TABLE OF CONTENTS

THE DATA MINING AND ANLYTICS FOR EDUCATION IN CURRENT ERA

Asmita R. Namjoshi
Asst. Prof. Tilak Maharshtra Vidyapeeth, Pune
asmita03@gmail.com

ABSTRACT

Educational data mining is a process by which we can identify the path of student and also it made a student predictable. This paper also gives light on different behaviour pattern of student in different social and educational area. It also gives focus on subject specialization for higher studies predictions.

KEYWORDS: *Data Mining, Data Mining Student, Student Performance Analysis, Student Ability Indexing.*

I. INTRODUCTION

Data Mining in terms is likely to be understood as uprooting data from past and analyze that for future use. The various books and research paper has been printed for discussing about different algorithm for doing data mining. This paper is projected to show how data mining and data analytics can be useful for educational domain.
The term is taken like EDM [1] Educational Data Mining [2].

This refers to nothing until you gather data and store. This refers how to use this data for better results of future students. For an instance- in a school there may be several students comprises of different classes and at the same different teachers with different subject to be taught. It is variety of combination like, some students want Physics teacher more comfortable and some took hindi or some may prefer marathi teacher. Now the problem is that why there is perception of liking. This is a common phenomenon of having good and bed perception about a teacher on basis of old student's experiences. Educational[3] data

mining is emerging as a research area with a suite of computational [4], [5], [6], [7] & [8] and psychological methods and research approaches for understanding how students learn. Now a day's computer supports interactive learning methods and tools, intelligent tutoring systems, simulations, games and has opened up opportunities to collect and analyze student data to discover patterns and trends in those data, and to make new discoveries and test hypotheses about how students learn. Data obtained from online learning systems can be aggregated over large numbers of students and can contain many variables that data mining algorithms can explore for model building

EDM ACHIEVEMENT

1. Know about the behavior and learning pattern
2. Findings specialization skills.
3. Identify lope holes, bridging that may student achieve bigger goals.
4. Head the system towards learning methodology enhancements.

5. Identification of human fast learning methods.
6. Select the process which can easily merge with traditional methods and gives better learned scholar.

II. NEED OF STUDY

My study is totally dedicated to the process of finding best student from the past performance of a candidate on two platforms i.e. Educational and Social.

In brief the study evolves in the process of creating a person who is not only a good scholar as well a predictable social human too. Analyzing students' data and information to classify students, or to create decision trees or association rules, to make better decisions or to enhance student's performance is an interesting field of research, which mainly focuses on collecting, analyzing and understanding student's educational data, it indicates their educational performance, and generates specific rules, classifications, and predictions to help students in their future educational performance.

III. RELATED WORK

1. Baradwaj and Pal [9] conducted a research on a group of students enrolled in a specific course program across a period of 4 years (2007-2010), with multiple performance indicators,
   • Previous Semester Marks,
   • Class Test Grades,
   • Seminar Performance,
   • Assignments,
   • General Proficiency,
   • Attendance,
   • Lab Work and
   • End Semester Marks"
2. Abeer and Elaraby [10] conducted a similar research.
3. Pandey and Pal [11] conducted a data mining research using some pattern for performance prediction.

IV. PROBLEM STATEMENT

Basic problem is very clear and spectacular. A question must be asked by every Indian that "What the reason that our student does not perform is best while study in INDIA but when go abroad they perform well and also involved in many research and development program.

It is clearly understood that whether the condition of ours are not up to the level or we do not understand the current scenario of education. If we go to 100 years back, at that time industrialization was on move, but here we were not in touch with and was only using material provided by them

Our education system was not coped up with current trends. Why? Unsolved till now. Why as a nation we are unable to find out best students to work. We are still unable to identify areas by emphasize on that we can predict student and its ability.

V. RESEARCH METHODOLOGY

Mining process is the process which is focused on creating relativity [11], [12], [13] & [14] of student in between his

1. Social
2. Educational
3. Performance issue due to above both factors.

For achieving goal we have done a survey on 100 students, enrolled in different course in same year of Tilak Maharashtra Vidyapeeth, Pune, and MH, INDIA.

Mix that questionnaire with their academic results sheets and then prepare relation on above mentioned relation.

Tabular data: A tabular data has been generated based upon their question answer by student. Questionnaire had some 100

question from different section related to our study. This was represented in following form-

