A study on social media addiction analysis on the people of Bangladesh using machine learning algorithms

Minjun Nahar Mim[1], Mehedi Firoz[1], Mohammad Monirul Islam[1], Mahady Hasan[2], Md. Tarek Habib[2]
[1]Department of Computer Science and Engineering, Faculty of Science and Information Technology, Daffodil International University, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, School of Engineering, Technology and Sciences, Independent University, Bangladesh

Article Info

Article history:

Received Jan 3, 2023
Revised Feb 8, 2024
Accepted Mar 6, 2024

Keywords:

Classification
Depression
Frustration
Illnesses
Machine learning
Social media addiction

ABSTRACT

Social media has become a fundamental element of contemporary life, providing countless benefits but also posing substantial concerns. While technology improves connectedness and information exchange, excessive use raises issues about social and personal well-being. The emergence of social media addiction emphasizes its influence on everyday routines and mental health, with many people favoring online activities above vital tasks, resulting in real repercussions. Twitter, Facebook, and Snapchat have a significant impact on emotional well-being, adding to global rates of despair and anxiety. To measure the frequency of social media reliance, we studied data from 1,417 individuals using machine learning methods such as decision tree (DT) classifier, random forest (RF) classifier, support vector classifier (SVC), k-nearest neighbors (K-NN), and multinomial naive Bayes (NB). Understanding the behavioral patterns that drive addiction allows us to create tailored therapies to encourage healthy digital behaviors. This study highlights the critical necessity to address social media addiction as a complicated societal issue. Our major goal is to determine the amount of people who are addicted to social media.

Corresponding Author:

Minjun Nahar Mim
Department of Computer Science and Engineering, Faculty of Science and Information Technology
Daffodil International University
Dhaka, Bangladesh
Email: minjun15-12974@diu.edu.bd

1. INTRODUCTION

Social media is now seen as a common way to pass the time. People get frustrated, depressed, and suffer from various ailments as a result of spending too much time on social media [1]. Identifying those who are underrepresented in the multitude of people who express their frustration on social media is simply too difficult. The perspectives of these people will be contingent and progressive [2]. Social media posts, outreach, comments, and criticism will all reveal their mental health [3].

The internet has expanded throughout the world. It is feasible to interact across the globe using social media, and social media is luring individuals to communicate from one end of their internet domain to another [4]. It will become much more widespread in the near future. It will gradually broaden people's emotions, thoughts, sentiments, and ambitions [5], and these social media platforms will be the sole method of expressing or sharing people's emotions, feelings, aspirations, wants, and thoughts [6]. Currently, social media functions as an addiction [7]. Eventually, this addiction brings people to the stage of suicide [8] and enables sensitive individuals in society commit themselves [9]. It should be important before being impacted by a psychiatric disorder [10].

Social media addiction was particularly prominent among school, college, and university students [11], [12]. In addition to social media addiction, the terms social media dependency, problematic network use, compulsive network use, effective social media addiction, ineffective social media use, and emotional uses of social media, are all closely connected to social media [13]. The most popular messaging platforms on social media include Messenger, WhatsApp, Telegram, and others [14]. In today's digital world, young people are surrounded by the internet, smartphones, computers, video games, television, and other devices, which have become habits for them [15]. Today's youth believe it is usual to utilize all social media sites [16]. Young individuals might be at risk when using social media [17]. It can be extremely frustrating for them [18], while also increasing their addiction [19]–[22]. Machine learning algorithms are frequently used to generate prospective features as instructions. Sadness, happiness, despair, frustration, emotion, worry, and pain [23], [24], are all linked to the mental state [25], [26] in all networks. Excessive social media use causes melancholy, frustration, and anxiety [27]. After assessing the intensity of depression and advanced indicators, there is a risk of suicide with cardiovascular disease [28].

This paper tries to identify the social media addiction of Bangladeshi people in advance using sentiment analysis then we will get another result using machine learning algorithm. Many Bangladeshi people waste time by misusing social media, and from there we take the initiative of this research. The main purpose of our research is to find out a proof that the people of Bangladesh are actually addicted to social media. We are the first for this research, collect some data from 1,418 people of Bangladesh, through survey form and manually. Then before using the machine learning algorithm from the collected data by performing sentiment analysis pie chart of the results, we get is added in the results section and this is the first step of our research. After this, from our collected full data set we applied the machine learning algorithm on 30% of the data we get more results whose observation table and graphical representation is also added to the result section.

