International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Student Grade Predictor using Machine Learning

Aneek Sultan[1], Marthi Balaji[2], Obed Jamir[3], Nahida Nazir[4]

[1]*Aneek Sultan, Lovely Professional University*
[2]*Marthi Balaji, Lovely Professional University*
[3]*Obed Jamir, Lovely Professional University*
[4]*Nahida Nazir, Lovely Professional University*
[1-4]*Dept. of Computer Science Engineering, Lovely Professional University, Punjab, India*

Abstract – *The existing infrastructure in educational institutions in India does not encourage the use of advanced data analysis and the use of predictive methods to improve the learning curriculum. Two factors address this lack of an existing system for monitoring student academic progress. Firstly, the education policy that has been focused on all these years has been to ensure compulsory education for all at primary and secondary level in schools and this shifts our attention from the actual education involved. Secondly, the complexity of predictive methods makes it difficult for its integration in an existing curriculum. Machine learning strategies can be used for predictive purposes in a variety of subjects. Such strategies are expected to enable students to improve their performance according to the predicted grades and will inform their student managers who need special care. In this paper, we go through a few machine learning algorithms which have been used to predict academic performance in the past and further implement linear regression and neural network algorithms to predict student performance ourselves. Of the two machine learning algorithms used, neural network algorithms tend to be a better option in predicting student performance.*

1. INTRODUCTION

As universities form a high profile of education administrators in India where university graduates bring new innovations, their performance needs to be improved through the poverty of all education in the country. Calendar and Feldman (2009) found that most students drop out of university in their first year due to poor academic support. The first year of an undergraduate student is therefore called the "make or break" year. If not for appropriate support from the admin, it may lead to the student drooping out early. Predicting the grades of the student at the university itself is one of a feasible solution that may help keep students focused at their grades by preemptively predicting their grade performances

Currently, there are many suggested ways to assess student performance. In this paper we will analyse and implement few of the most effective methods in doing so and compare our results.

The contributions to this field aimed by this paper include the following:

- We analysed actual data collected from undergraduate students of Computer science Engineering Department at LPU.
- We reviewed the literature about grade prediction systematically and systematically and tried describing the same briefly.
- We evaluated the existing prominent machine learning techniques.
- We include a distinct and comprehensive feedback system that would alert the student if he/she needs special attention in the course for which the grade is predicted.
- We include the factors that affect and deteriorate the student performance in various courses based on predefined knowledge of pedagogy.

2. RELATED WORK

In the application of the artificial intelligence to prediction of grades various methods have been employed in the field with moderate to adequate knowledge base and sampling sets varying from adequate to ample. This has been incorporated either to predict grades for improvement for identifying the students at risk of dropping out or failing consistently.

2.1 MATRIX FACTORIZATION

In 2011, Thai-Nghe et al. came up with certain models that were based on matrix factorization to help predict performance of students in elementary courses like algebra and when implemented were largely a success. Such techniques are implemented when the data is sparse and students' background knowledge and tasks are absent. The data is split into train set and test set. The data included is the log file of the interactions made between the student and the computer-aided tutoring system. This approach was very detailed and used multi way arrays which aided as a log of student performances thus giving it a temporal effect. They extended the research and used tensor-based factorization to predict student success. This auto generates logs when the platform is logged onto.

2.2 PERSONALISED MULTIPLE LINEAR REGRESSION

The sample size used for this project has been relatively small and considering the fact that the accuracy of the Grade prediction using the Matrix Factorization (MF) method is not up to the mark when using smaller samples. PLMR was put to test when applied in various universities to predict results in in campus tests in universities like Stanford university and University of Minnesota. Here it was concluded that both PLMR and Matrix method have low rates of errors and was feasible for scalable implementation. The university management systems (like ums in LPU) can also incorporate this method to incorporate in-situ prediction of grades.

