

## International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

IMPACT FACTOR: 6.017

*IJCSMC, Vol. 7, Issue. 12, December 2018, pg.158 – 162*

# Student Performance Prediction Techniques: A Review

**Shalini Kumari**

Research Scholar, Bahra University, Wagnaghat, Shimla Hills

[raghuwanshishalini186@gmail.com](mailto:raghuwanshishalini186@gmail.com)

**Parul Gazta**

Assistant Professor, Bahra University, Wagnaghat, Shimla Hills

[parulgazta@gmail.com](mailto:parulgazta@gmail.com)

*Abstract: The prediction analysis is the approach which can predict future possibilities based on the current information. The prediction analysis can be done using the technique of classification and neural networks. 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.*

**KEYWORDS:** *Student performance, prediction, Classification*

## Introduction

There is no limit on the data storage; the data stored is so large that it is almost impossible to fetch relevant and important data on time. The data can be stored in files, hard disks, CD drives, databases and several other applications. The fetching of relevant and important file from this infinite amount of data is called Data Mining. Educational Data Mining is the very new and advanced method being employed by the data mining and Knowledge Discovery in Databases (KDD). Its main objective is to focus mainly on the useful patterns and recognize the useful and relevant data from the educational information systems, like admission system, registration, course management, syllabus management and many other systems. These systems and projects deals with the students in different stages of educational institutions like schools, colleges and universities [1]. Researchers are working in this field to discover and determine the relevant knowledge to help the educational fields to manage their smooth working of schools, colleges and universities. This will also enhances the better and improved

functioning of the institution by managing their activities. It studies the students' data and information by classifying the student's data with the help of decision tree or association rules. It is the better and interesting field of research which most commonly focuses on the observation and understanding of student's data which is related to their educational field [2]. This involves the analyzing the performance data, managing the proper functioning of the educational institutes. It makes sure that all the students are getting proper and sufficient facilities and they are not facing any problem in their education. It also includes the classification, in which is the most common data mining technique used to classify the educational institutions according to their performance. Educational Data Mining (EDM) is used in this research to study the collected student's information by a survey and classification is performed on the basis of the gathered data [3]. This classification, classify and predicts the student's performance in their coming semesters or upcoming examinations. The main objective of this research is identifying the relationship between the student's personal and social factors as well as in their academic performance analysis. This provides facility to the student and the institutions to work in an organized and synchronized manner, by analyzing the performance of every student and then they decide who is underperforming and who is performing well in their academics. This will help the educational institutions to work in a very proper and uniform manner [4]. 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 system can able to evaluate students' as well as teachers' performance and it also recommend respective training to them. Every educational field has the main objective to provide better and advanced knowledge to each and every student. These institutions are using developed and updated modes of education like whiteboards, projectors, smart study and many other modes are being employed in the institutions [5]. Various researchers are working on this field and they have evaluated the student's performance but the improved performance of the student is not sufficient to have a proper functioning of the institutions. They also need to develop such technology which can analyze and manage the complete working of the educational industries. Therefore, the proposed system helps to study the student's as well as teachers overall performance and suggest proper training of the teachers. This system architecture has sub-components. Database contains the list of students, teachers and the type of courses the particular institution is offering. It is further divided into three categories that is; student information, teacher information and the course information. Various data mining techniques have been employed for the classification like decision tree algorithm [6], support vector machines, artificial neural network and discriminant analysis. This module makes use of Apache PredictionIO Machine Learning Server, in order to predict and evaluate the performance. Within any system, the most common components involved as users. The individuals that are going to utilize the system and for whom they are using it can be defined here. Mainly, the performance of teachers and students are predicted and evaluated in the designed model. Thus, the education quality, performance as well as the overall results is improved by taking required actions. The first user of system is known to be the students. The performance of students is evaluated and predicted by the system. Thus, the individual performance of students can be checked and it can also be evaluated by them whether they can perform well in exams or not. The system is largely affected by the teachers. The result that is generated by prediction tool is viewed by the teachers. The performance of students is improved and they are helped in passing the course by the teacher by making appropriate decisions. The class-wise results of students are viewed by the administrator or principal. Notices are arranged for teachers as per these results [7]. Also the course that needs to be offered to students can be decided by this administrator.

