International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 5, October 2022, pp. 5553~5561
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i5.pp5553-5561

Satisfaction prediction of online education in COVID-19 situation using data mining techniques: Bangladesh perspective

Lamisha Haque Poushy[1], Salauddin Ahmed Bhuiyan[1], Masuma Parvin[2], Refath Ara Hossain[1], Nazmun Nessa Moon[1], Jarin Nooder[1], Ashrarfi Mahbuba[1]
[1]Department of Computer Science and Engineering, Faculty of Science and Information Technology, Daffodil International University, Dhaka, Bangladesh
[2]Department of General Education Department, Faculty of Science and Information Technology, Daffodil international University, Dhaka, Bangladesh

Article Info

Article history:

Received Jun 5, 2022
Revised May 26, 2022
Accepted Jun 25, 2022

Keywords:

COVID-19
F1-score and accuracy
Machine learning
Online education
Prediction

ABSTRACT

This research focuses on the education-based online learning platform. Due to the coronavirus disease (COVID-19) epidemic, online education is gaining global popularity. It has shown how successful it is in investigating the quality of online education at the COVID-19 pandemic situation by 799 students from different academic institutions, schools, colleges, and universities. A Google web form has been utilized as the data gathering mechanism for this survey. This paper perused the prediction of online education through data mining and machine learning approaches in an online program. The data was collected through online questionnaires. To predict online education's satisfaction rate, four different types of classifiers are used e.g., logistic regression classifiers, k-nearest neighbors, support vector machine, naive Bayes classifiers. The key purpose of this research is to find out an answer to a question which is, "are the student's satisfied with starting the new online teaching system, or will it be an ambivalent effect for students in the future?".

Corresponding Author:

Nazmun Nessa Moon
Department of Computer Science and Engineering, Faculty of Science and Information Technology,
Daffodil International University
Dhaka 1207, Bangladesh
Email: moon@daffodilvarsity.edu.bd

1. INTRODUCTION

The year 2020 has been an unforgettable year in which coronavirus disease (COVID-19), the pestiferous ailment of COVID-19, is discovered by the city of Wuhan in the province of Hubei, China, for the first time on November 17, 2019 [1]. This epidemic has damaged the world's healthcare system and has an impact on the livelihood of all people. In December 2019, according to World Health Organization (WHO), in China, the pesticide distress COVID-19 was first detected [2]. On March 8, 2020, Bangladesh receives the first three cases of COVID-19 [3]. In order to obstruct the stretch of the disease, the government of Bangladesh decided that all of the communication shut off like museums, close to all educational institutions, restaurants, offices, markets, movie theatres, and maintain social distance. Shut down every country's border connection and travel. As a result of around more than 23.1 million students in the class were suspended in the COVID-19 situation. Many countries' education institutes started regular online teaching to students by Google Meet, Zoom, Ziteboard, Skype, Screencastify, and Facebook. In order to promote online education, keep running general education going on the university first started offering online education soon later started colleges and school institutions. Now google meet, Zoom and Facebook live

offer services deliver such as online teaching and classroom. However, we face some problem on online education platforms such as access to the internet limited, unavailability of electronic devices, lofty cost of the internet, speed of low internet, whether these virtual education platforms can meet the needs of students and teachers, whether network learning is capable of high-quality teaching and learning, online education can become an effective medium of special time education in Bangladesh and perspective of Bangladesh suggestion developed the network online education according to the research result. At present time, researchers in many countries are trying to figure out in their research how useful online learning methods are for students. According to a published research paper from Bangladesh Education Journal where they use 150 students' data (public and private university) [4]. The result of this study is that e-learning is satisfied urban and rural area students, measuring the satisfaction level of e-learning of public and private university students. Also, many colleges and universities provide various tutorials for better understanding. However, COVID-19 pandemic situation, online education is largely conducted by teachers in their own institutions. Subsequent studies discussed the satisfaction of online learning or education platforms and did not focus on the quality of interaction. This paper is evaluated based on all the past information of the data in online education platforms in Bangladesh students. Are students satisfied with online education platforms? We used a logistic regression algorithm, linear regression model and saw which one gave me a better accuracy output in this method.

