

Innovative Technique of Improving Student Performance Using Data Mining Algorithm

Rupali Chudasama ^[1], Prof. Amol Joglekar ^[2]

M.Sc Computer Science ^[1]
 Department of Computer Science ^[2]
 Mithibai College, Mumbai
 Maharashtra - India

ABSTRACT

Artificial neural network is a model which used to infer a function from observations. It is useful when data is complex and the design of such a function manually is impractical. Neural networks produce reasonable results compared to typical time series approaches which try to detect trends in a season but fail to give sensible results for unseasonable times. Neural Networks are effective in medical diagnosis (diagnosing various types of cancers), game-playing and decision making (backgammon, chess etc), pattern detection (radar systems, face identification, object recognition) and data mining. This paper aims to predict student's performance based on many factors including previous performance, other co-curricular activities, social life, health, etc.

Keywords:- Artificial Neural Networks, Student performance, ANN, Education, Predictive Model

I. INTRODUCTION

Several researches have been done to improve students' performance while establishing the factors affecting their academics. Artificial intelligence has the ability to develop effective and efficient student model which can make a calculated guess of their future academic performance. A student's performance depends on a number of factors that may or may not be directly related to academics. The factors could be past performance, participation in co-curricular activities, peer group, social activities, health status, financial background, family conditions, etc. Each factor can have a different level of effect on the student's performance.

The objectives of this study is to determine some suitable factors that affect a student's Performance and to transform these factors into forms suitable for an adaptive system coding and to model an Artificial neural network that can be used to predict a student's performance.

A. NEURAL NETWORK

Many neural network model, even biological neural network assume main simplification over actual biological neural network. Such simplification is necessary to understand the intended properties and to attempt any mathematics analysis. Even if all the properties of the neurons were known, simplification still needed for analytical purpose. All such models are known as artificial neural network, here after called as ANNs. In ANNs, all the neurons are operating at the

same time leading to parallel structure, which makes them to perform tasks at much faster rate compared to conventional computer.

An artificial neural network (ANN) is a computational model that attempts to account for the parallel nature of the human brain. An (ANN) is a network of highly interconnecting processing elements (neurons) operating in parallel.

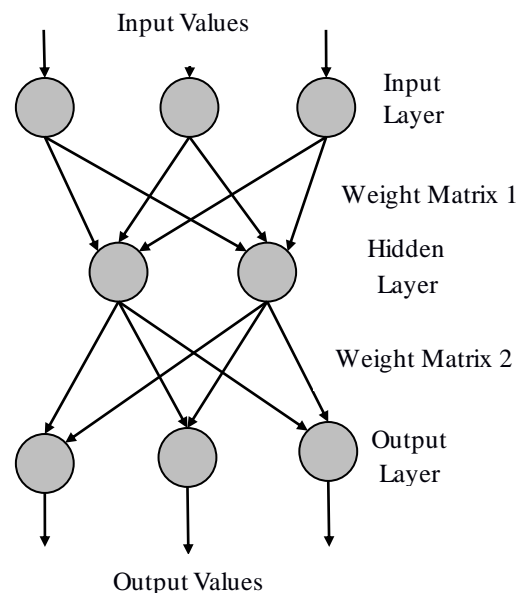


Fig 1: Artificial Neural Network

These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layer(s) of units, called hidden layer(s)[8].

II. LITERATURE REVIEW

Zainudin Zukhri and Khairuddin Omar [2] have reported successful application of Genetic Algorithm for solving difficult optimization problems in new students’ allocation problem.

Vuda Sreenivasa Rao[1] has developed a model for improving academic performance evaluation of students based on data warehousing and data mining techniques that use soft-computing intensively.

Bhardwaj and Pal [10] conducted study on the student performance based by selecting 300 students from 5 different degree college conducting BCA (Bachelor of Computer Application) course of Dr. R. M. L. Awadh University, Faizabad, India. By means of Bayesian classification method on 17 attributes, it was found that the factors like students’ grade in senior secondary exam, living location, medium of teaching, mother’s qualification, students other habit, family annual income and student’s family status were highly correlated with the student academic performance.