## Literature Review

Jie Xu, et.al (2016) proposed a new machine learning technique using which the performance of students in degree programs is predicted. There are two important features included within the proposed technique [8]. The multiple base predictors are included to develop a bi-layered structure which is the first feature of proposed method. Also, on the basis of the evolving performance states of students, a cascade of ensemble predictors is designed using which predictions can be done. For discovering the course relevance, a data-driven technique is designed as a second feature. It uses probabilistic matrix and latent factor models. Simulations are performed on a dataset collected at UCLA over three years and the results are achieved which show that the performance of proposed approach is better as compared to existing techniques.

S. M. Merchán, *et.al* (2016) proposed a predictive model to be applied to predict the academic performance of students. Several data mining methods are applied on the data of 932 students of a university of Columbia to evaluate and analyze their performances [9]. On the basis of input data given, the expected results and output characterization and other factors, the evaluation of results achieved is done. The prediction accuracy is an important parameter to evaluate the performances as well. Considering the specific details of the population examined and the requirements specified by the institution, the said pertinence is evaluated. For preventing any kinds of academic risk and desertion, timely decisions are considered important along with the accompaniment of students with their learning procedure. This research work is developed further by providing few recommendations and thoughts by different researchers.

Ishwank Singh, *et.al* (2016) proposed a simple clustering analysis through which the behavior of student is understood. To understand if there is a regular improvement in the performance of student, a good benchmark has been set up by the data mining algorithm [10]. During the admission and placement process, this analysis is very helpful. The projects, internships, skill sets, Xth, XIIth, and B.Tech marks are few parameters included for this analysis. Since the implementation is easy and the computational efficiency is high, K-means algorithm is used in clustering. In future, other clustering techniques can be applied to improve efficiency levels. Also, to achieve a better student performance analysis, the ranking or classification of objects present within the clusters can be done.

Ms. Tismy Devasia, *et.al* (2016) proposed classification within the information of student such that on the basis of previously existing information, the division of students can be predicted [11]. Naïve theorem is applied since several techniques are used for knowledge classification within the area unit. For the prediction of performance at the top of that particular semester, various types of information were collected from the previous information of the students available. To encourage the students of different categories to perform well, the lecturers and students can be benefitted through this study. The students who need any special guidance can be highlighted through this study. Also, the failure ratio can be reduced with the help of this. For the upcoming semester examination, acceptable actions can be taken through this.

Yuni Yamasari, *et.al* (2016) proposed feature extraction techniques in this paper. The student data was gathered in a serious game to perform extraction on the basis of category and Bloom's Taxonomy [12]. The proposed approach is implemented on this data to perform certain evaluations. It is seen that the level of accuracy is improved and the execution time is minimized through this method. In comparison to traditional FCM, the level of accuracy is improved up to 2.3-4.7%. Also, in comparison to traditional approach, the execution time is 2.2-2.7 seconds faster for the proposed FCM approach. The performance of clustering process on the achievement of student is enhanced when the features are extracted using CBE\_FCM and BTBE\_FCM. Weight is added to each feature for improving the proposed methods. The correlation level of student achievement is considered to be an important factor here.

Nurul 'Ulyani, *et.al* (2017) presented that the major factor that leaves a huge impact on the behavioral intentions of student is the service quality performance. Within seven Malaysian public and private universities the paper-and-pencil questionnaires were distributed [13]. The descriptive statistics and covariance-based structural equation modeling were used to analyze the data. The least likely execution of favorable behavioral intentions was influenced by the freedom, serenity, management dimensions as well as aesthetic factors. A positive behavior towards the student housing was seen as per the results achieved when students adapted to live in multi-cultural community in which they would have access to good hospitality, personal privacy and appropriate building ambiance.

