

# Educational support system using AHS and data analytics

**Pushkar Khedekar<sup>1</sup>, Swapnil Mandavkar<sup>2</sup>, Darshan Mhatre<sup>3</sup>, Rohit Mudaliar<sup>4</sup>, Jayshree Jha<sup>5</sup>**

Student, Information Technology, Atharva College of Engineering, Mumbai, India<sup>1, 2, 3, 4</sup>

Professor, Information Technology, Atharva College of Engineering, Mumbai, India<sup>5</sup>

**Abstract:** The paper is based on concept of smart database which deals with formal student guidance throughout his four years of engineering. The older versions made use of static assessment to generate generic student report on the other hand Academic Advisor makes use of various smart algorithms to generate a dynamic student guidance support system which help students to keep a track of their progress and helps them to make informed decisions throughout their academic session and also in their future job endeavors. It helps the IT firms to shortlist students for interview based on the criteria provided by Academic Advisor based on a special and smart grading system. Hence Academic Advisor will provide student with a virtual personalized tutor.

**Keywords:** Educational Data Mining(EDM); Guidance Support System; Adaptive Hypermedia System; Academic Analytics.

## I. INTRODUCTION

The recent developments in the field of information technology has led to enormous volumes of data being stored in various formats like files, records, documents, sound, images, videos, scientific data and various upcoming data formats. The data collected from various repositories is heterogeneous in nature and thus knowledge discovery and extraction becomes difficult. Knowledge discovery in databases (KDD), also known as data mining, aims at the discovery of useful information from large collections of data. The basic principle to be considered while incorporating a data mining system is to make use of different methods and algorithms which detect some kind of pattern and sequence in the data available.

Educational Data Mining [1] is an emerging field, concerned with developing techniques for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings to which they are adapted.

Whether educational data is taken from students' use of computer-supported collaborative learning, interactive learning environments or administrative data from schools, colleges and other institutions, it has multiple levels of meaningful hierarchy. Hierarchical patterns often need to be determined by characteristics of the data itself, rather than using predefined rules. Issues of time, context, and sequence also play important roles in the study of educational data.

Educational Data Mining uses many techniques such as Decision Trees, K- Nearest neighbor, Naive Bayes, Neural Networks and many others.

Using these techniques many kinds of knowledge can be discovered such as association rules, clustering and classification. The discovered knowledge can be used for prediction regarding enrolment of students in a particular course, detection of unfair means used in online examination, alienation of traditional classroom teaching model, detection of abnormal values in the result sheets of the students, prediction about student's performance and so on.

The main objective of this paper is to predict the trends of marks based on the historic data (datasets) and real time

data of students by conducting tests. The student's historic data will be collected from respective institutes which will be used to analyze predictive score using association rules of data mining. The prediction thus obtained by historic data analysis is complimented by real time data in the form of real time test in every subsequent semesters which allows for a more precise and concise prediction of marks along with graphic support which will enable for improved guidance to the individual student.

## II. ADAPTIVE HYPERMEDIA SYSTEM(AHS)

Adaptive Hypermedia [2] is a disputed research field where hypermedia is made adaptive according to user model. In contrast to traditional e learning systems whereby all users offered or even directed a standard series of hyperlinks, AHS tailors what the user sees from model of the user's goal, preferences and knowledge. There are five basic features which are used by existing AHS users:

- Goals
- Knowledge
- Background
- Preferences

## III. BACKGROUND AND RELATED WORK

Brijesh Kumar Bhardwaj and Saurabh Pal [3] in their paper state that in this research, the classification task is used for student performance evaluation. Since there are many approaches that are used for data classification, one them being decision tree method is used here. By this task we extract the knowledge that describes students' performance in each end semester examination. They used decision tree learning algorithms like ID3 and C4.5. It helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising/counseling as per requirement.

Mohammed M. Abu Tair and Alaa M. El-Halees [4] in their paper state that they used educational data mining to improve graduate students' performance, and also solve the issue of low grades of graduate students. In their case study they tried to extract useful knowledge from graduate

students data collected from the college of Science and Technology – Khanyounis. The data included fifteen years period [1993-2007]. After preprocessing the data, they applied data mining techniques to discover clustering, classification, association and outlier detection rules. In each of these four tasks, they presented the extracted knowledge and described its importance in educational domain.

S.B. Kotsiantis Educational Software Development Laboratory [5] in their paper state that they had cited the most current articles that used machine learning techniques for educational purposes and had presented a case study for predicting student's marks. Student's key demographic characteristics and their marks in small number of written assignments which constituted the training set for a regression method in order to predict the students' performance. Finally, a prototype version of software support tool for tutors was constructed.

V.Ramesh, P.Parkavi and P.Yasodha [6] in their paper state that the scope of this paper is to investigate the accuracy of data mining techniques in such an educational scenario. The first step of the study was to collect students' data. Authors collected records of 300 Under Graduate students of computer science course, from a private Educational Institution. The second step was to clean the data and choose the relevant attributes. In the third step, Naïve Bayes Simple, Multi Layer Perception, SMO, J48, REP Tree algorithms were constructed and their performances were evaluated. The study revealed that the Multi Layer Perception is more accurate than the other algorithms. This work would help the institute to have accurate prediction of the performance of the students.

Patrick D. Schalk, David P. Wick, Peter R. Turner and Michael W. Ramsdell [7] in their paper state that the authors had used well-accepted conceptual assessment instruments. Initial state data such as the SAT, and their ingeniously developed instruments designed to measure aptitude in mathematics to develop a machine learning-based predictive model for student performance. This analysis of theirs concluded that it exhibited a strong correlation between mathematics and physics performance. This analysis contributes to an integrated evaluation of the current programs, which led to an assessment-based initiative to offer strategic guidance to new students, better placing them for career and academic success in their selected STEM disciplines.