2. LITERATURE REVIEW

Different researchers and scholars from different countries have conducted research on the success of online education and the development of their online education techniques. Many scholars and researchers in Bangladesh worked on online education systems. Following are some common studies that have been done on online education, Dutta and Smita [5] recently work on an online education system to discourse the impact of the tertiary education system through the students in Bangladesh. They collected 50 university student's data semi-structured interviews and used data analysis methods. They find some problems of tertiary level education in Bangladesh and provide some essential information or steps that should be taken into this COVID-19 situation then it will be possible to give good education in future. Sultana and Nasring [6] identified the correlative factors that more affect students for higher education in Bangladesh, on the basis that they collected 182 students' data from several public universities and private universities in Bangladesh. This study used binary logistic regression to predict the importance of the factor of student satisfaction. Predict facilities of bus service to find random sampling methods, urban and rural area student's satisfaction level, they find accuracy level those students who are undergraduate. Also showed both male and female both university student's satisfaction level. Uddin *et al.* [7] discussed only the one public university that's Dhaka university students in Bangladesh among 417 data collected and only analyzed 388 students. The study's main focus is the quality of online service and information impact of online teaching and impacts online teaching and how to improve to deliver better online education teaching and platform to students. Uddin and his team used SMART-PLS 3.0 software, structural equation modelling (SEM) method to find more precise estimated values and used Delone and Mclean information system success model (DMISM) model. Uddin and his team interpretation that student's satisfaction is 44%. For higher education students who are getting graduated, Uddin used multiple regression to predict the satisfaction and adept 9-factor model in this paper. Mahonta *et al.* [8] shows a model study has been done on 250 students at a degree college Dinajpur, Bangladesh. Mahonta and team used a study model of RATER or SERVQUAL. And used SPSS (v, 23) software to find the mean of the result. Mahonta and team highlighted the limitations of this study. Many students do not share their right to counsel. Lack of personal understanding mistake. And one of the biggest limitations is data samples. Abdelkader *et al.* [9] indicated the best approach as well as the ideal dimensionality of the feature subset. The current study's findings clearly corroborate the well-known link between a small number of characteristics and higher predictive accuracy. The utility of feature selection (FS) for high-accuracy student satisfaction level (SSL) prediction is amazing, as the relevant set of attributes can effectively aid in the development of constructive teaching initiatives. This research results in an 80 percent reduction in feature size and a 100 percent increase in classification accuracy. Baturay and Yukselturk [10] collected 189 surveys on Dropout Students about Online Education Program. They use k-nearest neighbor (k-NN), decision tree (DT), naive Bayes (NB), and neural network (NN) algorithm for prediction. They got 87%, 79.7%, 76.8%, and 73.9% for 3-NN, DT, NN, and NB. They got 63% accuracy but after preprocessing they got 83% accuracy. Driscoll *et al.* [11] studied about 20 undergraduate students about the course Web-based Multimedia Development. They use k-NN, DT, and NB, and they got 78% accuracy. Amerieh *et al.* [12] predict student's online academic curricular. They propose a new model for data mining techniques. They use NB, DT, and NN algorithm for prediction. They got 25.8% accuracy but after testing the newcomer student they got more than 80% accuracy. Chen *et al.* [13] collected 800 surveys

on user satisfaction with online education. They use back propagation (BP) neural network algorithm for prediction data and the prediction accuracy reached 77.5%. Also, Saifuzzaman *et al.* [14] mentioned in their research the current situation and recent case studies of COVID-19 overall Bangladesh where Rahman *et al.* [15] found out the impact of mental health in this situation, and Shetu *et al.* [16] proposed an effective e-learning framework, from where we intend to do our research and predict a handsome accuracy and furthermore, Shetu *et al.* [17] found a way to predict student's academic performance through data mining technique.