Mukta and Usha[3] carried out an analysis to predict the academic performance of business school graduates using neural networks and traditional statistical techniques and the results were compared to evaluate the performance of these techniques. The underlying constructs in a traditional business school curriculum were also identified and its relevance with the various elements of admission process was presented.

Kanakana and Olanrewaju [4] used Artificial Neural Network and linear regression models to predict student performance after access to higher education. Data received from the Tshwane University of Technology was utilized for the study. The total Average Point Scores (APS) students obtained in grade 12 was employed as input variable. The results indicated a better agreement between ANN model prediction and observed values compared to those in the linear regression.

Pandey and Pal [9] conducted study on the student performance based by selecting 600 students from different colleges of Dr. R. M. L. Awadh University, Faizabad, India. By means of Bayes Classification on category, language and

background qualification, it was found that whether new comer students will performer or not.

III. RESEARCH METHODOLOGY

This study was conducted using an online survey form to collect the data in an Excel sheet. Further analysis was conducted on this data to infer a function that can compute and generate results showing a student’s performance. The computation data are categorized as input variables. The output variables on other hand represent performance of a candidate in terms of their final result.

1. INPUT VARIABLES

The input variables selected for this research are as below:

- 1) Student’s Gender.
- 2) Student’s SSC Score.
- 3) Student’s HSC Score.
- 4) Student’s Last exam average Score.
- 5) If the student has taken any gaps in his/her studies.
- 6) If the student has done any certifications along with his\her studies.
- 7) Location of the college.
- 8) Traveling time to reach college.
- 9) Whether the student has participated in any co-curricular activity such as NSS/NCC/College-festivals/Sports.
- 10) How attentive the student is in lectures and lab practical.
- 11) How much time the student spends on social media.
- 12) Student’s current health status.
- 13) Does the student indulge in alcohol or smoking?

These factors were transformed into a format suitable for neural network analysis. The domain of the input variables used in this study shown in Table1.

Table 1: Input variable

Sr. No	Input Variable	Domain
1	Gender	Female Male
2	SSC Score	More than 70 61-70 51-60 Less than 50

3	HSC Score	More than 70 61-70 51-60 Less than 50
4	Last exam	More than 70 61-70 51-60 Less than 50
5	Gap in studies	Yes No
6	Certification	Yes No
7	Location of the college	Same city of the college Outside the city
8	Travelling time	Less than 30min 30min-1hour 1hour-2hour More than 2hour
9	Co-curricular activity	NSS NCC College festival Sport
10	Attentive in lecture	Scale from 1-10
11	Social media	Less than 30min 30min-1hour 1hour-2hour 2hour-3hour More than 3hour
12	Any habits	Alcohol,Smoke

2. OUTPUT VARIABLE

The output variable represents the performance of student on future exam.

Table 2: Output variable

Sr. No	Result	Output variable
1	Good	60% or more
2	Poor	Less than 60%

3. ALGORITHM

Artificial neural network with feed-forward topology is called Feed-Forward artificial neural network and as such has only one condition: information must flow from input to output in

only one direction with no back-loops. There are no limitations on number of layers, type of transfer function used in individual artificial neuron or number of connections between individual artificial neurons.

The algorithmic steps are as follows:

1. Take input from users.
2. Initialize weights with some random values.
3. Transmit the data received to the first hidden layer.
4. In this hidden layer, calculate the net input (X), by using the following equation
5. Compute the output of the hidden layer by applying activation function over X and send it to the next hidden layer.
6. Repeat steps 4 & 5 for the rest of the hidden layers.
7. The final hidden layer will send its output to the output layer.
8. The output is computed by using the following equation

$$X_i = \sum_{i=1}^n I_i W_i \quad (1)$$

where I is input and W are the weights for the inputs.

$$Z_i = \sum_{i=1}^n X_i \quad (2)$$

where Z is the output

IV. PROPOSED SYSTEM ARCHITECTURE

IV. CONCLUSION AND RESULT

An Artificial Neural Network for predicting student performance model was conceived which using feed forward algorithm for training. The factors for the model were obtained from students who filled the online survey. The model was tested and the overall result was 78.94%. This study showed the potential of the artificial neural network for predicating student performance.

