The Comparison of Forward and Backward Neural Network Model – A Study on the Prediction of Student Grade

SAIKAT GOCHHAIT[1], YAGYANATH RIMAL[2], SAKUNTALA PAGENI[3]
[1]Symbiosis Institute of Digital and Telecom Management, constituent of Symbiosis International (Deemed University), INDIA
[2,3]Pokhara University, NEPAL

Abstract :- A neural network model can be used effectively in predicting training accuracy using machine learning. Based on the comparison of forward and backward neural networks, coded to communicate their output in the requisite manner using machine language is the basis of the present study. With the help of students' background information, to predict the Grade Point Average (GPA) of 580 engineering students based on various parameters, including mental health. The study is based on the Boruta algorithm and the random forest methods for data preparation in the matrices (12 * 2 = 24) of single-layered, multiple-layers, and forward and reverse algorithms adopted to test the prediction and accuracy of the grade point average by analyzing histograms, confusion matrices, and regression analysis. This study suggests the best model for predictions with the help of artificial neuron network that has roughly half the number of single layers and with three hidden layers.

Keyword:- Grade point average, Hypothesis design, Multinomial regression, Deep neural network

Received: February 28, 2021. Revised: July 17, 2021. Accepted: July 28, 2021. Published: August 8, 2021.

1 Introduction

The process of learning a concept that is unfamiliar requires an understanding of brain development that is more practical and appropriate, which facilitates learning knowledge, skills, values, beliefs, and habits of human learning. Using multiple instance regression, we solve the problem of qualitatively and quantitatively predicting the flexible length of the major histocompatibility complex of student grades. It is not possible to develop reliable machine learning neural network model prediction tools without high-quality data sets, including student test results, final grades, and their relationship to past performance [4]. In contrast, the neural network model for GPA prediction indicates the shortcomings of these commonly used data sets that are often used to evaluate machine learning approaches in GPA prediction [12]. Lastly, the study proposes an enhanced similarity reduction procedure with the help of backward and forward neural networks for student grade prediction which is more stringent than the standard methods currently used.

In a similar study conducted by the department of education at the University of Maastricht, the researchers concluded that there are no relationships between students' perceptions of assessment and their performance on assessments [3]. Students prefer written assessments, including home exams and essays in which they are allowed to quote sources. Materials such as notes, books, and papers. Computerized tests and portfolios aren't among the students' preferred methods of evaluation. Oral testing isn't among their preferred methods of evaluation either. According to the Islamic University of Bahawalpur's grade prediction research, 16 students scored more than 80 in internal assignments, but failed the final examination [7]. However, this study only looked at master's in education students. Accordingly, the students receiving the highest score of 88.07 in assignments and the lowest score of 49.57 in the final examination stand out.

In a similar way to explore various attributes of success among a sample of 50 students using the ID3 algorithm [1]. The study to analyze 300 students from affiliated colleges of Panjab University using linear regression [8]. In a study conducted at a Nigerian university, from data collected on 1847 students and they discovered that the instrument's internal consistency reliability was 0.86 when administered in heterogeneous classes at such a university [5]. Based on the data warehouse of student records, the Naive Bayes method was applied to predict student performance. To achieve an accuracy of 86.66%, 175 records were used for training data and 45 records were used as test data for the study of 220 students. The study suggests in using algorithm to predict the outcome among a large class set of data, in such a way that it is at the university level requirement to run three tests of any subject [10]. The study claim to have achieved 82 percent accuracy by using an artificial neural network based on J48 and multiple linear regression along with regression analysis of 181 students' marks to predict student performance [13].

The study depicts the Bayesian and Combined Health Information Database (CHID) algorithms to predict the performance of students in the higher secondary school based on 35 contributing factors [11]. The results are reported by classifying 300 students using Na*ve Bayesian reasoning [1]. The study uses Bayesian classification to classify 600 students [9]. The previous study based on the analysis of 778 Portuguese language and mathematics students with the help of tree algorithms applied to study the most effectiveness in distinguishing between students who would pass or fail their courses [2].