3. RAISED MODEL

All researchers get below ninety percent of accuracy but in our research, we get up to ninety percent score. In our research to classify the student's satisfaction rate, we have to use linear regression, k-NN, support vector machine (SVM), and naive Bayes classifiers. Here use sixteen variables to classify the satisfaction rate. But we have prioritized some variables to classify. In this study, we get 799 data from students and around 160 pieces of data used to check to predict our classification. This research use tow set of data one is the tanning data set and other is the testing data set. Pre-process the data and fit in algorithm and finally evaluate the performance table; Figure 1 represents the model of procedure in our work.

Figure 1. Procedure of satisfaction on online education using machine in COVID-19 situation

3.1. Model of machine learning

An algorithm machine learning is a development area of research which is the perturbation of how many to set up computer programs with the question that automatically enhance with experience. At present, successfully developed many machine learning applications. Databases may apprehend valuable implicit regularities in problems that can be discovered automatically by specifically machine learning. Analyzing outcomes of given databases. A set of categories a new case belongs to a classification problem consists of recognizes to which given for training used a historical data, whose class or category membership is known which contains an example. In this study, four supervised machine learning algorithms have been used to a dataset having information from students who are satisfied or not and model able to predict their disease outcome.

3.1.1. Logistic regression classifier

Logistic regression is a predictive analysis. Logistic regressions are used to describe data and explain the link between one binary variable dependent and one or more independent variables named, ordinal, interval, or ratio-level. Input value "x" is linearly compounded to predict output value "y," by weight or coefficient value. The main difference between the output value model and the linear regression is that it is a binary $\{0\ or\ 1\}$ value instead of a numeric number. For one observation, here $z(y = 1 = no) = \alpha$ or $z(y = 0 = yes) = 1 - \alpha$. In n independent observations the number of successes p follows the binomial distribution of β and α-parameters as illustrated in (1):

$$z\,(p,\alpha) = \binom{n}{p}\alpha^{p}(1-\alpha)^{p} \quad (1)$$

where p=1, 2, 3….n and distribution the binomial expected value is F(y). Binomial data are described by (2).

$$V\ (\alpha(x) = \gamma + \beta x \quad (2)$$

Consider the regression logistic function is used to determine frequencies or calculate data. Such as data generally parade a function distribution for the probability to calculate a number α of events in a given transient ever and anon, knowing the average number ω of events in that interval. The prediction function is shown in (3).

$$f\ (\alpha) = \frac{1}{1+\omega^{-(\alpha)}} \quad (3)$$

3.1.2. k-nearest neighbors (k-NN) classification

k-NN is a classification and non-perimeter approach in which unfamiliar figures are compared to those from the training set and taken into class in line with the training illustration. Hence, elements of an n-dimensional space can be classified into K sets. It is defined by the user in order to obtain a better classification also represents the number of neighbors where K is a number of parameters. k-NN classification is counted based on a vote of the K-neighbors closest to each of the points. The best parameter k value is 5 or sometimes its value is 7 but a low value for K such as K=1, K=2, or K=3 can be noisy and lead to the effects of outliers in the k-NN model. Both classification and regression predictive problems k-NN can be used. In the industrial workplace in classification problems, it is used more widely.

3.1.3. Naive Bayes (NB) classification

NB classification represented as vectors of feature values where finite sets are drawn from the class labels and it is a simple technique that assigns class labels to problem instances. Therefore, a naive Bayes classifier is automatically acquired by only inclining the model of the numerical parameters. To the end, leading to a counting time complexity that is linear with reverence to the amount of training illustration. Only information about their resembling values and the variables is needed to estimate probabilities. Space efficient known as NB algorithm where necessary only the information provided by two-dimensional tables where each entry corresponds to a probability estimate of a particular variable for a given value.

3.1.4. Support vector machine (SVM) classification

SVM is a supervised learning algorithm, which is one of the popular and used for classification as well as regression both solving problems. In SVM algorithm, find the value or a point in *n*-dimensional space (where *n* is the number of features given dataset) with the value of every feature being the value of a particular coordinate. To the end SVM also provides satisfactory accuracy level.

