

# Predicting Students Yearly Performance using Neural Network: A Case Study of BSMRSTU

Md. Fahim Sikder

Department of Computer Science and Engineering  
Bangabandhu Sheikh Mujibur Rahman Science  
and Technology University, Bangladesh  
fahimsikder01@gmail.com

Md. Jamal Uddin and Sajal Halder

Department of Computer Science and Engineering  
Bangabandhu Sheikh Mujibur Rahman Science  
and Technology University, Bangladesh  
jamal.bsmrstu@gmail.com, sajal@bsmrstu.edu.bd

**Abstract**—Students academic performance is the reflection of both academic background and family support. This performance record is critical for the educational institution because they can learn from this to improve their quality. Educational data mining helps to analyze these data and extract information from it. We can determine the status of learners academic performance. For achieving this we can use techniques like decision tree, neural network, classification, data clustering, support vector machine and so on. In this paper, we will predict student's yearly performance in the form of Cumulative Grade Point Average (CGPA) using neural network and compare that with real CGPA. In this regard, a real dataset would be of great importance. We used real dataset from Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU) to perform the prediction.

**Index Terms**—Neural Network, Prediction, Backpropagation algorithm

## I. INTRODUCTION

Nowadays educational institutions are operating in a complex and highly competitive environment. In particular, the competition prevails mostly in the area of academic results. However, it is not an easy task to analyse students results to assess their existing levels and accordingly make strategies for the future development [1]. It has been well acknowledged by now, students academic performance, to a great extent, resonates their future place, in other way the future of any country or of any civilization. Also, academic performance is very crucial for the job market. That is why, students, teachers and guardians tend to give additional consideration to academic achievements. Amid this scenario, Educational Data Mining appears to show us some pathways.

Actually, Educational Data Mining refers the techniques to mine significant information from large repositories of educational dataset. We usually take data from educational institutions and apply different techniques to produce meaningful information.

Students academic performance, generally, is measured by examination results that are considered to hint where the learners can reach in future. In the prevalent evaluation system, examination result is measured by Cumulative Grade Point Average (CGPA) [2]. It is the average of grade point obtained by any individual learner in all the subjects as per the

scheme of studies. Hence, if any institute is able to predict the performance beforehand, it may take necessary steps to improve students performance.

Time and again, researchers used educational data mining to create students performance analysis model. It facilitates learning system, predicts the dropout rate and assists grading system and so on. These models embody different techniques for prediction such as regression, decision tree, clustering, classification and the like to produce results. However, the accuracy of those predictions is not yet satisfactory.

Bearing the already prevalent status of Educational Data Mining, this paper is initiated to gauge the yearly performance of students in a more authentic manner. For this we use a real dataset from the Department of Computer Science and Engineering at Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU). Here, we take the data of 120 students for results analysis. By using data mining technique we extract required information. We also use supervised neural network and MATLAB as the data mining tool to mine the data from the dataset. In this case, the Levenberg Marquardt back propagation algorithm is of use to train the neural network. Then, we divide the whole dataset into three parts, one is for training the network, another one is for validating the results and the last one is for testing the results. Therefore, the error percentage of the prediction will be lower.

The main contributions of this paper are as follows. We design a new technique based on supervised neural network which digs out information from students data. The experimental results show that the academic performance largely depends on family background and students engagements with social media interactions.

The remaining sections of the paper are organized here. In section II describes related work on predicting students performance. Section III introduces our proposed solution for the prediction model. Description of dataset are described in section IV. Implementation of our proposed technique has been described in section V. Result of the system are shown in section VI. The paper concludes with the direction of future work in section VII.

## II. RELATED WORK

Researchers are working to modernize the educational system by the aid of educational data mining. Researcher gives concentration to the Educational Data Mining (EDM) field for its' high potentiality. It could be used in the educational field to enhance the understanding of the learning process of a student.

San Pedro et. al [3] analysis a web based tutoring system for mathematics from 3747 school students and predicted whether a student will (5 years later) attend college. Authors learned the students are successful in middle school mathematics as measured by the tutoring system are more likely to enrol 5 years later in college. On the other hand the students who showed confusion, carelessness in the system have lower probability of college enrollment. For the prediction they used logistic regression classifier. Vihavainen et. al [4] worked with a snapshot data from Computer Science students programming course of Helsinki University and tried to predict whether a student will fail introductory mathematics course.