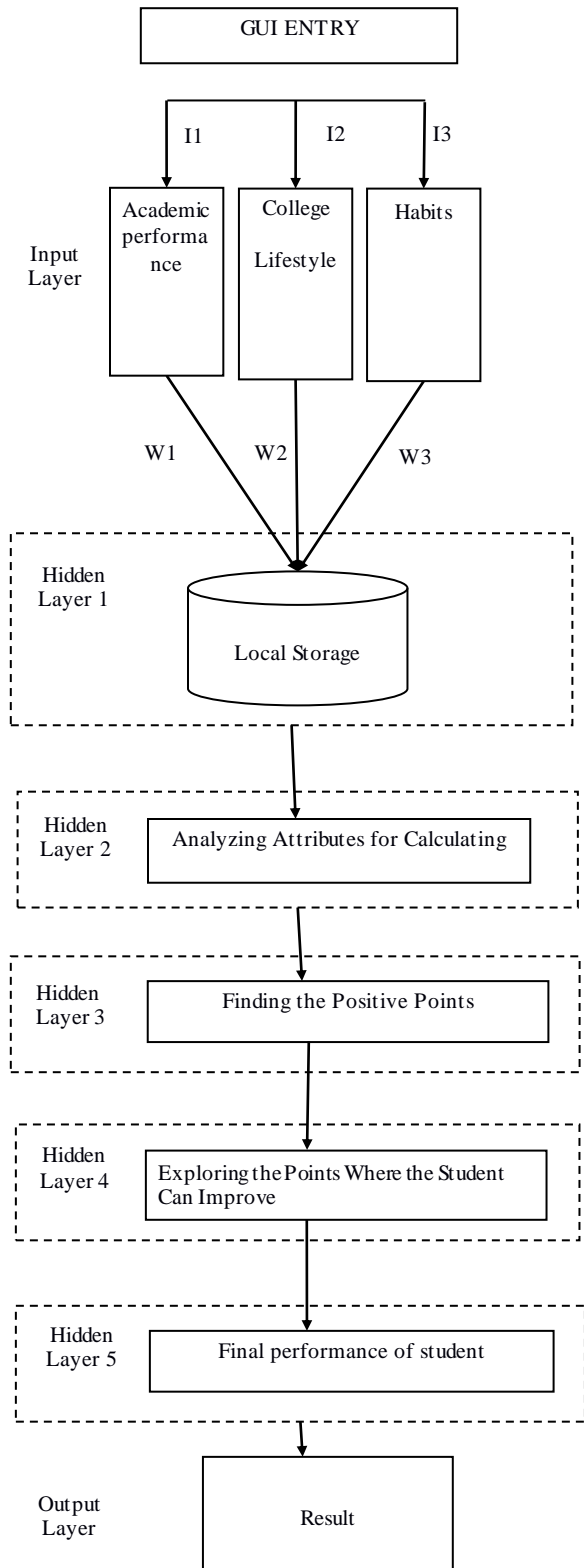


Fig 2. Proposed Architecture

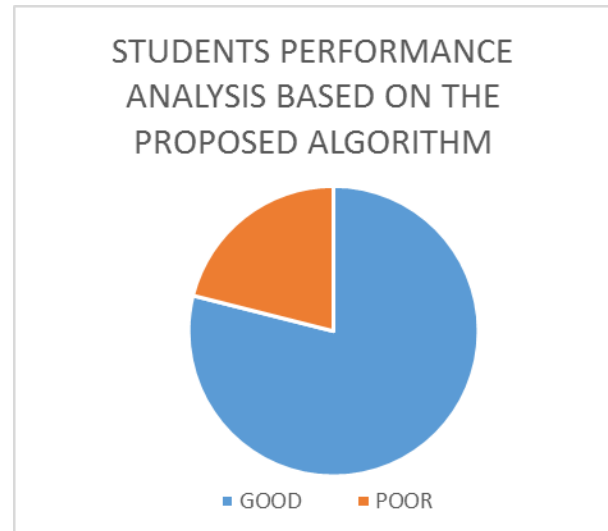


Fig 3: Student’s performance analysis based on the proposed algorithm

From the data collected, 32% of the students said that they chose their college because of course preference, 34% chose their college because of the reputation of the college, while 26% said that the college was near to the their home.