3.1.5. Selection

This research study uses our data set to classify the four classifications algorithm and see which algorithm is giving a better result and all the above algorithms discuss in this paper. All algorithms perform and give a good score where the accuracy level was minimum above 80%. To this end, SVM and logistic regression both are given the closest accuracy. In the result part, we will show the all-model result, tables, and discuss them.

3.2. Data pre-processing

Data pre-processing refers to the pre-phase of processing datasets. Generally, raw data sets are not able to perform according to the algorithm and generate expected outcomes. So according to our research data pre-processing is required. In this phase, we have collected 799 surveys. And to preprocess we sort our data in Google Spreadsheet. Then we converted each data to a numeric value then we organized each data to check the accurate prediction rate. Besides, we have preprocessed data on WEKA. First, we upload. CSV file on WEKA, then we convert it into *.arff* file which is denoted as attribute file format with the end goal to classify our accuracy of prediction as shows in Table 1. We have received the Table 1 data through various questionnaires mentioned in Table 2. We use basic few questions attempts and throw out the student via online which is clearly demonstrate by Table 2 survey questionaries.

3.3. Predict correlation coefficient

The relationship between the relative movements of two variables to measure known as the method is the correlation coefficient. The range is between 1.0 to -1.0 of the values. Calculated the number if there was an error in the correlation measurement that means the number greater than 1.0 or less than -1.0. A

perfect negative correlation shows -1.0, while 1.0 shows a perfect positive correlation. If a correlation of 0.0 shows no linear relationship between the movements of the two variables. Similarly, Figure 2 explains the heat-map of the high-frequency value during the analysis period. The darker shades of light blue in the heat-map indicate the satisfaction point that had high frequency in the online education platform. For Instance, age, platform, pcib, ocib, st indicating the how for this internally involve in the online learning. In this COVID-19 period online calling and video calling increasing day by day and its show in the platform table. And also increasing internet using or browsing which has been showing the heat map. Platform connected to age and institute to show the dark light blue color.

Table 1. Processing data by excel sheet

Gender	Age	Institute	Hour	IU	Cost	Net	platform	better	pcib	ocib	st	tc	syp	satisfied
1	22	3	6	0	700	3	3	1	0	0	2	0	0	1
0	23	3	3	0	700	1	3	1	0	0	2	3	2	1
0	24	3	3	0	500	3	3	1	0	1	2	4	0	1
1	22	3	6	0	600	5	3	1	0	0	7	1	0	1
0	24	3	3	0	500	3	3	1	1	1	3	3	0	1

Table 2. Inquisition of classification and subject of content based on survey questioners

Classification of inquisition	Subject of content
Which one is better?	Physical class, Online class
Which costs you more money for educational purposes?	Physical class, Online class
Why physical class is better?	Understood better, Good conversation, Concentrate better
Why online class is better?	Anytime attend class, the teacher can be contacted at any time, No transportation problem
Feedback of student understanding to teaching	Percentage of number
Teacher and student communication?	Percentage of number
Are you satisfied in online education	Yes, no