Bayer et. al [5] predicted whether a bachelor student will drop-out from university. They worked with the data of Applied Informatics bachelor students from Masaryk University and predicted student's studies, activities with other students via email or discussion. They found students who communicate with students having good grades can successfully graduate with a higher probability than students with similar performance but not communicating with successful students. In this case, J48 decision tree learner, IB1 lazy learner, PART rule learner, SMO support vector machines have been used.

Bhardwaj and Pal [6] predicts students performance and found out living location has high influence on students final grade. They used Purvanchal Universitys Department of Computer Applications student's data and used Bayesian Classifier for predicting.

Al-Radaideh, et al [7] guess the final grade of students who studied the C++ course in Yarmouk University, Jordan. Three different classification methods namely ID3, C4.5 and the NaiveBayes are used. The results indicated that Decision Tree model had better prediction than other models.

Nguyen and Peter [8] forecast the performance of the students and compared the efficiency of two classifiers namely Decision Tree and Bayesian networks using WEKA tool. They used two different groups of students of undergraduate and postgraduate level. The performance of Decision Tree was 3-12% more accurate than Bayesian networks. This research was helpful for identifying the weak students for guiding and selecting good students for scholarship.

S. Anupama and Vijayalakshmi [9] expect the performance of final exam of MCA students according to their internal marks. They used C4.5 Decision tree algorithm. They compare the predicted results and actual results which indicates, that there was a significant improvement in results as the prediction helped a lot to identify weak and good students and help them to score better marks. They also compared the model with ID3 Decision Tree algorithm and prove that the developed model is

better in terms of efficient and time taken to build the decision tree.

Smith et. al. [10] propose visible school security measures by student academic performance, attendance, and post secondary aspirations. Baris Cetin [11] introduce an approaches to learning and age in predicting college students' academic achievement. Enhanced higher order orthogonal iteration algorithm for student performance prediction has been proposed by Prema et. at [12].

In this paper we propose neural network based students performance analysis which is most efficient than the others. We show that the student's performance depends on student's live style, leaving place and social interactions.

## III. PROPOSED SOLUTION

The overview of the system is shown in Figure 1. It depicts the overall idea how the prediction will happen. First of all we collect data from educational institutes. After collecting the data we preprocess it for the data mining tool. In this case, we use MATLAB tool which is a fourth-generation multi paradigm programming language and data mining tool. Using this tool we organized the data. Then we design a supervised neural network which trains the data and extracts information from the data. In our experimental case, we predict the student's yearly performance. We also use Levenberg-Marquardt back propagation algorithm to train the neural network. In this system, we predict student's performance and then we find out the accuracy percentage. Each part of our propose system is described in the following subsections.

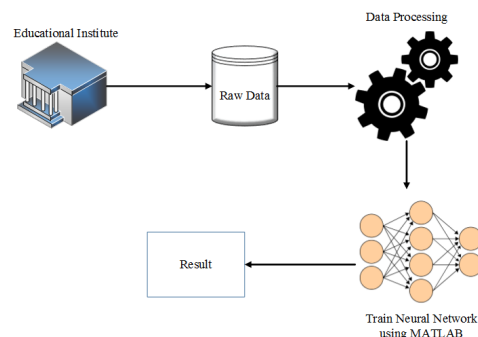


Fig. 1. System Overview

### A. Data Collection

Student's performance is especially significant for institutions because its raking depend on student's quality. Most of the time institutions need observe student's yearly performance for their future development. Despite the importance of predicting student's performance, collecting such data from the real-world is a challenging task. For predicting the student's yearly performance at first we collect student's relevant data. Here we use 120 students data for the research which is taken from the department of Computer Science and Engineering of Bangabandhu Sheikh Mujibur Rahman Science and Technology University. We collect these data from the students via an

online survey. In this survey we consider fourteen factors. In these factors, we regard as not only academic data but also others personal information such as family education, living area, social media interaction, drug addiction etc.

