

A Survey on Performance Evaluation and Prediction of Teachers', Students' Performance Using Data Mining Techniques

Sudarshan B. Wadkar¹, Dr. S. C. Dharmadhikari², Santosh Kumar Dwivedi³

ME Student, Department of Information Technology, PICT, Pune, India¹

Head of Department, Department of Information Technology, PICT, Pune, India²

Director, Anomaly Solutions Pvt. Ltd., Pune, India³

Abstract: Making use of Data mining (DM) in education is a rising interdisciplinary research field often referred as educational data mining (EDM). The principal objective of any educational institute is to provide excellent education to its students. One method to gain perfect level of quality in higher education is discovering knowledge that evaluates teachers' and students' performance and tries to predict areas of improvements and strengths. To satisfy this objective, EDM plays a vital role. The proposed work offers an effective approach for evaluation and prediction of teachers' and students' performance in institutions of learning using data mining technologies. The proposed system also acknowledges improvement areas of student as well as teachers' and is able to recommend respective training to them.

Keywords: Data Mining, Educational Data Mining, Performance Prediction, Evaluation, Training Recommendation.

I. INTRODUCTION

Data mining is the field of discovering knowledge and useful information from huge amounts of data. Data mining can be applied to various number of fields, including retail sales, bioinformatics, and counter-terrorism.

Education institute has faster growth in their data and it results into new problem that how to use this data for improving quality of education, services and making upper level decisions in institute. To solve this problem institute makes use of various formal and informal methods based on "qualitative and quantitative" approaches but these methods are not much helpful in achieving the quality objectives of institute.

However, methods used in higher education for quality purposes are mainly based on predefined queries and charts to analyse the data. In addition, these methods lack the ability to reveal useful hidden information [1].

To solve this problem more precisely data mining provides new emerging field called Educational Data Mining (EDM). EDM deals with developing methods that exploits data from educational institute with aim to provide quality education, service and making better decision.

The EDM helps to gain some additional insights from educational entities such as: teachers, staff, student, alumni. These entities are more helpful to distribute staff and resources more effectively and precisely that results in making better decision to increase students' success rates, increase learning outcomes, decreasing drop-out rate [2] [3].

To meet quality education, the higher education institute faces common problem of how to make evaluation of teachers' performance in respective course. A simple approach to evaluate teachers' performance is through questionnaire and students' responses about the course.

Teachers' performance evaluation through students' responses were first introduced in 1920 and then onwards an ongoing debate on validity and reliability of these evaluations. This happens because of mainly two reasons:

- 1) To evaluate teachers' performance and course, students don't have enough maturity and experience.
- 2) The popularity of teacher, grades given by teacher and popularity of course has direct impact on students' evaluation [4] [5].

Generally the course evaluation includes questions about teacher, course and overall assessment of the course, and this evaluation is used to observe and increase quality of teaching (Cronbach, 1963, Kulik, 2001, Nasser and Fresko, 2002).

Rest of the paper is organised as- section II is literature survey discusses about the previous study/ research made by some researchers and is helpful for understanding the history of the study. Literature survey tells the history and root of problem statement how the problem statement has formed. Section III gives proposed system architecture and it explains components of the system. It also tells how these components are inter-connected and how the communicate with each other. Section IV concludes the study.

II. LITERATURE SURVEY

Anwar Muhammad Abaidullah, Naseer Ahmed, and Edriss Ali [1] has proposed system that uses k-means clustering for analysis of students’ feedback data for making effective decision by upper level committee members responsible for observing and making assessment of educational quality and for improving teaching quality and improving quality of students’ learning experience. They have discussed model for using clustering to improve students’ learning experience that would result in improvement in quality of educational environment of institute.

Mustafa Agaoglu [6] discussed different data mining techniques and its use in Educational Data Mining (EDM) to predict teachers’ performance in higher education. In his study, he discussed four classification techniques - decision tree algorithm, artificial neural networks, support vector machines and discriminant analysis. He has used dataset consists of students’ responses to real course evaluation and applied these classifier models to the data set and results are compared according to precision, recall, accuracy and specificity. According to him performance of C5.0 classifier is best with respect to precision, recall, accuracy and specificity. The analysis of his study shows that teachers’ success using students’ perception depends on students’ interest in respective course.

