An Analytical Model to Track, Analyze and Predict Scholar's Academic Performance

G. Vigneshwaran[1*], S. Simritha[2], R. Rasika[3], S. Sangeetha[4]

[1,2,3,4]*Department of Computer Science and Engineering, SNS College of Technology, Coimbatore, India*

Abstract*: The objective of this machine learning project is to classify and predict the future academic grades and leadership scores of the students through building a convolution neural network to predict the scores. The application works as a platform for exposing the semester marks of the students through machine learning technique. The main goal is to predict the academic performance using machine learning to develop a model to predict the student's semester grade result. The ability to predict student performance in education is very significant in educational environments. The stored database contains student's information to improve student's perspective and behaviour. Using that information, we can analyse the performance, which will help for both students and mentors. The system learns the Attendance of the student, Difficulty of the future subjects and previous performance of a student to predict the future semester grades with the help of attendance and activities. An institution needs to know the case history of their registered students of their institute to predict their performance. This will help mentors consolidate the student on improving and developing each student's curriculum record. It refers to performing various data produced by students in order to evaluate learning process like, predict the future performance and identify probable problems. The mentor can identify student's performance and can counsel the students to perform better, likewise the student can improve their performance in examinations. This analysis will be performed to improve educational process. The main intention is to identify and support the students to score better marks.

Keywords*: Academic prediction, Detection of grades, Score analysis, Tracking student performance.

1. Introduction

To track and analyze the performance of students' academic we used the concept of machine learning. The poor performance of task has been affecting the secondary exam to step into higher education. Prediction and classification of student performance offers a warning and gives the skills for improving the poor performance. The aim is to deal with behaviour pattern of students. Analysis of database in education environments are learning, predictive, data mining and machine learning techniques. It is used to classify the students' performance and behaviours to grow up on their progress. The challenging task is to produce satisfying results. Machine learning algorithms such as artificial neural network, support vector machine, k-nearest neighbour is used to predict the academic performance. In this problem, a reason for declines

the performance is quality of data. It focused on the prediction and classification performance which is poor or best. Here the work applied is Principal component analysis (PCA). 10-fold cross-validation also introduced to predictive performance of the models and to judge in a dataset. The aim of this paper is proposed hybrid approach of four baseline machine learning algorithms.

2. Literature Survey

Alyahyan, [1] extracted values using high dimensional learning process to provide a step-by-step set of guidelines for educators willing to apply data mining techniques to predict student success.

Ahmed O, [2] presents a comprehensive review of related studies that deal with SAP and dropout predictions. To group the studies, this review proposes taxonomy of the methods and features used in the literature for SAP and dropout prediction.

Francis, [3] used new prediction algorithm for evaluating student's performance in academia has been developed based on both classification and clustering techniques and been tested on a real time basis with student dataset of various academic disciplines of higher educational institutions

Francis, [4] prepared a structure which will analyse the pupil's performance from their last performances using concepts of Data Mining under Classification. Classification Algorithms like Decision Tree, Naïve Bayes and Support Vector Machine can help us for predicting student's performance.

Hussain, [5] prepared set of attributes are first defined for a group of students majoring in Computer Science in some undergraduate colleges in Kolkata. Since the numbers of attributes are reasonably high, feature selection algorithms are applied on the data set to reduce the number of features.

Raihana, [6] applied support vector machine (SVM) method for classifying the students. The results for each quality of life domain showed that students with both low and high academic performance were classified into high academic performance class. The same result was obtained when all domains were combined.

Hasan, [7] proposed work aid in improving student's grades in the module. Helping stakeholders to analyze and evaluate the module delivery and results. Early detection and solution can

be made both at the institutional level and module level.

Yassein, [8] captured and arranged with the use of statistical package for social sciences (SPSS) and data mining tool (clementine). Developing an accurate student's performance prediction model is challenging task. Data mining based model were used to identify which of the known factors can give an early indicator of expected performance.

3. Proposed Method

The application predicts the grades and leadership scores of students based on their historical data, the upcoming difficulty level and ongoing attendance record.

1. This method uses Keras neural network design.
2. Preprocessing the data for the machine to learn.
3. Defining Keras model and compiling the model.
4. Fit the keras model and evaluate the model.
5. Analyse the developed Keras model and find the accuracy.