In this research [29] discuss the fundamental challenges of validity and validation, which should be regarded as a way to develop social media text mining (SMTM) for personality evaluation and, more broadly, personality assessment approaches using machine learning. On the validity of the tests and the validation of the construct. This research raises important problems about validity that have gotten little scholarly consideration. Rani and Kumar [30] employed three machine learning algorithms, including naive Bayes (NB), J48, and support vector machines (SVM). This suggested approach has been tried on platforms such as Facebook, Twitter, and LinkedIn. SVM outperformed other classifiers, with 68% accuracy for social reviews and 82% for tweets. Wang *et al.* [31] in this research they worked on SVM modeling, statistics, support vector classifier (SVC) models and support vector regression (SVR) models for other classifications. They were mounting neuroimaging studies that have significantly understanding the neurobiological mechanism underlying internet addiction. Pratama *et al.* [32] In this research have been investigated in 2019, the most social media platform, of which 130 million users come and they find that statistics on media and gadget usage not only have positive impacts, also negative impacts such as behavioral mental problems. In the survey, they used several techniques such as data science, partial least squares and structural equation modelling. Kim *et al.* [33] used this approach to examine how to employ machine learning in social media datasets. Kikin *et al.* [34] extract features from convolutional neural network pictures using ImageNet's pre-trained autoencoder. Convolutional neural network (CNN) were calculated and compared to each other on this dataset, and they conducted a study of a standard dataset taken from online social media, with detection based on a machine-learning method. It focuses on people's depressive mode, which may be used to prevent mental illness and suicide. Investigated in 2019 the largest social media platform 130 million useless of them come from Indonesia. Savci *et al.* [35] looked at fifteen predictor factors to Prediction of problematic social media use (PSU), including social media usage behaviors. Ahmad *et al.* [36] their proposed methodology is useful to overcome the above problems in feature coding, data acquisition pre-processing, feature selection, and they find average accuracy results from 89 to 91 with the perceptron neural network. Farjana *et al.* [37] in this article, there were two types of assessment course-based assessment and gender-based assessment also to determine the best accuracy using these assessments, provide a mental disorder detection system that provides online social extraction. Their proposed methodology is helpful in overcoming the above-mentioned issues in feature encoding, data pre-processing, data acquisition, feature selection, and classification. In these papers, there are two types of evaluation courses based evaluation and gender-based evaluation also to find out the best accuracy using these evaluations. Losada *et al.* [8] every year almost 25% of people in the United Kingdom suffer from mental disorders. 350 million people suffer from mental illness across society.

In that sense, our future work of research will have to analyze the addiction to social networking sites and help to fix the problems of the next generation. Our paper's major goal is to gather views from the public by conducting surveys and learning from the data using machine learning algorithms to determine the prevalence of social media addiction.

The paper is arranged as follows: section 1 includes an introduction and a brief literature review. Section 2 provides a brief discussion of the study methodology. Section 3 describes the outcome and debate. Section 4 includes the conclusion.

2. METHOD

Gaining personal information through an interview is one of the most direct and effective ways to gain a respondent's trust and cooperation while gathering survey data. An online survey approach is essentially one of the most often used survey formats. We chose this approach because, when it comes to social media-related themes, individuals can connect with online means more readily and we receive speedier replies. Although we nearly had 1,418 respondents to our survey, several of them provided valuable information that aided in our efforts. The six most widely used methods are ultimately included in our dataset: random forest (RF) classifier, SVC, k-nearest neighbors (K-NN) classifier, decision tree (DT) classifier, and multinomial NB. We defined each survey's questions very carefully which is described below.

2.1. Data collection

Through an online survey, we got 1,417 responses from people who shared a lot of information that we converted into our dataset to find out a final output that defines how many people are actually addicted to social media. First of all, we thought of a huge variety of questions for the survey that helped attract people to give their information, and then finally, we fixed almost 23 questions. Specifically, each question is very effective and more efficient for reading the person's hidden mind.

2.2. Selected features

Before doing the data survey, we read and analyzed different kinds of papers and reports. Then we find out some points that helped us short out the socially addicted people. In this paper, we mark the points as features. There are 23 features we got from many research analyses, which are given in Table 1.

Table 1. Selected features for the machine learning algorithms in this research

Features	Citation	Features	Citation
In the era of social media, enjoy reading books	[35]	Open the internet for the sake of time, but it will take much longer than that	[17], [22]
Due to its huge effect on contemporary society, living without the internet would undoubtedly be dull	[5]	The internet still gives chase to offline or imagine that you are online	[5], [9], [17]
Rather than hanging out with friends, spend time on the internet	[15]	Attempting to spend a few more minutes on the internet after work, necessity, or reply to someone else's call	[14], [20]
Think about social media making unsocial	[7], [35]	Spend on the internet, trying to keep it a secret from others	[6], [26]
More internet uses	[22]	Try to limit time spent on the internet. However, in actuality, it is impossible	[35], [37]
Try to hide internet uses	[11], [18]	Depressed or upset without the internet	[14], [35]
Send more time on internet rather than other work	[8], [36]	Family and acquaintances remark about the amount of time spent online	[8], [28]
Spend time with the internet till late at night	[2], [19]	Be envious of other people's allegedly luxurious lives on social media	[3], [16]
Create new relationships in the virtual world	[2], [33], [35]	Excited or irritated when someone makes a humiliating comment	[2], [22], [35]
ignoring uninteresting ideas from reality and focusing on the internet	[29]	Felt alone after utilizing social media	[11]
Internet attraction greater than that of your loved one	[8], [9], [11]	Social media addiction changes your personality	[8], [37]
Using social media create negative effect on personal relationships	[11], [25]		

