

Predicting the impact of different Variables on Students Academic Performance using Artificial Intelligence

Pratiksha R.Shetgaonkar

*Asst. Professor, Department of Computer Engineering,
Shree Rayeshwar Institute of Engineering & Information technology, Goa-403103*

Abstract- Undergraduate student's academic performance is a long standing issue in higher education and a great deal of research over the past many years. At the end of each semester, students' result will be analyzed in order to evaluate students' academic performance.

Thus there is a need for the prediction models which can effectively predict the student's academic performance based on some relevant attributes which can be useful to guide the students with low academic record initially.

Most of the existing research on the performance prediction uses various techniques such as classification, decision tree etc as a prediction methodology which basically classify the students rather than predicting the continuous grades of the students.

In this paper, an attempt will be made to apply artificial neural network which is trained using backpropogation algorithm in predicting the academic performance of the students, which has some noticeable benefits over the other suggested prediction models in the educational domain.

The benefits of the proposed system include:

The more accuracy in prediction and can provide the continuous as well as classified discrete output.

Keywords— performance prediction, artificial neural network, artificial intelligence, backpropogation.

I. INTRODUCTION

Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in.

Students' academic performance is based on diverse factors, attributes and variables such as:

- Causes resident in society.
- Causes resident in school/college.
- Causes resident in the family.
- Causes resident in the student.

Prediction models that include all these variables are necessary for the effective prediction of the performance of the students.

The prediction of student performance with high accuracy is helpful in identifying the students with low academic achievements initially. The identified students can be individually assisted by the educators with their effective guidance so that their performance is better in future which has a great impact in building their professional life.

The primary purpose behind carrying out this research project is

- To predict the academic performance of the students based on some identified attributes or factors and determine which factors affect the most, the academic performance of the students.

The other secondary objectives are

- Learning about the effectiveness of the artificial neural networks in prediction.
- Determining some suitable factors that affect a student's performance and to transform these factors into forms suitable for an adaptive system coding
- To model an artificial neural network that can be used to predict a candidate's performance based some data for a given student.

II. IMPLEMENTATION

A. Performance Prediction Model

The following fig.1 gives the block diagram for the proposed performance prediction model.

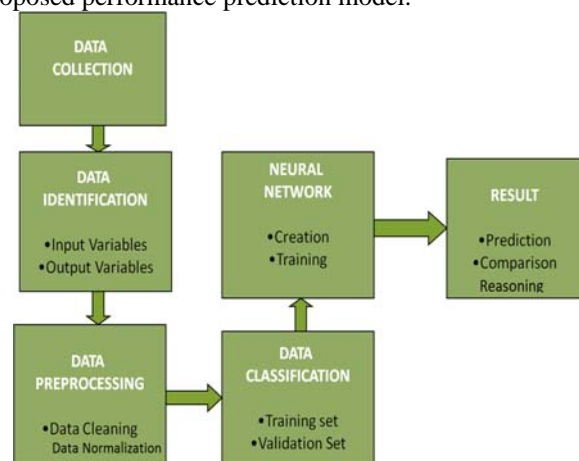


Fig. 1 Performance Prediction Block diagram

i) Data Collection:

It included the following steps:

- i. Preparation of the questionnaire
- ii. Distributing the questionnaire among the students and collecting the data

ii) *Data Identification:*

Once the data is collected, based on the requirement of the NN the data variables were grouped into the Input and output Variables. The following input variables were used:

1. Gender
2. Single Child or not
3. Category
4. Employed Family members
5. Handicapped or not
6. SSCE percentage
7. HSSCE percentage
8. Health problems
9. Attendance
10. The following output variable was obtained:
11. % obtained in exam

iii) *Data Preprocessing:*

“Data preprocessing” is a fundamental building block of data mining. To be more precise it includes, modifying the source data in to a different format which:

It basically includes steps such as:

- i) Data cleaning
- ii) Data normalization

Data Normalization was done using the following methods:

Main Max Normalization

Min-max normalization: Performs a linear transformation on the original data values. Suppose that minX and maxX are the minimum and maximum of feature X. We would like to map interval [minX, maxX] into a new interval [new_minX, new_maxX]. Consequently, every value v from the original interval will be mapped into value new_v using the following formula:

$$new_v = \frac{v - min_x}{max_x - min_x} \cdot (new_max_x - new_min_x) + new_min_x$$

