Bulletin of Electrical Engineering and Informatics
Vol. 11, No. 2, April 2022, pp. 926~935
ISSN: 2302-9285, DOI: 10.11591/eei.v11i2.3617

Prediction of internet user satisfaction levels in Bangladesh using data mining and analysis of influential factors

Md. Hasan Imam Bijoy, Sumiya Alam Akhi, Md. Ali Ashraf Nayeem, Md. Mahbubur Rahman, Md. Jueal Mia
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Article Info

Article history:

Received Oct 14, 2021
Revised Jan 11, 2022
Accepted Mar 12, 2022

Keywords:

Bangladesh
Data mining
Internet
Internet users
User satisfaction level

ABSTRACT

Today the world has already acknowledged as a global village by the inter-net which has technologically evolved into a significant performance instrument for individuals, businesses, and countries seeking to achieve betterment. This study is based on data mining techniques to predict the satisfaction level of internet users from the context of Bangladesh. After conducting a public survey with 18 questions, we were able to acquire 451 responses from participants. Data for user satisfaction was associated with end-user characteristics including certain getting high speed, internet packages, cable type of Wi-Fi connection with targeting various age groups and occupations. The research's most key conceptual breakthrough was the reliability of magnitude predictions of user satisfaction level based on their experience with internet use. The empirical findings indicate that people in Bangladesh have high expectations in existing internet technology, and they are very dissatisfied with their facilities of internet use and to measure satisfaction level related with monthly limit of the Wi-Fi packages and the elements affecting internet speed. Several classifier models were applied to our dataset and among them, Random Forest (RF) performance reaches the top position with 91.53% accuracy.

Corresponding Author:

Md. Hasan Imam Bijoy
Department of Computer Science Engineering, Daffodil International University
Dhanmondi, Dhaka-1207, Bangladesh
Email: hasan15-11743@diu.edu.bd

1. INTRODUCTION

The internet is a global computer network that gives a scope of data and correspondence administrations. It includes connected organizations that utilize characterized resemblance conversations. The internet has become an indispensable tool virtually for all Bangladeshis. Especially for achieving or advancing their careers and globalizing their ideas as well as creativity, the younger generation see the internet as a useful instrument. A completely different world that has accessed another wellspring of knowledge and has made numerous new facilities is only the internet.

The usage of the internet has increased dramatically over the last few decades, and it has become an important part of people's everyday lives, with several good implications [1]. But there is confusion if the internet service of Bangladesh satisfies its generations properly with the accessibility, prices, speed or even connectivity or not. That’s why we have wanted to measure the satisfaction of users by analysing some data regarding internet. Internet users are basically those who have been using the internet at least once in the last 90 days. Because of the digital revolution, the number of active users is growing day by day mainly in Bangladesh and by the end of January 2021, it had risen to 112.713 millions [2]. There are many sorts of operators in between them, such as mobile internet, internet service provider (ISP) and public switched telephone network (PSTN). The number of internet users escalated by 7.7 million between 2020 and 2021, bringing the country's internet penetration rate to 28.8% in January 2021 [3].

At the end of January 2021, Bangladesh had 171.854 million mobile phone subscribers. Among them, the major variety of sim cards are four and their subscribers, like 79.758 million for GrameenPhone Ltd. (GP), 51.122 million for Robi Axiata Limited (Robi), 35.555 million for Banglalink Digital Communications Limited and 5.419 for Teletalk Bangladesh Ltd [4]. About 1.01 crore broadband internet connections were available by the end of June 2021. Despite a considerable rise in internet users, in June 2021, the country's versatile web speed was positioned 135th out of 137 nations and was distinctly in front of Afghanistan and Venezuela as far as web speed [5].

We looked at a lot of publications in this section to determine if there was anything missing from prior studies. Then, Wang and Chen [6] basically introduced a process to measure the standard of the apparatus, the user satisfaction and how they are benefited. They had followed an approach of the measurement scale of web services. Their study was only for the benefit of the 3.5 G network which was not so well established by that time in Taiwan. They conducted a web-based survey of 426 questionnaires. There are many regression analyses for quality measurement, customer satisfaction and benefits and the results were good enough. For determining quality satisfaction from users, Yen [7] used an attribute-based model for internet self-service technology in his paper. With some internet-based purchasing experience, a survey was introduced and gathered 459 datasets. The effects of ISST attributes on user satisfaction in various technologies displayed a good fit to their six-factor model. Each scale's alpha coefficients ranged from 0.76 to 0.82. Chase *et al.* [8] presented a four-dimensional generic model of information quality expanding past information quality-based works. There are six constructs for the proposed model based on their categories. They had collected a total of 10,329 data from 16 companies related to internet services and the final 8,761 data was selected for further analysis. The general model showed reasonable bounds and these findings point to a good model fit, showing that the proposed model well describes the connections between latent components. Bruce [9] delivered in his study how satisfied users are in information searching on the internet from Australians' perspectives. An interview sample had been used as the data for this study where they were able to know if any internet-based course they completed or not. There is a reduction of the p-value for different data analyses and their ideas can be processed for further research.

