Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 8, July 2021: 463-471

Deep Learning in Students' Performance Prediction

S. Saranya[1],C.Priya[2],M.Vidhya Lakshmi[3],R.Premkumar[4]

[1-4]Department of Electronics and Communication Engineering,

[1]Easwari Engineering College, Ramapuram, Chennai, Tamil Nadu, India.

[2]Mohamed Sathak Engineering College,Kilakarai,Tamil Nadu, India.

[3]SRM Institute of Science and Technology,Ramapuram,Chennai, Tamil Nadu, India.

[4]Siddharth Institute of Engineering and Technology, Chittoor,Andhra Pradesh, India.

Abstract:

Many students endure increased tension and anxiety in the days leading up to and during their exams. Test anxiety, on the other hand, is more severe and can hinder learning and test performance. Exam anxiety is more than just apprehension prior to a test. A little exam anxiousness can escalate into crippling feelings of concern, dread, and panic for a student who suffers from test anxiety, which can have a poor impact on their performance. At any age, a student suffers from test anxiety. Test anxiety begins to rise in students in the second to fourth grades and continues to be a problem throughout the middle school yearsage and makes a huge impact on their performance. The anxiety is caused due to several things like timed examination, taking the test in a crowded classroom, and the fear of the unknown. The prediction of the student's performance can be highly beneficial for reducing student's anxiety for the examination. A Deep learning model that uses predictive analytics can be used for the prediction.

Keywords: Anxiety, Prediction,Deep learning model, Students' performance.

1. Introduction

Education is an imperative aspect for the people because it provides people with knowledge, enables them to know their rights and duties towards the family, society as well as the nation. Education is also a salient criterion for sophomore's well-being like employment income and social status and is a strong predictor of attitudes. The knowledge of students that they acquire from education is tested through competitive exams, practical exams, etc. A forthcoming test or exam can be a stressful for any student at times.

Many students endure increased tension and anxiety in the days leading up to and during their exams. Test anxiety, on the other hand, is more severe and can hinder learning and test performance. Exam anxiety is more than just apprehension prior to a test. A little exam anxiousness can escalate into crippling feelings of concern, dread, and panic for a student who suffers from test anxiety, which can have a poor impact on their performance. At any age, a student suffers from test anxiety. Test anxiety in students increases dramatically between the second and fourth grades, stays obstructive though middle age, and has a significant impact on their performance. The anxiety is caused due to several things like timed examination, taking the test in a crowded classroom, and the fear of the unknown. Recent surveys give us the information that the students are suffering from mental illness due to the anxiety of facing the exams. They take drastic actions from distracting from the anxiety such as self-harming, eating disorder, and in extreme cases, suicide. As we know facts that computers can predict anything from the data given to them. The prediction of the student's performance can be highly beneficial for reducing student's anxiety for the examination. A Deep learning model that uses predictive analytics can be used for the prediction [1] [2]. The following are the steps involved in the process:

- Collecting data that impacts the student's performance and will be used to predict the student's performance with the model.
- Pre-processing the data and normalizing the data for more accuracy.
- Implementation of Artificial neural networks (ANN) is a tool for predicting student performance.
- The model must be validated before it can be used by students.

Data mining technique is used for collecting the data which can be used by the machine learning model for the prediction. Through data mining technique data like attendance percentage of the students, internet connection available for the students, previous exam performance of the students, students' traveling time, etc. can be collected by which the model will be trained [3]. Data mining is a useful technique that helps to spot the various factors that affect the students' performance [4]-[8]. It helps us to explore the education system and factors that affect the students' health.

2. Literature Survey

VladimirL.Uskov, Jeffrey P.Bakken, Adam Byerly, Ashok Shah

Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education Journal Name: IEEE Year: 2019. Linear regression, logistic regression, k-nearest neighbour classification, nave Bayes classification, artificial neural network regression and classification, decision tree classification, random forest classification, and support vector machine classification were all compared in this study.Best accuracy was demonstrated by linear regression with 3.07% average error between predicted and actual total student final scores.