Z score Normalization

Z-score normalization: also called zero-mean normalization. The values of attribute X are normalized using the mean and standard deviation of X. A new value new_v is obtained using the following expression:

$$new_v = \frac{v - \mu_x}{\sigma_x}$$

where mX and sX are the mean and standard deviation of attribute X, respectively.

iv) *Data Classification:*

Once the data is preprocessed, it needs to be classified into two sets for inputting into an neural networks as

- a) Training Set
- b) Validation set

v) *Creation of Neural Network and Training:*

This is the most difficult stage of the entire process. It includes steps such as

- a) Selecting of the neural network architecture

- b) Training the neural network with the training data set

vi) *Prediction of Results :*

This is the stage where the Neural Network predicts the result based on different parameters.

vii) *Prediction of Reason:*

This is the stage where we tried to predict the probable reason behind the result predicted by the NN.

B) Architectural Design

The following fig.2 shows the architectural design of the proposed system.

III. CODING AND IMPLEMENTATION

The Backpropogation algorithm was used for training the neural network. The C#. NET technology was used to implement backpropogation algorithm.

IV. OBSERVATIONS AND RESULTS

i) *Observations:*

The data set collected for the application comprised of 250 undergraduate students details about the selected variables. The data set was divided into the training and the testing set. All the operations as mentioned in previously that is preprocessing, normalization etc were carried out on to the data and the corresponding predicted output is computed. The trend is complemented with the supporting reason reflecting the probable cause for the change in trend as shown in fig 2.

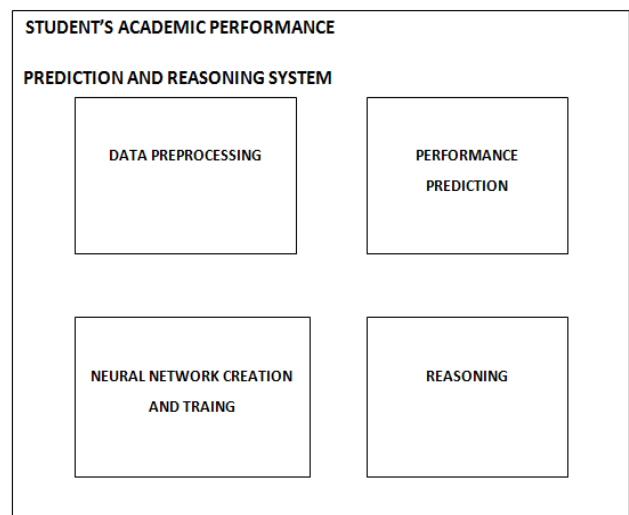


Fig.2 architectural design

ii) *Neural Network Prediction*

By varying the configuration of the neural network and its associated parameters the following observations were made:

- a) The Artificial neural network with the one input layer, one hidden layer and one output layer were observed to give better results.
- b) The configuration with the number of neurons in the hidden layer slightly more than that of the

input layer improved the prediction, but the larger the number of neurons in the hidden layer, longer time it took to train the neural network.

- c) The higher the learning rates the lesser the time for neural network training and vice versa. The learning rate around 0.02 was found to be good for training the network.
- d) Lower the threshold for the error rate (equal to 0.01), longer the time it took to train the network but improved the prediction accuracy.

The fig.3 below shows the interface on which the above observations were made:

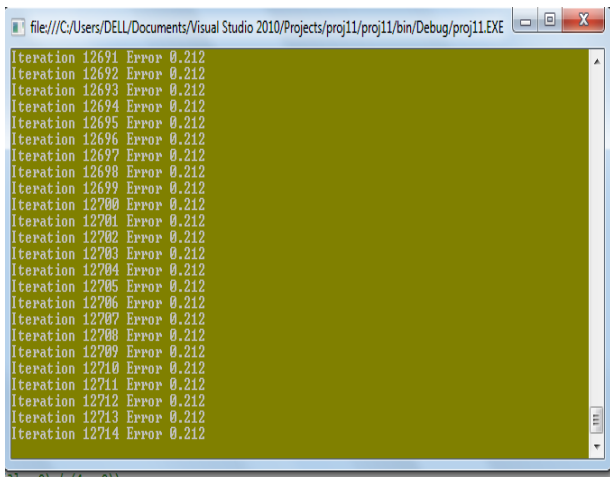


Fig 3.Training NN

Based on the above observations, the appropriate configuration and parameters of the neural network were selected to test the prediction accuracy. It was observed that the difference between the predicted percentage and the actual percentage was found to be in the range of $\pm 2.5\%$ for most of the records. This is being shown in the graph in Fig.4. The series1 represents the predicted percentage and Series 2 represents the actual percentage. The testing was carried out for almost 25 students among which the result of 10 students are as shown in the table below.

iii) Results

The following table 1 shows the difference in the actual and predicted results for some of the records tested using the implemented methodology.

