

Predictive Evaluation of Student Marks using Machine Learning

M Chalapathi Rao¹, B Ramji², B P Deepak Kumar³

^{1,2,3}Assistant Professor, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India.

Abstract - This paper describes a predictive model that is used to predict the performance of student marks based on study hours. In this paper, regression classifier is implemented to predict the student's academic performance and the characteristics that influence a student's academic achievement are examined in order to forecast the student's performance. It helps them who need more attention. Some novel attributes such as Hours, scores etc are used for understanding their learning environment. In addition, an ensemble method is applied to improve the performance of the classifier. The obtained result shows that the proposed model performs better with credible accuracy rate among others.

Key Words: Classification algorithm, Grade, Marks, Machine Learning, Prediction, Linear Regression.

1. INTRODUCTION

Machine learning based methods are more useful to automate student marks prediction using linear regression technique. Early stage prediction is not possible in this way. Using machine learning automation the process of prediction can be done. Machine Learning is a subset of artificial intelligence that can be described as the application of computer algorithms to allow the computers learn automatically based on data provided and use it to learn for themselves. Various models can be proposed to predict students' performance in the registered courses, which gives valuable information to help students' and will lead to improving the students' learning process based on predicted marks. Machine Learning methods are used to extract knowledge from big data and analyse it for a specific reason. It is the analysis of examples and experience without the use of machines explicitly programmed. Machine learning is concerned with the development of computer programs that can access data and learn from it in order to derive valuable functions. Machine learning techniques can be used to analyse data from various perspectives and extract useful information from developers. To access the machine learning algorithms with other approaches, we may be able to improve students' performance outcomes in a more equitable way. Research groups, students and teachers may also have an adverse effect. Predicting student academic achievement and taking early actions to improve student performance and teaching quality is one strategy to accomplish this. The goal of this method is to use linear regression to predict a student's grade. The goal of this study is to forecast students' grades based on their study hours.

The rest of this paper is structured as follows. Section 2 discusses background and related work. Section 3 is elicited methodology. A brief study of classification algorithm is presented in section 4. Results and discussion is presented in section 5 and conclusion will be in section 6.

2. BACKGROUND AND RELATED WORKING

A new framework for predicting the raw data was pre processed in terms of missing value filling, converting values from one form to another, and selecting pertinent attributes and variables [1]. Data Mining, Educational Data Mining, Predictive Model, and Classification are some of the terms used in this paper.

Classification algorithms such as decision trees and Bayesian networks can be used on educational data to predict a student's exam success. This forecast will aid in identifying poor students and assisting them in improving their grades [2].

Particularly the data set is then subjected to five classes of Machine Learning Algorithms (MLA), with the decision tree class of algorithms yielding the best results. It was also discovered that the outcomes of this model's prediction are equivalent to those of other previously created models [3].

The study [4] aims to examine, a model for predicting student success in an academic organization is proposed. Neural Networks are a machine learning technology that was used in the algorithm.

The paper provided a series of classifiers, including Artificial Neural Networks, Nave Bayesian, and Decision Tree, are used to evaluate the performance of the student's predictive model [5]. In addition, to boost the performance of these classifiers, they used ensemble approaches.

A system review is presented in [6] for the academic performance of ITU students who were accepted into the Electrical Engineering department's bachelor's degree program. The RBM technique has been found to be superior to other techniques for forecasting students' success in a given course.

[7]. In an education research study, various elements such as social, economic, personal, cultural, geographical, institute environment, and others are assessed. Such factors may either assist a student in shining during the academic era or cause a student's academic program to come to a halt. Drop-out is the term for this type of failure.

In particular findings have been confirmed from student performance using their attendance and other grades [10]. In addition, data is analyzed and forecasted.

The current study will go over a well-known machine learning technique called linear regression. It also explains how to evaluate learning algorithms using a variety of performance metrics.

3. METHODOLOGY

The primary goal of this work is to create a model that is more efficient and correct. We strive to implement the model in such a way that the error rate is as low as possible, we also employ the ensemble learning technique to improve machine learning results. As a result, the model's correctness supervised machine learning algorithm that always operates is ensured. This work will also have an impact on curriculum design in educational fields in general.

This section provides student classification models that use students' fundamental characteristics and academic records from previous academic periods to forecast student achievement based on study Hrs. The proposed model is shown in Fig 1.