2 Theoretical Frameworks

The present study with the help of neural networks algorithms in order to forecast student achievement in schools, colleges, and online multiple-choice courses [6]. Using artificial neural networks, researchers estimated the future student performance based on the forward and backward neural network models to predict student grades based on students' family, social, health and wealth characteristics.

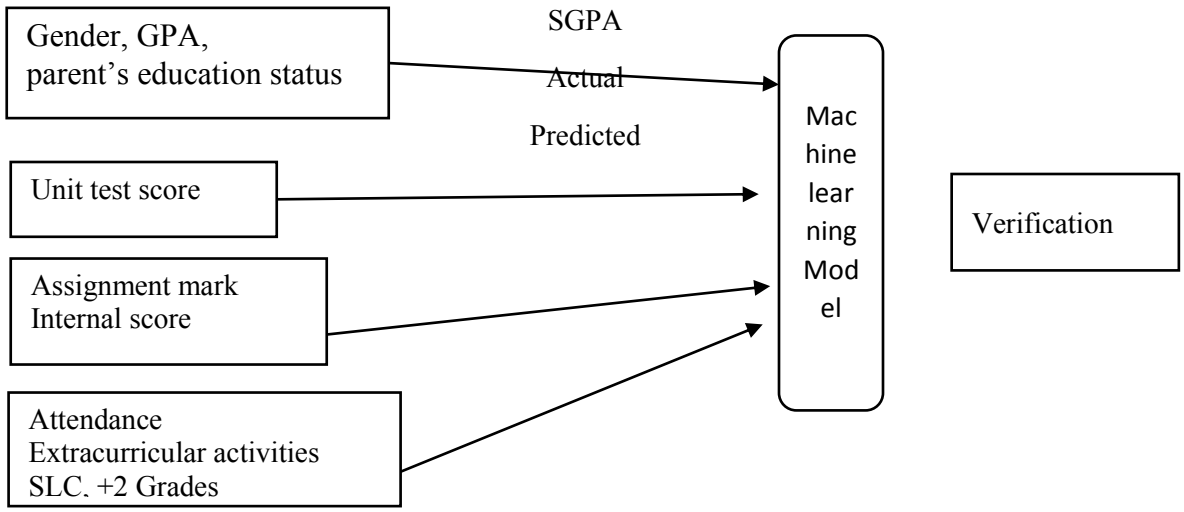

Figure 1: Conceptual Framework of Research Design

The figure 1 depicts the teaching-learning process with the combination of various dimensions, through the use of neural networks in predicting the SGPAs verified against the final university grading for model validation and to reduce errors of misclassification.

3 Objectives

The specific purpose of this study is to analyze the student grades for SOE for the fall of 2019.

The purpose of this assessment is to determine whether the performance of continuous assessment to determine the final grading.

To determine the correlation between internal marks and final grades achieved by the students.

4 Model Design and Setup

Students' predictions were scored using the powerful R open source tools. The tool offer a wide range of classification and prediction methods. The algorithm for predicting bachelor's degree grades using a neural network. The holdout approach for evaluating classifiers refers to training the models using 70% of data while evaluating them on 30% of data called testing data or unseen data. The reason for this is to avoid over fitting, a phenomenon that occurs when analyzing the same set of records from which the model was constructed [11]. Boruta algorithm and random forest model were also analyzed sufficiently to separate unwanted columns of research into 49 columns. Figure 2 shows the validation of a data mining classification model considered as the most important phase in data mining process. The process of validation helps in assessing how well the data mining models perform against real time data.

.

