Effect of Social Media on Mental Health

Tasfia Haque Sababa
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
tasfia.haque.sababa@g.bracu.ac.bd

Tajish Farhan
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
md.tajish.farhan@g.bracu.ac.bd

Syed Thakdir Ahmed Turjo
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
syed.thakdir.ahmed.turjo@g.bracu.ac.bd

Nusrat Jaman Pretty
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
nusrat.jaman.pretty@g.bracu.ac.bd

Tahshinul islam
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
tahshinul.islam@g.bracu.ac.bd

Dr. Amitabha Chakrabarty
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
amitabha@bracu.ac.bd

***Abstract*—An online community where people work together to produce, share and change their ideas and comments about any data is known as web-based social networking. Over the past ten years, long-distance social networking purpose of connecting has profoundly changed how people interact and collaborate. This study aims to evaluate how social media expanded and opened doors to making social correlations that might add to psychological wellness challenges. It included inquiries for demographic data, an example of long-range interpersonal communication use, social relationships, and well-being impacts. A descriptive study examined responses from 793 adults to a self-created questionnaire with four sections that was sent using the Google survey tool. In order to develop a paradigm concerning the Chi-Squared test has been used to know about the relationships between social networking site use and the three different categories of psychological distress which are depression, anxiety, loneliness. Previously, different methodologies and theories have been used to know about the effects of social media on mental health. We have tried using the supervised learning classifiers such as Random forest, Decision Tree, SVN, KNN, ADA Boost, Naïve bayes and Voting classifier on the data set pertaining to the use of social networking sites and mental health issues which was dynamically examined. Among those classifiers we chose the best results with the highest accuracy and f1 scores as these results can tell us how much false positives and false negatives we get. Depending on the scores, we found that SVN classifier delivered the best results with an average accuracy score of 38% for depression, 48% for anxiety and 65% for loneliness and an F1 scores of 28%, 42% and 44% respectively.**

I. Introduction

The use of social media has significantly increased over the last 10 years and now affects practically every aspect of modern day life. A collection of web-based services that let users communicate with one another in various ways make up these social media websites. On numerous social media platforms, people can communicate their thoughts, interests, and even actions with friends and family. These websites are used by users to stay updated on news and trends. Use of social media encourages the release of dopamine in the brain. These platforms are often compulsive and might have interconnections to anxiety and depression. People now a days publishes everything they do on social media, so when someone is left out, they feel terrible. A 2018 British study found that the growing usage of social media results in sleep deprivation, which affects mood, memory, and academic performance. According to the Pew Research Center, 81 percent of teenagers and 69 percent of adults use social media. According to BTRC, there are currently 162.920 million mobile phone users in Bangladesh. Multiple social media usage has also been found to be associated with worsening anxiety symptoms in adults between the ages of 19 and 32. Users' anxiety levels are significantly impacted by the quantity of likes or comments. Social media users frequently publish pictures of their lavish lifestyles and sometimes phony posts, which leads them to portray a deceptively favorable self-image. Additionally, other individuals may view those posts, which could lead to their sadness and depression. According to a poll by The Bangladesh Institute of ICT in Development, almost 80% of women in Bangladesh experience online harassment which could be a another reason of the worsening mental health issues. Girls receive sexually explicit videos, messages, and pictures in 64 percent of cases. Social media users experience extreme despair and anxiety as a result of this.

This paper begins with a quick overview of mental health issues and some background research on the potential links between social media use and these issues. Following this,

other researchers' works were shared, which encouraged us to continue our research. The description of the data set and its collection process follows. It concludes with a thorough explanation of how the data set was employed and applied to various machine learning algorithms to create a system that can predict mental health states and outcomes.

II. Research Objectives

The aim of this descriptive research is to assume the effects of social media on mental health using decision tree and random forest algorithms. A data-set consists of 793 participants across Bangladesh aged between 15-40 years was collected. The data-set is processed through feature selection(Chi Square), splitting data and training data. Using machine learning algorithms such as decision tree, random forest, SVM we create a system to predict about three major mental health states- depression, anxiety and loneliness.

The objectives of this research are:

- To predict the impacts of social media on mental health
- To get the most efficient accuracy rate about social media's impact on mental health using algorithms.
- To evaluate and offer recommendations on improving the research.

