Performance Prediction for Post Graduate Students using Artificial Neural Network

Vinod Kumar Pal, Vimal Kamlesh Kumar Bhatt

Abstract: Understanding university students' performance and identifying factors that affect university students' performance are very important issues for educational institutions, educators, and students. To extract some meaningful information from these large volumes of data, academic organizations have to mine the data. The current system of evaluating student performance is not feasible and it has been observed that it often leads to dissatisfaction among the students, as in the absence of correct predictors of success in educational institutes, students and institutions put emphasis on incorrect predictors and invest time and resources in. This paper presents a comprehensive study on predicting student performance in R Programming for postgraduate students using deep learning (which is a small part of the artificial neural network). The study's objective is 1) to study the prediction accuracy rate using R programming 2) to analyze the factors that affect academic achievement that contribute to predicting academic performance of students. The researchers evaluated the proposed methods on a dataset consisting of 395 student's records with 30 attributes of students has been gathered from UCI Repository after the collected data has been preprocessed, cleaned, and filtered using R Programming.

Keywords: Artificial Neural Network, Educational Data Mining, Prediction of Students' Performance, University Education

I. INTRODUCTION

The Higher Education Institutions should be motivated by providing the students with quality education because education is a very important issue for a country's development. To improve the quality and competitiveness of students, institutions or universities must have some strategy to achieve their objectives, and for the implementation of these strategies, institutions would require drawing their attention towards finding solutions to the existing problems. For that matter, the factors affecting student success need to be identified early. The success of a particular student can be seen with a final grade in the particular subject.

It has been observed that the application of using the artificial neural network in the education field has grown very traditionally.

The artificial neural network enables us to discover new and useful student knowledge. Institutes have a wide scope to use BI application because there are a wide range of data sources, i.e. traditional databases, web pages, etc., and groups with varied interests such as students, teachers, administrators, alumni, etc. (Ma, Liu, Wong, Yu, & Lee, 2000).

There are many noteworthy things that can be recognized in this domain that can be answered using BI techniques (Luan, 2002), for example, what are the factors that affect students ' performance? Can students ' performance be predicted well in advance? This paper addresses the latter question. Student performance modeling is an important tool not only for students, but for educators as well. Since the tool improvises the better understanding of students, professionals can execute corrective measures for weak students, well in advance. Though there is high amount of data accumulation in education institutes, it is not utilized properly and the hidden information within is not discovered properly. The hidden information within the databases offers great opportunity to explore potential knowledge about students (Altaher & Barukab, 2017). Especially analytics and data mining techniques permits institutions access the large data sets available to identify the hidden patters, which reflects students' behavior and learning (Zacharis, 2015). The hidden knowledge present in the large data sets can benefit institutions not only in understanding students in a better way but can also help in improving teaching (Romero & Ventura, 2010).

Predicting the performance of the student in the field of educational data mining is an important and hot topic, which is very useful for students and educational institutions. (Oladokun, Adebanjo, & Charles-Owaba, 2008) proposed an artificial neural network model for university engineering students to predict performance. Hamsa, Indiradevi and Kizhakkethottam (2016) have proposed a student performance prediction model for each topic that will help the lecturer take more care of those students at risk in the results. The present paper is planned to explore the various factors affecting the success of the performance prediction of the student and student based on the final grade of the exam. Deep learning method (which is an essential part of the artificial neural network) would be used to predict the grade of students through R Programming.

Revised Manuscript Received on May 28, 2019.

Vinod Kumar Pal, Balaji Institute of International Business, Pune (India)

Vimal Kamleshkumar Bhatt, Balaji Institute of Modern Management, Pune (India)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

II. RELATED WORK

A number of researchers have been attempting to study data from students over the past decade to predict student performance (Karamouzis & Vrettos, 2008). Various methodologies were used to including regression analysis, classification, Data mining approaches and correlation analysis. Each of these approaches generates different levels of success but are not as effective as the potential for deep learning in terms of discovering existing and undetected patterns.

It was suggested that using the Generalized Regression Neural Network (GRNN) to predict student's Academic performance, the results showed that secondary school performance is the single most important variable associated with university graduation overall performance (El-Refae & Al-Shayea, 2010). (Sonja & Milija, (2014) developed a multilayer neural network model, has the ability to predict student success, and it would be helpful if students were to work harder and improve their knowledge in specific branches. (Sebastian, 2015) presented a multilayer perceptron neural network used to implement prediction strategy and discovered attributes that affect student performance as a unit test, assignment mark and class attendance.