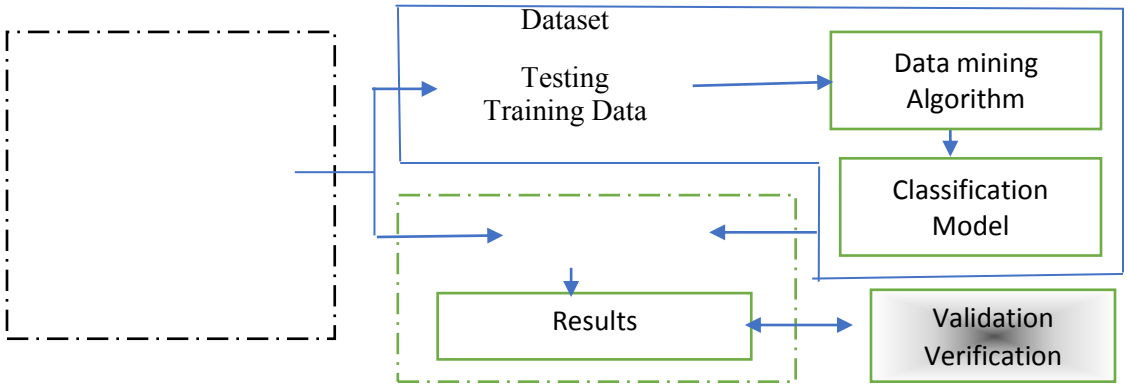

Figure 2: Data Organization and Management

5 Model Performances and Evaluation

Students grading is a collective effort of the students' previous and ongoing studies during each semester, with the help of neural network in predicting the items with high confidence with the help of input weight of each matrix element of a variable to another complete neuron that's formed using the outputs of previous neurons, weighted with bias terms, pushed for creating signal weights for students at the beginning of every semester by applying that information to 580 undergraduate engineers. To predict the best outcome, both full and half

models of the neural network were created based on a variety of parameters. While the full model includes 580 student records, the half model includes 290 students whose exam results were valid in fall 2019. Using the 22 input signals as input, the neural network first passes the data over a single hidden layer whose confusion matrix, the output neuron, and the accuracy of the both training and testing are calculated. Based on the dependent variables, the neural network output is calculated into the number of subdivision grades. As a result of some unimportant variables with similar patterns of marks to the model, the prediction accuracy determined through the forward and backward propagation summarized in the following table.

S. N	Predicator Variable	No of Input	No Hidden	Training Accuracy	Test Accuracy	Average
1	P/F (0/1)	22 Full	1	91	81	86
	P/F ()/1)	22 Half	1	98	95	96.5
2	GPA(Deci.)	22 Full	1	38	41	39.5
	GPA(Deci.)	22 Half	1	82	83	82.5
3	GPAX(Cat.)	22 Full	1	98	95	96.5
	GPAX(Cat.)	22 Half	1	98	95	96.5
4	GPA	13 Full	1	83	84	83.5
	GPA	13 Half	1	99	93	96
5	GPA	13 Full	5	91	87	89
	GPA	13 Half	5	92	93	92.5
6	GPA	13 Full	7,3	85	88	86.5
	GPA	13 Half	7,3	75	80	77.5
7	GPA	13 Full	7,3,2	91	89	90
	GPA	13 Half	7,3,2	81	82	81.5
	Back Propagation					
1	GPA	13 Full	1	48	52	50
	GPA	13 Half	1	82	82	82
2	GPA	13 Full	5	50	91	70.5
	GPA	13 Half	5	82	82	82
3	GPA	13 Full	7,3	52	52	52
	GPA	13 Half	7,3	76	76	76
4	GPA	13 Full	7,3,2	88	80	84
	GPA	13 Half	7,3,2	83	84	83.5
5	GPA	13 Full	25,12,7,3	82	82	82
	GPA	13 Half	25,12,7,3	73	73	73

Table 1: Neural Network Model Accuracy

Table 1 depicts how best neural network model can be used in order to determine the accuracy for both training and test with the sets of samples as part of this study. Therefore 24 different neural network models created.