2.3 DROPOUT EARLY WARNING SYSTEM(DEWS)

This system of warning has been out to use in Wisconsin since 2015 when Knowles designed the advanced data mining model to predict grades. The system is being scaled to be extended statewide. The prominent feature of the model is the dropout rates that it helps calculate. This model is very versatile as it can be customized for the requirements of the institutions in other countries. This can be very useful in countries like India where the dropout rates are at two digit of 18% (Educational statistics at a glance, MHRD, 2016) when compared to the average in developed countries

2.4 FACTORISATION MACHINES

Future grade prediction methods are designed to predict the marks a learner will achieve in the upcoming term. Some GMU based samples were used by Sweeney in 2015 to predict marks student may obtain using methods based on Matrix factorization. The research concluded that the method when implemented showed very minimalistic error in both cold and non-cold problems. When further extending their work, they used various other methods like PMLR and RF. They concluded that these methods when used in a combination yield better error results.

3 DATASET DESCRIPTION

The dataset has been collected through real world performance of students' at Lovely Professional University, pursuing Bachelors in Computer Science Technology(Btech Cse). The dataset consists of seven attributes which were filtered from a set attributes based on their significance in affecting student performance

The dataset mentioned below is the training dataset with a total of 260 entries, in which the attributes are all independently related to the attribute "Grade". Each attribute in the dataset is further categorized on the basis of 'attribute type' and 'data type'. 'Attribute type' defines attributes as either continuous or categorical where, continuous attributes can have any value within a set range while categorical attributes can only have a set number of values which divide our dataset into two or more categories. 'Data type' defines the type of value for the attributes in our dataset like Integer, Boolean, float, double and string etc.

ATTRIBUTE	ATTRIBUTE TYPE	DATA TYPE
CA	CONTINOUS	INTEGER
CA+MIDTERM	CONTINOUS	INTEGER
ATTENDANCE	CONTINOUS	INTEGER
MIDTERM	CONTINOUS	INTEGER
S/T RELATION	CONTINOUS	BOOLEAN
LAST MINUTE PREPARATION	CONTINOUS	BOOLEAN
GRADE	CONTINOUS	FLOAT

3 MACHINE LEARNING ALGORITHMS USED

3.1 LINEAR REGRESSION

The primary machine learning algorithms used to predict our result is a form of linear regression that has been modified to better fit the requirements of the task in hand. Linear Regression is the basic form of regression analysis. It assumes that there is a linear relationship between the dependent variable and the predictor(s). In regression, we try to calculate the best fit line which describes the relationship between the predictors and predictive/dependent variable. (Montgomery,(2015) Pattern Recognition and machine learning. Chapter 2:Linear Regression) By performing linear regression on each attribute independently with the final grade, we are able to train the dataset to produce most relevant predictions based on provided attributes.

To do this we first fit a line to our graph such that the overall distance of every point on the graph is at a minimum distance from the line. This is done by using least-squares method.* (Douglas C. Montgomery (2013) et al)

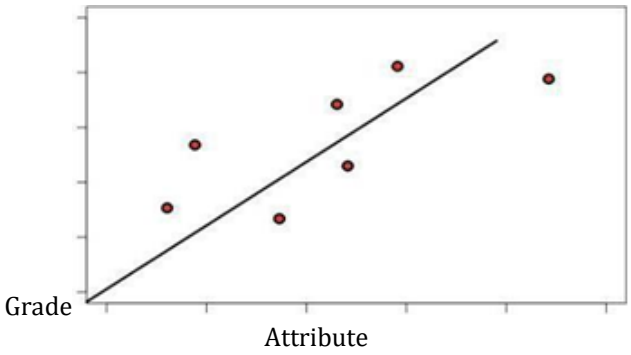

Figure 1. Fitting a line to the graph.

3.2 NEURAL NETWORKS

Neural network is another classification algorithm that is based on the working of the human brain. It consists of nodes which behave like the neurons in the brain. A neural network consists of an input layer, multiple hidden layers and an output layer. These layers further consist of variable number of nodes based on the dataset. Every node in each layer is connected to every other node in the next layer with a random predefined weight "w" assigned to each connection and a bias "b" assigned to each node in the hidden layer. Each node in the hidden layer is also known as an activation function "a". (Charu C. Aggarwal (2018) Neural Networks and Deep Learning. CHapter1.2:Basic Architecture of Neural network)

The attributes present in our dataset acts as the nodes in the input layer and its respective values are then fed into each node of the hidden layer, where they are manipulated by its respected weights and biases to produce the respective values of the hidden layer. These values further get manipulated by the weights and biases of the output layer to produce the values of the nodes in the output layer. The error in the output values is then further used to adjust our weights and biases via backpropagation over a number of iterations. (Raul Rojas (2013) Neural Networks: A Systematic Introduction. Chap 7: The Backpropagation algorithm)

a1= sigmoid (w1*x1 + w2*x2 + ... + wn*xn + b)