2.3. Data preprocessing

Data processing is the manipulation of data by a computer. However, data processing may be defined as the collection and transformation of information into a useful format. Data preprocessing is a subset of data preprocessing that refers to any type of processing performed on raw data in order to prepare it for another data processing approach. It has always been an important initial step in the data mining process. The researcher have performed seven stages of preprocessing on the data. To begin, we use web sources to obtain raw data on diabetic patients. After the data have been collected, they need to be cleaned, nulls need to be removed, noise needs to be removed, and any missing values need to be solved before the final preprocessing result can be obtained by labelling the data. Which may be seen in the flowchart shown in Figure 1.

2.4. Proposed work

The major purpose of this study is to analyze social media addiction among Bangladeshis using a machine learning algorithm. Firstly, use sentiment analysis in initial prediction the use 5 algorithm such as DT classifier, RF classifier, SVC, K-NN classifier, and multinomial NB prediction technique is implemented in the proposed work. A workflow has been created for the social media addicted Bangladeshi person identification.

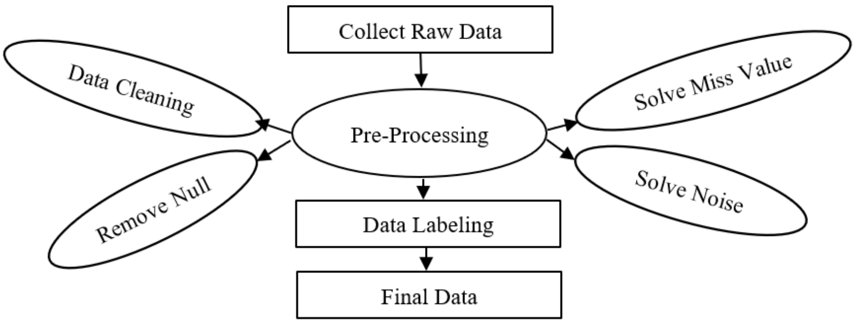

Figure 1. Data pre-processing

Description steps of Figure 1 are mentioned as: i) data collection from student, Job holder, Businessman through google form and manually; ii) at the end of data collection, we preprocess the data like Null value remove; iii) after pre-processing, we process the data by labeling that to train the matching learning algorithm; and iv) then train 5 matching learning algorithms on the pressed data to get the final prediction result.

A workflow framework for machine learning categorization is shown in Figure 2. Before analysis can begin, the data set must first go through data pre-processing. When the data is processed, it undergoes additional refinement and organization after undergoing preprocessing. These classifiers, which include DT, RF, SVC, k-neighbors classifiers, and multinomial NB, are trained on the processed data using the sklearn package. A positive (Yes) or negative (No) result is then obtained depending on the input data by using the trained models for detection or prediction. Several classifiers are trained and evaluated in a methodical manner for best results, thanks to this methodology.

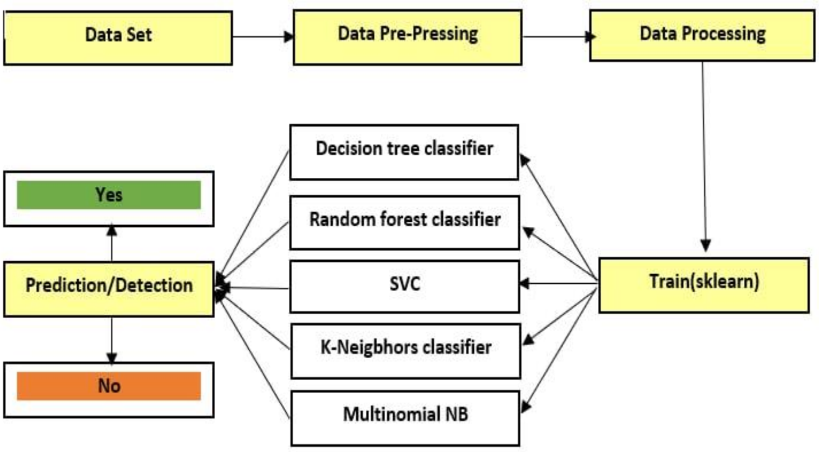

Figure 2. Workflow structure

2.5. Classifier description

A classifier is a machine learning system that automatically organizes or categorizes data into one or more "classes”. To reduce the volume of data in a way that makes it easy to recognize similarities and differences. Thus, millions of figures may be grouped into a few classes that share characteristics.