To predict the student grades for those passing results were excellent, 22 variables were analyzed when previous grades of students were combined with historical information. As half the model represents 290 students while a full model passes all 580 student information. As a single neuron in the hidden layer with the output of 96 and 86 percent accuracy for a variable with two states is the university's final result for a variable with only two states. The student with GPA of 4.0 is at least 82 percent more accurate than the student who has a GPA of 2.0. In the same way, when grade point average is defined as categorical variables of grades such as A, A-, B+, B... using the same value, the best average precision is 96 percent: 96 percent for both full and half models.

To predict student grades based on the significant variables (previous grades and internal evaluation of ongoing six subjects' grades), therefore neural model needed to be simplified rather than predictions that were accurate by 96 percent and 83 percent, respectively. Based on five neurons in the hidden layer, this model has an accuracy of 92:89 percent. Using seven hidden layers in both half and full models produce the best accuracy of 77:86 percent and with the help of seven hidden layers in both half and full models produces the best accuracy of 81:90 percent. Three hidden layers were used by the model to determine both the accuracy models; the decreasing pattern produces over 80% accuracy.

Similar to an algorithm used to predict student grades by reverse-backward traversing neurons' weights after first passing to them. The accuracy of the Model with 50% when a single hidden layer conversed at 82. In the same way, five neuron models with 82:70 percent accuracy can be considered to be favorable with the forward network with the same inputs. As well, the backward traversing model of 76:52 is at least as good as the forwarding model of 77:83 when two hidden layers (7, 3) are placed between independent and dependent variables. The best model, using the same input in a forward neural network, has an accuracy of 77:90 percent with three hidden layers (7,3,2) between 13 inputs to a single variable and a backward model, 83:84 percent. When 73 to 82 percent accuracy is obtained from a backward model with four hidden layers (25,12,7,3) of 13 input signals, then the reversed accurate full model is the best algorithm.

In any case, neural communication to predict student grades can be improved based on input data and weight; the three-layer back propagation of original student score prediction is greater than 90%, which is excellent. The GPAX with the categorical numeric model with a single neuron considered to be best with the help of forward neural network when there are more variables as independent.

