue University	This project aim is to identify the students who were at risk.	The assessment of a student requires so much time so that they could not assisted in a proper way in the early in the semester. More over which ever the procedure they were following that are general in nature but not according to the course.
5	Institutional Research, Planning, and Effectiveness (IRP&E)	In this university the full information about the students who are at risk is shared with the faculties	Colorado State University (CSU).	The students who are at risk is shared with the faculties	No recommendation from the tool. Individual teacher has to take care of the students who are at risk in the corresponding subject.
6	Middle Tennessee State University (MTSU)	The first year students retention increased by 3%.	Middle Tennessee State University (MTSU)	The student can improve the academic performance in the first year itself.	They is no consideration of 2nd, 3rd and 4th year students data.

71

7	Students Success Network	Students' academic performance increased from 75% to 90%.	Grinnell College	It only finds the similarities in the student success rate based on the psychological and social factors influence	Find the similarities it would not suggest any recommendation who are at risk.
8	GPS (Grade Performance System)	Teacher – Student interaction made easy.	Northern Arizona University	Teacher can interact with Students directly.	Only teacher can post the messages. Student can receive the messages posted by the teacher but he cannot post message to the teacher.
09	Blackboard CMS	Information regarding the student performance is available to the teacher up-to-date	University of Maryland Baltimore County (UMBC)	Follow the performance and Predict student success.	No interaction with the student regarding his performance.
10	BIRT (Business Intelligence and Reporting Tools)	Information regarding the student performance is available to the teacher up-to-date	Graduate School of Medicine, University of Wollongong	To find the students at risk	Less security on student data.
11	PACE (Progress and Course Engagement)	Information regarding the student performance is available to the teacher up-to-date	Rio Salado Community College	To find the students at risk	No interaction with the student regarding his performance.
12	EIS (Executive Information System)	Information regarding the student performance is available to the teacher up to date	University of Central Florida	To find the students at risk	No interaction with the student regarding his poor performance in the specific subject.
13	Student Success Plan (SSP)	Information regarding the student performance is available to the teacher up-to-date	Sinclair Community College	For advising and retention	No prediction regarding the performance of the student.

V. COMPARISON OF RESEARCH WORK CARRIED OUT

Table III

COMPARISON OF RESEARCH WORK CARRIED OUT IN THIS FILED

Sl No	Title	Name of the journal	Techniques used	Outcome
1	Edifice an Educational Framework using Educational Data Mining and Visual Analytics[1]	*I.J. Education and Management Engineering*	Classification and Association Rule mining	Analysis of the students' performance.
2	Analysis of Girls Vocational High School Students Academic Failure Causes with Data Mining Techniques[17]	www.tandfonline.com	Clustering Technique	Analysis of failure causes of high school students

3	Educational Data Mining Classifier For Semester One Performance to Improve Engineering Students Achievement [18]	Middle-East Journal of Scientific Research	Classification Technique and Neuro-Fuzzy classifier	Assessment of students at early in the beginning of the course
4	Students' Employability Prediction Model through Data Mining [19]	International Journal of Applied Engineering Research	ClassificationBayesian methods, Multilayer Perceptions and Sequential Minimal Optimization (SMO), Ensemble Methods and Decision Trees	Predict the employability of the PG students
5	Mining Educational Data to Analyze Students' Performance [20]	International Journal of Advanced Computer Science and Applications	Classification decision tree method	Predict the enrolment of students in a particular course for higher education system in the university and to identify the dropouts.
6	Data Mining Approach for Predicting Student Performance. [21]	Journal of Economics and Business		First year student data taken at the time of enrollment. Based on the previous academic marks, The success was evaluated with the passing grade at the exam
7	Students Academic Failure Prediction using Data Mining [24]	International Journal of Advanced Research in Computer and Communication Engineering	Classification and Feature selection Algorithm	Designed a model to predict student failure at the end of the semester.
8	Cause of School failure from Teacher and Student's Perspective. [25]	International Journal on New Trends in Education and Their Implications	Classification chy-square test	The reasons are quite different for the student failure

VI. RESULTS & DISCUSSIONS

Data is increasing day by day in educational field, so there is a need of EDLA to find hidden knowledge from the data. Even though the teaching learning process is existing since the ancient age and it is changing its process since the beginning, it needs some more improvements for the benefit of stack holders. Educational system has to adapt according to the course, as well as student capability. By this comparison study we are trying to explain the importance of Learning Analytics in the education sector. EDLA increase the student success and retention in the institute/university. This is also used as early alert system for the students who are at risk. Purdue University is benefitting with the help of EDLA. Table 1 tells where the LA can be applied on all the level of learning of students. Table 2 summarizes all the existing tools used for the improvement of the student and faculty learning success. Table 3 explains how the different methods are using for the student performance improvement. By using of above survey we can conclude that by using of EDLA the educational stack holders are benefiting and improving more than 50% than before. By applying EDLA effectively on data we can use predictive analytics to improve student success and completion of course. EDLA techniques are also useful for the researchers to find the new methods and techniques to improve the teaching and learning process. EDLA finds insights and helps to improve and excellence the learning – teaching process.