Tenzin Doleck, David Lemay, Ram Basnet, Paul Bazelais

Predictive analytics in education: a comparison of deep learning frameworks Journal Name: Education and Information Technologies 25(3):1-13 Year:2020.A comparison of the predicted accuracy of prominent deep learning frameworks/libraries and machine learning algorithms, such as Keras, Theano, Tensorflow, fast.ai, and Pytorch. The optimizer employed in deep learning libraries has a different effect on performance as measured by prediction accuracy. Furthermore, we discover that deep learning is equivalent to other machine learning methods in terms of performance.

TismyDevasia, Vinushree TP, VinayakHegde

Prediction of Students' Performance Using Educational Data Mining,An investigation on the classification used in predicting students' performance. This study informs us about the most important factors that must be considered when predicting a student's success.

Student Academic Performance Prediction using Artificial Neural Networks: A Case Study Mubarak Albarka Umar School of Computer Science and Technology Changchun University of Science and Technology, Jilin, China. This study gives us the knowledge about the data mining.

Predicting Students Yearly Performance using Neural Network: A Case Study of BSMRSTU, Md. Fahim Sikder, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh,This research shows that students' annual performance is influenced not only by academics but also by external factors such as social media, living conditions, and so on. A neural network predicts student performance more accurately than other methods.

3. The Artificial Neural Networks

In the structure of the brain, a neural network consists of highly interconnected Processing Elements or units. Each of these processing elements is designed to impersonate its biological counterpart, the neuron. Each unit accepts a set of weighted inputs and provides the related output. Neural Networks Algorithm is used for the computers to solve the problem that traditional computers cant for example cancer detection, fraud detection, speech recognition, etc.

A neural network is the general form of regression analysis. Advantages of neural networks are:

a) There is no need for a function to which the data are fitted.

b) The networks can capture non-linear relationships.

c) Bayesian method helps to estimate the uncertainty of extrapolation.

Another significant feature of the neural networks is that they readily get updated as more past data becomes available; that is the model continues to learn and extend its knowledge. Thus this model is referred to as adaptive systems. The similarity with the brain makes a neural network model simulate a wide range of functions both linear and non-linear.

Neural networks are made up of (n) layers of neurons, two of which are input and output layers. The input layer receives and transmits data from external sources, while the output layer displays the computation's findings. The hidden layer between the input and output layers derive relevant patterns or features from the received signals in the relay. Those features which are considered predominant from the above process are sent to the output layer.

Sophisticated neural networks will have delay elements, feedback loops in addition to the hidden layer. The ability of Neural Networks to solve the complex problem depends on the number of hidden layers although the recent studies suggest that three hidden layers are more sufficient for solving the most complex problem. The test validation set evaluates the performance of the neural networks. About 30% of the total sample of the data is severed as test data and remaining is severed as training data to train the model.Ann can be either supervised or unsupervised. In supervised learning, the networks are given with the actual data as the training set; that is the model is trained with the input with the output given. The actual output is made available to the network along with the input data. Processing of the input and result is then done by the model to get errors which are then back propagated to the system causing it to adjust the weights of the network.

In unsupervised learning, inputs are only given to the model and the outputs are not made available. The training in this model is completed when the model reaches user-defined performance level. Internally, such networks monitor their performance by checking for trends in the input signals and making adjustments in accordance with the network's function; this information is then stored in the network topology.

4. Methodology

Data mining is the process of extracting information from a large amount of data.This technique is used to analyse the end-of-semester examinations of students in a large dataset.

(A) Data Preparation

From session 2016 to 2020, data on students was collected from the college using a sampling method from the electronics and communication department.This step involves joining data into a single set which is stored as different tables.

(B) Data Selection and Pre-Processing

In this step, the data which is more significant for the prediction is selected. The students' register number, 10th, 12th, degree marks, attendance, traveling time, internet connection, etc. is taken as the important factor for the predictions.

(C) Input Variables

The input data are available with the students which are easily obtained through schools or colleges. The input variables are:

1. 10th score.
2. 12th score.
3. UG score up to the previous semester.
4. Attendance percentage.
5. Average study hour per unit.
6. Fail percentage.
7. Network availability.
8. Social media.
9. No.of revisions done.