### B. Data Processing

In recent years, education institute collect a vast amount of data. As the volume of data increase day by day the relationship underneath the data is becoming more challenging. Hence, we need to filter out some unnecessary data. For this research, we used MATLAB as the data mining tool for filtering these unnecessary data. MATLAB works with numerical value that is why we convert the data into different numerical values. Then we divide the fourteen factors in two parts. One part is the input part which uses thirteen factors among these fourteen factors and remaining one factor is called target value. This target value is the most essential part of the data because the neural network will find the pattern of the data and predict the result based on the target value. Figure 2 shows the steps of data reprocessing.

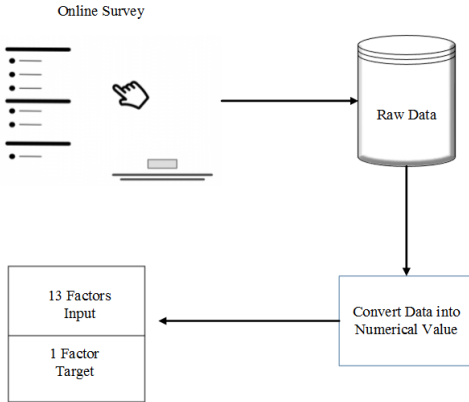


Fig. 2. Preprocess the data

### C. Neural Network

Recently, research in the field of neural networks has been attracting rising concentration. Neural network has solved some of the mysteries posed and makes some improvement for future predictions. To get the prediction first we make a neural network. Then we train the neural network with the pre-processed data. We used supervised neural network and train the network using Levenberg-Marquardt back propagation algorithm. This method is the part of the neural network toolbox integrated in the MATLAB. Supervised learning is the machine learning system of inferring a function from labeled training data. In supervised learning system two kinds of data are provided which called inputs and outputs. Then the system processes the inputs and compares its verdict outputs against the required outputs. Sometimes systems find some errors. These errors are back propagating to the system. Then the system adjusts the weights which control the network. This process happened recursively as the weights are frequently

tweaked. The dataset which permit the training process is called the training set. In the training phase of the network same dataset is processed frequently as the connection weights are always experienced. Sometimes training process continues some days. This process is stopped only when the system reaches some statistically desired output or expected accuracy. However, some networks never learn because the input data does not contain the specific information. That is why we cannot find the desired output. Networks cannot manipulate essential learning if there is not enough data. There should be an adequate amount of data so that part of the data can be held back as a testing phase. This network consists many layered including multiple nodes in each layer are capable of storing information.

Levenberg-Marquardt algorithm [13] propose second-order training speed method without having to compute the Hessian matrix. When the performance function has the form of a sum of squares, the Hessian matrix can be approximated by equation 1 and the gradient can be computed by equation 2.

$$H = J^T J \quad (1)$$

and the gradient can be computed as

$$g = J^T e \quad (2)$$

where  $J$  is the Jacobian matrix [14] that contains first derivatives of the network errors with respect to the weights and biases,  $J^T$  is transpose Jacobian matrix and  $e$  is a vector of network errors. The Jacobian matrix can be computed through a standard back propagation technique that is much less complex than computing the Hessian matrix.

## IV. DESCRIPTION OF DATA-SET

In our dataset we use fourteen factors. We divide these factors into two parts for prediction. The factors are described in below.

### A. Class Test Marks

Class test Marks is a significant element for student's performance analysis. Generally, the students who get good marks in class test having a good chance to do good results in final exam. We take the class test marks of the students and divide this into three divisions which are Poor, Average and Good. If the marks is below forty percent, that is poor. If the mark is between forty and eighty percent then it is average and if the marks is above eighty percent then it is good.

### B. Class Performance

Class test marks is one of the elements of student's performance measurement. There are some other elements. One of these elements is class performance. In our experimental analysis, we consider two divisions of class performance: good and poor. If the students class performance is satisfactory level then we called it is good otherwise poor.

### C. Class Attendance

Class Attendance is a vital factor for student's performance analysis. In our observation, we show that the students who attended class regularly have a great chance to understand teachers lecture than the others who do not attend classes. That is why these students find good marks for regular attendance and can answer the questions fully in the exam hall. We consider class attendance mark is ten. If students mark is less than four then it is poor if it is between five and seven then it is average and above eight is good.

### D. Assignment

Assignment is a common factor for students. Every teacher likes to assign assignment for measuring students understanding and new thinking capability. It has three categories. If the marks is below forty percent then is poor, if it is between forty and eighty then it is average and if the marks is above eighty percent then it is good.