Table 1: Questionaries' Elements

| Keys	Details
Sex	Gender details
Nationality	Country residents status
Lang	Talking language
Teaching Lang	Teaching language in Vidyapeeth
First Year %	Percentage acquired
Status of study	Result status at final
Address	Address of locality where he / she lives.
Parent Occupation	Parents work.
Parent employed in university	Yes/no
Caste	Caste belongs to.
Caste discount	Discount type availed or not
Siblings	Sister/brother details
Family Income	Income in total and earning member details
Father education	Father qualification
Mother Education	Mother qualification
Friends	No of friends and the status of friendship.
No of hours in home study	Hours counting only
No of hours outdoor games	Hours counting only
No of hours spends with relatives	Hours counting only
If working then no of hours of working	Hours counting only
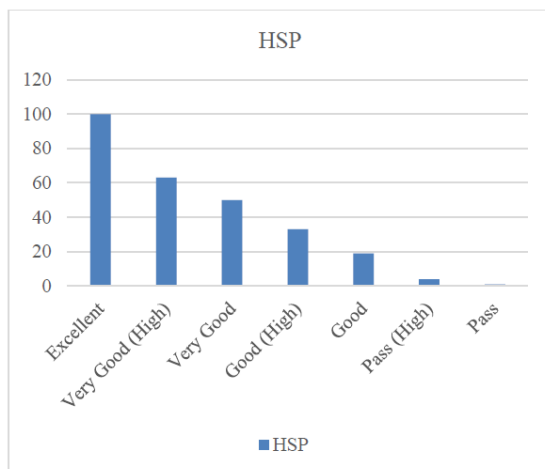
Figure 1: Histogram of HSP attribute
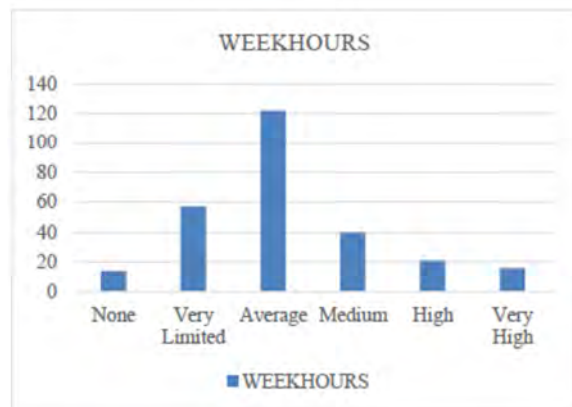

Figure 2: Semester performance attribute

Figure 3: Week hours distribution (based on last four questions)

VI. CONCLUSION

In this paper, several data mining algorithm were used to create models which could efficiently able to predict the student's.

1. First, a survey was taken from targeted Vidyapeeth students and collected multiple personal, social, and academic data.
2. Second, the collected tabular data was preprocessed and explored to become appropriate for the data mining.
3. Third, the implementation of mining tasks was presented on the tabular data in hand to generate classification models and testing them.
4. Finally, interesting results were drawn from the models, as well as, interesting patterns too.
5. Decision algorithms have been implemented, in our current study, it was sought that student's performance is not totally dependent on their only

academic efforts, in addition to that, there are many other factors that have equal to greater influences as well.

At end, this study can motivate and help all Indian universities to perform tasks on their students' regularly to find out interesting results and which can help both the university as well as the students in many ways.

VII. FUTURE SCOPE

Similarly with the same tabular data we can do more data mining like applying clustering algorithm to find out list of clustered student then creating clustered based questionnaire. Which may give more accurate and interesting result about the student learning pattern?

REFERENCES

[1] The State of Educational Data Mining in 2009: A Review and Future Visions. Baker, Ryan S. J. D.; Yacef, Kalina, Journal of Educational Data Mining, v1 n1 p3-16 2009.

[2] Top 10 algorithms in data mining, Xindong Wu,Vipin Kumar, J. Ross Quinlan,Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan Angus Ng, Bing Liu, Phlip S. Yu, Zhi-Hua Zhou, Michael Steinbach, David J. Hand, Dan Steinberg., Received: 9 July 2007 / Revised: 28 September 2007 / Accepted: 8 October 2007 Published online: 4 December 2007 © Springer-Verlag London Limited 2007

[3] http://edtechreview.in/dictionary/394-what-is-educational-data-mining [Accessed Feb 27, 2018].

[4] ALEVEN, V. and KOEDINGER, K.R. 2001. Investigations into help seeking and learning with a Cognitive Tutor. In Proceedings of the AIED-2001 Workshop on Help Provision and Help Seeking in Interactive Learning Environments, 47-58. R. LUCKIN Ed.

[5] BAKER, R.S., CORBETT, A.T. and KOEDINGER, K.R. 2004. Detecting Student Misuse of Intelligent Tutoring Systems. In Proceedings of the 7th International Conference on Intelligent Tutoring Systems, Maceio, Brazil, 531-540

[6] Baker, R.S.J.D., BARNES, T. and BECK, J.E. 2008. 1 st International Conference on Educational Data Mining, Montreal, Quebec, Canada.

[7] BARNES, T. 2005. The q-matrix method: Mining student response data for knowledge. In Proceedings of the AAAI-2005 Workshop on Educational Data Mining.

[8] BARNES, T., DESMARAIS, M., ROMERO, C. and VENTURA, S. 2009. Educational Data Mining 2009: 2nd International Conference on Educational Data Mining, Proceedings, Cordoba, Spain.

[9] Baradwaj, B.K. and Pal, S., 2011. Mining Educational Data to Analyze Students' Performance. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.

[10] Ahmed, A.B.E.D. and Elaraby, I.S., 2014. Data Mining: A prediction for Student's Performance Using Classification Method. World Journal of Computer Application and Technology, 2(2), pp.43-47 .

[11] Pandey, U.K. and Pal, S., 2011. Data Mining: A prediction of performer or underperformer using classification. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2), 2011, 686-690.

[12] Pandey, U.K. and Pal, S., 2011. Data Mining: A prediction of performer or underperformer using classification. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2), 2011, 686-690.

[13] ZAÏANE, O. 2002. Building a recommender agent for e-learning systems. In Proceedings of the International Conference on Computers in Education, 55–59.

[14] WITTEN, I.H. and FRANK, E. 1999.Data mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann, San Fransisco, CA. ZAÏANE, O. 2001. Web usage mining for a better web-based learning environment. In Proceedings of conference on advanced technology for education, 60-64.