

Academic Advising Using Data Analytics and Adaptive Hypermedia System

Swapnil S Mandavkar¹, Rohit A Mudaliar², Pushkar P Khedekar³, Darshan P Mhatre⁴, Jayshree Jha⁵
^{1,2,3,4} Student, Dept. of IT, Atharva College of Engineering, academicadvisor93@gmail.com
⁵ Professor, Dept. of IT, Atharva College of Engineering

Abstract—Educational Data mining is concerned with developing techniques to extract unique and meaningful information from large database. It is also used to identify patterns and to cluster them together with various association rules. The older versions made use of static assessment to generate generic student report, which proved to be inaccurate. In literature review we found that most of the current systems were dependent completely on historical data and ignored the caliber of the current students. So to overcome this limitation our software made use of three kinds of data viz. large historical data set, students current academic performance and periodic tests combined to produce a much better result.

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 student's 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

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. APRIORI AND ASSOCIATION RULE MINING

Association rule mining is one of the most well renowned data mining tasks. It discovers connections among characteristics in databases, producing statements concerning characteristics values. Here we find the similarities in item sets belonging to a large item sets. The two important steps in association rule mining are: First, apply minimum support to all frequent itemsets in a database. Second, use these frequent itemsets and the minimum confidence constraint to form association rules. There are a lot of different association rule algorithms. We will use Apriori algorithm.

IV. BACKGROUND AND RELATED WORK

Peter Brusilovsky [10] in 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.

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.

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 student's 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.

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.

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.

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.

V. IMPLEMENTATION PROCEDURE

Software which we have developed has two module viz. Admin and Student. Admin module will be responsible for managing students and Student module will aid students to enter marks, manage their profile, analyze the result etc.

For ease of prediction we have segregated all the semester subjects into three categories viz. Logical, Technical and Communicational.

The basic concept of our implementation consist of using historical data, their academic scores and periodic test combined to give near accurate prediction. Once we get the predictions, we use Adaptive Hypermedia System (AHS) to provide tailor made suggestions.

Our historical data consist of a large database of passed out students categorically segregated marks on which predictions can be made using Apriori algorithm. The periodic test consists of questions divided into logical, technical and communicational sections

The historic data score is calculated using the following algorithm:

Step 1: The students current marks are segregated into the three categories mentioned above.

Step 2: The historic dataset is scanned for entries which lie in the range of student's score (Where range is some user defined threshold) for that particular semester.

Step 3: Now the next semester scores of all the short-listed entries are aggregated.

Step 4: Now this aggregated value called the Historical Data Score (HDS).

The software makes use of the following formula:

$$\text{Prediction} = \text{NSP} \pm 3$$

$$\text{NSP} = 3/8 * \text{HDS} + 3/8 * \text{CSS} + 2/8 * \text{PTS}$$

Where,

NSP: Next Semester Prediction

CSS: Current Semester Score

PTS: Periodic Test Score

VI. RESULT

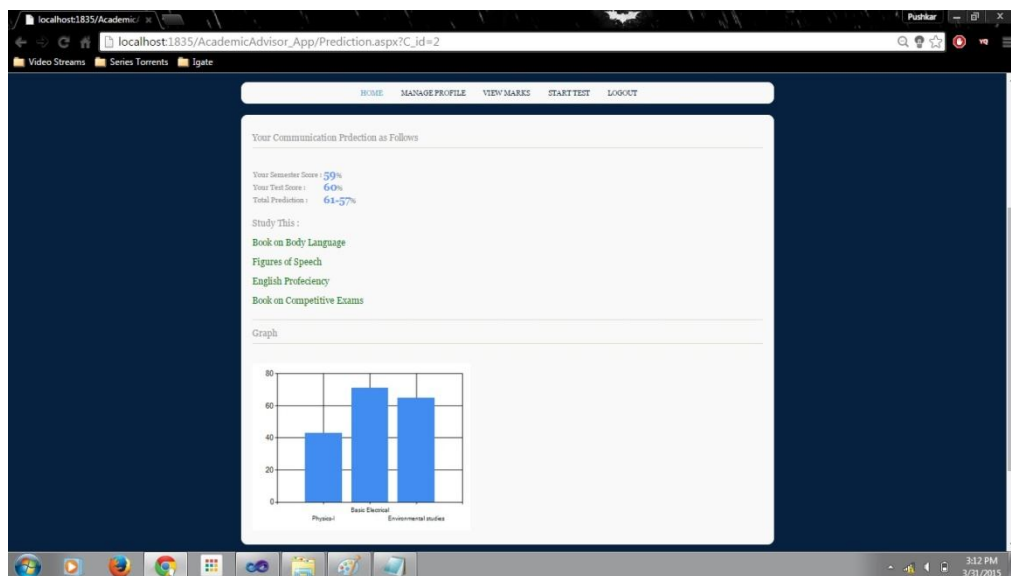


Figure 1. Prediction and suggestion

As per the above screenshot we can see that our software provides personalized categorical predictions and suggestions. The student can also track his progress with the help of inbuilt graph.

CONCLUSION

Thus Academic Advisor acts as a real time tutor which will help students to access 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 their current understanding and then to guide appropriately, so as to take necessary steps in order to improve their logical, lingual and communicational proficiency. Since the prediction will be carried out on the

basis of a large data set, along with the help of current data combined will produce a more realistic predictions which will increase the probability of accurate prediction. Academic advisor will prove to be a boon to many educational institutes in order to identify different category of students and also try and improve the quality and content of teaching accordingly. Hence our software takes the field of educational data mining one step ahead to accurate predictions and hence try and change the concept of self learning and self assessment to a great extent. Hence Academic Advisor will overcome the shortcomings of other prediction software by filling in the loop hole by using all kinds of data including historic, current and test based data.

ACKNOWLEDGEMENTS

Our sincere thanks to all the people who have contributed in anyway in carrying out this work and help us with our software.

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", 41st 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 Ragma, "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.