### E. Lab Performance

Practical is very essential for permanent learning. It also helps others very well. Using practical knowledge we contribute our economic and social development very well. Lab performance is best for measuring practical knowledge. Three categories of lab performance are considered in this paper. First one is poor performance; second one is average and lastly good performance.

### F. Previous Semester Result

The student's who get good result will inspire to continue his/her performance for the next time. On the other hand, the student's do not achieve expected results will be disappointed. That is why, future results depend on current result. Similarly current results depend on previous results. Hence, we consider last semester result for our experimental analysis.

### G. Study Time

Study time is main factors for students who study more time have an opportunity to learn more things. We divide this study time into two categories. Those students who study less than three hours per week is in the average category and those study more than ten hours per week is in the good category. This hours are expects class time hours.

### H. Family Education

Learning phase is started at family. If parents are educated then they can inspired of their children to be educated. If parents are able to see her children future goal then the children try to reach the goal. That is why; family education is very important factors student's quality. Student's parental education statuses are divided into three categories and they are poor, average and good.

### I. Living area

Our country, Bangladesh, is a developing country. The development of this country is not equally distributed. The students who live in town area are more concern about their future compare to who live in remote area. On the other hand, student's who get higher education from his/her home will get special care from their parents. Regarding this issues, this factor is divided into two categories. Student's living in hall/mess are in one category and those are living at home are different category.

### J. Social Media Interaction

Present time is internet era. Students are connected by the internet for their study, family and personal purpose. Most of the youth are habituated to use social media like Facebook, Youtube, Linkedin and so on spend the huge time. The student's spend more time will get less time to study. Hence, social media interaction is considered a factor of student's performance analysis. Student's who spend more than thirty minute and less than two hours per day with social media are in average category and those who spend more than two hours per day are in Exceed category.

### K. Extracurricular Activity

Extracurricular activity helps student's refreshment and increase thinking power. It is also good for health. Since extracurricular activities are helpful for students performance. In our research paper we consider the student's who are associated with extracurricular activities are in yes category and who are not associated are in no category.

### L. Drug Addiction

Addiction is very harmful for health and mental improvement. The student's who take drug is not serious about their study. They are involved different violence work in our society. Student's performance depends on drug addiction. That is why, we consider drug addicted students are in yes category and no drug addicted student's are in no category.

### M. Affair

Affair is a relationship among boys and girls. Recent research shows that affair could affect the academic performance. So, the student's who have affair are in the yes category and vice versa are in the no category.

### N. Year Final Result

Above discussion factors are input factors. Based on these factors we measure student's performance that is called year final result. This factor is the target factor. In this case, the final year CGPA of students is stored. Neural network will target this factor to train the networks. After the prediction, predicted result will be compared to the original results and will be find the accuracy percentage.

The overall view of these fourteen categories are given in Table I.

TABLE I  
DATA SET

Factors	Value	Factors	Value
Class Test Mark	Poor Average Good	Family Education	Poor Average Good
Class Performance	Average Good	Living Area	Hall/Mess Home
Class Attendance	Poor Average Good	Social Media Interaction	Average Exceed
Assignment	Poor Average Good	Extra Curricular Activity	Yes No
Lab Performance	Poor Average Good	Drug Addiction	Yes No
Study Time	Average Good	Affair	Yes No
Previous Result	CGPA	Year Final Result	CGPA

## V. IMPLEMENTATION

In the implementation of the system first we take the preprocessed dataset and use them in the MATLAB. Then we make the neural network with ten hidden layers. To train the network we add the training algorithm, in this case we use Levenberg-Marquardt backpropagation algorithm. For applying this in MATLAB we use trainlm function.

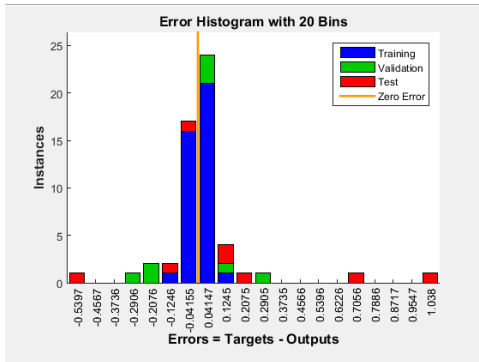


Fig. 3. Error Histogram

The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update in the equation 3.

$$X_{k+1} = X_k [J^T J + \mu I]^{-1} J^T e \quad (3)$$

Where,  $X$  is bias variable,  $J$  is Jacobian matrix,  $\mu$  is scalar,  $I$  is identity matrix,  $e$  is vector of network errors. In our experimental analysis we divide the dataset in three parts. First part is training dataset where we use 70 % data. Second part is testing dataset where we use only 15 %. And most importantly, another 15 % data is used for validate the results. After taking all these steps we train the network using train function. We find the errors from the difference between targets results and output results. In the figure 3, shows the different errors histogram with 20 bins.