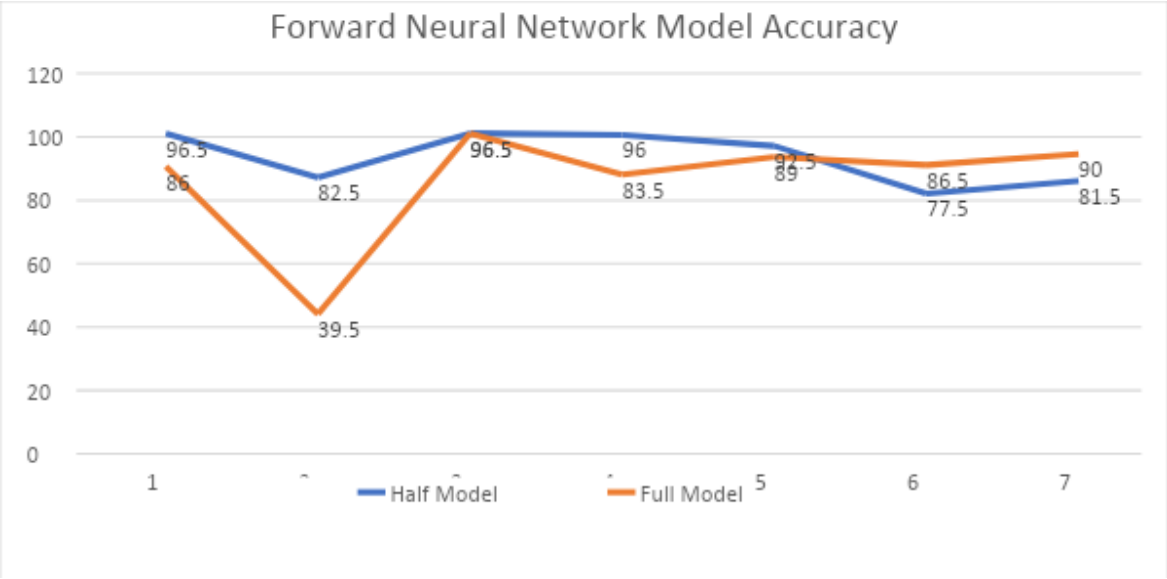

Figure 3: A line chart comparison

In considering the above line graph, it appears the prediction of GPA when results pass or fail is more robust when independent predicting variables have large numbers as compared to the scores they scored and Grade Point Averages they are awarded. In predicting variables with the fewest layers, input signals considered as the best model for predicting student grades, whereas the half and full models are considered as more time consuming between input and output.

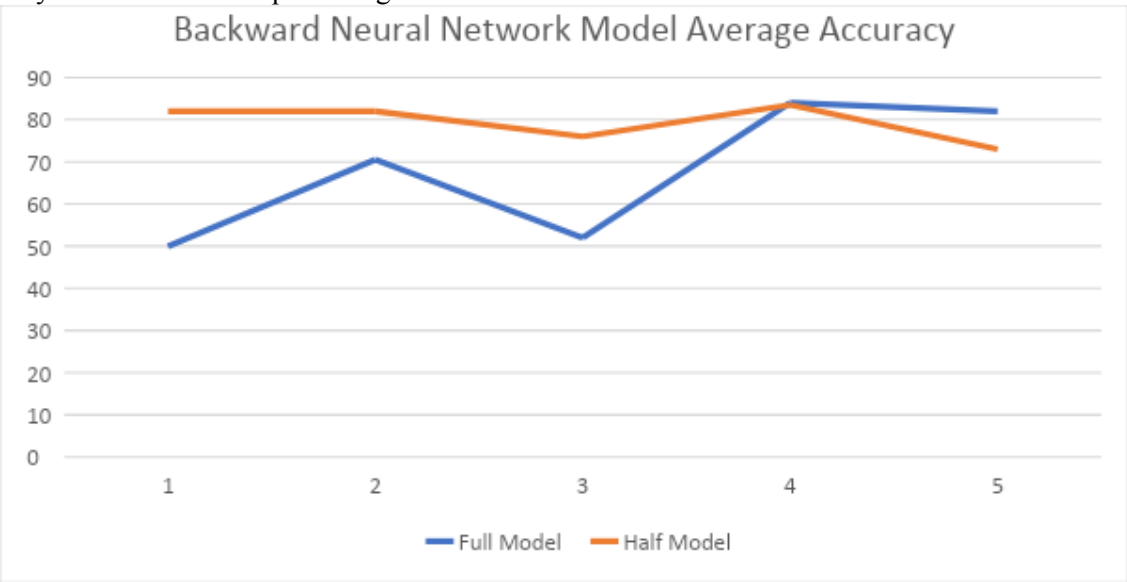

Figure 4: Line Chart Backward Comparison

Accordingly, the GPA grade for the single layer output of either a half or a full model with 13 independent variables will be least accurate compared to two and three hidden layers. It is therefore recommended to use a three-layer neural network whose average output in both full and half is greater than 80 percent as a backward model.

6 Conclusion

In the direct application of multiple linear feed forward networks of student grade prediction, neural network models are used to solve very large daily life activities. Study is designed to develop a neural network model that can accurately predict student grades based on background information compared to total accuracy under different mental health conditions, internal assessments, and final grades. As a result of Boruta and the random forest model test, 22 input signals are used to convert the neural network model. For forward passes, it is recommended that for large enough predicting variables, a single hidden layer with fewer neurons be used, while for reversed backward passes, three hidden layers with reduced input signals should be used.