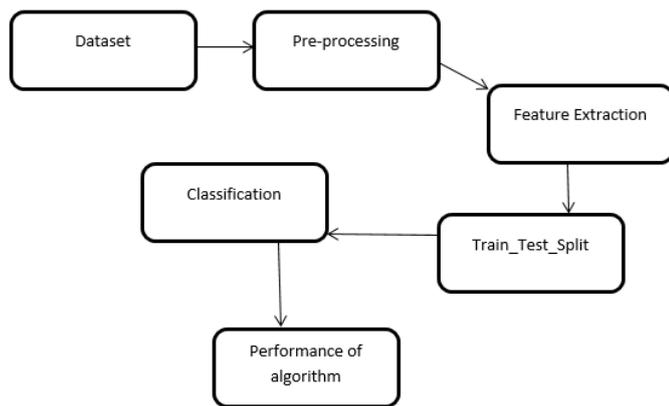


Fig 1: Proposed Model

The system also effectively assists trainers in conducting programs that benefit the student community.

4. CLASSIFICATION ALGORITHM

The primary goal of linear regression is to determine the best value for dependent and independent variables in a simple regression problem, the model would take the following form:

$$Y = mX + c$$

When we have more than one input in higher dimensions that is called a plane or a hyper-plane. The form of the equation and the specific values used for the coefficients are thus the representation. In regression different preparations of data using these heuristics like linear assumption, remove noise, remove co linearity, Gaussian distributions and rescale inputs. Three major uses for regression analysis are

determining the strength of predictors, forecasting an effect and trend forecasting. The main objective of linear regression is to find or predict the best fit line using least square method to do so the regression line would take the following formula to check how good a modulus performing using R-square method.

$$m = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$c = \bar{y} - m\bar{x}$$

So, in order to reduce the mistake, we must first figure out how to calculate it. A loss function is a measure of how far the anticipated value differs from the actual value in machine learning. To determine the loss or mistake in our model, we used the Quadratic Loss Function.

$$L(x) = \sum_{i=1}^n (y_i - p_i)^2$$

5. RESULTS AND DISCUSSION

The variables have been found to be positively correlated. A low mean absolute error value indicates that the model's chances of making errors or forecasting incorrectly are extremely low.

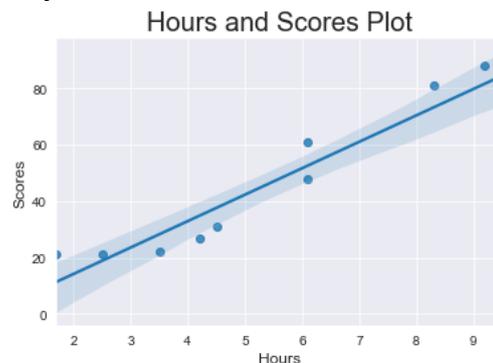


Fig2: Visually Comparing the Predicted Marks and Actual Marks

Hours	Predicted Output
4.2	30
8.3	81
4.5	41

Table1: Predicting the Percentage of Marks

According to the regression model, a student who studies for 8.3 hours per day is likely to receive 81.00 marks. Students are given grade 81.00 marks, as shown in the table above; however, this regulation applies to other grades too.

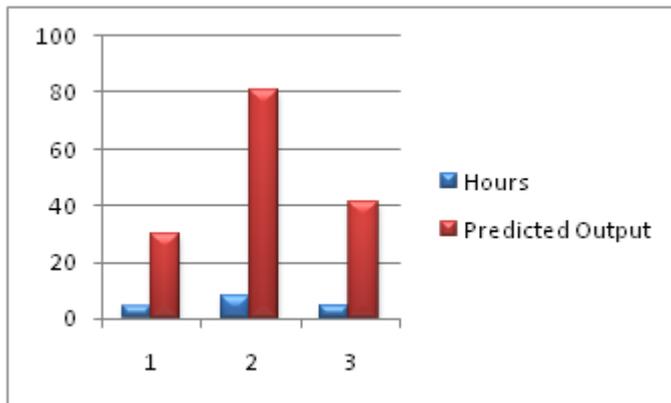


Fig3: Predicting the Percentage of Marks

The Fig 3 shows attributes for the used machine learning algorithm in the data source. From figure it is noticed that hours and predicted output.