Ruby & David, (2017) proposed to use various classification algorithms such as ID3, MLP, and a hybrid model to predict accuracy of students ' academic performance. The results help the institutions know the students ' academic status in advance and can focus on weak students to improve their academic outcomes. (Altaher & Barukab, 2017) proposed for student academic performance prediction an adaptive neuro-fuzzy inference system. In a literature review of data mining techniques used to predict student performance (Shahiri, Husain, & Rashid, 2015), it was found that cumulative grade point average (CGPA) is the most influential attribute as it determines future mobility in education and career. The neural network has the highest prediction accuracy, according to their findings, by (98 percent) followed by decision tree by (91 percent). Support vector machine and neighbor k-nearest had the same accuracy (83 %), while naive Bayes gave less accuracy of prediction (76 %). (Ahamed, Mahmood, & Rahman, 2017) used the K-Nearest neighbor and supported the student performance prediction vector machine. The result shows 86 percent accuracy. (Yadav, Bharadwaj, & Pal, 2012) used student attendance, seminar, class test and assignment marks to predict the final grade of a semester for students. (Kouser, Joann, & Suganya, 2016) proposed a model for predicting student performance through artificial neural networks in online mixed learning courses. The proposed model was trained by learning from the decision tree to predict the ability of the student to successfully pass the course. The proposed model's accuracy rate was very high at 98.5 percent. (Livieris, Drakopoulou, & Pintelas, 2012) developed a neural network classification tool for the' Mathematics' course. MSP–Trained neural networks show more consistent behavior and better classifiers. (Ibrahim & Rusli, 2007) compared three predictive models of artificial neural network, tree decision & linear regression to predict academic performance of students. The first-semester demographic profile and CGPA used as the predictor variable. The results showed that all three predictive models achieved a precision rate of over 80 %. (Jishan, Rashu, Haque, & Rahman, 2015) compared the naive classification of Bayes with the neural network models. Any prediction system's accuracy improves with the use of SMOTE oversampling and optimum equal binning width to preprocess dataset that is small in size and contains continuous attributes. (Ramesh, Parkavi, & Ramar, 2013) proposed a student performance prediction model, suggesting that the multilayer perceptron algorithm is best suited to predict grades. The proposed model achieved a 72 percent accuracy, which demonstrates the high potential efficiency. (Naser, Zaqout, Ghosh, Atallah, & Alajrami, 2015) proposed a model for predicting IT & Engineering students ' performance using feedforward Backpropagation algorithms. The proposed model achieved 84.6 percent accuracy, which shows the comparison of high impact compared to other algorithms. The two powerful meta-heuristic algorithms CS and COA (J. F. Chen, Hsieh, & Do, 2014) were used to predict student academic performance. The researchers compared the ANN-COA to ANN-CS, the ANN-COA found somewhat better results to predict the performance of the student. (Herawan, 2012) developed an artificial neural network (ANN) GPA prediction model. To predict the success of the student, the researchers used correlation analysis, main component analysis, and prediction modeling.

The researchers proposed that student's academic achievement only affected by the process of study in the university. Differences in demographics, social and economic background, as well as differences of national test scores and entrance exams as academic background, cannot be a predictor of the learning achievement of students at the university. (Ibrahim & Rusli, 2007) compared the three predictive models: artificial neural network (ANN), decision tree and linear regression to predict the academic performance of students. The researchers found that all three predictive models achieved 80 percent accuracy and the model of artificial neural network (ANN) was better than the other two predictive models. (O&P, 2017) proposed a framework for administering student academic performance prediction using the techniques of learning analysis. Researchers found that from the analysis students performing well in courses in mathematics had better chances of gaining excellence in other courses in computer science. Researchers have tried using different methodologies and compared Artificial Neural Network results.

Table. 1 Comparison of Various methods (Amrieh, Hamtini, & Aljarah, 2016)

Evaluation	Decision Tree	Naïve Byes	Artificial Neural Network
Measure	75.8	67.7	79.1
Accuracy	75.8	67.7	79.2
Recall	76.0	67.5	79.1
Precision	75.9	67.1	79.1

III. RESEARCH METHODOLOGY

The goals of this study are: 1) preparing a model to predict student performance (Final Grades), 2) comparing prediction accuracy with traditional prediction tools such as regression. For the purpose of the study, we have used the data available in UCI repository. The number of data cases is 395 with 30 variables, which is an integer in nature with no missing values. The brief description of the variables used in the study is as mentioned below.