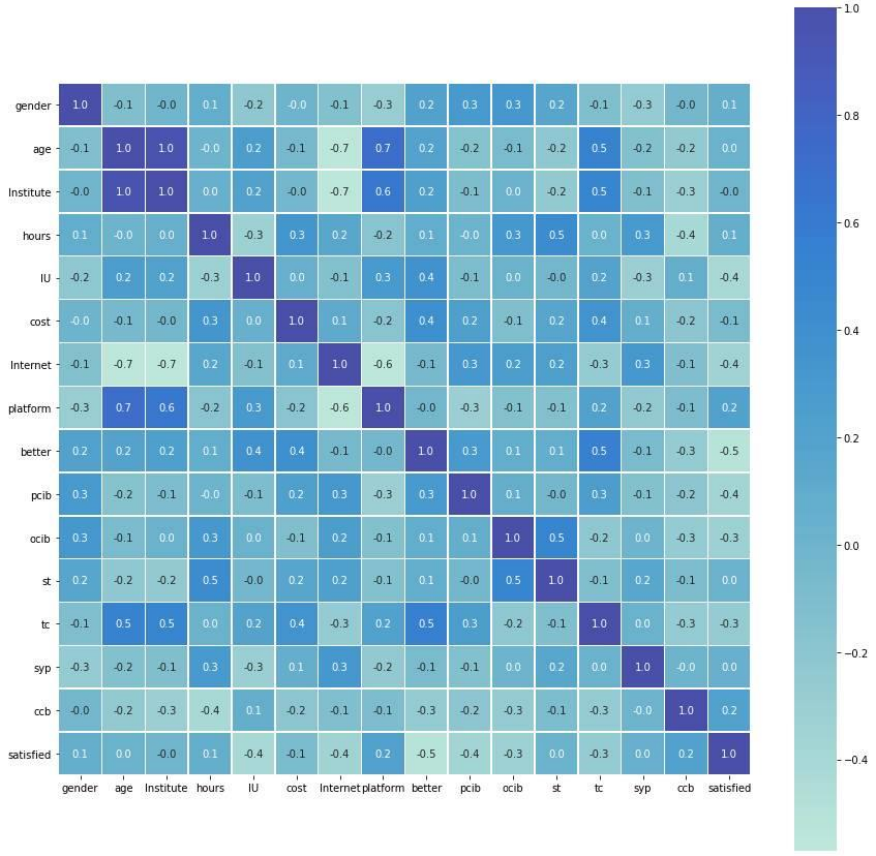

Figure 2. Show high-frequency key-word correlation coefficient representation the satisfaction value on this during online classes in COVID-19

4. DATA AND RESULT ANALYSIS

In this modern era technology e.g., internet of things (IoT) [18], data mining [19], [20], neural network [21] and machine learning [22]–[25] and so on plays a vital role where in our study we worked on data mining to predict our desire result to contribute the research area. In this part, we are analysis result data which check into cross-validation then convert to binary formation and the learning algorithms in terms of accuracy of evaluating the performance. We see recall, precision F1-score and accuracy table, and also receiver operating characteristic (ROC) curve. These all metrics are detailed below.

4.1. Performance evaluation metrics

A confusion matrix is a technique of a classification algorithm for abbreviating the performance. Accuracy of classification alone can be confusing if you have an unequal number of observations in each class or if you have more than two classes in a dataset [24] which a column and row for every class. The actual class score is the predicted class and the row in the column of each and every matrix shows the number of the test instance. Table 3 shows the confusing matrix which is we get from the value of the machine. Students are satisfied or not represent by {0, 1} here 0=Yes and 1=No.

We get 2*2 confusion matrix for {0, 1} classes to explore the result of our research. Table 4 represents the confusion binary matrix. The confusion matrix is shown in Table 4 of binary representation. In this study over 20%, which means 160 individuals' data are uses for testing and calculate the TN=true negative, FP=false positive FN=false negative, TP=true positive base on the test dataset and also find out the value of each of the particular model.

Table 3. Cross validation check based on confusion matrix

Model name	Actual score	Predicted score	
Logistic regression	0	0	1
		81	8
	1	2	69
Naive Bayes	0	73	16
	1	8	63
k-Nearest Neighbors	0	84	5
	1	8	63
Support vector machine	0	81	8
	1	2	69

Table 4. Representation of binary form of confusion metric

Model name	Actual class	TP	TN	FP	FN
Logistic regression	0	81	69	2	8
	1	69	81	8	2
Naive Bayes	0	73	63	8	16
	1	63	73	16	8
k-nearest neighbors	0	84	63	8	5
	1	63	84	5	8
Support vector machine	0	81	69	2	8
	1	69	81	8	2

4.2. Precision metric

Positive predictive value is also known as precision. Precision is a truly positive value and defined as the ratio of non-negative examples. A precise model predicted only the non-negative class, in cases precise model to be non-negative very likely. The precision matrix can be enumerated by the below formula. Based on this formula create a table and calculate the precision value is show Table 5.