Authentic Output	Predicted Output
27	30
81	81
31	41

Table 2: Comparing the Authentic Output and Predicted Output

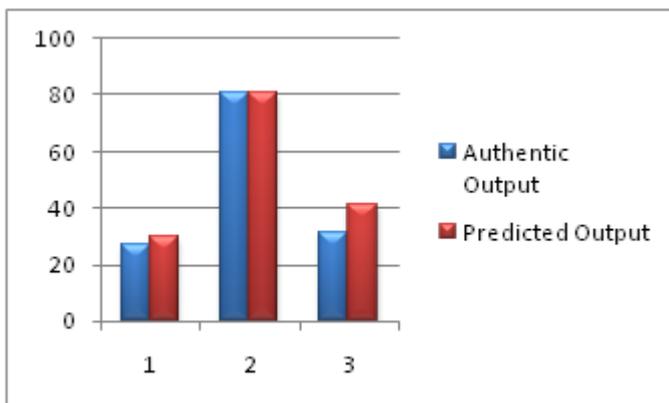


Fig 4: Comparing the Authentic Output and Predicted Output

The Fig 4 shows the accuracy of linear regression algorithm with existing data to emphasize the authentic output as well as predicted output. It may be made multivariate by adding other factors to get more precise results. The marks are then transformed to a uniformity benchmark of 100 (%). Furthermore, the marks are transformed into percentages to ensure that the system is accurate.

6. CONCLUSION

The Prediction of student grade performance assists tutors in making a successful understanding about how well or badly their students will perform in their learning process. We used real-time data in the training and testing stages of the proposed system to provide real-time implementation of the prediction using deep analytics based on students' databases. LR model is designed to make the most accurate predictions possible. This study' accuracy was higher than the original research's. Finally we observed that regression model gives optimum results in a structured way.

ACKNOWLEDGEMENT

The authors send their warm thanks to CMR Technical Campus, Hyderabad for providing extensive support for carrying this research work.

REFERENCES

- [1] B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.
- [2] S.K. Yadav, S. Pal, "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification" World of Computer Science and Information Technology Journal (WCSIT) 2012.
- [3] Anal Acharya, Devadatta Sinha, "Early Prediction of student performance using machine learning technique", International Journal of Computer Applications, December 2014.
- [4] H. Agrawal, H. Mavani, "Student Performance Prediction using Machine Learning" International Journal of Engineering Research & Technology Vol 4, Issue 3 (2015).
- [5] Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue, PoojaThakar Assistant Professor VIPS, GGSIPU Delhi, India, Anil Mehta, Ph.D Professor University of RajasthanJaipur,India, International Journal of Computer Applications, 2015.
- [6] Amrieh, E.A., Hamtini, T. and Aljarah, I., 2016. Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. International Journal of Database Theory and Application, 9(8), pp.119-136.
- [7] Iqbal, Z, Qadir, J., and Kamiran, F., "Machine Learning based student grade prediction: A case study," 17 Aug 2017.
- [8] Using Machine Learning to Predict Student Performance, Murat Pojon, University of Tampere Faculty of Natural Sciences Software Development M. Sc. Thesis Supervisor: JormaLaurikkala June 2017.

- [9] XU, J., MOON, K. H., AND VAN DER SCHAAR, M. 2017. A machine learning approach for tracking and predicting student performance in degree programs. IEEE Journal of Selected Topics in Signal Processing.
- [10] Machine Learning Approach to Predict Student Academic Performance-International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887 Volume 6 Issue IV, April 2018.
- [11] Solomon, D. (2018) 'Predicting Performance and Potential Difficulties of University Student using Classification: Survey Paper', International Journal of Pure and Applied Mathematics, 118(18), pp. 2703-2707.
- [12] Jeromie Reggie Ebenezer, R Venkatesan, K Ramalakshmi, Jewel Johnson, Glen P I, Vibin Vinod International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6S, April 2019.
- [13] Altabrawee, H., O. A. J. Ali, and S. Q. Ajmi. (2019) 'Predicting Students' Performance Using Machine Learning Techniques,' J. Univ. BABYLON pure Appl. Sci., vol. 27, no. 1, pp. 194-205.
- [14] Adekitan, A. I. and Salau, O. (2019) 'The impact of engineering students' performance in the first three years on their graduation result using educational data mining'. Heliyon 5 e01250.ggf22n.



Mr. B.P. Deepak Kumar received his M.Tech. in Computer science and Engineering from JNTUH. Currently, he is working as Assistant Professor in computer science and engineering department at CMR Technical Campus, Hyderabad. His research interests include IOT, Machine Learning and AI.

BIOGRAPHIES



Mr. M Chalapathi Rao received the M.Tech. (2008) in Software Engineering from JNTUH. Currently, he is working as Assistant Professor in computer science and engineering department at CMR Technical Campus, Hyderabad. During his carrier, he has been author or co-author of numerous technical papers. His research interests include Machine Learning, Information Security and Bug related issues in SE.



Mr. BANOTHU RAMJI received his B.Tech Degree in Computer Science & Engineering from KLR College Engineering & Technology affiliated by JNTU Hyderabad, Telangana, India in 2012, Master of Technology in Computer Science & Engineering from Campus college JNTU Hyderabad in 2015. His research areas include Internet of Things (IOT)/Data Mining.