REFERENCES

1. Dr. S. Anupam Kumar (2016) "Edifice an Educational Framework using Educational Data Mining and Visual Analytics" *I.J. Education and Management Engineering,* 2, 24-30
2. Omer Deperlioglu & Fatma Sibel Birtil (2017) Analysis of Girls Vocational High School Students' Academic Failure Causes with Data Mining Techniques, The Anthropologist, 23:3, 505-512, DOI: 10.1080/09720073.2014.11891970
3. Usamah Bin Mat, Norlida Buniyamin and Pauziah Mohd Arshad (2016) *"*Educational Data Mining Classifier for Semester One Performance to Improve Engineering Students Achievement" Middle-East Journal of Scientific Research 24 (2): 338-346.
4. Tripti Mishra, Dharminder Kumar, Sangeeta Gupta (2016) "Students' Employability Prediction Model through Data Mining" International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 4 (2016) pp 2275-2282.

5. Brijesh Kumar Baradwaj, Saurabh Pal (2011) "Mining Educational Data to Analyze Students' Performance"*(IJACSA)* International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.

6. Edin Osmanbegovic, Mirza Suljic (2012) "Data Mining Approach For Predicting Student Performance" Economic Review – Journal of Economics and Business, Vol. X, Issue 1, May 2012

7. DavidJ.Finch, Leah K, Hamilton, Riley Baldwin, Mark Zehner, (2013) "An exploratory study of factors affecting undergraduate employability", Education + Training, Vol. 55 Issue: 7, pp.681-704, https://doi.org/10.1108/ET-07-2012-0077

8. Bangsuk Jantawan, Cheng-Fa Tsai (2013) "The Application of Data Mining to Build classification Model for Predicting Graduate Employment " *International Journal of Computer Science and Information Security, Vol. 11, No. 10, October 2013*

9. Lumbini P. Khobragade, Prof. Pravin Mahadik (2015) "Students' Academic Failure Prediction Using Data Mining "*International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 11, November 2015.*

10. Dr. Emine ÖNDER (2016) –"Cause of School Failure from Teacher and Students perspective" "International Journal on New Trends in Education and Their Implications April 2016 Volume: 7 Issue: 2 Article: 02 ISSN 1309-6249

11. Griffiths, M.E., & Graham, C.R. (2009). "Patterns of user activity in the different features of the Blackboard CMS across all courses for an academic year at Brigham Young University". Journal of Online Learning and Teaching, 5(2), 285-292.

12. Nirit Raichel, (2013),"Computerized class management software: teachers' views and assessment", Journal of Organizational Change Management, Vol. 26 Issue 5 pp. 830 – 852.

13. http://www.usf.edu/ucm/marketing/cms-implementation. aspx

14. C., Ventura, S., & Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial. Computers & Education, 51(1), 368-384

15. Purdue Early Warning System (PAWS) Signals Project

16. Griffiths, M.E., & Graham, C.R. (29). Patterns of user activity in the different features of the Blackboard CMS across all courses for an academic year at Brigham Young University. Journal of Online Learning and Teaching, 5(2), 285-292.

17. Deperlioglu, Omer & Sibel Birtil, Fatma. (2016). Analysis of Girls Vocational High School Students' Academic Failure Causes with Data Mining Techniques. The Anthropologist.

18. Usamah Bin Mat, Norlida Buniyamin and Pauziah Mohd Arshad (2016) Engineering & Technical Education Research Group (Enter), Faculty of Electrical Engineering, Universiti Teknologi MARA, Middle-East Journal of Scientific Research 24 (2): 338-346, 2016.

19. Mishra, Tripti & Kumar, Dharminder & Gupta, Sangeeta. (2016). Students' employability prediction model through data mining. 11. 2275-2282.

20. Brijesh Kumar Baradwaj, Saurabh Pal (2011) International Journal of Advanced Computer Science and Applications, (IJACSA) Vol. 2, No. 6

21. Osmanbegović, Edin & Suljic, Mirza. (2012). DATA MINING APPROACH FOR PREDICTING STUDENT PERFORMANCE. Journal of Economics & Business/Economic Review. 10. 3-12.

22. Finch, David & Hamilton, Leah & Baldwin, Riley & Zehner, Mark. (2013). an exploratory study of factors affecting undergraduate employability. Education + Training. 55. 10.1108/ET-07-2012-0077.

23. Jantawan, Bangsuk & Tsai, Cheng-Fa. (2013) the Application of Data Mining to Build Classification Model for Predicting Graduate Employment."

24. Jantawan, Bangsuk & Tsai, Cheng-Fa. (2013). The Application of Data Mining to Build Classification Model for Predicting Graduate Employment.

25. Lumbini P. Khobragade1, Prof. Pravin Mahadik (2015) International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 11, November 2015

26. Dr. Emine ÖNDER (2016) International Journal on New Trends in Education and Their Implications April 2016 Volume: 7 Issue: 2 Article: 02 ISSN 1309-6249.

27. D.Fatima, Dr. Sameen Fatima, Dr.A.V.Krishna Prasad "A Survey on Research work in Educational Data Mining", IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661,p-ISSN: 2278-872, Vol-17, Issue-2, 43-49.

28. D.Fatima, Dr. Sameen Fatima, Dr.A.V.Krishna Prasad "Integrating WEKA into Web Application: Predicting Student's Performance", International Journal of Engineering Research and Applications, ISSN: 2248-9622, IJERA,Dec 2017.

29. A Chapter entitled "Mining on Social Media" in Web Data Mining and the Development of Knowledge Based Decision Support Systems Editor(s): G. Sreedhar (Deemed University, India), IGI Global Publications.