INPUT VARIABLES		
S.NO	**DESCRIPTION**	**POSSIBLE VALUES**
1	10th score	In percentage (0-100)
2	12th score	In percentage (0-100)
3	Attendance	In percentage (0-100)
4	Average study hour per unit	In hours (0-24)

5	Is the student is provided with network connection?	Yes(1) / No(0)
6	What is the student traveling time from college to home?	In hours(0-24)
7	Is the student there in any social media?	Yes(1) / No(0)
8	How many revisions does the student do?	In numbers (0-10or20)
9	Previous years semester Percentage	In percentage (0-100)

Fig.1. Input Variable

	cs41	cs42	cs43	alhpu	ashpu	10th	12th	att	fp/aari	travh	netcon	rev	semgpa
2	82.17	89	85.59	12	15.6	99	95	82.05	1.5	0.8	1	1	80
3	35.5	40.33	37.92	12	14.4	89	79	95.81	13.5	0.91	0	0	68.6
4	78.5	80.67	79.59	12	14.4	82	94	96.48	0	0.99	1	0	70.83
5	78	82	80	12	13	96	89	83.2	0	0.14	1	0	74.6
6	67.17	70.17	68.67	12	11	80	77	77.07	3	0.58	1	0	64.2
7	82.83	87.17	85	12	17	92	83	94.49	0	0.36	1	1	75.4
8	62.83	67.6	65.22	12	11	98	85	77.01	6	0.04	1	1	72.5
9	75	75.17	75.09	12	16	99.6	94.5	82.72	1.5	0.47	1	1	75
10	74.5	80	77.25	12	15.3	91	95.5	82.51	1.5	0.67	1	1	77.09
11	86.83	88.17	87.5	12	19	96	92.08	94.01	0	0.92	1	1	83.75
12	80.5	86.83	83.67	12	16.6	82	72	74.35	0	0.08	1	1	77.5
13	59	64.67	61.84	12	14.4	92	80	95.97	4.5	0.94	1	0	67
14	44.83	58.5	51.67	12	15	91	81	88.32	9	0.68	1	0	69
15	87.17	87.5	87.34	12	20	99	92	83.53	0	0.2	0	1	81.6
16	40	50.17	45.09	12	19	88	62	72.46	9	0.14	1	1	72
17	75	75.33	75.17	12	13	95	94	70.05	0	0.8	1	0	70
18	53.5	64.83	59.17	12	22	90	78	89.76	6	0.05	1	1	88.57
19	80.17	83.67	81.92	12	17	86	95.16	80.64	0	0.07	1	1	72
20	43	57.83	50.42	12	18.7	96	79.08	87.98	9	0.64	1	0	67.8
21	82	88.17	85.09	12	23.8	90	94.36	73.6	0	0.09	0	1	89.2
22	62.67	57	59.84	12	16.8	100	87	73.42	9	0.59	0	0	66
23	81.33	80.67	81	12	8.4	97	91	93.84	0	0.56	0	0	66
24	81.67	83.83	82.75	12	17.4	97	92.75	70.1	0	0.85	1	1	74.58
25	87.17	89.5	88.34	12	18.3	98.6	96.5	72.68	0	0.98	1	1	74.58
26	63.67	81.2	72.44	12	15.2	96	79	77.08	4.5	0.08	1	0	70.42
27	82.5	85.83	84.17	12	17.8	100	92.5	76.13	0	0.95	1	0	70.4
28	72.76	89.83	81.3	12	11.8	98	92	80.53	0	0.85	1	1	68.33
29	75	86.83	80.92	12	12.6	95	87	94.04	0	0.04	1	0	69
30	33.33	57.83	45.58	12	5	97	92	70.26	9	0.14	1	1	62
31	28.67	38.4	33.54	12	4.9	93	87	72.59	12	0.98	1	0	61.2