Isaac *et al*. [10] aimed in their study to expound the consequences of Yemeni Government employees as well as internet users' gratification [11], combined with the DeLone and McLean IS success model and task-technology fit. About 530 data was collected from a questionnaire survey of employees. Existing scales were used to measure the four components in the proposed model. The link between actual usage and performance effect is mediated by both user happiness and task-technology-fit (TTF) which is basically a proper fit towards the model. The paper could be stronger if all aspects of internet usages were analyzed rather than only the Yemen Employees. Davis and Hantula [12] analyzed the pleasure of the users and the lateness of downloading with the help of the internet. There was a competition held between 82 graduate and undergraduate volunteers to evaluate the internet speed and the information had been recorded through a simulated tool. By analyzing the records, they discerned that download delay, as well as academic sagacity, had been affected by the internet. Finally, the accuracy showed 0.92 for end-user satisfaction.

Sarawagi and Nagaralu [13] set an aim to estimate a discussion on the utility of offering data mining methods as internet services. If various portals assign different types to similar content, the user will have to choose between the forecasts of different portals. So, in this paper, they had just elaborated on the usefulness and limitations of data mining as a service. Bala *et al*. [14] introduced work on customer satisfaction on mobile networks from Bangladesh's perspectives. There were about 9000 students from which they had selected 400 samples for their study. After analyzing their data, they had concluded that the Teletalk operator gave customers more satisfaction than other mobile operators. Because of internet communication and the ease of retrieving information via the internet allow for greater development of critical thinking and problem solving, foster independence and autonomy, and allow for greater interaction, the high speed of internet technology in education should increase student learning process and retention [15], [16].

In our paper, we try to find out the satisfaction of internet users who have used one of them between mobile internet or broadband internet connection. We have tested our dataset with a total of nine data mining classifiers and identified a few execution assessment metrics and ran a result correlation to find the best classifier in the functioning scenario. Based on the evaluation of the obtained results, it is claimed that the Random Forest classifier gets the best results when compared to estimates.

2. RESEARCH METHOD

This section is part of our step-by-step working process. We all know that data collection, data preprocessing, model implementation, and results analysis are all key aspects of every machine learning project. As a consequence, three parts: data description, classifier description, and model procedure are all described in subsection.

2.1. Data description and analysis

Every research study is dependent on a dataset, and an ideal dataset aid in the study's success. A public survey was used to gather data for this investigation. The public survey we performed consisted of 18 questions, and we were able to obtain 451 responses from anonymous participants. We tried to obtain data from internet users by asking 18 questions that covered all aspects of internet usage. All of the questions, i.e. 18 variables, were used to implement the model. There are 17 independent variables in our dataset, however, only one variable is used as a dependent variable. Table 1, contains all variables, descriptions, variable kinds, and potential values. We separated our dataset into two halves, with 80% of the data being used for model training and the remaining 20% of data utilized for testing.

Table 1. Description of attributes and their possible values

Variable	Description	Variable type	Possible values
GT	Gender type	Independent	Male (0), Female (1)
UT	User type	Independent	Mobile data (0), Wi-Fi (1)
UAG	Age group of user	Independent	10-20 (0), 20-30 (1), 30-40 (2), 40-50(3), 50-60(4)
UOT	Occupation type of user	Independent	Govt employee (0), Private employee (1), Student (2), Engineer (3), Doctor (4), Lawyer (5), Teacher (6), Business (7), Banker (8), Unemployment (9), Others (10)
UAL	Residential area of user	Independent	Village (0), Town (1)
UDL	Divisional location of user	Independent	Dhaka (0), Rajshahi (1), Chattagram (2), Sylhet (3), Rangpur (4), Khulna (5), Barishal (6), Mymensingh (7)
UDT	Device type of user	Independent	Mobile (1), Tab (2), Laptop (3), Computer (4), Mobile + Tab (5), Mobile + Laptop (6), Mobile + Computer (7), Tab + Laptop (8), Tab + Computer (9), Laptop + Computer (10), Mobile + Tab + Laptop (11), Mobile + Tab + Computer (12), Mobile + Laptop + Computer (13), Tab + Laptop + Computer (14), Mobile + Tab + Laptop + Computer (15)
UST	Sim type of user	Independent	Grameen Phone (0), Airtel (1), Banglalink (2), Robi (3), Teletalk (4)
USNT	Sim network type of user	Independent	2G (0), 3G (1), 4G (2)
UEDPPM	Expanses of data-package per month by user	Independent	up to 1024 MB (0), 1-3 GB (1), 3-5GB (2), 5-10GB (3), 10-15GB (4), above 15GB (6)
UEMPM	Expenses of money per month by user	Independent	upto100 BDT (0), 100-200 BDT (1), 200-500 BDT (2), 500-1000 BDT (3), 1000-2000 BDT (4), Above 2000 BDT (5)
UPU	Use of purpose by user	Independent	YouTube (0), WEB (1), Social-Media (2), Others (3)
MTSSM	Most time spend on social-media	Independent	Facebook (0), Instagram (1), TikTok (2), Twitter (3), Others (4)
MSIGU	Maximum speed of internet get by user	Independent	Up to 500 kbps (0), 500 kbps – 1Mbps (1), 1 Mbps - 3 Mbps (2), 3 Mbps - 5 Mbps (3), Above 5 Mbps (4)
BSGTU	Best speed get on time by user	Independent	Morning (0), Noon (1), Night (2)
CTWFU	Cable type of Wi-Fi user	Independent	CAT 5 Cable (1), CAT 6 Cable (1), CAT 7 Cable (2), Optical Fiber (3), Other (4)
IPLWUM	Internet package limit of Wi-Fi user	Independent	1 Mbps (0), 2 Mbps (1), 3-5 Mbps (2), 5-10 Mbps (3), 10-15 Mbps (4), Above 15 Mbps (5)
USL	Satisfaction level of user	Dependent	Very dissatisfied (0), Dissatisfied (1), Average (2), Partially satisfied (3), Very satisfied (4)