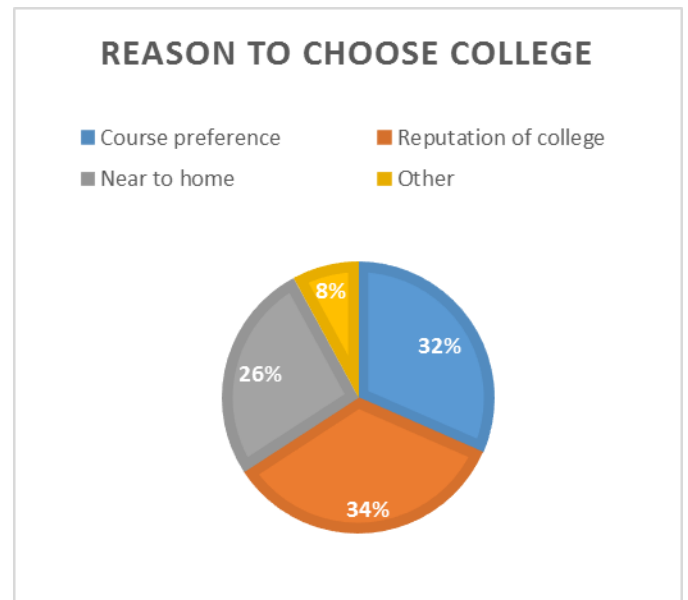


Fig 4: Reason to choose college

74% of the students said that they had an attention span varying between 70-100% of the time.

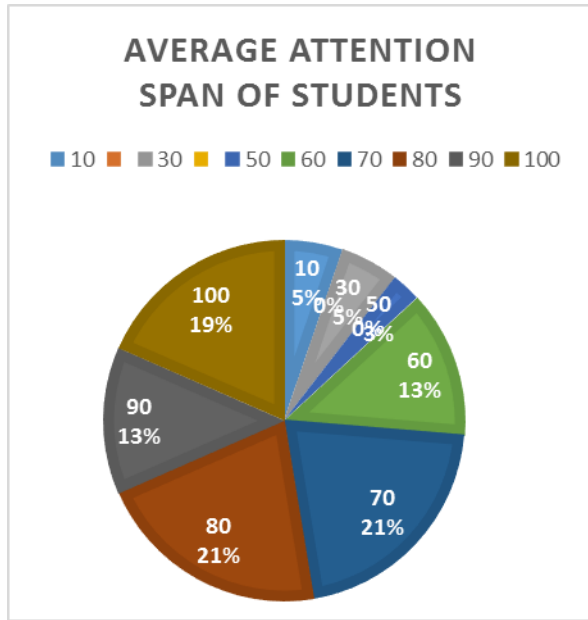


Figure 5: Average attention span of student's

Of the data collected, 55% of the students said that they have average percentage of more than 70% in their last exam, 21% of the students said that they have average percentage between 61-70%, 11% of the students said that they have average percentage between 51-60% while 13% said that the average of last exam percentage less than 50%.

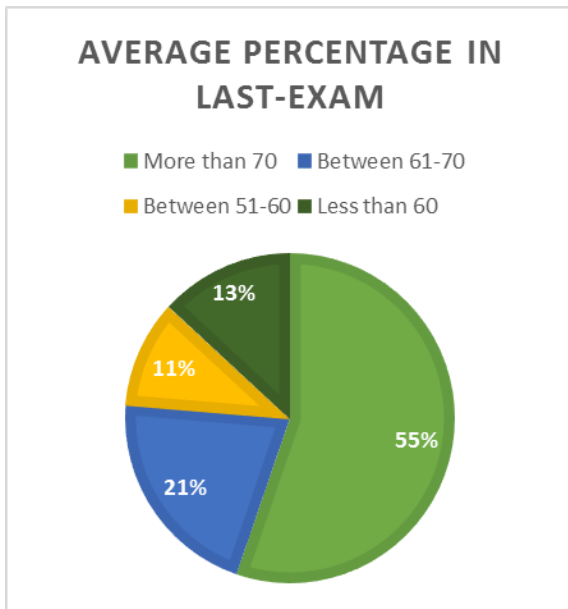


Fig 6: Average percentage in last-exam

48% of the students said that they spent approximately 2 hours on social media, 26% of the students said that they spent

approximately 1 hour on social media, 21% of the students said that they spent approximately 3 hour on social media, while the remaining 5% spent more than 3 hours on social media.