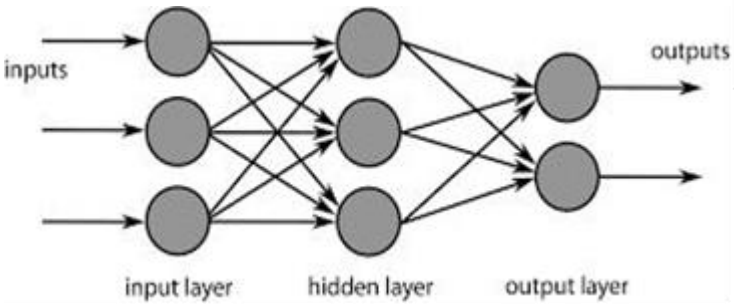

Figure 2. A simple neural network.

4. PREDICTION OF STUDENT GRADES

In our study we would like to predict subject grades on a scale from 0 to 4 where,
0.6-1.0 = D
1.1-1.5 = C
1.6-2.0 = B
2.1-2.5 = B+
2.6-3.0 = A
3.1-3.5 = A+
3.6-4.0 = O
Representing our grades in numeric form makes it easier to compare our predicted grades to the original grades. Also the ranged representation of the grades helps to account for the fluctuations that may occur in our predicted grades. Finally, the predicted grades can further be used to predict student GPA.

4.1 RESULTS

The following graph represents the result produced after using our trained data on our test set. Each point on the graph represents the predicted grade acquired by a student in the test set.

Number of Students in the test set= 60

1. LINEAR REGRESSION

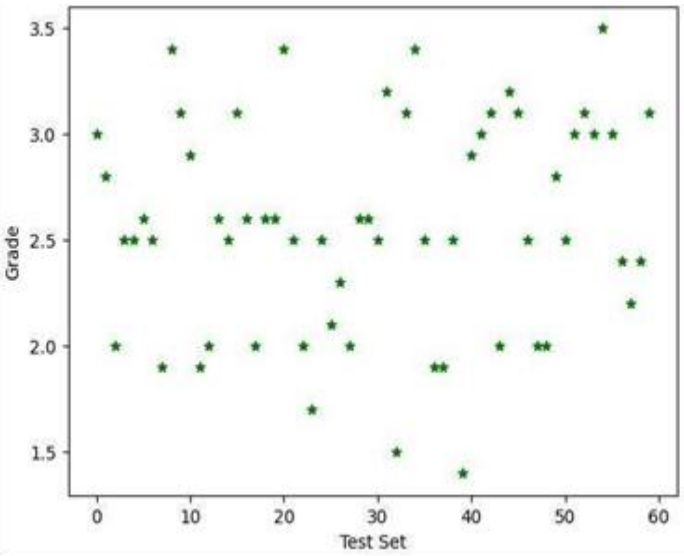

Figure 3. Predicted grades for test set using linear regression algorithm.

2. NEURAL NETWORK

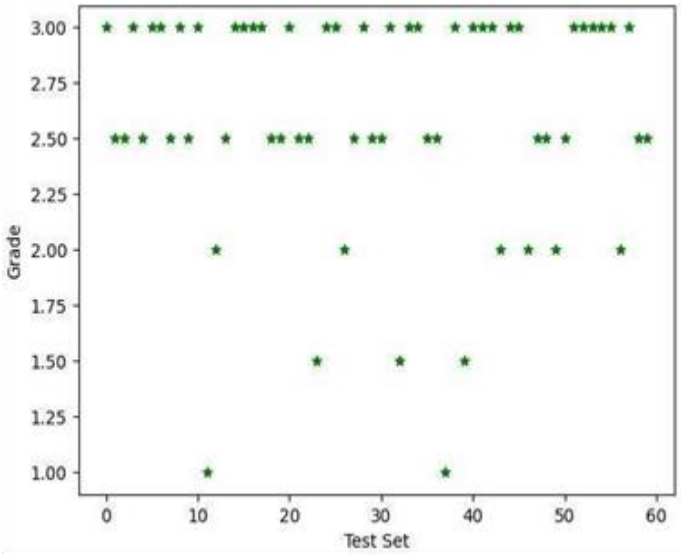

Figure 4. Predicted grades for test set using neural network algorithm.