The correlation matrix is a statistical matrix that depicts the correlation coefficient for the dataset's variables. The matrix indicates the relationship between the independent variable and all possible pairs of values for dependent variables in the classification problem. This correlation matrix demonstrates the linear relationship between the variables in our acquired data. For our dataset, the correlation matrix is shown in Figure 1. We sought to show the correlation between seventeen independent variables (GT, UT, UAG, UOT, UAL, UDL, UDT, UST, USNT, UEDPPM, UEMPM, UPU, MTSSM, MSIGU, BSGTU, CTWFU, IPLWUM) and instead just one dependent variable (USL). According to the matrix's findings, internet package limit of Wi-Fi user (IPLWUM) has a significant positive relationship with user satisfaction level (USL), whereas best speed get on time by user (BSGTU) has a marginally negative correlation with USL. As a result, the monthly limit of the Wi-Fi packages and the elements affecting internet speed are the deciding factors in this case.

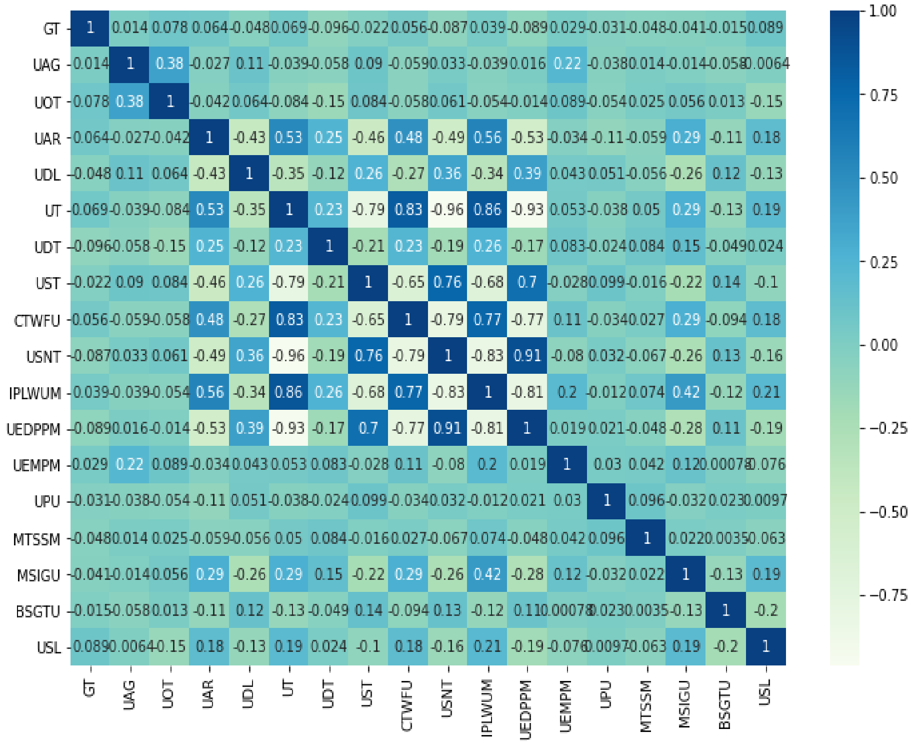

Figure 1. Correlation matrix of our working dataset

2.2. Classifier description

In machine learning, a classifier is a mechanism for predicting the target attribute based on feature data points. The dataset was examined to nine classifiers, with the relevant theory stated as:

KStar is an instance-based classifier that employs an entropy-based distance function, which sets it different from other instance-based classifiers. It is a variation of knearest neighbours (KNN), which is also known as a lazy learner. This classifier does not learn; instead, it memorizes the training data, performs some preprocessing, and then waits for the test tuple, which it detects and classifies based on its resemblance to the predefined training tuples. When training input, this type of classifier performs less and when a test tuple is classified, it works more [17]-[19].

A multilayer perceptron (MLP) is a neural network with an input, an output, and one or more hidden layers [20]. A single-layer perceptron can only learn linear functions; a multilayer perceptron, on the other hand, can learn nonlinear functions. MLP's learning technique is known as the backpropagation algorithm. The signal is received by the input layer, and a decision is predicted by the output layer depending on the input. To approximate continuous functions, the hidden layers perform as a computational engine. In MLPs, the previous layer's output is employ as the input to the next layer. MLPs [21] are feedforward networks with a forward pass in which signals flow from the input layer to the output layer via hidden layers, and a backward pass in which backpropagation is used to reduce error by tweaking model parameters (weights and biases) using stochastic gradient descent enhancement.