2.5.1. Decision tree classifier

A DT is a flowchart-like structure in which each interior node assesses a feature and the leaf nodes represent class labels. Branches indicate the combination of traits that leads to various classifications. It visualizes potential solutions depending on certain variables to facilitate decision-making. In classification, each node seeks to categorize data by establishing conditions that optimize class purity. This iterative approach will continue until the dataset is completely separated. Thus, DT provide a comprehensive representation of prospective options, allowing for detailed analysis and educated decision-making across a wide range of domains.

$$\text{Gini (D)}=1-\sum_{i=1}^{m} p_i^2$$

2.5.2. Random forest classifier

RF, a prominent machine-learning approach, leverages ensemble learning to solve classification and regression problems by mixing many DT. It is well-known for its simplicity and adaptability, and while hyper-parameter adjustment is not always necessary, it produces good results. It is widely used in a variety of domains and stands out for its capacity to efficiently handle both classification and regression problems, making it a popular option among data scientists and machine learning practitioners.

$$\text{error } (M_i)=\sum_{j=1}^{d} w_j \times err\left(X_j\right)$$

2.5.3. Support vector classifier

SVC, a statistical clustering algorithm, performs well in low-dimensional information without making assumptions about cluster features. However, in high-dimensional data circumstances, preprocessing such as incorrectly performing principal component analysis may be required. In transmission applications, SVC manages grid voltage by using thyristor-controlled reactors to absorb VARs and potentially lower system voltage if both the algorithm and the SVC are capacitive.

$$A.B=|A| \cos\theta\times|B|$$

2.5.4. K-nearest neighbors classifier

The K-NN classifier determines the five nearest neighbors by explicitly using geometric distance to measure closeness. K-NN calculates the distance between a query and all samples in the dataset, selects the K closest instances, and then determines the most frequent label or average labels. This method allows for efficient categorization by prioritizing local neighborhood information, which aids in tasks like as pattern recognition and grouping based on the properties of surrounding data points.

$$\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{\frac{1}{q}}$$

2.5.5. Multinomial naive Bayes

The multinomial NB classifier is appropriate for different alternatives and incomplete counts, including term frequency-inverse document frequency (TF-IDF). It normally requires feature counts, although it can also handle incomplete counts. The alpha float parameter, which has a default value of one, is utilized in the multinomial distribution.

$$P(A|B)=P(A)\times P(B|A)/P(B)$$

In order to ascertain if different categorization algorithms were effective in forecasting our objective variable, we conducted an evaluation of their performance. The RF, DT, SVC, K-NN, and multinomial NB classifiers were among the techniques evaluated. To guarantee comparability, the same dataset was used for training and testing each method. Table 2 provides a summary of the evaluation's findings. Each algorithm's testing and training scores are shown in the table, giving a clear picture of how effectively each model generalizes to new data.

Table 2. Train and test score for diffirent classifications

Algorithms	Train	Test
DT classifier	0.96	0.77
RF classifier	0.96	0.82
SVC	0.87	0.81
K-NN classifier	0.85	0.81
Multinomial NB	0.73	0.74

2.5.6. Classifier performance

Important measures including accuracy, precision, recall, and F1 score are utilized to assess a classifier's efficacy. By comparing the number of samples properly categorized to the total number of samples, accuracy calculates the classifier's overall correctness. By determining the percentage of genuine positives among all positive forecasts, precision evaluates the accuracy of positive predictions. By calculating the percentage of true positives among all real positives, recall assesses the classifier's capacity to recognize positive examples. Combining recall and accuracy into a single metric, the F1 score offers a balanced assessment that is particularly helpful when class distributions are unbalanced. When combined, these indicators provide a thorough understanding of the classifier's effectiveness.

$$\text{Accuracy}=\frac{No.\ of\ correctly\ classified\ samples}{No.of\ tested\ samples}\times 100\%$$

$$\text{Precision}=\frac{TP}{TP+FP}\times 100\%$$

$$\text{Recall}=\frac{TP}{TP+FN}\times 100\%$$

$$F_1 score = \frac{2\times precision\times recall}{precision+recall}\times 1$$

The basic goal of this study is to generate a classifier report that contains precision, recall, f1-score, support, accuracy, macro average, and weighted averages. Additionally, Table 3 displays the classifier report for each and every classifier.

Table 3. Classifier report for five algorithms

RF classifier	Precision	Recall	F1-score	Support
0	0.81	0.83	0.82	179
1	0.83	0.81	0.82	180
Accuracy			0.82	359
Macro avg	0.82	0.82	0.82	359
Weighted avg	0.82	0.82	0.82	359

DT classifier	Precision	Recall	F1-score	Support
0	0.74	0.81	0.78	179
1	0.79	0.72	0.76	180
Accuracy			0.77	359
Macro avg	0.77	0.77	0.77	359
Weighted avg	0.77	0.77	0.77	359

SVC	Precision	Recall	F1-score	Support
0	0.74	0.82	0.78	179
1	0.80	0.72	0.76	180
Accuracy			0.77	359
Macro avg	0.77	0.77	0.77	359
Weighted avg	0.77	0.77	0.77	359