Author's Names	Year	Description	Outcomes
Jie Xu, Kyeong Ho Moon, and Mihaela van der Schaar	2016	A new machine learning technique is proposed using which the performance of students in degree programs is predicted.	Simulations are performed on a dataset collected at UCLA over three years and the results are achieved which show that the performance of proposed approach is better as compared to existing techniques.
S. M. Merchán, J. A. Duarte	2016	A predictive model to be applied to predict the academic performance of students is proposed. Several data mining methods are applied on the data of 932 students of a university of Columbia to evaluate and analyze their performances	On the basis of input data given, the expected results and output characterization and other factors, the evaluation of results achieved is done. The prediction accuracy is an important parameter to evaluate the performances as well.
Ishwank Singh, A Sai Sabitha, Abhay Bansal	2016	A simple clustering analysis is proposed through which the behavior of student is understood. To understand if there is a regular improvement in the performance of student, a good benchmark has been set up by the data mining algorithm.	Since the implementation is easy and the computational efficiency is high, K-means algorithm is used in clustering.
Ms. Tismy Devasia, Ms. Vinushree T P, Mr. Vinayak Hegde	2016	Classification is proposed within the information of student such that on the basis of previously existing information, the division of students can be predicted.	The students who need any special guidance can be highlighted through this study. Also, the failure ratio can be reduced with the help of this. For the upcoming semester examination, acceptable actions can be taken through this.
Yuni Yamasari, Supeno M. S. Nugroho, I N. Sukajaya, Mauridhi H. Purnomo	2016	Feature extraction techniques are proposed in this paper. The student data was gathered in a serious game to perform extraction on the basis of category and Bloom's Taxonomy	The performance of clustering process on the achievement of student is enhanced when the features are extracted using CBE_FCM and BTBE_FCM. Weight is added to each feature for improving the proposed methods.
Nurul 'Ulyani, Mohd Najib, Nor'Aini Yusof, Amin Akhavan Tabassi	2017	The descriptive statistics and covariance-based structural equation modeling were used to analyze the data. The least likely execution of favorable behavioral intentions was influenced by the freedom, serenity, management dimensions as well as aesthetic factors.	A positive behavior towards the student housing was seen as per the results achieved when students adapted to live in multi-cultural community in which they would have access to good hospitality, personal privacy and appropriate building ambiance.

## Conclusion

The prediction analysis is the approach which can predict future possibilities from the current information. The prediction analysis can be done with the techniques of classification. This review paper is based on the student performance prediction. The various classification techniques are reviewed in paper for the student performance prediction.

## References

- [1] Amin Zollanvari, Refik Caglar Kizilirmak, Yau Hee Kho, and Daniel Hernandez-Torrano, “Predicting Students’ GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviors”, 2017, IEEE
- [2] Sneha Chandra, Maneet Kaur,” Enhancement of Classification Accuracy of our Adaptive Classifier using Image Processing Techniques in the Field of Medical Data Mining”, 2015, IEEE
- [3] Yomna M. ElBarawy, Ramadan F. Mohamedt and Neveen I. Ghali,” Improving Social Network Community Detection Using DBSCAN Algorithm”, 2014, IEEE
- [4] Dominik Fisch, Edgar Kalkowski, Bernhard Sick,” Knowledge Fusion for Probabilistic Generative Classifiers with Data Mining Applications”, 2013, IEEE
- [5] Dianwei Han, Ankit Agrawal, Wei-keng Liao, Alok Choudhary,” A novel scalable DBSCAN algorithm with Spark”, 2016 IEEE International Parallel and Distributed Processing Symposium Workshops
- [6] Md. Rejaul Karim, and Dewan Md. Farid,” An Adaptive Ensemble Classifier for Mining Complex Noisy Instances in Data Streams”, 2014, 3rd INTERNATIONAL CONFERENCE ON INFORMATICS, ELECTRONICS & VISION
- [7] Karlina Khiyarin Nisa, Hari Agung Andrianto, Rahmah Mardhiyyah,” Hotspot Clustering Using DBSCAN Algorithm and Shiny Web Framework”, 2014, IEEE
- [8] Jie Xu, Kyeong Ho Moon, and Mihaela van der Schaar, “A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs”, 2016, IEEE
- [9] S. M. Merchán, and J. A. Duarte, “Analysis of Data Mining Techniques for Constructing a Predictive Model for Academic Performance”, IEEE Latin America Transactions, Vol. 14, No. 6, June 2016
- [10] Ishwank Singh, A Sai Sabitha, Abhay Bansal, “Student Performance Analysis Using Clustering Algorithm”, 2016, IEEE
- [11] Ms.Tismy Devasia, Ms.Vinushree T P, Mr.Vinayak Hegde, “Prediction of Students Performance using Educational Data Mining”, 2016, IEEE
- [12] Yuni Yamasari, Supeno M. S. Nugroho, I N. Sukajaya, Mauridhi H. Purnomo, “Features Extraction to Improve Performance of Clustering Process on Student Achievement”, 2016, IEEE
- [13] Nurul ‘Ulyani Mohd Najib, Nor’Aini Yusof, Amin Akhavan Tabassi, “Service Quality Performance of Student Housing: The Effects on Students’ Behavioural Intentions”, 2017 IEEE 15th Student Conference on Research and Development (SCORED)