

Analyzing Educational Data through EDM Process: A Survey

Tripti Dwivedi
Department of Computer Engineering
BUIIT
Bhopal, India

Diwakar Singh
Department of Computer Engineering
BUIIT
Bhopal, India

ABSTRACT

This paper presents the survey of student background history which helps academic planners in institute to give right direction to student. If the class of student is predicted in mid-session of institute in final year then it will be easy for the academic planner to plan some important workshop for the enhancement of performance of student which helps it in placement at the end of academic session. In educational institute data mining techniques play an important role in each activities of institute whether it is academic, cultural, examination and training and placement etc. in which Educational Data Mining which is a field of data mining helps a lot to find the actual filtered data in various field of department in institute. Hidden knowledge through data mining techniques is extracted from large database which helps to predict the pattern in such activities. It plays a great role in predictions of student data for placement and performance.

Keywords

EDM (Educational data mining), Classification, Decision tree, Association Rule

1. INTRODUCTION

Data mining is also known as knowledge discovery in databases is a technology used in different discipline to search for significant relationship among variables in large database. It extract knowledge form hidden data base and find the necessary information according to the requirement. These tasks are done by applying various algorithm of data mining on training dataset. These dataset passes through various steps of data mining and finally the filtered data is used by the user according to the requirement. It plays an important role in education institute. Educational data mining (EDM) is a field of data mining which work in the field of education to extract knowledge .EDM community [14] defines EDM as follows EDM is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational setting and using those method to better understand students and the setting which they learn in[14].It also help academic planners to give intimation in advance to student that in which direction of studies or other activities they lack[15].Classification methods like rule mining, decision tree, association rule with other techniques ,Bayesian network etc are uses to predict the student performance which helps in placement.

There are some fact arises when we apply data mining techniques using statistical model [16] Data Mining is a vast process that includes many techniques which is based on many models in which statistical work is based on a hill climbing algorithm in combination with a verisimilitude

ratio test based hypothesis which usually not well for large database. So we prefer algorithm like decision tree induction, association, logistic regression, naive Bayes etc. to filter educational data for the appropriate placement prediction.

The main objective of this survey paper is to find in advance the area or field for student in last year for placement. In this survey it is find that various factor like student performance in last exam, his background, the language known, family income, communication skill, technical skill, leadership quality, mothers education ,family income and many more criteria helps to academic planners for analysis of student data. This analysis help a lot in deciding the placement area in which to some extent it is predicted that if they will go for the same then they will do much better in future. The main emphasis of placement officer is to guide students during course in institute so that they exposed themselves to the corporate industry at right time.

The paper is organized as: section II describes the EDM section III gives brief description of literature survey with comparison between different techniques; section IV concludes the paper.

2. EDM

Educational data mining is emerging as a research area with a suite of computational and psychological methods and research approaches for understanding how students learn. Education Data Mining in data mining retrieve hidden knowledge by applying various techniques of data mining like clustering, rule mining, web based mining, test mining, neural network,baysian network, and many others which gives us a final result and if it need some requirement to changes then again the raw data if filtered according to need[4]. Data collected from education institute can be aggregated over large numbers of students and can contain many variables that data mining algorithms can explore for model building.

A. Goals Of EDM

EDM is used to predict student s behavior, future enhancement skill, predicting for scholarship and many more areas of student which help in to intensify the quality of skill and attitudes [17].

New model can be discovered by application of EDM and a improvement of previous model is also possible. Discovering or improving domain models that characterize the content to be learned and optimal instructional sequences [17];

It supports the effect of study for education institute by number of learning software [17].

In EDM process users, educators, administrators, and researchers are involved because at each level of EDM from t all these four stakeholders play major role may be in decision, changes in learning process, development and allocating the resource for institute so that the learning process in institute may give good result in these areas.

B. EDM Phases

EDM generally consists of four phases [17]:

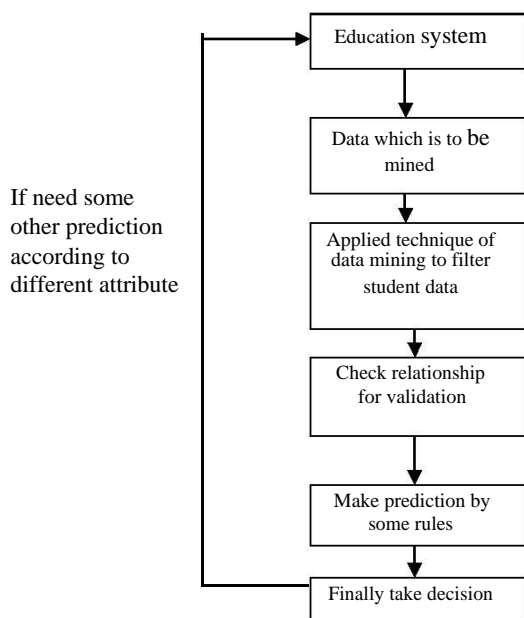
During first phase of the EDM process relationships in between data is find by searching through training set of data .This can be done through several algorithm of data mining like clustering,regression,association rule mining.

During second phases of EDM whatever relationship was arised in first phase is checked for validation.

These Validated relationships are used for prediction in learning environment..

Finally decisions are taken on the basis of prediction results to make some policy for institute.