K-NN classifier	Precision	Recall	F1-score	Support
0	0.79	0.82	0.80	179
1	0.81	0.78	0.80	180
Accuracy			0.80	359
Macro avg	0.80	0.80	0.80	359
Weighted avg	0.80	0.80	0.80	359

Multinomial NB	Precision	Recall	F1-score	Support
0	0.73	0.72	0.72	179
1	0.72	0.74	0.73	180
Accuracy			0.73	359
Macro avg	0.73	0.73	0.73	359
Weighted avg	0.73	0.73	0.73	359

The performance of a classifier may be judged by how near its output curve is to the top left corner. Points on the diagonal (false positive rate (FPR)=true positive rate (TPR)) are the minimum acceptable output from a random classifier. The lower the test's accuracy, the closer the curve is to the receiver operating characteristics (ROC) space's 45-degree diagonal. Figure 3 should also include a depiction of the ROC curve for each of the classifiers.

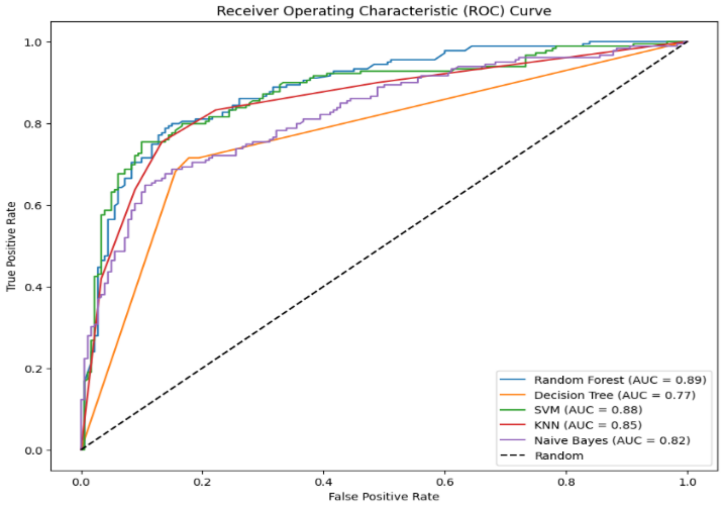

Figure 3. ROC curve plot

3. RESULTS AND DISCUSSION

In this study, we analysed the data using five distinct types of algorithms, and the results showed that the scores and accuracy can vary significantly. When compared to other classifiers, the RF classifier has the highest accuracy, coming in at 82%. Following that, the K-NN classifier has an accuracy of 80%, making it the second most accurate method.

3.1. Comparative analysis

Many times, people are unaware of topics that they may learn about via social media. Furthermore, depending on how often a person uses the internet, social media may either enhance or diminish a person's personality at any one time. When individuals utilize the internet for good purposes, it may be highly beneficial to them; yet, when people use it for malicious purposes, it can be quite harmful. However, those who use the internet more excessively are frequently those who shouldn't be using it. People are becoming more and more negative every day due to internet addiction.

The table labeled "Table 4. Classification performance of the five algorithms used" compares several algorithms employed in various research studies focusing on social media and internet-related problems. It comprises information about the publishing year, work content, problem domain, and algorithm utilized. The table shows the highest accuracy attained by each algorithm. For example, in 2023, a RF algorithm was employed for social media analysis with an accuracy of 82%, but in 2021, a hybrid suggested algorithm reached an accuracy of 92%. Other algorithms discussed include SVC, Mixed, ANN, SVM, and TSVM, each having its own accuracies and distinct metrics for predicting eRisk, such as the average number of days or words submitted.

Table 4. Classification performance of the five algorithms used

Work done	Year	Content of work	Problem domain	Algorithm	Best accuracy
Our work	2023	Social media	Analysis	RF	82
[1]	2021	-	-	Hybrid proposed	92
[30]	2019	Social media	Sentiment analysis	SVC	82
[34]	2020	Social media	Detection	Mixedd	Analytical result
[35]	2020	Social media	Prediction	ANN	62
				SVM	63
[8]	2017	Internet	eRisk predection	Avg num. of days from first to last submission	608.31
				Avg num. words per submission	26.9
[38]	2017	Social media	Detection	TSVM	84.2

Using survey data, table 5 compares the identification of social media addiction by age and gender. Responses are divided into three age groups: under 18, between 18 and 30 years old, and above 30. The table displays the overall proportion of replies for each age group along with the percentages of "Yes" and "No" responses for men and women. A comparison column that shows whether "Yes" replies outweigh "No" responses is also included. According to the statistics, a greater proportion of men and women across all age categories ("Yes" replies) than those who did not ("No" responses) reported being addicted to social media, with the largest difference shown in the under-18 age group.