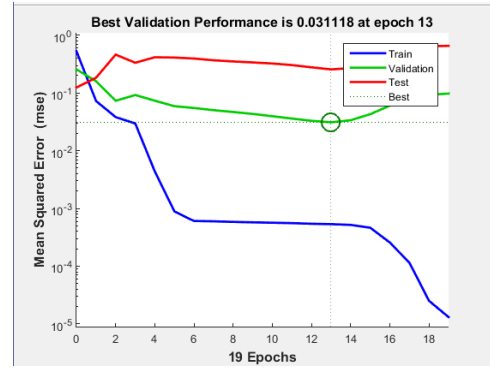


Fig. 4. Best Performance Validation

## VI. RESULT

After running the neural network we have the predicted result. The result is best validated at 13 epochs out of 19 epochs. Figure 4 shows the performance of the neural network.

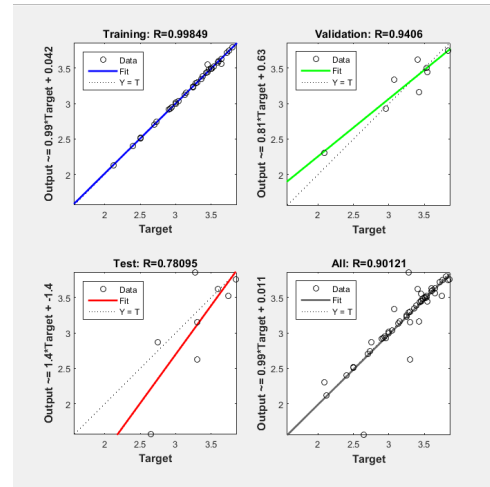


Fig. 5. Regression

In the figure 4 dot represent best results. We observe validation results touch our dot line at 13 epochs and the performance is 0.031118 which represent performance accuracy is around 97 %.

We use three parts of dataset regression. In the figure 5 shows the training, Testing, validation and all phase regression analysis of the network. Where dash line represents the target and solid line represents the best fit linear regression line between outputs and targets.

Our experiment we use seven students results prediction randomly. In the table II shows these students original result and predicted results. Based on these results we find the accuracy percentage. In the table we show the highest is 99.986 % when students original result is 2.7 and we get lowest accuracy 91.547 % when original result is 3.83. In this results, we say that our system is better when students result is poor. Because good results depends many factors. That is why, its accuracy is small than the others. Oladokun et. al [15] also

TABLE II  
RESULT

Original Result (CGPA)	Predicted Result (CGPA)	Accuracy Percentage
2.125	2.116	99.555%
3	3.192	93.973%
3.83	3.506	91.547%
3.43	3.358	97.9%
3.535	3.604	98.064%
3.5	3.479	99.416%
2.7	2.664	99.986%

did a study on predicting students performance using artificial neural network. They used multilayer perception which was trained by static back propagation. The accuracy of the results was at best 88 % and the average result was 74%. But our studys result shows that we have accuracy more than 90% which is better than the existing study.

The Mean Squared Error (MSE) is a measure of how close a fitted line is to data points. The smaller the MSE, the closer the fit is to the data. And the Root Mean Squared Error (RMSE) is just square Root of MSE. We calculated the RMSE with same data set and shows the comparison result in the figure 6.

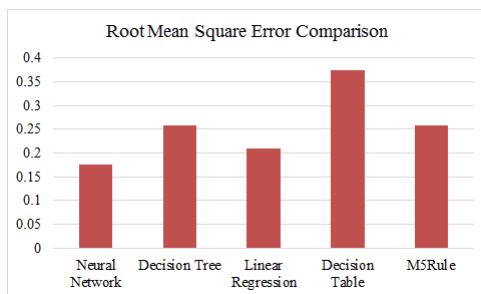


Fig. 6. Comparison of RMSE among different techniques.