Instance-based k (IBk) is a lazy classifier category version of the k-nearest neighbours technique (KNN). Instead of building a model, the IBk method provides a forecast for a test case just-in-time. For each test instance, the IBk method uses a distance measure to choose k "near" examples from the training data, and then makes a prediction based on those selected instances. The IBk approach is a k-nearest neighbour classifier that has been demonstrated to perform well enough for activity categorization in terms of classification accuracy (>90%) [22].

The RandomCommittee [23] is a weka-meta classifier that includes building a number of Base classifiers with distinct random number seed values, and then computing the average of the predictions given by the various base classifiers to get the final classifier performance. If a batch prediction is being done, batchSize is the recommended number of instances to investigate. It is possible to supply more or fewer instances, however, this allows implementations to define a specific batch size.

Random Forest [24] is a supervised learning approach, which is a basic machine learning algorithm that, in the majority of cases, produces great results even without hyper-parameter tuning. It put together a "forest" out of a collection of decision trees trained by the "bagging" approach. The bagging approach's core notion is that combining many learning models improves the end outcome. Because of its simplicity and versatility, it is also one of the most often used algorithms (it can be used for both classification and regression tasks).

The partial decision tree algorithm (PART) [18] is a rule-based classifier that uses partial decision trees to extract rules. It builds the tree using the same user-defined parameters as J4.8 and C4.5's heuristics. As a consequence, J4.8 and the component classifier can both produce identical results for a given dataset. Logistic model tree (LMT) [25] is a tree-based classifier that employs logistic regression functions and classification trees. The LMT approach can handle numeric, nominal, and missing values, as well as binary and multi-class target variables. LMT is a supervised classification technique that combines decision tree learning with logistic regression. The categorical dependent variable is predicted using a set of independent variables utilizing the supervised learning technique of logistic regression [26]. A decision tree can be used to graphically and succinctly depict decisions and decision making in decision analysis. The decision tree paradigm is used, as the name implies. Cross-validation is used in the basic LMT induction technique to select a number of LogitBoost iterations that do not overfit the training data.

Randomizable filtered classifier [27] is a weka-meta FilteredClassifier variation that uses a randomizable filter, in this case, RandomProjection, as well as IBk as the basic classifier. Apart from that, and ensuring that at least one of the two base schemes implements the Randomizable interface, it performs the same functions as FilteredClassifier, which now also implements Randomizable. Bagging (bootstrap aggression) [28] is a powerful ensemble technique. An ensemble approach is a strategy for making more accurate predictions by combining results from different machine learning algorithms. Bootstrap aggregation is a generic method that may be used to minimize variation in algorithms with a lot of it. Bagging has a high variance as hybrid techniques like classification and regression (CART). The Bootstrap technique is applied to a high-variance machine learning system, such as decision trees, in the process of bagging.

2.3. Implementation procedure

The main goal of this research is to determine the degree of internet users satisfaction, as well as to look into the elements that impact the status of internet service and users. Satisfaction of users with internet service is influenced by a number of important elements such as monthly packages, time, location, and so on. To model implementation, we go through the steps shown in Figure 2.

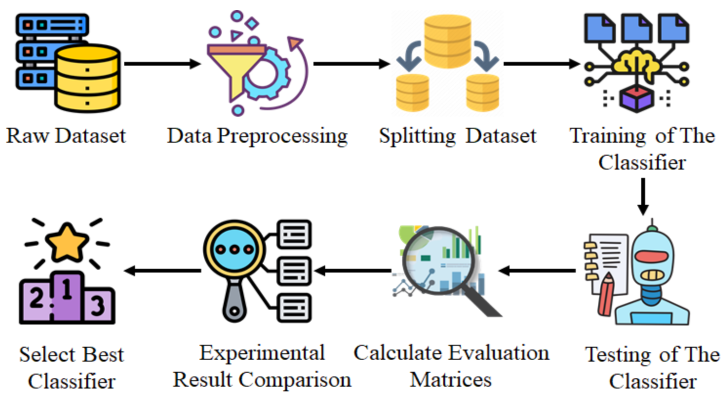

Figure 2. Working procedure of internet user satisfaction using data mining technique

First and foremost, we prepared a public survey questionnaire with three sections: the first section was the conceptual explanation and their agreement to participate in our survey, the second portion with some identical questions, then at that point, the third and significant segment with 18 questions covering all possible factors for an internet user, as we have previously demonstrated in the data description. After that, we attempted to get responses (raw data) from a variety of people. Following data collection, we proceed to data preprocessing and apply certain preprocessing techniques in order to input the data into the classifier. In data preprocessing we handle missing values, data in the wrong format by several python libraries and then category or nominal data that has to be converted to numerical data using Label Encoding. To label variables or attributes, the numeric number used in our dataset used; for example, “user satisfaction level (USL)” has five possible outcomes as very dissatisfied (0), dissatisfied (1), average (2), partially satisfied (3), very satisfied (4).