Table 5. Comparison of age and gender for social media addiction detection through survey data

Age	Total response (%)	Female		Compare	Male		Compare
		Yes (%)	No (%)		Yes (%)	No (%)	
Less than 18	8	81	19	Yes>No	97	3	Yes>No
18-30	89	63	37	Yes>No	59	41	Yes>No
More than 30	3	100	0	Yes	100	0	Yes>No

4. CONCLUSION

Like an all-pervasive force, social media has spread across almost every facet of our lives. From having our first cup of morning coffee to traveling across the world, everything can be done or at least experienced through social media and its many aspects. It is almost as if these days, we are just as much social media beings as we are social beings. Having known all that, it should not come as a surprise to anyone that social media addiction is a major problem for many out there. Through the implementation of features such as news feed scrolling and sharing, anybody who spends enough time online cannot help but associate every part of their lives with social media. Therefore, addiction to these often mindless and pointless activities on social media platforms has become inevitable. Also, another factor that is worsening this eventual addiction is the sheer number of social media platforms available online these days. Their variety and usage alone will be enough to compel one to spend hours scouring through them and learning all about their mind-numbingly addictive features. As more and more people keep flocking to these platforms, their owners eventually make enough money to keep optimizing their performance and inventing new features to keep the users addicted to their app or website. This, in the process, is only making it more difficult for long-term users to quit social media and focus on more productive activities. It has gotten to the point where this never-ending loop of feeding off of website traffic and using it to generate more users has become a sort of vicious cycle with some truly vicious consequences that risk-taking away a person's focus, mental drive, and freedom. If we do not become aware of this pandemic now, it might soon become too late for us to even save our future generations from the nefarious clutches of social media addiction. This work has open up several questions that need of further investigation. Further work needs to be done to establish whether as limitations such as need investigated more deeply and compared with other works. It is advised that further research be done in the following areas: better model construction and increased data collecting.

REFERENCES

[1] D. S. Smys and D. J. S. Raj, "Analysis of deep learning techniques for early detection of depression on social media network a comparative study," *Journal of Trends in Computer Science and Smart Technology*, vol. 3, no. 1, pp. 24–39, 2021.

[2] H.-N. Le and R. C. Boyd, "Prevention of major depression: early detection and early intervention in the general population," *Clinical Neuropsychiatry*, vol. 3, no. 1, pp. 6–22, 2006.

[3] C. S. Egbe, "Mental health–a slow paradigm shift in stigma, diagnostics and treatment," 2020.

[4] A. Halfin, "Depression: the benefits of early and appropriate treatment," *American Journal of Managed Care*, vol. 13, no. 4, pp. 92–97, 2007.

[5] I. M. Cameron *et al.*, "Measuring depression severity in general practice: Discriminatory performance of the PHQ-9, HADS-D, and BDI-II," *British Journal of General Practice*, vol. 61, no. 588, pp. 419–426, 2011, doi: 10.3399/bjgp11X583209.

[6] A. Picardi *et al.*, "A randomised controlled trial of the effectiveness of a program for early detection and treatment of depression in primary care," *Journal of Affective Disorders*, vol. 198, pp. 96–101, 2016.

[7] K. L. Smarr and A. L. Keefer, "Measures of depression and depressive symptoms: beck depression Inventory-II (BDI-II), center for epidemiologic studies depression scale (CES-D), geriatric depression scale (GDS), hospital anxiety and depression scale (HADS), and patient health Questionna," *National Library of Medicine*, vol. 63, no. 11, pp. 454–466, 2011.

[8] D. E. Losada, F. Crestani, and J. Parapar, "eRISK 2017: CLEF lab on early risk prediction on the internet: experimental foundations," in *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 8th International Conference of the CLEF Association, CLEF 2017, Dublin, Ireland, September 11--14, 2017, Proceedings 8*, Springer, 2017, pp. 346–360.

[9] M. Park, D. W. McDonald, and M. Cha, "Perception differences between the depressed and non-depressed users in Twitter," in *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, 2013, vol. 7, no. 1, pp. 476–485, doi: 10.1609/icwsm.v7i1.14425.

[10] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," in *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, 2013, pp. 128–137, doi: 10.1609/icwsm.v7i1.14432.

[11] P. Wallace, "Internet addiction disorder and youth: there are growing concerns about compulsive online activity and that this could impede students' performance and social lives," *EMBO Reports*, vol. 15, no. 1, pp. 12–16, 2014, doi: 10.1002/embr.201338222.

[12] J. J. Al-Menayes, "Dimensions of social media addiction among university students in Kuwait," *Psychology and Behavioral Sciences*, vol. 4, no. 1, p. 23, 2015, doi: 10.11648/j.pbs.20150401.14.