4.2 RESULT ANALYSIS

Let us first analyse the original data for our test set.

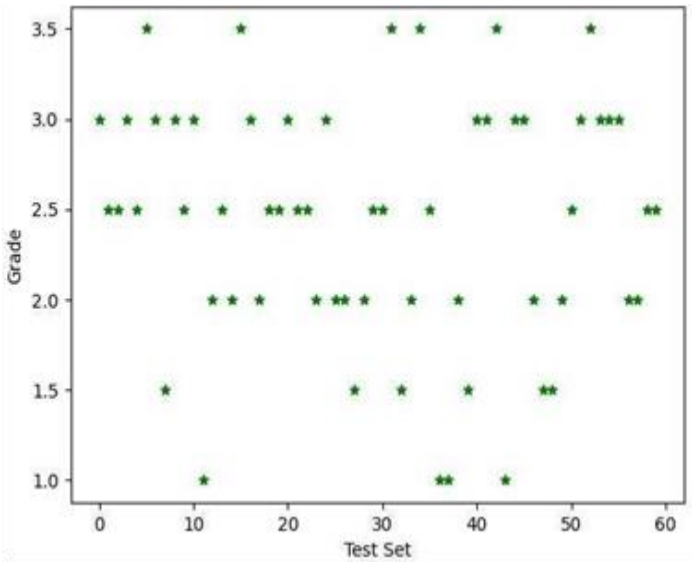

Figure 5. Original grades for the test set.

Now, we derive the R^2 value for our linear regression model and hit count for our neural network from the predicted and actual results.

R^2= 1- sum squared regression (SSR)/Total sum of squares (SST) = **0.6272.** From this, we can say that our linear regression module accounts for at least **62.72 %** of the actual result.

Hit Count = Number of correct predictions = 40

Accuracy = (40/(Total No. of predictions))*100 = 66.66

Therefore, our neural network module accounts for at least 66.66% of the actual result.

5. CONCLUSION

We discussed many machine learning techniques including linear regression and neural networks which we eventually used in our prediction module and observed how various independent attributes affect our final result. By making use of this modules we can observe that by making specific changes in our daily lives can significantly improve our academic performance. Keeping a check on the academic performances helps a student maintain his/her focus and improve on specific fields where they find the need to do so. Grades assist in understanding the different topics and subjects that must be taught in each section, as well as how each subject's level must be raised. Grades provide a proper learning approach that must be adhered to for the student to grasp the definition fully. Grades serve as a benchmark against which we measure our potential results. It tells us whether we're making progress or not. Grading is a criterion by which we will determine how much we have progressed and how much more hard work we will need to achieve a certain degree of expertise. Grades often assist students in comprehending the various variations of a specific theorem or subject, allowing them to gain a deeper understanding of the material they are studying. Teachers can better understand which subjects their students have grasped and which topics they need to be taught again using grades. It also gives teachers an indication of and students are struggling in class and therefore need extra attention. Our main aim is to simplify the method and get an accurate data of the students in order to keep a check on their academic performance.

REFERENCES

[1] MONTGOMERY (2015) Pattern Recognition and machine learning. Chapter 2: Linear Regression

[2] CHARU C. AGGARWAL (2018) Neural Networks and Deep Learning. Chapter1.2:Basic Architecture of Neural network

[3] DOUGLAS C. MONTGOMERY (2013) Introduction to Linear Regression Analysis Chap 2.1: Simple linear regression model

[4] RAUL ROJAS (2013) Neural Networks: A Systematic Introduction. Chap 7: The Backpropagation algorithm

[5] DOUGLAS C. MONTGOMERY (2013) Introduction to Linear Regression Analysis Chap 2.2: Least- squares estimation of parameters

[6] ANKIT CHAUDHARY (2018) Introductory Guide to Factorization Machines & their application

[7] SWAPNA BHATTACHARYA et al (2016) Statistics at a glance; Table 10; MHRD

[8] KAVITA SRINAVASA (2021) View from India; the decade and of and for AI