B. K. Baradwaj and S. Pal [7], proposed system that describes performance of student in semester end examination which is helpful in identifying possible dropouts and students who needs special attention, so that teacher can appropriate training/advising to them. The decision tree method algorithm was applied by the authors on students’ database to predict the division of respective student. The study made by them will help to both students and teachers to improve individual division of student as well as to improve overall result of a class. This study will also work to identify those students who need special attention to reduce drop-out rate and taking necessary and proper action for next semester examination.

Chin-Chia Hsu and Tao Huang [8] conducted a study on the use of data mining technology to evaluate student’s academic achievement via multiple channels of enrolment like joint recruitment enrolment, athletic enrolment and application enrolment. A similar study was carried out by Osofisan and Olamiti [9] where they investigated the academic background in relationship with the performance

of students in a computer science programme in a Nigerian university. Their study showed that the grade obtained from senior secondary school examination (SSCE) in mathematics is the highest determinant of students’ performance using the C4.5 learning algorithm in building the model of the student’s performance.

To investigate factors associated with the evaluation of teachers’ performance, Mardikyan S., and Badur B. [10] conducted a study that uses two different data mining techniques: I) Decision tree and II) Regression analysis. For decision tree two algorithms was used CART and CHAID and for regression analysis, stepwise regression method applied. Their study showed that the teachers’, who have well prepared the outline of course, having satisfactory materials, who give grades equally and help students outside the class, attracts more students and receives higher evaluation. The result of this study can also be used while assigning teachers to the course or while awarding teacher.

Surjeet Kumar Yadav And Saurabh Pal [11] used ID3, C4.5 and CART decision tree algorithms on data set consists of engineering student to predict students performance in final exam. Prediction models that included all psychological, personal, social and other environmental variables were necessary for the effective performance prediction. The result of this study shows that C4.5 has more accuracy of 67.7778: as compared to ID3 and CART algorithms. The result of the study provides guideline to the students’ performances who were predicted to promote next year and who were fail. They have made comparative analysis of result gained by these algorithms and result shows that prediction helped weaker students to improve and brought out enhancement results. Typically course evaluations include questions about the course and the teacher and an overall evaluation, are widely used to supervise and improve teaching quality (Cronbach, 1963, Kulik, 2001, Nasser and Fresko, 2002). This topic has a lengthy literature, including survey design (Aleamoni, 1978, Bangert, 2004, Bangert, 2006, Cashin, 1995), and data-driven studies on which teacher and course attributes have direct impact on the overall assessment (Feldman, 1976, Feldman, 1978, Feldman, 1978, Marsh, 1982, Marsh, 1980, Cohen, 1981, Feldman, 2007, Bedard and Kuhn, 2008, Badur and Mardikyan, 2011). More researchers has shown that high ratings are given to upper-year courses, small classes, and teachers’ having more experience, passionate and energetic and who responds to questions more effectively [12].

Title	Publication & Year	Author	Merit	Demerit
Identifying hidden patterns in students' feedback through cluster analysis	Int. J. Comput. Theory 2015	A. M. Abaidullah, N. Ahmed, E. Ali	Enhance learning experience of students that would ultimately improve the quality of educational environment of an educational institution.	-

Data mining application in higher learning institutions.	Inform. Edu.-Int. J. 2007	N. Delavari, S. Phon-Amnuaisuk, M. R. Beikzadeh	It presents the capabilities of data mining in the context of higher educational system by proposing an analytical guideline for higher education institutions to enhance their current decision processes. The DM-HEDU is used to analyse the current works of data mining in education and identify the existing gaps.	-
Mining educational data to analyse students' performance.	Int. J. Adv. Comput. Sci. 2011	B. K. Baradwaj S. Pal	Describes students' performance in end semester examination. Identifies those students which need special attention to reduce fail ration and taking appropriate action for the next semester examination.	Analyses only students' performance but not teachers'. Doesn't tell what to do to students to increase their performance.
Teachers' Performance Evaluation In Higher Educational Institution Using Data Mining Technique.	IJAIS 2016	M.O., Asanbe, Osofisan A.O., William W.F	It shows that data mining techniques can be applied in the prediction of teachers' performance. It also shows some of the factors that contribute mostly on the performance of teacher.	Has focus only on teachers' performance. Doesn't provide any suggestion to increase teacher's performance.

III. PROPOSED SYSTEM ARCHITECTURE

Every educational institute aims at delivering quality education to their students, to meet this institute must able to evaluate teachers' as well as students' performance so that they can provide appropriate guideline to student and can able to arrange proper training for teachers also. Many researchers have developed systems which able to evaluate students' performance but improving students' performance is not the sufficient to provide quality education as teacher plays an important role in educating student. So, the proposed system can able to evaluate students' as well as teachers' performance and it also recommend respective training to them.