[13] D. J. Kuss and M. D. Griffiths, "Internet gaming addiction: a systematic review of empirical research," *International Journal of*

Mental Health and Addiction, vol. 10, no. 2, pp. 278–296, 2012, doi: 10.1007/s11469-011-9318-5.
[14] D. J. Kuss and M. D. Griffiths, "Adolescent social media addiction," *International Journal of Environmental Research and Public Health*, vol. 35, no. 3, pp. 49–52, 2017.
[15] F. Bányai *et al.*, "Problematic social media use: results from a large-scale nationally representative adolescent sample," *PLoS ONE*, vol. 12, no. 1, 2017, doi: 10.1371/journal.pone.0169839.
[16] A. Lenhart, K. Purcell, A. Smith, and K. Zickuhr, "Social media & mobile internet use among teens and young adults. Millennials," *Pew Internet & American Life Project*, vol. 1, pp. 1–16, 2010.
[17] A. O. Folaranmi, "A survey of Facebook addiction level among selected Nigerian University undergraduates," *New Media & Mass Communication*, vol. 10, no. 2012, pp. 70–80, 2013.
[18] P. A. Atroszko, J. M. Balcerowska, P. Bereznowski, A. Biernatowska, S. Pallesen, and C. S. Andreassen, "Facebook addiction among Polish undergraduate students: validity of measurement and relationship with personality and well-being," *Computers in Human Behavior*, vol. 85, pp. 329–338, 2018, doi: 10.1016/j.chb.2018.04.001.
[19] C. S. Andreassen *et al.*, "The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: a large-scale cross-sectional study," *Psychology of Addictive Behaviors*, vol. 30, no. 2, pp. 252–262, 2016, doi: 10.1037/adb0000160.
[20] E. L. Pelling and K. M. White, "The theory of planned behavior applied to young people's use of social networking web sites," *Cyberpsychology and Behavior*, vol. 12, no. 6, pp. 755–759, 2009, doi: 10.1089/cpb.2009.0109.
[21] C. Marino, G. Gini, A. Vieno, and M. M. Spada, "A comprehensive meta-analysis on problematic Facebook use," *Computers in Human Behavior*, vol. 83, pp. 262–277, 2018, doi: 10.1016/j.chb.2018.02.009.
[22] K. Wilson, S. Fornasier, and K. M. White, "Psychological predictors of young adults' use of social networking sites," *Cyberpsychology, Behavior, and Social Networking*, vol. 13, no. 2, pp. 173–177, 2010, doi: 10.1089/cyber.2009.0094.
[23] A. Wongkoblap, M. A. Vadillo, and V. Curcin, "Researching mental health disorders in the era of social media: systematic review," *Journal of Medical Internet Research*, vol. 19, no. 6, 2017, doi: 10.2196/jmir.7215.
[24] A. E. Aladag, S. Muderrisoglu, N. B. Akbas, O. Zahmacioglu, and H. O. Bingol, "Detecting suicidal ideation on forums: proof-of-concept study," *Journal of Medical Internet Research*, vol. 20, no. 6, 2018, doi: 10.2196/jmir.9840.
[25] S. M. Rice *et al.*, "Online and social networking interventions for the treatment of depression in young people: a systematic review," *Journal of Medical Internet Research*, vol. 16, no. 9, 2014, doi: 10.2196/jmir.3304.
[26] S. Balani and M. De Choudhury, "Detecting and characterizing mental health related self-disclosure in social media," in *Conference on Human Factors in Computing Systems - Proceedings*, 2015, vol. 18, pp. 1373–1378, doi: 10.1145/2702613.2732733.
[27] M. L. Birnbaum, S. K. Ernala, A. F. Rizvi, M. De Choudhury, and J. M. Kane, "A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals," *Journal of Medical Internet Research*, vol. 19, no. 8, 2017, doi: 10.2196/jmir.7956.
[28] M. Conway and D. O'Connor, "Social media, big data, and mental health: current advances and ethical implications," *Current Opinion in Psychology*, vol. 9, pp. 77–82, 2016, doi: 10.1016/j.copsyc.2016.01.004.
[29] L. Tay, S. E. Woo, L. Hickman, and R. M. Saef, "Psychometric and validity issues in machine learning approaches to personality assessment: a focus on social media text mining," *European Journal of Personality*, vol. 34, no. 5, pp. 826–844, 2020, doi: 10.1002/per.2290.
[30] S. Rani and P. Kumar, "A sentiment analysis system for social media using machine learning techniques: social enablement," *Digital Scholarship in the Humanities*, vol. 34, no. 3, pp. 569–581, 2019, doi: 10.1093/llc/fqy037.
[31] Y. Wang *et al.*, "Identifying internet addiction and evaluating the efficacy of treatment based on functional connectivity density: a machine learning study," *Frontiers in Neuroscience*, vol. 15, 2021, doi: 10.3389/fnins.2021.665578.
[32] M. O. Pratama, D. Harinitha, S. Indriani, B. Denov, and D. Mahayana, "Influence factors of social media and gadget addiction of adolescent in Indonesia," *Jurnal Sistem Informasi*, vol. 16, no. 1, pp. 16–24, 2020, doi: 10.21609/jsi.v16i1.918.
[33] J. Kim, D. Lee, and E. Park, "Machine learning for mental health in social media: bibliometric study," *Journal of Medical Internet Research*, vol. 23, no. 3, 2021, doi: 10.2196/24870.
[34] P. M. Kikin, A. A. Kolesnikov, and E. A. Panidi, "Social media data processing and analysis by means of machine learning for rapid detection, assessment and mapping the impact of disasters," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 43, pp. 1237–1241, 2020, doi: 10.5194/isprs-archives-XLIII-B3-2020-1237-2020.
[35] M. Savci, A. Tekin, and J. D. Elhai, "Prediction of problematic social media use (PSU) using machine learning approaches," *Current Psychology*, vol. 41, no. 5, pp. 2755–2764, 2022, doi: 10.1007/s12144-020-00794-1.
[36] R. I. Ahmad, M. J. Iqbal, and M. Bakhsh, "Sentiment analysis of social media contents using machine learning algorithms," *Technical Journal*, vol. 24, no. 4, pp. 47–55, 2020.
[37] P. Farjana, G. N. Nikhila, K. K. Nandini, M. Afrin, and A. Papasani, "Social network mental disorder detection via online social media mining using machine learning framework," 2019.
[38] H. H. Shuai *et al.*, "A comprehensive study on social network mental disorders detection via online social media mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 7, pp. 1212–1225, 2018, doi: 10.1109/TKDE.2017.2786695.