The figure shows that our propose Neural Network method RMSE is 0.1765. The other techniques like Decision Tree, Linear Regression, Decision Table and M5Rule RMSE values is greater than the Neural Networks. That is why, we said that our method is better than the others.

## VII. CONCLUSIONS

The study has unveiled, students yearly performance greatly depends on not only academic activities but also on external activities such as, social media interaction, living area condition and so on. Artificial neural network can be used efficiently to predict students performance better than other methods. This study will help the educational institute in terms of selecting students for serious and rigorous academic researches, and also will help them identify those who need extra care towards better growth.

In our experiment, we have used only 120 students data. In our future research, we will use larger datasets and will try to engage neural networks towards predicting consumer behavior analysis and many other like this.

## ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their comments that improved the quality of our paper. This research was supported by the research fund of Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh. Sajal Halder is the corresponding author.

## REFERENCES

- [1] S. Sebastian and J. J. Puthiyidam, "Evaluating students performance by artificial neural network using weka," *International Journal of Computer Applications*, vol. 119, no. 23, pp. 36–39, 2015.
- [2] S. K. Yadav, B. Bharadwaj, and S. Pal, "Data mining applications: A comparative study for predicting student's performance," *arXiv preprint arXiv:1202.4815*, 2012.
- [3] M. O. Pedro, R. Baker, A. Bowers, and N. Heffernan, "Predicting college enrollment from student interaction with an intelligent tutoring system in middle school," in *Educational Data Mining 2013*, 2013.
- [4] A. Vihavainen, M. Luukkainen, and J. Kurhila, "Using students' programming behavior to predict success in an introductory mathematics course," in *Educational Data Mining 2013*, 2013.
- [5] J. Bayer, H. Bydzovská, J. Géryk, T. Obsivac, and L. Popelinsky, "Predicting drop-out from social behaviour of students." *International Educational Data Mining Society*, 2012.
- [6] B. K. Baradwaj and S. Pal, "Mining educational data to analyze students' performance," *arXiv preprint arXiv:1201.3417*, 2012.
- [7] Q. A. Al-Radaideh, E. M. Al-Shawakfa, and M. I. Al-Najjar, "Mining student data using decision trees," in *International Arab Conference on Information Technology (ACIT'2006)*, Yarmouk University, Jordan, 2006.
- [8] N. T. Nghe, P. Janecek, and P. Haddawy, "A comparative analysis of techniques for predicting academic performance," in *Frontiers In Education Conference-Global Engineering: Knowledge Without Borders, Opportunities Without Passports, 2007. FIE'07. 37th Annual. IEEE*, 2007, pp. T2G–7.
- [9] S. A. Kumar and M. Vijayalakshmi, "Efficiency of decision trees in predicting students academic performance," in *First International Conference on Computer Science, Engineering and Applications, CS and IT*, vol. 2, 2011, pp. 335–343.
- [10] E. E. Tanner-Smith and B. W. Fisher, "Visible school security measures and student academic performance, attendance, and postsecondary aspirations," *Journal of youth and adolescence*, vol. 45, no. 1, pp. 195–210, 2016.
- [11] B. Cetin, "Approaches to learning and age in predicting college students' academic achievement," *Journal of College Teaching & Learning (Online)*, vol. 13, no. 1, p. 21, 2016.
- [12] P. Nedungadi and T. Smruthy, "Enhanced higher order orthogonal iteration algorithm for student performance prediction," in *Proceedings of the Second International Conference on Computer and Communication Technologies*. Springer, 2016, pp. 639–649.
- [13] A. A. Suratgar, M. B. Tavakoli, and A. Hoseinabadi, "Modified levenberg-marquardt method for neural networks training," *World Acad Sci Eng Technol*, vol. 6, pp. 46–48, 2005.
- [14] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *Neural Networks, IEEE Transactions on*, vol. 5, no. 6, pp. 989–993, 1994.
- [15] V. Oladokun, A. Adebajo, and O. Charles-Owaba, "Predicting students academic performance using artificial neural network: A case study of an engineering course," *The Pacific Journal of Science and Technology*, vol. 9, no. 1, pp. 72–79, 2008.