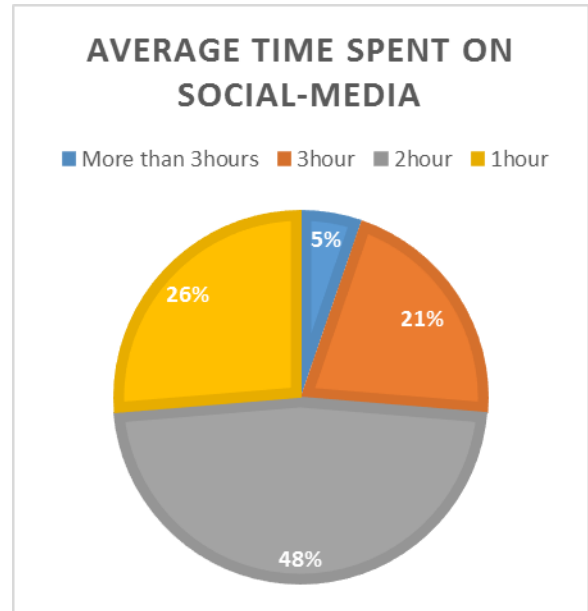


Fig 7: Average time spent on Social-media

V. FUTURE SCOPE

Model will be tested with some real life data from colleges after completing the training to NN. More accuracy can be achieved by improving algorithm and considering some parameters. The number of samples can be tested at run time.

REFERENCES

- [1] Zukhri& Omar(2008) Solving New Student Allocation Problem with Genetic Algorithm: A Hard Problem for Partition Based Approach. International Journal of Soft Computing Applications, Euro Journal Publishing Inc., 6-15.
- [2] Sreenivasarao& Yohannes (2012) Improving Academic Performance of Students of Defense University Based on data Warehousing and Data Mining. Global Journal of Computer Science and Technology. 12(2), 201-209.
- [3] P.Mukta and A.Usha, "A study of academic performance of business school graduates using neural network and statistical techniques",

- Expert Systems with Applications, Elsevier Ltd., vol. 36, no. 4, (2009).
- [4] G. Kanakanal and A. Olanrewaju, “Predicting student performance in Engineering Education using an artificial neural network at Tshwane university of technology”, ISEM 2011 Proceedings, (2011) September 21-23, Stellenbosch, South Africa.
- [5] V.O. Oladokun, A.T. Adebajo, O.E. Charles-Owaba, “Predicting Students’ Academic Performance using Artificial Neural Network:A Case Study of an Engineering Course”, *The Pacific Journal of Science and Technology*, Volume 9. Number 1. May-June 2008
- [6] S.N.Sivanandam and S.N.Deepa, “Principles of Soft Computing”, Wiley India Edition, 2007
- [7] DR. YASHPAL SINGH, ALOK SINGH CHAUHAN “NEURAL NETWORKS IN DATA MINING”, *Journal of Theoretical and Applied Information Technology*, 2005-2009
- [8] Irfan Y. Khan, P.H. Zope, S.R. Suralkar, “Importance of Artificial Neural Network in Medical Diagnosis disease like acute nephritis disease and heart disease”, *International Journal of Engineering Science and Innovative Technology (IJESIT)* Volume 2, Issue 2, March 2013
- [9] U. K. Pandey, and S. Pal, “Data Mining: A prediction of performer or underperformer using classification”, *(IJCSIT) International Journal of Computer Science and Information Technology*, Vol. 2(2), pp.686-690, ISSN:0975-9646, 2011.
- [10] 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.