During phases 3 and 4, various human judgment are taken to filter data.[17]



3. RELATED WORK

In the field of education data mining ,it is found by analyzing different literature that for predicting the performance ,placement prediction ,scholarship prediction for recruitment process different factor plays a vast role to measures the accurate result which include : student grade in secondary exam , mother' s education , living location, medium of teaching , family income , student other habits , student family status, student learning activities, learning behavior, language known to him , decision making capability , grades in last year exam and many more criteria related to student habits are co related to analyze the prediction of various fields[7][12].Some literature survey areas of different author s are mention below

Vandamme.et.al in 2007 work on decision tree and neural network for Academic performance prediction of students [1].

Alaa el-Halees in 2009 described data mining as Data Mining can be used in educational field to enhance our understanding of learning process to focus on identifying, extracting and evaluating variables related to the learning process of students [18].

In 2010 Zhang etal[1] work to improve student retention in

higher education by using Naïve Bayes and decision tree techniques of data mining.

In 2011 Bhardwaj and Pal[7] describe the student performance analysis by using the classification techniques.

In february 2012 Mohd m Abul Tair;Alaa M.El-Hales [19]shows for the improvement of performance of graduate student by techniques like association rule,classification,clustering.

Umamaheswari.ks.Niraimathi [12] in 2011 categorise Student for Recruitment Process by applying classification techniques,association rule and clustering.

Oktariani Nurul pratiwi[8]2013 work for predicting the students for placement.

Irfan Ajmal Khan and Jin Tak choi[15] in 2014 work for the finding the student class for scholarship prediction. Sayali Rajesh Suyal Mohini Mukund mohad Aril[5] in the same year shows the Quality Improvisation of Student Performance.

Some study also show the percentage of accuracy in predicting the placement of problem statement

Table 1 Comparison Between Different Methods For Predicting Accuracybased On Percentage

S.N o.	Author	Techniques	Result
1	Oktariani Nurul pratiwi 2013[8]	OneR J48 K.Star Naïve Bayes	78.66% 79.61% 74.52% 76.75%
2	Ajay Shiv Sharma,S.S 2014[20]	Logistic Regression	83.33%
3	Vikas Chirumamilla, B. S. (2014).[13]	Naïve bayes C4.5	66.18% 77.78%

4. CONCLUSION

Partial In this survey paper it is analyzed that different techniques are used in literature survey to find the performance and find the class of students for placement prediction .This predictive model predicts the future scope of each student. Performance of student helps placement officer to guide student to choose their right career in which their talent or commands are good. Mostly companies hire the student according to the skill like technical, communication skill, CGPA marks, his behavior and many more features. And if the academic planners organize workshop in advance in pre final year for student then student will update them. This will improve the overall productivity.

5. FUTURE SCOPE

As per as the future aspect it will be better if we are revise our curriculum system and other activities for every semester in each year according to the public ,private or government sectors companies requirement. These major changes will improve the overall performance such as choosing the right careers option after completion of studies. By considering such the parameters of students would result in better accuracy of prediction.

6. REFERENCES

- [1] Dutta, r. j. (june 2013). "A survey on educational data mining and research trends department". *International Journal of Database Management system* .
- [2] K.Yacef, R. a. (2009). "The State of Educational Data Mining". *Journal of Educational Dta Mining* .
- [3] Kamber, J. a. (2000). "*Data Mining: Concepts and techniques*
- [4] Kumar, J. (2015). "A Comprehensive Study of educational Data mining". *IJEECSE* .
- [5] Mohod, S. R. (April 2014). "Quality Improvisation of Student Performance using Data Mining Techniques". *International Journal of Scientific and Research Publication* Vol 4 .
- [6] Naqvi, S. a. (2006). "Factor affecting student's performance:ACase of Private College". *Bangladesh e-Journal of Sociology*,vol 3,no1 .
- [7] Pal, B. K. (2011). "Mining Educational Data to Analyze Students Performance". *IJACSA* .
- [8] Pratiwi, O. N. (2013). "Predicting student Placement Class using Data mining". 2013 IEEE International Conference on Teaching ,assessment and learning for Engineering(TALE), (p. 618). Bali Dynasty Bali Resort,Kutta Indonesia.
- [9] Q.A.Al-Radaideh, E. a.-N. (2006). "Mining student data using decision trees". *International Arab Conference on Information Technology(ACIT 2006)*. Yarmouk University Jordan.
- [10] Rakesh Kumar Arora and Dr, D. B. (jul 2014). "Placement Prediction Through Data Mining". *IJARCSSE*
- [11] S.Niramathi, U. a. (August 2013). "A Study on Student Data Analysis Using Data Mining Techniques". *IJAECSSSE*
- [12] S.Pal, U. a. (2011). "Data Mining :Aprediction of performer or underperformer using classification"(IJCSIT). *IJCSIT* .
- [13] Vikas Chirumamilla, B. S. (2014). "A Novel approach to predict student placement Chance with Decision Tree induction". *IJST* .
- [14] www.educationaldatamining.org. (n.d.).
- [15] Chori, I. A. (2014). "An application of Educational Data Mining Techniquesfor scholarship Prediction". *IJSEIA* , 31-42.
- [16] C.M.Vera, C. a. (feb 2013). "Predicting School Failure and drop out by using Data Mining Techniques". *IEEE Journal of Latin America Learnin g Technology* .
- [17] http://en.wikipedia.org/wiki/Educational_data_mining.
- [18] el-Halees, A. (2009). "Mining student's data to analyze e learning behaviour:A case Study".
- [19] Jay Ruby, K. (2014). "Predicting the performance of students in higher education using Data Mining Classification:A Case Study. *IJRASET* .
- [20] Ajay Shiv Sharma, S. P. (2014). "Placement preidction system using Logistic Regression". *IEEE* .