Table. 2 Student Attributes and their Description

Sr. No.	Variable	Description of Variable with Property
1	school	Student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira)
2	sex	Student's sex (binary: 'F' - female or 'M' - male)
3	age	Student's age (numeric: from 15 to 22)
4	address	Student's home address type (binary: 'U' - urban or 'R' - rural)
5	famsize	Family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
6	Pstatus	Parent's cohabitation status (binary: 'T' - living together or 'A' - apart)
7	Medu	Mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
8	Fedu	Father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
9	Mjob	Mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
10	Fjob	Father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
11	reason	Reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
12	guardian	Student's guardian (nominal: 'mother', 'father' or 'other')
13	traveltime	Home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
14	studytime	Weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
15	failures	Number of past class failures (numeric: n if 1<=n<3, else 4)
16	schoolsup	Extra educational support (binary: yes or no)
17	famsup	Family educational support (binary: yes or no)
18	paid	Extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
19	activities	Extra-curricular activities (binary: yes or no)
20	nursery	Attended nursery school (binary: yes or no)
21	higher	Wants to take higher education (binary: yes or no)
22	internet	Internet access at home (binary: yes or no)
23	romantic	With a romantic relationship (binary: yes or no)

24	famrel	Quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
25	freetime	Free time after school (numeric: from 1 - very low to 5 - very high)
26	goout	Going out with friends (numeric: from 1 - very low to 5 - very high)
27	Dalc	Workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
28	Walc	Weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
29	health	Current health status (numeric: from 1 - very bad to 5 - very good)
30	absences	Number of school absences (numeric: from 0 to 93)

Sampling Universe

The students from any University, Institution or any Educational Institute.

Sampling Unit

The sampling unit shall be students from postgraduate courses.

Sampling Instrument

For the purpose of analysis, we have used the Dataset (Cortez & Silva, 2008) available in UCI repository. Data available in the UCI repository is to be first screened for any missing values and multivariate outliers as multivariate analysis are to be used.

Data Preprocessing

Before using the data for analysis, data must be checked and data preprocessing must be done. As the data that we are using have no missing data, no missing values needed to be replaced. We checked for the multivariate outliers with the help of Mahalanobis Distance (Mahalanobis, 1936), and there were only six cases having Mahalanobis Distance greater than 60 and the corresponding significant value less than 0.001, which was calculated using the Cumulative Distribution of Chi-Square distribution.

Data Analysis

The entire dataset was first preprocessed and then was used for the analysis. It is important to check whether there are any differences among gender in obtaining a final G3 grade. For finding out the gender wise differentiation, Chi-Square Test was used. The following Figure 1 represents the gender feature into 184 male and 205 female students.

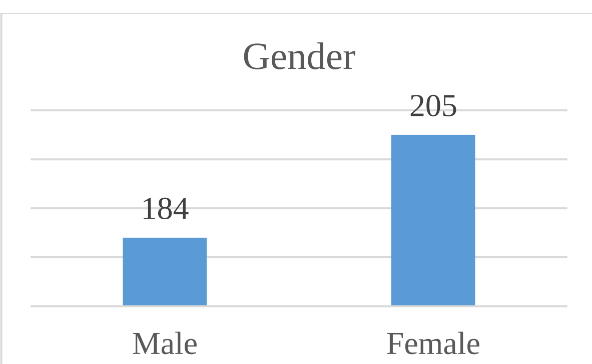

Fig. 1 Gender Ratio

The total marks obtained are plotted against gender in Figure 2, as shown below.