Our prepared data is divided into two sets after preprocessing: training and testing. For training purposes, 80% of the whole data set was employed. The remaining 20% of the data set has been utilized for testing. This is a completely random process. We used test data to estimate the level of satisfaction among internet users after training the classifiers. Some of the performance evaluation measures have been calculated here. We identified the best classifier to predict in this scenario using these criteria. Using (1)–(7), many performance measures in percentage have been derived based on the confusion matrix created by the classifier.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \times 100\%) \quad (1)$$

$$Sensitivity\ or\ Recall\ or\ True\ Positive\ Rate\ (TPR) = \frac{TP}{TP+FN} \times 100\% \quad (2)$$

$$Specificity\ or\ True\ Negative\ Rate\ (TNR) = \frac{TN}{FP+TN} \times 100\% \quad (3)$$

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{FP+TN} \times 100\% \quad (4)$$

$$False\ Negative\ Rate\ (FNR) = \frac{FN}{FN+TP} \times 100\% \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (6)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \times 100\% \quad (7)$$

3. RESULTS AND DISCUSSION

Several classifiers are used in this paper to analyze the satisfaction level of internet users in Bangladesh. Very dissatisfied (0), dissatisfied (1), average (2), partially satisfied (3), and very satisfied (4) are the five classes in our label column, indicating that our work is a multiclass problem. As a result of the applied classifier, construct a 5×5 confusion matrix as stated in [29], [30], [31]. Table 2 (see Appendix) shows the confusion matrix created by each of the classifiers.

To track down the best model for our work and assess this work, accuracy, TPR, TNR, FPR, FNR, precision, and F1 score from the above confusion matrix is processed. The consequence of a few exhibition assessment measurements is introduced in Table 3. In the overall examination of the results, Table 3 (see Appendix) shows that the Random Forest classifier beats the other eight classifiers. The accuracy of the Random Forest classifier is 90.47, 90.46, 90.24, 93.57, and 92.90% for the very dissatisfied (0), dissatisfied (1), average (2), partially satisfied (3), and very satisfied (4) classes, respectively. The Random Forest classifier's F1 Score for the very dissatisfied (0), dissatisfied (1), average (2), partially satisfied (3), and very satisfied (4) classes is 70.34, 87.89, 84.51, 21.62, and 23.81% respectively, which is outrageous of all the classifiers. Furthermore, the result of other data in Table 3 corroborates the random forest classifier.

4. CONCLUSION

The main focuses of this paper are to measure the impression or satisfaction level of internet users and to give a review about the condition of Bangladesh's internet in data mining approaches. Basically, the internet plays a significant role in the field of economy. As Bangladesh is a developing country, if the internet issues cannot be resolved now, we cannot hope for a better future for our country. Besides, the education system of any country is mainly dependent on the internet nowadays. So, if our country cannot provide us the internet facility in a correct way, it will be difficult for the whole nation to be educated like other developed countries. This paper can assist a policy maker with settling on legitimate choices which can help the entire age of Bangladesh to be an appropriate digitized country later on world. This work can support the internet providers of Bangladesh to enhance the quality of the internet according to the user's satisfaction. We assessed many performance assessment indicators to evaluate the working classifier. We discovered that the Random Forest classifier beats all other data mining approaches. We will work with extra datasets with more provisions later on, and we will utilize more data mining strategies.

APPENDIX

Table 2. Confusion matrix for applied nine classifiers

Model	Class	TP	FN	FP	TN
KStar	Very dissatisfied	47	26	31	347
	Dissatisfied	119	59	63	210
	Average	74	72	66	239
	Partially satisfied	5	24	20	402
	Very satisfied	6	19	20	406
Multilayer perceptron	Very dissatisfied	40	33	31	347
	Dissatisfied	129	49	73	200
	Average	76	70	68	237
	Partially satisfied	6	23	13	409
	Very satisfied	6	19	12	414
Instance based K	Very dissatisfied	45	28	33	345
	Dissatisfied	116	62	57	216
	Average	77	69	72	233
	Partially satisfied	7	22	20	402
	Very satisfied	6	19	18	408
Random committee	Very dissatisfied	40	33	31	347
	Dissatisfied	129	49	73	200
	Average	76	70	68	237
	Partially satisfied	6	23	7	415
	Very satisfied	6	19	15	411
Random Forest	Very dissatisfied	51	22	21	357
	Dissatisfied	156	17	26	252
	Average	120	13	31	287
	Partially satisfied	4	25	4	418
	Very satisfied	5	20	12	414
PART	Very dissatisfied	36	37	34	344
	Dissatisfied	104	47	42	258
	Average	76	44	46	285
	Partially satisfied	6	23	23	399
	Very satisfied	3	22	14	412
Logistic model tree	Very dissatisfied	39	34	32	346
	Dissatisfied	102	76	65	208
	Average	70	76	80	225
	Partially satisfied	7	22	26	396
	Very satisfied	7	23	18	403
Randomizable filtered classifier	Very dissatisfied	42	29	38	342
	Dissatisfied	112	56	47	236
	Average	73	53	36	289
	Partially satisfied	6	23	20	402
	Very satisfied	5	22	22	402
Bagging	Very dissatisfied	22	51	24	354
	Dissatisfied	136	31	87	197
	Average	76	53	43	279
	Partially satisfied	6	23	2	420
	Very satisfied	1	24	1	425

Table 3. Performance evaluation metrices and comparison of nine classifier's performance