32	77.83	84.83	81.33	12	15.8	97	90	95.15	0	0.38	1	1	76
33	69.83	76	72.92	12	13.9	98.4	91.16	83.08	0	0.78	1	1	70.5
34	46	54.67	50.34	12	14.6	94	83	77.43	12	0.68	1	0	63
35	73	73.33	73.17	12	22.8	93.2	87.5	74.34	3	0.36	1	1	75
36	49.67	69.5	59.59	12	19.9	95.4	93.83	89.44	7.5	0.04	1	1	77.6
37	86.33	92.67	89.5	12	24.8	98	91.25	89.27	0	0.14	0	1	82.5
38	77.33	83.67	80.5	12	21.9	74	80	89.48	1.5	0.95	1	1	75.9
39	62.5	69.83	66.17	12	16.4	86.66	83.8	94.39	4.5	0.17	0	0	66.7
40	54	57.17	55.59	12	5.2	93	82.5	78.86	9	0.37	0	0	60.9
41	75	78.5	76.75	12	9	96	91	75.8	0	0.42	0	1	73.3
42	83.67	91	87.34	12	18.5	95	90	71.87	0	0.43	1	1	84.4
43	80.83	83.83	82.33	12	17	94	89.25	91.22	0	0.95	1	1	75.2
44	90.5	93.67	92.09	12	22.8	98.6	94.25	89.69	0	0.14	1	1	80.8
45	86.67	89.33	88	12	15	96	88	84.87	0	0.8	1	1	72.08
46	87.17	88.67	87.92	12	25.6	90	93	73.36	0	0.56	1	1	86.8
47	42	54.33	48.17	12	12.9	82	85	69.02	12	0.87	1	1	74.5
48	75.33	79.83	77.58	12	14	99	96.6	88	0	0.16	1	1	80
49	77.67	78.83	78.25	12	17.4	96	92	82.74	1.5	0.68	1	1	79.17
50	69.67	73.5	71.59	12	15.2	93.8	93.21	91.43	1.5	0.85	1	0	72.5
51	74.67	70.67	72.67	12	13.9	95.8	92.8	84.06	1.5	0.68	1	0	65.8
52	77.33	84.33	80.83	12	18.4	88	81.8	91.81	0	0.32	1	1	73.3
53	74.67	73.67	74.17	12	10.7	98.8	94.5	90	1.5	0.02	1	1	71.67
54	22.33	32.4	27.37	12	14.4	96.8	94.25	91	15	0.69	0	1	76.67
55	80.67	85.33	83	12	9.6	75	60	93	0	0.28	1	1	68
56	0	20	10	12	16.4	90	90.25	81	15	0.98	0	1	78.4
57	85.17	88.17	86.67	12	21.8	94	92	94	0	0.93	1	1	86
58	60.67	66.17	63.42	12	17.4	94	78	97	6	0.33	1	1	72.8
59	64.67	82.5	73.59	12	11.5	80	75	76	1.5	0.42	1	1	74.9
60	77.33	79.33	78.33	12	20.7	89	79	76	0	0.71	1	1	76.7
61	61.33	55.2	58.27	12	4.9	95	92.75	94	7.5	0.22	1	1	79.2
62	82.5	85.8	84.15	12	11.2	92	72	85	1.5	0.7	1	0	70
63	34	48.6	41.3	12	9.2	96	93.2	70	12	0.66	1	1	85.9
64	90.17	92.5	91.34	12	6.8	96	96	71	0	0.87	0	1	70.4
65	74.83	81.83	78.33	12	3	96.4	94.6	75	0	0.97	0	0	65
66	47.67	38.33	43	12	4.8	99	93	75	13.5	0.41	1	1	72.2
67	67	67.83	67.42	12	9.5	95	89	78	4.5	0.26	1	1	69.8
68	63.33	62.33	62.83	12	15.4	84	92.08	79	4.5	0.78	1	1	69
69	72.83	83.83	78.33	12	5.4	92	86	87	0	0.5	1	1	64
70	43.17	40.17	41.67	12	16.4	96	92	92	12	0.41	0	1	73.3
71	74.5	72.17	73.34	12	14.7	95	89.5	89	0	0.12	1	1	80.42