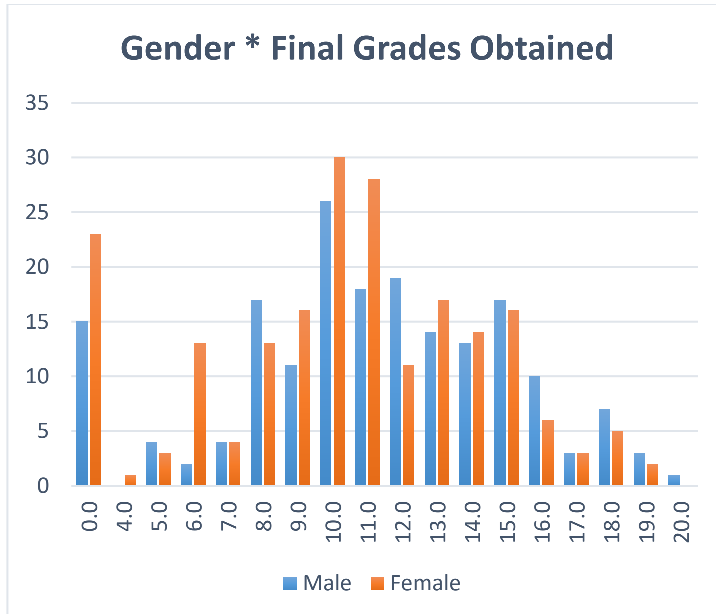

Fig. 2 Gender * Final Grades Obtained

The Chi-Square Test results are based on the data as shown in Table 4 and as the Asymptotic Sig. The value exceeds 0.05 (05 percent), which shows that there is no significant relationship between the final marks obtained and gender.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table. 3 Chi-Square Test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18.758a	17	.343
Likelihood Ratio	20.501	17	.249
Linear-by-Linear Association	4.296	1	.038
N of Valid Cases	389		
a. 12 cells (33.3%) have expected count less than 5. The minimum expected count is .47.			

IV. EXPERIMENT AND RESULTS

Environment

We run the experiments on the 4 GB RAM PC, with 1.90GHz of Intel i3 Processor. In evaluating the Artificial Neural Network (Deep Learning), we used R Programming. We split the data into two parts, train data set containing 70% of the data and test data set containing the remaining 30%.

Evaluation Measures

We used common classification quality assessment measures in our experiments: accuracy, accuracy, recall, and F-measure (T. Y. Chen, Kuo, & Merkel, 2006; POWERS, 2011).

Table. 4 Confusion Matrix

		Detected	
		Positive	Negative
Actual	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

A ratio of the total number of cases correctly predicted, i.e. True Positive and True Negative represent Accuracy in all cases, while Precision refers to the True Positive case ratio to the True Positive and False Positive sum. Recall refers to True Positive's ratio to True Positive and False Negative's sum (Fawcett, 2005). Furthermore, we used F-Measure to combine the recall and precision that is considered to be a good indicator of their relationship (T. Y. Chen et al., 2006). Out of the total data available, we have divided the data set into two parts, i.e. training containing approximately 80% of the data i.e. 311 cases (total is 389 after removing six multivariate outliers) and testing containing approximately 20% of the data i.e. 78 cases. First, on the training data, the model was trained and tested on the test data set.

Evaluation Results

We used linear regression to compare the baseline by making Final Grades (G3) as the dependent variable and rest all variables treated as independent variables. The R^2 Value came to be 0.261 and the model summary exhibited the significance of the model at 5% significance level. Based on the Coefficients we calculated the estimated Grades (Estimated G3) and found Root Mean Square Error (RMSE), representing an absolute measure of fit, which came to be 3.957. When coefficients obtained from the linear regression were used to predict the Final Grades (G3), the model could only predict 48 cases correctly and rest 341 cases were predicted incorrectly. The values of Accuracy, Precision, Recall and F-Measure are 12.339 percent, 21.918 percent, 22.018 percent, and 21.968 percent.

Random Forest, an ensemble learning classification method, works by building a multitude of decision trees at training time and producing output that is the class mode or mean prediction of the individual trees. (Ho, 1995). For R Programming (Liaw & Wiener, 2018), which is an extension of the Random Forest algorithm (Breiman, 2001), an algorithm was developed that was first used to develop the model and then predict the final grades (G3). For predicting the accuracy of prediction of Final Grades (G3) Random Forest Package of R was used and was compared with the Actual Grades (G3) and RMSE value came to be 1.766. The values of accuracy, precision, recall and F-measure were 28.101 percent, 43.874 percent, 43.874 percent, 43.874 percent and 43.874 percent.

We used the artificial neural network and deep learning to improve prediction accuracy and use Confusion Matrix to find out prediction accuracy. When tested, the model produced the predicted Final Grade value, which is a dependent variable, and the Confusion Matrix was prepared using both the Actual value of Final Grade G3 and the predicted value of Final Grade G3. The following table 5 represents the training dataset confusion matrix.