The primary method of this paper is to provide human brain like abilities to our machine through neural network. First we have an input layer which includes data set then there are hidden layers, choose number of hidden layers wisely which may disturb the accuracy of our model to some extent. Lastly, we have our final layer which is the output layer to give results. We train our data again and again till then it learns all the features that are required.
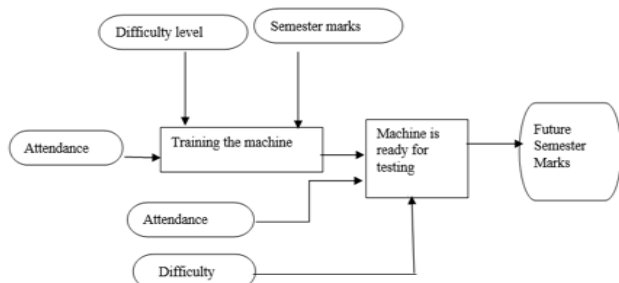

Fig. 1. Block diagram

Data pre-processing is a data mining technique that involves transforming the raw data into an understandable format. Import important libraries such as pandas, NumPy all these Python libraries used for working with arrays. Datasets are in CSV format. Create a matrix of the dataset (X) and a dependent vector (Y).

Create a Sequential model and add layers one at a time until we get a better network architecture. Specifying the number of neurons or nodes and the activation function. Execute the model on the data. We can train or fit our model on our loaded data. Training occurs over epochs and each epoch is split into batches.

- Epoch: One pass through all of the rows in the training dataset.
- Batch: One or more samples considered by the model

within an epoch before weights are updated.

Evaluated our model using the evaluate() function will return the loss of the model on the dataset and the second will be the accuracy of the model on the dataset.

Analyzing the efficiency of the system is done using the Root Mean Square Error(RMSE). RMSE is the square root of theaverage of the square of all of the error.

4. Experimental Work

- *Data description:* We have collected the data of particular student. We have collected the historical academic data from first semester to seventh semester. Attendance percentage, subject wise marks, and difficulty level of the subject upto seventh semester.
- *Sequence data:* We got 42 sequences from first semester to seventh semester. From these data set we used 39 samples for training purpose and 14 samples for validation purpose.
- *Training detail:* To training the model we used the optimizer and normalized. We have used a various set of parameters with a different number of epochs to measure the RMSE of Training and Testing dataset.

5. Conclusion

The overall prediction accuracy in our analysis varies from 85.71% to 100%, learning a discrete class that takes three values. Results are more than promising and enable the future implementation of a student performance prediction. The predicting technique is not only helping the tutors but it also helps students. In order to help predict the academic scores, a predicting model with good accuracy is required.

References

[1] Alyahyan, E., Düştegör, D. Predicting academic success in higher education: literature review and best practices. Int J Educ Technol High Educ, 17, 3 (2020).
[2] Ahmed O. Ameen, Moshood A. Alarape, Kayode S. Adewole- Students' Academic Performance and Dropout Predictions: A Review, Year – 2019.
[3] Francis, B.K., Babu, S.S. - Predicting Academic Performance of Students Using a Hybrid Data Mining Approach. J Med Syst, 43, 162 (2019).
[4] Francis, Bindhia K., and Suvanam Sasidhar Babu. "Predicting academic performance of students using a hybrid data mining approach." Journal of medical systems, 43, no. 6 (2019): 1-15.
[5] Hussain, Mushtaq, Wenhao Zhu, Wu Zhang, Syed Muhammad Raza Abidi, and Sadaqat Ali. "Using machine learning to predict student difficulties from learning session data." Artificial Intelligence Review, 52, no. 1 (2019): 381-407.
[6] Raihana, Z., and Farah Nabilah, A. M., Classification of students based on quality of life and academic performance by using support vector machine. Journal of Academia UiTM Negeri Sembilan, 6(1):45---52, 2018.
[7] Hasan, R., Palaniappan, S., Raziff, A. R. A., Mahmood, S., and Sarker, K. U. Student Academic Performance Prediction by using Decision Tree Algorithm, 4th International Conference on Computer and Information Sciences, 2018.
[8] Yassein, N. A., Helall, R. G. M., and Mohomad, S. B., Predicting student academic performance in KSA using data mining techniques. Journal of Information Technology and Software Engineering, 7(5):1-5, 2017.