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

4.3. Recall metric

Recall is also known as sensitivity. This metric is the result of recall measure how complete. An algorithm-wide breadth meaning that it has high recall takes a massive segment of the positives instance. The recall is enumerated as (5).

$$Recall = \frac{TP}{TP+FP} \tag{5}$$

4.4. F1-score and accuracy

F1 represents an interesting metric and seeks a balance between precision and recall when there is an uneven class distribution. Accuracy is the total number of samples to the ratio of the total number of correct value predictions. When it comes to a good pursuit, each class has an equal amount.

$$F1 = 2 * \frac{precision \times recall}{precision + recall} * 100\% \quad (6)$$

$$Accuracy = (TN + TP)/(TP + FP + TN + FN) \quad (7)$$

In this scenario of the part enumerate the F1 score by the following the method and find out the metric value for 0 and 1 both of each category. Already we have figure out the value of precision Table 5 and recall Table 6. Then enumerate the accuracy of the dataset which is means that how much accurate our data. Show in Table 7 logistic regression classification and SVM give 93.75% accuracy.

Table 5. Result of precision value for each class 0 and 1

Model name	Actual class	Precision
Logistic regression	0	0.975904
	1	0.896104
Naive Bayes	0	0.901235
	1	0.7975
k-nearest neighbors	0	0.9130435
	1	0.9265
Support vector machine	0	0.975904
	1	0.971831

Table 6. Result of recall value for each class 0 and 1

Model name	Actual Class	Recall
Logistic regression	0	0.9101124
	1	0.972
Naive Bayes	0	0.820225
	1	0.887324
k-Nearest Neighbors	0	0.944
	1	0.887324
Support vector machine	0	0.9101124
	1	0.971831

According to Table 7 the same value of logistic regression and SVM but if we are beholden to the F1 score actual class (0, 1) then see the difference of the percent's value, for the actual class value of logistic regression 0 is 94.1860 which is approximately 94.19 and 1 is 93.2479 which is approximately 93.35 and the value of SVM 0 is 94.1804 which is approximately 94.18 and 1 is 93.2374 which is approximately 93.24. To the end, find the logistic model give out the highest score, and classified the dataset.

Table 7. Analysis the result of performance evaluation metrics and compare with F1 score by 0 and 1

Classifier name		Logistic regression	Naive Bayes	K-nearest neighbors	Support vector machine
Accuracy		93.75%	85.0%	91.88%	93.75%
F1 score	0	94.19%	85.88%	92.83%	94.18%
	1	93.25%	84.002%	90.65%	93.24%

5. CONCLUSION

In this study, we collected student own experience data and a survey on online education platforms in Bangladesh during the COVID-19 pandemic situation. Thorough review and analysis of online student data, we have concluded that zoom and google meet provide high-quality service and some colleges take live classes to students through Facebook. Students pace some problems such as the inability to submit the education time, fall behind, and a video delay on the class time. We found a scientific ecological model index with elements that affect satisfaction and a measure of satisfaction by examining the questionnaires from it is based on the personal satisfaction of the students that they have realized through the online platform. In our examining result we get 93.75% accuracy in our dataset which predicted the majority of the student they are not satisfied in this running online classes and internet speed.

REFERENCES

[1] P. Romagnani *et al.*, "The COVID-19 infection: lessons from the Italian experience," *Journal of Public Health Policy*, vol. 41, no. 3, pp. 238–244, Sep. 2020, doi: 10.1057/s41271-020-00229-y.

[2] Y. Krishnamoorthy, R. Nagarajan, G. K. Saya, and V. Menon, "Prevalence of psychological morbidities among general population, healthcare workers and COVID-19 patients amidst the COVID-19 pandemic: a systematic review and meta-analysis," *Psychiatry Research*, vol. 293, Nov. 2020, doi: 10.1016/j.psychres.2020.113382.

[3] S. Huq and R. K. Biswas, "COVID-19 in Bangladesh: data deficiency to delayed decision," *Journal of Global Health*, vol. 10, no. 1, Apr. 2020, doi: 10.7189/jogh.10.010342.