Table. 5 Confusion Matrix for Training Data Set

	Predicted Class																	Total Cases	Total Correct Cases	Percentage Correct	Error	Error Rate	
		0	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**					
Actual Class	**0**	25																	25	25	100.00%	0	0.00%
	4		1																1	1	100.00%	0	0.00%
	5			6															6	6	100.00%	0	0.00%
	6				13														13	13	100.00%	0	0.00%
	7					6									2				8	6	75.00%	2	25.00%
	8						22			1									23	22	95.65%	1	4.35%
	9							25											25	25	100.00%	0	0.00%
	10								42										42	42	100.00%	0	0.00%
	11									42							1		43	42	97.67%	1	2.33%
	12										25								25	25	100.00%	0	0.00%
	13											25							25	25	100.00%	0	0.00%
	14								2				20						22	20	90.91%	2	9.09%
	15				1									20					21	20	95.24%	1	4.76%
	16														12				12	12	100.00%	0	0.00%
	17															6			6	6	100.00%	0	0.00%
	18																9		9	9	100.00%	0	0.00%
	19																	5	5	5	100.00%	0	0.00%
Total	**25**	**1**	**6**	**14**	**6**	**22**	**25**	**44**	**43**	**25**	**25**	**20**	**20**	**14**	**6**	**10**	**5**	**311**	304	97.749%	**7**	**2.251%**	

The above table makes it very clear that out of 311 cases only seven cases could not be correctly predicted and rest 304 cases were correctly predicted, meaning the model trained on training data provided the accuracy of 97.749% and an error of 2.251%.

The model was then applied to the test data set with 78 cases and the resulting confusion matrix is shown in Table 6 below

Table. 6 Confusion Matrix for Testing Data Set

		Predicted Class																Total Cases	Total Correct Cases	Percentage Correct	Error	Error Rate	
		0	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**					
	0	2																	2	2	100.00%	0	0.00%
	4		3																3	3	100.00%	0	0.00%
	5			2															2	2	100.00%	0	0.00%
	6				2														2	2	100.00%	0	0.00%
	7					4													4	4	100.00%	0	0.00%
	8						6												6	6	100.00%	0	0.00%
	9							6											6	6	100.00%	0	0.00%
	10			1					9										10	9	90.00%	1	10.00%
Actual Class	**11**									9						2			11	9	81.82%	2	18.18%
	12										6								6	6	100.00%	0	0.00%
	13											6							6	6	100.00%	0	0.00%
	14												6						6	6	100.00%	0	0.00%
	15													5					5	5	100.00%	0	0.00%
	16								1						3				4	3	75.00%	1	25.00%
	17															2			2	2	100.00%	0	0.00%
	18																2		2	2	100.00%	0	0.00%
	19																	1	1	1	100.00%	0	0.00%
Total	**2**	**3**	**3**	**2**	**4**	**6**	**6**	**10**	**9**	**6**	**6**	**6**	**5**	**3**	**4**	**2**	**1**	**78**	**74**	**94.872%**	**4**	**2**	

As it is very much evident from the above table, the accuracy of the confusion matrix for testing data set is 97.749% and the corresponding error rate is 2.251%. The model was then applied to the entire data set to predict the final grades (G3), which was trained on the training data set.

We used accuracy, accuracy, accuracy, recall, and F-measure to compare the actual and predicted values of Training Data Set, Testing DataSet, and Total Data, as shown in Table 7 below. In addition, Random Forest's prediction accuracy is also displayed.

Table. 7 Confusion Matrix Results

	Deep Learning (Artificial Neural Network)			Random Forest	Linear Regression
	Training Data	Testing Data	Total Data		
Accuracy	97.749%	94.872%	97.429%	28.101%	12.339%
Precision	99.023%	97.368%	98.698%	43.874%	21.918%
Recall	98.701%	97.368%	98.698%	43.874%	22.018%
F-Measure	98.862%	97.368%	98.698%	43.874%	21.968%

The above table makes it very clear that the prediction accuracy is 97,749 percent for the training data set, 94,872 percent for the Testing data set and 97,429 percent for the same model for predicting the output of the whole data set. When using R program's Random Forest package, Accuracy came to 28.101 percent, after predicting the model-based Final Grades (G3).