Classifier	Class name	Accuracy (%)	TPR (%)	TNR (%)	FPR (%)	FNR (%)	Precision (%)	F1 Score (%)
KStar	Very dissatisfied	87.36	64.38	91.80	8.20	35.62	60.26	62.25
	Dissatisfied	72.95	66.85	76.92	23.08	33.15	65.38	66.11
	Average	69.40	50.68	78.36	21.64	49.32	52.86	51.75
	Partially satisfied	90.24	17.24	95.26	4.74	82.76	20.00	18.52
	Very satisfied	91.35	24.00	95.31	4.69	76.00	23.08	23.53
Multilayer perceptron	Very dissatisfied	85.81	54.79	91.80	8.20	45.21	56.34	55.56
	Dissatisfied	72.95	72.47	73.26	26.74	27.53	63.86	67.89
	Average	69.40	52.05	77.70	22.30	47.95	52.78	52.41
	Partially satisfied	92.02	20.69	96.92	3.08	79.31	31.58	25.00
	Very satisfied	93.13	24.00	97.18	2.82	76.00	33.33	27.91
Instance based K	Very dissatisfied	86.47	61.64	91.27	8.73	38.36	57.69	59.60
	Dissatisfied	73.61	65.17	79.12	20.88	34.83	67.05	66.10
	Average	68.74	52.74	76.39	23.61	47.26	51.68	52.20
	Partially satisfied	90.69	24.14	95.26	4.74	75.86	25.93	25.00
	Very satisfied	91.80	24.00	95.77	4.23	76.00	25.00	24.49

Table 3. Performance evaluation metrices and comparison of nine classifier's performance *(continue)*

Classifier	Class name	Accuracy (%)	TPR (%)	TNR (%)	FPR (%)	FNR (%)	Precision (%)	F1 Score (%)
Random Committee	Very dissatisfied	85.81	54.79	91.80	8.20	45.21	56.34	55.56
	Dissatisfied	72.95	72.47	73.26	26.74	27.53	63.86	67.89
	Average	69.40	52.05	77.70	22.30	47.95	52.78	52.41
	Partially satisfied	93.35	20.69	98.34	1.66	79.31	46.15	28.57
	Very satisfied	92.46	24.00	96.48	3.52	76.00	28.57	26.09
Random Forest	Very dissatisfied	90.47	69.86	94.44	5.56	30.14	70.83	70.34
	Dissatisfied	90.47	90.17	90.65	9.35	9.83	85.71	87.89
	Average	90.24	90.23	90.25	9.75	9.77	79.47	84.51
	Partially satisfied	93.57	13.79	99.05	0.95	86.21	50.00	21.62
	Very satisfied	92.90	20.00	97.18	2.82	80.00	29.41	23.81
PART	Very dissatisfied	84.26	49.32	91.01	8.99	50.68	51.43	50.35
	Dissatisfied	80.27	68.87	86.00	14.00	31.13	71.23	70.03
	Average	80.04	63.33	86.10	13.90	36.67	62.30	62.81
	Partially satisfied	89.80	20.69	94.55	5.45	79.31	20.69	20.69
	Very satisfied	92.02	12.00	96.71	3.29	88.00	17.65	14.29
Logistic model tree	Very dissatisfied	85.37	53.42	91.53	8.47	46.58	54.93	54.17
	Dissatisfied	68.74	57.30	76.19	23.81	42.70	61.08	59.13
	Average	65.41	47.95	73.77	26.23	52.05	46.67	47.30
	Partially satisfied	89.36	24.14	93.84	6.16	75.86	21.21	22.58
	Very satisfied	90.91	23.33	95.72	4.28	76.67	28.00	25.45
Randomizable filtered classifier	Very dissatisfied	85.14	59.15	90.00	10.00	40.85	52.50	55.63
	Dissatisfied	77.16	66.67	83.39	16.61	33.33	70.44	68.50
	Average	80.27	57.94	88.92	11.08	42.06	66.97	62.13
	Partially satisfied	90.47	20.69	95.26	4.74	79.31	23.08	21.82
	Very satisfied	90.24	18.52	94.81	5.19	81.48	18.52	18.52
Bagging	Very dissatisfied	83.37	30.14	93.65	6.35	69.86	47.83	36.97
	Dissatisfied	73.84	81.44	69.37	30.63	18.56	60.99	69.74
	Average	78.71	58.91	86.65	13.35	41.09	63.87	61.29
	Partially satisfied	94.46	20.69	99.53	0.47	79.31	75.00	32.43
	Very satisfied	94.46	4.00	99.77	0.23	96.00	50.00	7.41

REFERENCES

[1] Y. K. Dwivedi, N. Khan, and A. Papazafeiropoulou, "Consumer adoption and usage of broadband in Bangladesh," *Electronic Government, an International Journal*, vol. 4, no. 3, pp.299-313, 2007, doi: 10.1504/EG.2007.014164.

[2] "Internet Subscribers in Bangladesh June, 2021 | BTRC", *Btrc.gov.bd*, 2021. [Online]. Available: http://www.btrc.gov.bd/content/internet-subscribers-bangladesh-june-2021. [Accessed: 25- Jan- 2022].

[3] "Bangladesh charts 9m new social media users," *Dhaka Tribune*, 2021. [Online]. Available: https://www.dhakatribune.com/bangladesh/2021/04/26/bangladesh-charts-9m-new-social-media-users. [Accessed: 25- Jan- 2022].