Sr.No	Actual Percentage	Predicted Percentage	Difference
1	61	60.858	0.142
2	53	52.369	0.631
3	54	53.474	0.536
4	54	53.421	0.589
5	62	61.651	0.359
6	48.53	45.955	2.574
7	54	53.163	0.837
8	49.5	48.572	0.928
9	48	47.318	0.682
10	50	49.277	0.723

Table 1

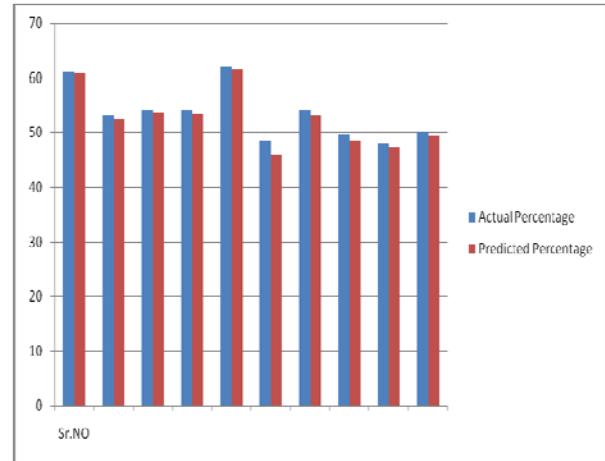


Fig.4 Comparison graph between Actual and Predicted Percentages

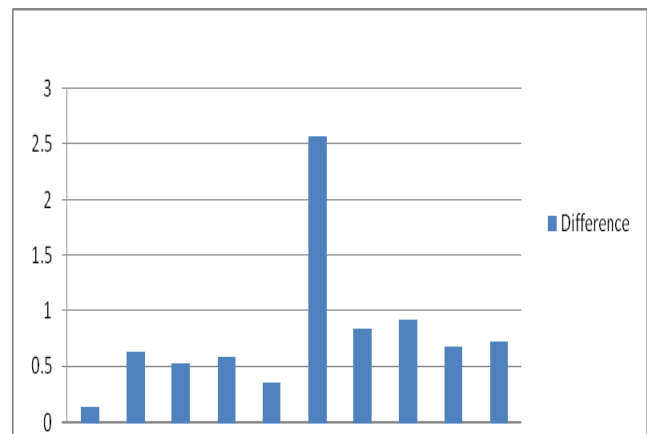


Fig.5 Graph showing the range of change between actual and predicted percentage

V. CONCLUSION

This research work presented in this paper was an attempt to develop a methodology to predict the performance of the undergraduate engineering students using the learning capability of the Artificial Neural Network which will be able to help us to identify the students with the low academic records initially.

The proposed system uses ANN and reasoning techniques similar to apriori algorithm to produce forecasts about the student's academic performance. The proposed model did fairly good prediction as far as the for the records being tested.

REFERENCES

- [1] Data Mining: A prediction for performance improvement using classification By Brijesh Kumar Bhardwaj and Saurabh Pal
- [2] Predicting NDUM Student's Academic Performance Using Data Mining
- [3] Techniques, By Yuhanim Hani Yahaya, Norshahriah Wahab
- [4] Mining Students data to analyse Learning Behavior:Case Study By ALAA EL-HALEES
- [5] <http://www.heatonresearch.com/>
- [6] <http://en.wikipedia.org/wiki/Backpropagation>
- [7] http://www.cs.ccsu.edu/~markov/ccsu_courses/DataMining-3.html
- [8] http://en.wikiversity.org/wiki/Learning_and_Neural_Networks
- [9] <http://www.neural-forecasting.com>

- [11] http://www.cse.iitm.ac.in/~cs672/Lectures/Data_Preprocessing.pdf
- [12] http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html
- [13] <http://neurondotnet.freehostia.com/index.html>
- [14] <http://takinginitiative.net/category/artificial-intelligence/neural-networks/>
- [15] <http://intelligencemining.blogspot.com/2009/07/data-preprocessing-normalization.html>
- [16] http://en.wikipedia.org/wiki/Standard_score
- [17] Data Mining, Concepts and Techniques, Jiawei Han, Micheline Kamber, Morgan Kaufmann Publishers.
- [18] Principles of Data Mining, Hand, Mannila and Smyth.