In the future research it is very important to indentify which students' characteristics are associated with test results, and which school characteristics are associated with the added value of the school [14]. In this regard it is useful for researchers to apply machine learning applications to acquire knowledge about students' learning in different subjects, develop optimal warning models, and discover behavioral indicators from learning analytical reports.

Funding : This research was completed under the financial support of University Grant Commission of Nepal as the proposal selected on topic mini research.

Conflicts of interest/Competing interests : This research has no conflict of interest however

researcher has taken ethical letter from the school of engineering Pokhara university. Similarly, the final examination result of student of all colleges open-source pdf available under website pu.edu.np.

Availability of data and material the data of source code could be provided, if necessary, on demand freely form the primary author.

Code availability the source code of this research analysis having around 1500 lines of code on r programming. This could be available without fees if any researcher on demand from primary author.

Authors' contributions as the first author primary researcher could have responsible on data collection production source code and interpretation of this research. This research is primary research conducted under university grades new way using machine learning neural network with 18 models with many data training.

References:

[1] Brijesh Kumar, B., & Saurabha, P. (2011). E preceding of 4th Global summit. Mining education Data Analyze student performance.

[2] Cohen, W. (1995). Effective rule induction. Twelves international conference on machine larning.

[3] David, G., Rijt, G. v., Filip, D., & Janine, v. d. Students' assessment preferences, perceptions. High Educ (2008) 56:645–658.

[4] Hussain, M. Z., & Abidi, S. M. Student engagement predictions in an e-learning system and their impact on student course assessment scores. Computational intelligence and neuroscience. vol. 2018, Article ID 6347186, 21 pages, 2018. https://doi.org/10.1155/2018/6347186

[5] John, A., & Manabete, S. Procedural Influence on Internal and External Assessment score of Undergraduage. Journal of Education and Practices;2015.

[6] Junemann, M. L. Neural networks to predict schooling failure/success. Lecture Notes in Computer Science, Nature Inspired Problem-Solving Methods in Knowledge Engineering. Berlin: Sp;2007

[7] Muhammad, A., & Aijaz Ahmed, G. CORRELATION BETWEEN MARKS OBTAINED BY STUDENTS. Turkish Online Journal of Distance Education-TOJDE April 2007 ISSN 1302–6488, Volume: 8 Number: 2 Article: 10.

[8] Hijazi, S. T., & Naqvi, S. M. Factors affecting students performance: A case of private colleges. Bangladesh E-Journal of Sociology;2006, 3 (1), 65-99.

[9] Pandey, U.K., & Pal, S. (2011). Data Mining : A prediction of performer or underperformer using classification. ArXiv, abs/1104.4163.

[10] Gadhavi, M., & Patel, C. Student final grade prediction based on linear regression. Indian J. Comput. Sci. Eng; 2017,8(3), 274-279

[11] Ramaswami, M. and R. Bhaskaran. "A CHAID Based Performance Prediction Model in Educational Data Mining." ArXiv abs/1002.1144 (2010): n. pag.

[11] Rimal, Y. (2019). Machine Learning Random Forest Cluster Analysis for Large Overfitting Data: using R Programming. In 2019 6th International Conference on Computing for

Sustainable Global Development. (INDIACom) (pp. 1265-1271). IEEE.

[12] Rimal, Y. P. et al (2021). Hyperparameter Determines the Best Learning Curve on Single, Multi-Layer and Deep Neural Network of Student Grade Prediction of Pokhara University Nepal. 2021. In Journal of Phys.

[13] Sorour, S., Mine, K., & Hirokawa, S. (2014). Predicting students' grades based on free style comments data by artificial neural network. In 2014 IEEE Frontiers in Education Conference (FIE) Proceedings, page 19,.

[14] Masci, C., Johnes, G., Agasisti, T. (2018) Student and school performance across countries: A machine learning approach. Eur. J. Oper. Res. 269, 1072–1085. [Google Scholar] [CrossRef]