[4] "Mobile Phone Subscribers in Bangladesh January, 2021 | BTRC," *Btrc.gov.bd*, 2021. [Online]. Available: http://www.btrc.gov.bd/content/mobile-phone-subscribers-bangladesh-january-2021. [Accessed: 25- Jan- 2022].

[5] "Number of internet connections soars by 1.75cr in FY21," *New Age | The Most Popular Outspoken English Daily in Bangladesh*, 2021. [Online]. Available: https://www.newagebd.net/article/145444/number-of-internet-connections-soars-by-175cr-in-fy21. [Accessed: 25- Jan- 2022].

[6] H. H. Wang, and C. C. Yu, "System quality, user satisfaction and perceived net benefits of mobile broadband services," In *Proceedings of 8th International Telecommunication Society Asia-Pacific Regional Conference Taiwan*, 2011, pp. 26-29.

[7] H. R. Yen, "An attribute-based model of quality satisfaction for internet self-service technology," *The Service Industries Journal, vol.* 25, no. 5, pp. 641-659, 2005, doi: 10.1080/02642060500100833.

[8] M. Chae, J. Kim, H. Kim, and H. Ryu, "Information quality for mobile internet services: A theoretical model with empirical validation," *Electronic markets* vol. 12, no. 1, pp. 38-46, 2002.

[9] H. Bruce, "User satisfaction with information seeking on the Internet," *Journal of the American Society for Information Science*, vol. 49, no. 6, pp. 541-556, 1998, doi: 10.1002/(SICI)1097-4571(19980501)49:6%3C541::AID-ASI6%3E3.0.CO;2-1.

[10] O. Isaac, Z. Abdullah, T. Ramayah, and A. M. Mutahar, "Internet usage, user satisfaction, task-technology fit, and performance impact among public sector employees in Yemen," *The International Journal of Information and Learning Technology*, vol. 34, no. 3, pp. 210-241, 2017, doi: 10.1108/IJILT-11-2016-0051.

[11] O. Isaac, Z. Abdullah, T. Ramayah, and A. M. Mutahar, "Factors determining user satisfaction of internet usage among public sector employees in Yemen," *International Journal of Technological Learning, Innovation and Development*, vol. 10, no. 1, pp. 37-68, 2018.

[12] E. S. Davis, and D. A. Hantula, "The effects of download delay on performance and end-user satisfaction in an Internet tutorial," *Computers in Human Behavior*, vol. 17, no. 3, pp. 249-268, 2001, doi: 10.1016/S0747-5632(01)00007-3.

[13] S. Sarawagi, and S. H. Nagaralu. "Data mining models as services on the internet." *ACM SIGKDD Explorations Newsletter*, vol. 2, no. 1, pp. 24-28, 2000, doi: 10.1145/360402.360412.

[14] T Bala, I. Hossain, Dr. AKM G. R. Mondal, "Measurement of Customer Satisfaction of Different Mobile Operators in Bangladesh; a Study on Bangabandhu Sheikh Mujibur Rahman Science and Technology Univeristy, Gopalganj, Bangladesh," *IOSR Journal of Business and Management (IOSR-JBM)*, vol. 20, no 3, ver. 3, pp. 38-47, March. 2018.

[15] M. Hasan, Md. H. I. Bijoy and S. Akhi, "Refute the Decision of Auto-Promotion and Real Facts of Digital Online Classes During the Pandemic in Bangladesh," *2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, 2020, pp. 1-6, doi: 10.1109/ICATMRI51801.2020.9398326.

[16] Md. R. Hasan, Md. H. I. Bijoy, S. A. Khushbu, S. Akter and S. A. Hossain, "Supervised Method Pursued For Overall Impact of Online Class During Lockdown in Bangladesh," *2021 12th International Conference on Computing Communication and*