Ajay Kumar Pal and Saurabh Pal [8] in their paper described the use of data mining techniques to improve the efficiency of academic performance in the educational institutions. In their paper, they presented a real-world experiment conducted in VBS Purvanchal University, Jaunpur, India. This method helped to identify the students who needed special advising or counseling by the teacher who gave higher quality of education.

Jayshree Jha and Leena Ragha [9] paper surveyed the most relevant studies carried out in EDM using Apriori algorithm. Based on the Apriori algorithm analysis and research, their paper pointed out the main problems on the application Apriori algorithm in EDM and presented an

improved support-matrix based Apriori algorithm. The improved Apriori algorithm proposed in this research used bottom up approach along with standard deviation functional model to mine frequent educational data pattern.

Peter Brusilovsky [10] is his paper discussed the idea behind the Adaptive Hypermedia. He described various set of identified methods and techniques. He also described various applications of Adaptive Hypermedia in various domain including Educational Data Mining. Finally he enlisted the existing Adaptive Hypermedia Systems.

#### IV. PROPOSED SYSTEM

Previous systems in the domain of Educational Data Mining (EDM) used to assist students to navigate through different courses depending on their marks or exam scores. Since, in India we don't have such choices to make regarding subjects, we can neglect that part and focus on student's performance. Few systems used to identify students who are at the risk of failure and suggest institutes a course of action to negate or reduce this risk. This is also one of the drawback of old system, since student under concern isn't directly connecting to the system. In our proposed system we are allowing students to interact with the system as much as possible. Most of the system uses data of previous students to predict performance and neglects any real time data of the students. However the trends in students' academic performance may or may not be the same. So we will use real time data that includes student marks and a test that will be taken at the end of every semester, aiding system to predict more accurately.

In our system we have segregated all semester's subjects under three major domain viz.

- Logical
- Technical
- Communicational

Students will be evaluated on the basis of aggregate of all subjects falling under corresponding domain which he/she has passed till date.

Student will have to attend a test generated by system at the end of each semester, which will have questions related to domain mentioned above.

For first two semester we will use previous student's data to predict the performance of current students. For this we will use Apriori algorithm to extract student performance patterns. Once students' two semester marks and two system generated test scores are obtained we will use students real time data to predict the students' performance and provide specific guidance.

For purpose of guiding students for better performance in exams we are proposing relatively new system in the domain of EDM known as Adaptive Educational Hypermedia System which is a subset if Adaptive Hypermedia System (AHS). In contrast to conventional e-learning (and face-to-face education) systems, in which all learners are offered or directed a predefined series of hyperlinks, adaptive educational hypermedia tailors what the learner sees to that learner's goals, needs,

abilities, interests, and knowledge of the subject, by providing hyperlinks that are most appropriate and relevant to the user. Essentially, the teaching tools "adapt" to the learner.

## V. CONCLUSION

The paper addresses the importance of knowledge mining from educational dataset and overview of existing systems used in education data mining and its flaws and innovative solution with a use of previous students dataset and students real time data. Thus Academic Advisor acts as a real time tutor which will help students to assess their logical, technological and communicational understanding at various stages throughout their academic tenure with the help of various predicting algorithms and providing periodic tests to learn about their current understanding and then to guide them appropriately.

## ACKNOWLEDGMENTS

Our sincere thanks to all the people who have contributed in anyway in carrying out this work.

## REFERENCES

- [1] <http://www.educationaldatamining.org/>
- [2] A white paper from DV X / DeVry Education Group and Integrated Education Solutions, Adaptive Learning Systems, February 2014.
- [3] Brijesh Kumar Bhardwaj and Saurabh Pal, "Mining Educational Data to Analyze Student's Performance", International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.
- [4] Mohammed M. Abu Tair and Alaa M. El-Halees, "Mining Educational Data to Improve Students' Performance A Case Study", International Journal of Information and Communication Technology Research, ISSN 2223-4985, Volume 2 No. 2, February 2012.
- [5] S.B. Kotsiantis Educational Software Development Laboratory, "Use of Machine Learning Techniques for Educational Purposes : A Decision Support System for Forecasting Student's Grade", Journal Artificial Intelligence Review, Volume 37 Issue 4, Pages 331-344, April 2012.
- [6] V.Ramesh, P.Parkavi and P.Yasodha, "Performance Analysis of Data Mining Techniques for Placement Chance Prediction", International Journal of Scientific & Engineering Research Volume 2, Issue 8, ISSN 2229-5518, August-2011.
- [7] Patrick D. Schalk, David P. Wick, Peter R. Turner and Michael W. Ramsdell, "Predictive Assessment of Student Performance for Early Strategic Guidance", 41<sup>st</sup> ASEE/IEEE Frontiers in Education Conference, October 12-15,2011
- [8] Ajay Kumar Pal and Saurabh Pal, "Analysis and Mining of Educational Data for Predicting the Performance of Students", International Journal of Electronics Communication and Computer Engineering, Volume 4 Issue 5, ISSN(Online) – 2249-071X, 2013
- [9] Jayshree Jha and Leena Ragha, "Educational Data Mining using Improved Apriori Algorithm", International Journal of Information and Computation Technology, ISSN 0974-2239 Volume 3, Number 5 (2013), pp 411-418
- [10] Peter Brusilovsky, "Methods and Techniques of Adaptive Hypermedia", User Modeling and User Adapted Interaction, v6 pp87-129, 1996.