The proposed system architecture consists of following components:

- A. Database (Students information, Teachers' information and course information).
- B. Data mining Techniques (Attribute selection and classification algorithm).
- C. Prediction and Recommendation Engine.
- D. Users

A. Database
 Database consists of engineering student, teacher and courses offered in college. The database is divided into three categories- student information, teachers' information, and course information. Students' information contains all data about student including personal information (roll number, name, address, gender

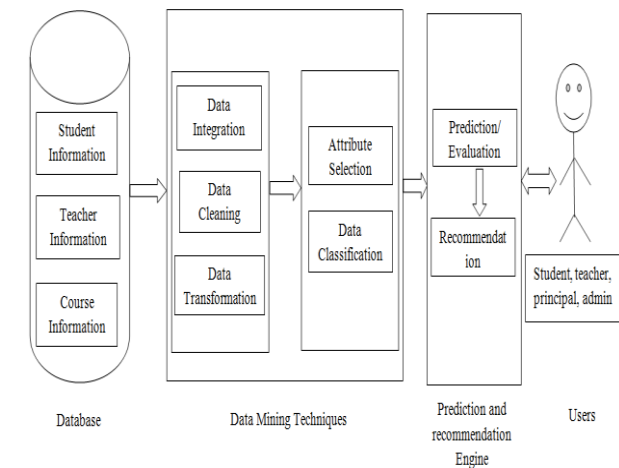


Figure 1: Proposed System Architecture

etc.), academic details includes marks of SSC, HSC, Diploma and marks obtained by student in each semester and information about project and other courses if any. Teachers' information includes his/her personal details, educational details, work experience etc. and rewards and achievements earned by teacher if any. Course information includes some details about teacher to whom course is assigned and students' performance in associated course, number of enrolled students to course etc.

B. Data Mining Techniques
 Various data mining techniques are available for classification such as - decision tree algorithms, support vector machines (SVM), artificial neural networks (ANN),

and discriminant analysis (DA). Agaoglu, Mustafa has shown in his study “Instructor Performance Using Data Mining Techniques In Higher Education” that the decision tree algorithm (C4.5) gives better result in terms of accuracy, precision and recall. So, the proposed system uses C4.5 decision tree algorithm for classification.

C. Prediction and Recommendation Engine

This module uses Apache PredictionIO Machine Learning Server for prediction and evaluation. PredictionIO consist of the following components:

- i. PredictionIO platform – It is an open source machine learning stack used for building, evaluating and deploying engines with machine learning algorithms.
- ii. Event Server – It is an open source machine learning analytics layer for unifying events from multiple platforms.

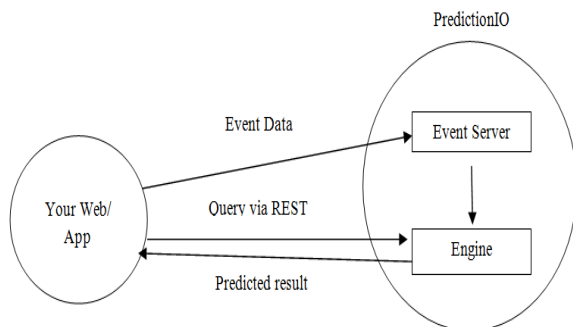


Figure 2: Prediction IO system.

D. Users

Users are basic components of any System, it defines who are going to use the system and to whom it is useful. The proposed system in this study focuses mainly on predicting and evaluating teachers’ and students’ performance so that appropriate action should be taken by authority to increase education quality, performance and overall results. Users in proposed system can be categorized into three categories: Student, Teacher and administrator/principal.

Students are the first user of the system, as system focuses on evaluating and predicting students’ performance student’s can see their own performance and check whether they can perform well or not in upcoming examination. If student is going to fail then an appropriate guideline can be provided by the teacher so that student can improve their performance which results in helping them from dropping out. Students have the authority only to see his/her won result and the notices arranged by teacher. Students can identify their learning tasks, resources and activities to improve their learning/ performance.

Teacher’s plays a key role in proposed system. Teachers can view the result generated by prediction tool and take proper decision to improve the performance of student and help them from dropping out. Depending on the result

teacher can give proper attention; can able to suggest suitable training to that particular student.

Administrator/principal can view the class wise result of students and accordingly arranges the notices for the teachers. He can also use them to decide which course to offer. Administrator can also evaluate teachers’ performance which results in assigning course to the teacher who is more effective to teach that course and it results in improving overall results of class; increasing education quality.