Networking Technologies (ICCCNT), 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9579827.
[17] J. G. Cleary, and L. E. Trigg, "K*: An instance-based learner using an entropic distance measure," In *Machine Learning Proceedings (1995)*, 1995, pp. 108-114, doi: 10.1016/b978-1-55860-377-6.50022-0.
[18] P. N. R., and J. Lakhmi, eds. *Advanced techniques in data mining and knowledge discovery*. London: Springer, 2005, doi: 10.1007/1-84628-183-0.
[19] D. C. T. Hernández, "An experimental study of K* algorithm," *IJ Information Engineering and Electronic Business.* vol. 2, pp. 14-19, 2015, doi: 10.5815/ijieeb.2015.02.03.
[20] J. Anderson, *An introduction to neural networks*. Cambridge, Mass.: MIT Press, 1995.
[21] F. Murtagh, "Multilayer perceptrons for classification and regression," *Neurocomputing*, vol. 2, no. 5-6, pp. 183-197, 1991. doi: 10.1016/0925-2312(91)90023-5.
[22] M. A. Ayu, S. A. Ismail, A. F. A. Matin, and T. Mantoro, "A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition," *Procedia Engineering*, vol. 41, pp. 224-229, 2012, doi: 10.1016/j.proeng.2012.07.166.
[23] A. Niranjan, D. H. Nutan, A. Nitish, P. D. Shenoy and K. R. Venugopal, "ERCR TV: Ensemble of Random Committee and Random Tree for Efficient Anomaly Classification Using Voting," *2018 3rd International Conference for Convergence in Technology (I2CT)*, 2018, pp. 1-5, doi: 10.1109/I2CT.2018.8529797.
[24] A. Liaw and M. Wiener. "Classification and regression by randomForest," *R news* vol. 2, no. 3 pp. 18-22, 2002.
[25] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine learning*, vol. 59, no. 1-2, pp. 161-205, 2005, doi: 10.1007/s10994-005-0466-3.
[26] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *The journal of educational research*, vol. 96, no. 1, pp. 3-14, 2002, doi: 10.1080/00220670209598786.
[27] D. Ferreira, H. Peixoto, J. Machado and A. Abelha, "Predictive Data Mining in Nutrition Therapy," *2018 13th APCA International Conference on Automatic Control and Soft Computing (CONTROLO)*, 2018, pp. 137-142, doi: 10.1109/CONTROLO.2018.8516413.
[28] A. Lemmens and C. Croux, "Bagging and Boosting Classification Trees to Predict Churn," *Journal of Marketing Research*, vol. 43, no. 2, pp. 276-286, 2006. doi: 10.1509/jmkr.43.2.276.
[29] Md. J. Mia, A. A. Biswas, A. Sattar, and Md. T. Habib, "Registration status prediction of students using machine learning in the context of Private University of Bangladesh," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 1, pp. 2594-2600, 2019, doi: 10.35940/ijitee.A5292.119119.
[30] A. Ahmmed, T. Sultan, Sk. H. I. Shad, M. J. Mia, and S. Mazumder, "An Automated Visa Prediction Technique for Higher Studies Using Machine Learning in the Context of Bangladesh," *Data Engineering for Smart Systems. Lecture Notes in Networks and Systems*, Springer, Singapore, 2022, vol. 238, doi: 10.1007/978-981-16-2641-8_53
[31] B. Adhikari, S. I. Era, M. D. Mohamed and J. Mia, "Depression Level Prediction For Students Using Machine Learning In The Context of Bangladesh," *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1-7, doi: 10.1109/ICCCNT51525.2021.9579475.

BIOGRAPHIES OF AUTHORS

Md. Hasan Imam Bijoy currently pursuing his bachelor's degree (B. Sc) in Computer Science and Engineering (CSE) at Daffodil International University (DIU), Dhaka-1207, Bangladesh. Student member at IEEE and ACM. He worked as a Teaching Assistant (TA) for courses data structure, digital electronics and algorithms under the faculties of the computer science and engineering department at DIU. He is a convener of the Virtual Multidisciplinary Research Lab (VMDRL). He is a research zealot, having published over 9 conference papers, two journals paper and one programming book "A Handbook of C Programming with Example". He has performed the role of reviewer at ICECET2021, and ICECCME2021. His area of interest includes machine learning, deep learning, computer vision, natural language processing, image processing, internet of things, and so many field. He can be contacted at email: hasan15-11743@diu.edu.bd.

Sumiya Alam Akhi is currently studying her bachelor's degree in Computer Science and Engineering (CSE) from Daffodil International University, Dhaka-1207, Bangladesh. She is the COM (Chief Operation Manager) at Virtual Multidisciplinary Research Lab. To perform the role of COM she does guide junior researcher to learn the research study. She has a number of articles in international conference proceedings to her name. Machine learning, natural language processing, deep learning, and data mining are among her research interests. She does computer programming in addition to research. She was a previous Executive Member of Daffodil International University's Computer and Programming Club. She is the first runners-up in Android Hackathon Contest 2020, organized by CPC, DIU. She can be contacted at email: sumiya15-11423@diu.edu.bd.

Md. Ali Ashraf Nayeem is a current student in the Computer Science and Engineering (CSE) department at Daffodil International University (DIU), Dhaka-1207, Bangladesh. He is mainly interested in a mathematical subject. He got 7th position in the inter-university math olympiad. He is also a member of the DIU Computer Programming Club. He also has a strong knowledge of both programming and problem-solving. He is trying to build up his career as a full-stack web developer. His interested areas are python, Django, Bootstrap, HTML, CSS and javascript and research on Machine Learning, AI-based Web application and so many. He can be contacted at email: ashraf15-11626@diu.edu.bd.

Md. Mahbubur Rahman is currently studying in Computer Science and Engineering (CSE) at Daffodil International University (DIU), Dhaka-1207, Bangladesh. He worked as a Lab Prefect (LP) for courses data structure under the faculties of the computer science and engineering department at DIU. He is a Chief Press Executive of the Virtual Multidisciplinary Research Lab (VMDRL). In addition to being a researcher, he is an IT expert. He has a long list of conference papers. His area of interest includes machine learning, natural language processing, image processing, internet of things, and so many field. He can be contacted at email: mahbubur15-11742@diu.edu.bd.

Md. Jueal Mia received his bachelor's degree (B.Sc.) and the master's degree (M. Sc) in Computer Science and Engineering (CSE) from Jahangirnagar University, Dhaka, Bangladesh in 2014 and 2015 respectively. Currently, he works as a Senior Lecturer in the Department of Computer Science and Engineering (CSE), Daffodil International University, Dhaka, Bangladesh. He has published many articles in international journals and conference proceedings. His research interest is mainly on artificial intelligence, computer vision, machine learning, expert systems, data mining, and natural language processing. He can be contacted at email: mjueal02@gmail.com.