V. CONCLUSIONS

Academic achievement of a student is of the highest priority for any institute or university across the globe. Using various methods to predict the performance of the student accurately would be highly required. Predicting the performance would also enable the institutions to focus more on students having more probability of performing lower in order to improve their performance. Earlier researchers studied common methods of ensembles such as Bagging, Boosting, and Random Forest (RF). In order to improve the performance of the model i.e. the accuracy of prediction, Artificial Neural Method was used. Comparing the accuracy of various methods like Linear Regression, Random Forest and Deep Leaning (ANN) we learn that Linear Regression could only produce an Accuracy of 12.339%. Random Forest produced a slightly higher accuracy of 28.101%, while on the other hand, Deep Leaning produced an Accuracy of 97.429% on Total Dataset. Different methods evaluate the performance of students ' predictive model, namely Artificial Neural Network, Random Forest, and Linear Regression. Based on the prediction accuracy, it can be stated in this paper that Artificial Neural Networks exhibit more consistent behavior and illustrate better classification outcomes than other traditional classifiers. Since it is very evident that Deep Learning can predict the final grades (G3) with an accuracy rate of 97.749 percent on the test dataset that was not used to prepare the model, another dataset with approximately the same accuracy can also be predicted. The technique of ANN modeling has many favorable features such as efficiency, generalization, and simplicity. These features make ANN an attractive choice for more accurate modeling of complex systems.

REFERENCES

1. Ahamed, A. T. M. S., Mahmood, N. T., & Rahman, R. M. (2017). An intelligent system to predict academic performance based on different factors during adolescence. *Journal of Information and Telecommunication*, *1*(2), 155–175. https://doi.org/10.1080/24751839.2017.1323488
2. Altaher, A., & Barukab, O. (2017). Prediction of Student's Academic Performance Based on Adaptive Neuro-Fuzzy Inference. *IJCSNS International Journal of Computer Science and Network Security*, *17*(1), 165–169. Retrieved from http://paper.ijcsns.org/07_book/201701/20170125.pdf
3. Amrieh, E. A., Hamtini, T., & Aljarah, I. (2016). Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. *International Journal of Database Theory and Application*, *9*(8), 119–136. https://doi.org/10.14257/ijdta.2016.9.8.13
4. Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
5. Chen, J. F., Hsieh, H. N., & Do, Q. H. (2014). Predicting student academic performance: A comparison of two meta-heuristic algorithms inspired by cuckoo birds for training neural networks. *Algorithms*, *7*(4), 538–553. https://doi.org/10.3390/a7040538
6. Chen, T. Y., Kuo, F. C., & Merkel, R. (2006). On the statistical properties of testing effectiveness measures. *Journal of Systems and Software*, *79*(5), 591–601. https://doi.org/10.1016/j.jss.2005.05.029
7. Cortez, P., & Silva, A. (2008). Using Data Mining To Predict Secondary School Student Performance. In *5th Annual Future Business Technology Conference* (pp. 5–12). https://doi.org/https://doi.org/10.13140/RG.2.1.1465.8328
8. El-Refae, G., & Al-Shayea, Q. (2010). Predicting Students' Academic Performance Using Artificial Neural Networks: A Case Study. *International Journal of Computer Science and Information Security*, *8*.
9. Fawcett, T. (2005). An introduction to ROC analysis. https://doi.org/10.1016/j.patrec.2005.10.010
10. Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm. *Procedia Technology*. https://doi.org/10.1016/j.protcy.2016.08.114
11. Herawan, A. R. (2012). *Predicting Students' Academic Success Using Artificial Neural Network*. Telkom University. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwj5rcv_wsbdAhWMTX0KHSdbApwQFjACegQICRAC

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

&url=http%3A%2F%2Fopenlibrary.telkomuniversity.ac.id%2Fpustaka%2Ffiles%2F96352%2Fabstract%2Fpredicting-students-academic-success-using-artific