IV. CONCLUSION

Aim of every educational institute is to improve its education quality and improve overall result of the institute. This study/ paper have a focus on how educational quality can be improved and what factors must be considered to increase quality of education. The study not only focuses on improving individual performance of student but also improving teachers’ performance. To improve performance it is necessary to predict current status of student and teacher so to predict status prediction and recommendation engine is used.

In this study C4.5 algorithm used because of its better result in terms of accuracy, precision, recall. To meet more accuracy C4.5 algorithm applied recursively on all attribute. To predict the result an open source PREDICTIONIO tool is used. PREDICTIONIO is used with recommendation to recommend necessary training or guideline to the student as well as teacher.

This study can be used as a real-time application in any educational institute for improving their educational quality, to improve performance of student and teacher, for making course evaluation. The scope of this study is not only predicting performance but also recommending proper training or guidance to students. However, it can be concluded that this method can be used to reduce drop-out rate by taking proper and necessary steps at right time to improve quality of learning.

REFERENCES

- [1] A. M. Abaidullah, N. Ahmed, and E. Ali, “Identifying hidden patterns in students’ feedback through cluster analysis,” *Int. J. Comput. Theory Eng.*, vol. 7, no. 1, pp. 16-20, 2015.
- [2] N. Delavari, S. Phon-Amnuaisuk, and M. R. Beikzadeh, “Data mining application in higher learning institutions,” *Inform. Edu.-Int. J.*, vol. 7, no. 1, pp. 3154, 2007.
- [3] M. Goyal and R. Vohra, “Applications of data mining in higher education,” *Int. J. Comput. Sci. Issue*, vol. 9, no. 2, pp. 113120, 2012.
- [4] L. Coburn., (1984), “Student Evaluation of Teacher Performance, ERIC/TME Update Series,” [Online]. Available: <http://ericae.net/edo/ED289887.htm>
- [5] S. A. Radmacher and D. J. Martin, “Identifying signicant predictors of student evaluations of faculty through hierarchical regression analysis,” *J. Psychol.*, vol. 135, no. 3, pp. 259-269, 2001.
- [6] Agaoglu, Mustafa. "Predicting Instructor Performance Using Data Mining Techniques In Higher Education". *IEEE Access* 4 (2016): 2379-2387. Web. 16 Nov. 2016.



- [7] B. K. Baradwaj and S. Pal, "Mining educational data to analyze students' performance," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 6, pp. 63-69, 2011.
- [8] Chin Chia Hsu and Tao Huang (2006): "The use of Data Mining Technology to Evaluate Student's Academic Achievement via multiple Channels of Enrolment." An empirical analysis of St. John's University of Technology.
- [9] Osofisan A.O. and Olamiti A.O. (2009): "Academic Background of Students and Performance in Computer Science Programme in a Nigerian University." *European Journal of Social Science*. Vol. 33 Issues 4. 2009.
- [10] MardikyanS., and Badur B. (2011). "Analyzing teaching Performance of Instructors Using Data Mining techniques. *Informatics in Education*," 2011, Vol. 10, No. 2, pp 245 – 257.
- [11] Surjeet K.Y and Saurabh P (2012): "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification," *World of Computer Science and Information Technology Journal (WCSIT)* ISSN: 2221-0741 Vol. 2, No. 2, 51-56, 2012.
- [12] JIANG, Yuheng Helen, Sohail Syed JAVAAD, and Lukasz GOLAB. "Data Mining Of Undergraduate Course Evaluations". *INFORMATICS IN EDUCATION* 15.1 (2016): 85-102. Web. 11 Nov. 2016.

BIOGRAPHIES

Sudarshan B. Wadkar, received B. Tech degree in computer science and Technology, from Shivaji University, Kolhapur in 2014. He is currently pursuing M.E. degree in Department of Information Technology from Pune Institute of Computer Technology, Pune. His area of interest includes Data Mining and Machine Learning.

Dr. S. C. Dharmadhikari, Head of Information Technology Department, Pune Institute of Computer Technology, Pune. She has received M.E. degree in Computer Engineering from Bharati Vidyapeeth, Pune and completed P.hD in computer from Devi Ahilya Vishwa Vidyalaya, Indore. Her research interest is Machine Learning, Data Mining and System Programming.

Santosh Kumar Dwivedi, he is a Director and founder of Anomaly Solutions Pvt. Ltd., Pune, India.