[4] T. Sultana and R. H. Khan, "Investigating university students' satisfaction on online class: Bangladesh perspective," *Bangladesh*

Education Journal, vol. 18, no. 2, 2019
[5] S. Dutta and M. K. Smita, "The impact of COVID-19 pandemic on tertiary education in Bangladesh: students' perspectives," *Open Journal of Social Sciences*, vol. 08, no. 9, pp. 53–68, 2020, doi: 10.4236/jss.2020.89004.
[6] F. Sultana and S. Nasrinq, "Students' satisfaction on academic services in higher education: public and private Universities of Bangladesh," *Journal of Economics, Management and Trade*, pp. 29–41, Mar. 2021, doi: 10.9734/jemt/2021/v27i130320.
[7] M. D. M. Uddin, A. Ghosh, and O. Isaac, "Impact of the system, information, and service quality of online learning on user satisfaction among public universities students in Bangladesh," *International Journal of Management and Human Science (IJMHS)*, vol. 3, no. 2, pp. 1–10, 2019.
[8] H. C. Mahonta, M. S. H. Sarker, M. M. Rahman, and S. K. Pramanik, "Service standardand students' satisfaction: a study on setabganj degree college, Bochaganj, Dinajpur, Bangladesh," *Asian Journal of Sustainable Business Research*, vol. 2, no. 1, pp. 29–39, 2020.
[9] H. E. Abdelkader, A. G. Gad, A. A. Abohany, and S. E. Sorour, "An efficient data mining technique for assessing satisfaction level with online learning for higher education students during the COVID-19," *IEEE Access*, vol. 10, pp. 6286–6303, 2022, doi: 10.1109/ACCESS.2022.3143035.
[10] M. H. Baturay and E. Yukselturk, "The role of online education preferences on student's achievement," *Turkish Online Journal of Distance Education*, vol. 16, no. 3, pp. 3–12, Jul. 2015, doi: 10.17718/tojde.47810.
[11] A. Driscoll, K. Jicha, A. N. Hunt, L. Tichavsky, and G. Thompson, "Can Online Courses Deliver In-class Results?: A Comparison of Student Performance and Satisfaction in an Online versus a Face-to-face Introductory Sociology Course," *Teaching Sociology*, vol. 40, no. 4, pp. 312–331, May 2012, doi: 10.1177/0092055X12446624.
[12] E. A. Amrieh, T. Hamtini, and I. Aljarah, "Mining educational data to predict student's academic performance using ensemble methods," *International Journal of Database Theory and Application*, vol. 9, no. 8, pp. 119–136, Aug. 2016, doi: 10.14257/ijdta.2016.9.8.13.
[13] T. Chen, L. Peng, X. Yin, J. Rong, J. Yang, and G. Cong, "Analysis of user satisfaction with online education platforms in China during the COVID-19 pandemic," *Healthcare*, vol. 8, no. 3, Jul. 2020, doi: 10.3390/healthcare8030200.
[14] M. Saifuzzaman, M. M. Rahman, S. F. Shetu, and N. N. Moon, "COVID-19 and Bangladesh: situation report, comparative analysis, and case study," *Current Research in Behavioral Sciences*, vol. 2, Nov. 2021, doi: 10.1016/j.crbeha.2021.100034.
[15] M. M. Rahman, M. Saifuzzaman, A. Ahmed, M. F. Mahin, and S. F. Shetu, "Impact of COVID-19 on mental health: A quantitative analysis of anxiety and depression based on regular life and internet use," *Current Research in Behavioral Sciences*, vol. 2, Nov. 2021, doi: 10.1016/j.crbeha.2021.100037.
[16] S. F. Shetu, M. M. Rahman, A. Ahmed, M. F. Mahin, M. A. U. Akib, and M. Saifuzzaman, "Impactful e-learning framework: a new hybrid form of education," *Current Research in Behavioral Sciences*, vol. 2, Nov. 2021, doi: 10.1016/j.crbeha.2021.100038.
[17] S. F. Shetu, M. Saifuzzaman, N. N. Moon, S. Sultana, and R. Yousuf, "Student's performance prediction using data mining technique depending on overall academic status and environmental attributes," in *Advances in Intelligent Systems and Computing*, vol. 1166, Springer Singapore, 2021, pp. 757–769., doi: 10.1007/978-981-15-5148-2_66.
[18] M. Saifuzzaman, S. F. Shetu, N. N. Moon, F. N. Nur, and M. H. Ali, "IoT based street lighting using dual axis solar tracker and effective traffic management system using deep learning: Bangladesh context," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Jul. 2020, pp. 1–5., doi: 10.1109/ICCCNT49239.2020.9225590.
[19] X. Chu, F. Cao, L. Jiao, J. Wang, and Y. Jiao, "Optimal allocation of higher education resources based on data mining and cloud computing," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–9, Mar. 2022, doi: 10.1155/2022/7067676.
[20] Z. Kanetaki, C. Stergiou, G. Bekas, C. Troussas, and C. Sgouropoulou, "Evaluating remote task assignment of an online engineering module through data mining in a virtual communication platform environment," *Electronics (Switzerland)*, vol. 11, no. 1, Jan. 2022, doi: 10.3390/electronics11010158.
[21] G. Wen, Q. Guan, X. Wu, and W. Luo, "Building a neural network model to analyze teachers' satisfaction with online teaching during the COVID-19 ravages," *Journal of Computer and Communications*, vol. 10, no. 01, pp. 91–114, 2022, doi: 10.4236/jcc.2022.101005.
[22] R. Yang, "Machine learning and deep learning for sentiment analysis over students' reviews: an overview study," *Preprints*, 2021, doi: 10.20944/preprints202102. 0108.v1.
[23] W. Villegas-Ch, M. Román-Cañizares, and X. Palacios-Pacheco, "Improvementof an online education model with the integration of machine learning and data analysis in an LMS," *Applied Sciences*, vol. 10, no. 15, Aug. 2020, doi: 10.3390/app10155371.
[24] Z. A. Shaikh *et al.*, "Blockchain hyperledger with non-linear machine learning: a novel and secure educational accreditation registration and distributed ledger preservation architecture," *Applied Sciences*, vol. 12, no. 5, Feb. 2022, doi: 10.3390/app12052534.
[25] B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student' performance prediction using machine learning techniques," *Education Sciences*, vol. 11, no. 9, Sep. 2021, doi: 10.3390/educsci11090552.