12. Ho, T. K. (1995). Random Decision Forest. In Proceedings of the 3rd International Conference on Document Analysis and Recognition (pp. 278–282). Montreal, QC. Retrieved from http://ect.bell-labs.com/who/tkh/publications/papers/odt.pdf
13. Ibrahim, Z., & Rusli, D. (2007). Predicting Students' Academic Performance: Comparing Artificial Neural Network, Decision tree And Linear Regression. Proceedings of the 21st Annual SAS Malaysia Forum, (September), 1–6. Retrieved from https://www.researchgate.net/profile/Daliela_Rusli/publication/228894873_Predicting_Students'_Academic_Performance_Comparing_Artificial_Neural_Network_Decision_Tree_and_Linear_Regression/links/0deec51bb04e76ed93000000.pdf
14. Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. Decision Analytics, 2(1), 1. https://doi.org/10.1186/s40165-014-0010-2
15. Karamouzis, S., & Vrettos, A. (2008). An Artificial Neural Network for Predicting Student Graduation Outcomes. In World Congress on Engineering and Computer Science. San Francisco, USA. Retrieved from http://www.iaeng.org/publication/WCECS2008/WCECS2008_pp991-994.pdf
16. Kouser, R. R., Joann, J. D., & Suganya, K. (2016). FORECASTING STUDENT ACADEMIC PERFORMANCE BY DECISION TREE LEARNING USING ARTIFICIAL NEURAL NETWORKS, 955–961.
17. Liaw, A., & Wiener, M. (2018). Breiman and Cutler's Random Forests for Classification and Regression. https://doi.org/10.1023/A:1010933404324>
18. Livieris, I. E., Drakopoulou, K., & Pintelas, P. (2012). Predicting students' performance using artificial neural networks, (March 2017), 28–30.
19. Luan, J. (2002). Data Mining and Its Applications in Higher Education. New Directions for Institutional Research. https://doi.org/10.1002/ir.35
20. Ma, Y., Liu, B., Wong, C. K., Yu, P. S., & Lee, S. M. (2000). Targeting the right students using data mining. In Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '00 (pp. 457–464). New York, New York, USA: ACM Press. https://doi.org/10.1145/347090.347184
21. Mahalanobis, P. C. (1936). National Institute of Sciences of India. In Proceedings of the National Institute of Sciences of India (pp. 49–55).
22. Naser, S. A., Zaqout, I., Ghosh, M. A., Atallah, R., & Alajrami, E. (2015). Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology. International Journal of Hybrid Information Technology. https://doi.org/10.14257/ijhit.2015.8.2.20
23. O, O., & P, C. (2017). Predicting Students' Academic Performances – A Learning Analytics Approach using Multiple Linear Regression. International Journal of Computer Applications, 157(4), 37–44. https://doi.org/10.5120/ijca2017912671
24. Oladokun, V. O., Adebanjo, A. T., & Charles-Owaba, O. E. (2008). Predicting students' academic performance using artificial neural network: A case study of an engineering course. The Pacific Journal of Science and Technology, 9(1), 72–79.
25. POWERS, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure To Roc, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies, 2(1), 37–63. https://doi.org/10.1.1.214.9232
26. Ramesh, V., Parkavi, P., & Ramar, K. (2013). Predicting Student Performance : A Statistical and Data Mining Approach. International Journal of Computer Applications, 63(8), 35–39. https://doi.org/10.1504/IJTEL.2012.051816
27. Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews. https://doi.org/10.1109/TSMCC.2010.2053532
28. Ruby, J., & David, K. (2017). An Analysis on Academic Performance of Students using a Hybrid Model for Higher Education, 9(3), 2175–2182. https://doi.org/10.21817/ijet/2017/v9i3/1709030146
29. Sebastian, S. (2015). Evaluating Students Performance by Artificial Neural Network using WEKA. *International Journal of Computer Applications*, *119*(23), 975–8887.
30. Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, *72*, 414–422. https://doi.org/10.1016/j.procs.2015.12.157
31. Sonja, I., & Milija, S. (2014). PREDICTING STUDENTS' ACADEMIC PERFORMANCE USING ARTIFICIAL NEURAL NETWORK : A CASE STUDY FROM FACULTY OF ORGANIZATIONAL SCIENCES. In *International Conference on Education in Mathematics, Science & Technology*.
32. Yadav, S., Bharadwaj, B., & Pal, S. (2012). Data mining applications: A comparative study for predicting student's performance. *International Journal of Innovative Technology & Creative Engineering*, *1*(12), 13–19. Retrieved from http://arxiv.org/abs/1202.4815
33. Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*. https://doi.org/10.1016/j.iheduc.2015.05.002