BIOGRAPHIES OF AUTHORS

Minjun Nahar Mim is currently pursuing his M.Sc. in Computer Science and Engineering (CSE) at Jahangirnagar University, Bangladesh & she is completed his B.Sc. from Daffodil International University. In a few days, she will be completing her Master of Science Degree in Computer Science and Engineering at Jahangirnagar University. She has interests in machine learning, deep learning, data science, and image processing. She is currently doing machine learning. Her passion for leveraging advanced computational techniques to tackle real-world challenges underscores his potential as a valuable contributor to the field. With a solid foundation and a thirst for knowledge, she is poised to make significant contributions to the ever-evolving landscape of machine learning research. She is quite active in his research work. She can be contacted at email: minjun15-12974@diu.edu.bd.

Mehedi Firoz is a Master of Science in Computer Science and Engineering (CSE) student at Jahangirnagar University in Bangladesh. He earned a bachelor's degree in computer science and engineering from Daffodil International University. With a keen interest in data science, machine learning, deep learning, and image processing, he is driven by a passion for leveraging advanced computational techniques to extract insights from complex datasets. His academic journey showcases a blend of theoretical knowledge and hands-on experience, reflecting his dedication to mastering the intricacies of these fields. He is doing machine learning right now. He devotes a great deal of time to his inquiries. He can be contacted at email: mehedi15-2475@diu.edu.bd.

Mohammad Monirul Islam earned an M.Sc. in computing from London Metropolitan University in the United Kingdom and a B.Sc. in computer Science from the University of Pune in India. He is currently pursuing Ph.D. degree at Jahangirnagar University's Department of Computer Science and Engineering. He is an Assistant Professor at Daffodil International University in Bangladesh's Department of Computer Science and Engineering. He has made major contributions to his profession, with multiple publications in prestigious international journals and conference proceedings. His research interests include machine learning and deep learning, data mining, embedded systems and robotics, internet of things (IoT), and intelligent systems. He can be contacted at email: monirul.shemul@gmail.com.

Mahady Hasan completed his Bachelor degree in Computer Science from Independent University, Bangladesh in 2001. He received his Master's degree in Information Technology from University of Sydney in 2004 and Ph.D. in Spatial Database from Computer Science and Engineering department of University of New South Wales (UNSW) in 2010. Prior to joining IUB as Assistant Professor, he served UNSW, Australian Pacific College and James Cook University, Australia. He has been working at IUB since 2004. Currently, he is the head of the department of Computer Science and Engineering and working in the Department as Associate Professor. Conducting research in the area of software engineering, big data, data analytics and engineering education is the key interest of him. He has published several Book Chapters, Journals and Conference Papers in many renowned venues. He can be contacted at email: mahady@iub.edu.bd.

Md. Tarek Habib achieved his Ph.D. degree at the Department of Computer Science and Engineering at Jahangirnagar University in 2022. He obtained his B.Sc. degree in Computer Science from BRAC University in 2006. Then he got an M.S. degree in Computer Science and Engineering from North South University in 2009. He is an Assistant Professor at the Department of Computer Science and Engineering at Independent University, Bangladesh. Earlier, he was an Associate Professor at the Department of Computer Science and Engineering at Daffodil International University. He is much fond of research. He has published many articles in international journals and conference proceedings. His research interest is in artificial intelligence, especially machine learning, artificial neural networks, computer vision, and natural language processing. Recently, he has started working on IoT too. He can be contacted at email: tarek@iub.edu.bd.