BIOGRAPHIES OF AUTHORS

Lamisha Haque Poushy studied Computer Science and Engineering from Daffodil International University. Her main interested research fields are machine learning and data mining. She can be contacted at email: lamisha15-8958@diu.edu.bd.

Salauddin Ahmed Bhuiyan studied Computer Science and Engineering from Daffodil International University. His main interested research fields are machine learning. He can be contacted at email: salauddin15-9130@diu.edu.bd.

Masuma Parvin is the Senior Lecturer in Mathematics, Department of General Educational Development, Daffodil International University. She completed my B.Sc. and M.S. in Mathematics from Jahangirnagar University. Her research interest goes on data modeling, mathematical modeling and population dynamics. She can be contacted at email: parvin.ged@daffodilvarsity.edu.bd.

Refath Ara Hossain is a lecturer of the Department of Computer Science and Engineering at Daffodil International University. Her interested research fields are data mining, digital image processing and machine learning. She can be contacted at email: refath.cse@diu.edu.bd.

Nazmun Nessa Moon is an assistant professor of the Department of Computer Science and Engineering at Daffodil International University. She received the B.Sc. degree in Computer Science and Engineering from Rajshahi University of Engineering and Technology and M.Sc. in Information and Communication Technology from Bangladesh University of Engineering and Technology (BUET). Her interested research fields are internet of things (IoT), digital image processing and machine learning. She can be contacted at email: moon@daffodilvarsity.edu.bd.

Jarin Nooder studied Computer Science and Engineering from Daffodil International University. Her main interested research fields are image processing, machine learning and data mining. She can be contacted at email: jarin15-9553@diu.edu.bd.

Ashrarfi Mahbuba studied Computer Science and Engineering from Daffodil International University. Her main interested research fields are machine learning and data mining. She can be contacted at email: ashrarfi15-9203@diu.edu.bd.