Fig.2. Executed Input Variable

(D) Output Variable

The output variable which is the dependent variable is the semester CGPA which is converted to percentage.

OUTPUT VARIABLE		
S.NO	**DESCRIPTION**	**POSSIBLE VALUES**
1	Semester CGPA	Converted to percentage (0-100)

Fig.3. Output Variable

(a) Topology

After the data is chosen and preprocessed it is important to choose the topology for the neural network. The network topology will describe the arrangement of the neural networks. The network topology that is available for us is numerous; each with its unique advantages and disadvantages. For instance, some will compensate speed for accuracy, while others can handle static variables, not continuous ones. Considering the nature of our data which is static and not so large multilayer perceptron was selected.

(b) Multilayer Perceptron

A feed-forward artificial neural network is a multilayer perceptron. A simple MLP has at least three layers, which are the input layer, the hidden layer, and the output layer. Back propagation is a supervised learning approach used by MLP to train the dataset. Because the MLP is fully connected, each node in one layer is connected to every other node in the next layer with a specific weight. The rectifier linear unit (ReLU) is utilised in this model to process the output straight from the positive input as well as to solve numerical issues with sigmoids. They have the disadvantage of training slowly and requiring a large amount of data for the training set.

(c) The Data Set Grouping

The data is divided into two categories the training set and the test set. The training set is used to train the model to observe relationships between the input data and expected output data. The data in the training set should be 80%of the entire data collected from the students. Test set is the data in which the trained model uses the above observed relationship and predicts the output. The test set should be 20% of the entire data collected from the students.

5. Constructing the Network Layer and Processing Element

In the process of building the neural network, the most significant step is to choose the number of the hidden layer. The selection of the processing unit and a hidden layer is a wispy one because having a small number of hidden layers will lower the processing of the network or sometimes it affects the accuracy. This process can be done using two ways; one can begin with a small layer and increase the size further according to the output this is called Growing Method; while the other is having a complex network and reducing its size by removing not so important component this process is called Pruning Method. In this model, Growth Method was used to build the neural networks. Trade-off should be made in determining PE (processing elements). This is because a large number of PE can lead to weights which can lead to poor generalization. Analytically, it is not possible to set the number of Processing Element so it is also varied in the model from 1-5 nodes to arrive at the best performance of the neural network model. The optimizer and loss function must be specified in the final step. By adjusting the features of the neural networks, such as weights and learning rate, an optimizer is used to shape and mould the model into its most correct form. They join the loss function and the model parameters in response to the loss function's output. The optimizer utilised in this model is adam, and the loss function is mean squared error.

(A) Mean Squared Error

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad (1)$$

Where, MSE = mean squared error, n =number of data points, Y_i = observed values and $\hat{Y}_i$ =predicated values.

S.NO	Test data	Predicted data
1	36.97	37.39
2	66.20	64.42
3	83.75	82.80
4	89.25	89.92
5	71.32	71.23
6	75.00	75.00
7	73.30	73.29
8	68.00	71.41

Fig.4. Obtained Output

6. Conclusion and Recommendation

This study has shown the budding artificial neural networks for improving the effectiveness of the education system and student's mindset towards the examination. The model was developed from the input variables of past academic performance and attained an accuracy of 96% which shows the potential of neural networks as a prediction tools for the candidates seeking any type of examination.

The model can be future developed by giving the respective students personalized learning methodology based on the predictions and personality interview questions in the form of a Questionnaire. One disadvantage of the model stem is that not all performance influencing factor can be obtained from the students.

References

[1] Osmanbegovic, M. Sulji, Data miningapproach for predicting student performance, Economic Rev iew 10.

[2] S. T. Jishan, R. I. Rashu, N. Haque, R. M. Rahman, Improving accuracy of student's final grade pr ediction model using optimal equal width binning and synthetic minority over sampling technique, DecisionAnalytics 2 (1) (2015).

[3] D. M. S.Anupama Kumar, Appraising the significance of self regulated learning in higher educatio n using neural networks, International J. of Engineering Research and Development Volume 1 (Issue 1) (2012).

[4] V. Oladokun, A. Adebanjo, 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 9 (1) (2008).

[5] V. Ramesh, P. Parkavi, K. Ramar, Predicting student performance: a statistical and data mining approach, International Journal of Computer Applications 63 (8) (2013).

[6] P. M. Arsad, N. Buniyamin, J.-l. A. Manan, A neural network students' performance prediction model (NNSPPM), in: Smart Instrumentation, Measurement and Applications (ICSIMA), 2013 IEEE International Conference on, IEEE, 2013.

[7] G. Gray, C. McGuinness, P. Owende, An application of classification models to predict learner progression in tertiary education, in: Advance Computing Conference (IACC), 2014 IEEE International, IEEE, 2014.

[8] T. Wang, A. Mitrovic, Using neural networks to predict student's performance, in: Computers in Education, 2002. Proceedings of International Conference on, IEEE, 2002.