III. Problem Statement

Our study aims to predict how social media might influence the mental health of individuals between the ages of 15 and 40. However we had to face many interesting problems while doing this research. Collecting such kind of huge dataset was one of the big challenges . Moreover many people didn't have the proper idea about mental health issues as depression anxiety and loneliness. Sometimes people don't even understand properly if they are facing mental health problems for which it was getting harder for getting the appropriate answers from the participants. Another big challenge we had to face was to get a good accuracy rate. We had to use more algorithms like SVM, ADA boost ,KNN, random forest, decision tree etc to increase the accuracy rate. Additionally, we had to use the Chi-Square to find out the relation between social networking site use and the three different categories of psychological distress which was becoming hard to implement in such a long dataset. Hyper parameter was also implemented on the algorithms in order to get a good accuracy rate but it didn't work much for which we had to move to the idea of combining the algorithms for the dataset. Therefore, doing this research, collecting dataset and getting a proper accuracy rate was quite challenging for us.

IV. Literature Review

The social media platform has become a part of our life nowadays. From children's to adults everyone is engaged in at least one or multiple social media platforms. While these platforms have some positive use cases, the negatives far outweigh the positives. From spreading false news on the internet to organizing hate crimes, social media is used everywhere. This section attempts to assess previous significant work in the field of social media, particularly in regard to mental health. We looked at the various methods used to obtain the main findings, and we demonstrated how social media can have a variety of effects on people's mental health, as well as how it comes with its own set of challenges due to excessive use, limited computational capabilities, and a large number of connected devices, all of which exacerbate mental health.

The research paper [6] suggests using SPSSv (16.0) software which was mainly used for the analysis purpose and some statistical tools like Crosstabs, Anova and Correlation were used to carry out the process. This study took descriptive research to take into consideration. For this study, both primary and secondary datasets were gathered. The questionnaire provided all the primary datas, while secondary datas was taken through websites, publications, and journals. A total of 90 MBA students were surveyed. And then a sample was chosen at random. The participants were given a standard questionnaire. The first section related to demographic questions and the second section does include multiple-choice questions about social networking usage patterns, time spent on social networking sites, daily activity on social networking sites, SNS engagement, and health-related issues like difficulty sleeping and eating, depression, and anxiety.

Another study [8] advocates using PRISMA to learn more about the connection between social media use and mental health impacts.After examining many databases, published papers from January 2010 to June 2020 were selected. To investigate the association between social media sites and mental health, articles from PubMed and Google Scholar were chosen. Due to the database's publicly accessible resources in terms of MEDLINE and other NLM resources, PubMed was utilized to complement and improve Google Scholar, which covers a wide spectrum of interdisciplinary scientific publications. The search criteria were created to encompass a wide range of definitions and literature on social media and mental health.

In the research paper [1] relationships between Facebook addiction and depression, insomnia or even positive sites were evaluated in a longitudinal study. A total of 349 inpatients from a psychosomatic rehabilitation facility in Germany were examined over the course of six weeks. "Beck Depression Inventory-II " German version was used to measure depression, with 21 items representing various symptoms of depression. A Likert Scale of four points was used to assess the degree of depression. And readings of the 21 Factors combined represented a person's depressiveness. According to the findings, the score of addictive Facebook usage was strongly positively connected with depression, and the duration of use was associated with sleeplessness. Which corresponds with the hypothesis of this research paper.

In the research paper [9] association between Facebook engagement and psychological distress, along with underlying mechanisms such as self-esteem and communication overload were examined using structural equation modeling. The study was conducted based on a survey consisting of respondents

from 513 college students. Respondents were asked how often they felt unhappy, apprehensive, restless, or fidgety .Based on the answers the psychological distress was measured using Kessler Psychological Distress Scale. As a result the index had five components and was trustworthy. Whereas eight items of a seven-point Likert scale were used to determine the frequency of Facebook activity. Their findings back up other recent research that also shows us Facebook and psychological wellbeing is negatively correlated. It also shows us that while communication overload does not have any association with Facebook interaction and distress. It opens the door for it through low self-esteem. They do acknowledge, however, that there is a chance that psychological discomfort may also induce individuals to use Social media.

The research paper [10] uses two major datasets. The first one gives the dates of facebook's rollout into 775 collages. The second one consisted of 17 successive waves of the national college health assessment's thorough survey regarding mental health conducted at the time of Facebook's expansion. This study is based on a generalized difference in survey respondents before and after Facebook was launched. And, by employing their technique, they were able to rule out a number of variables that may have influenced the conclusion, making the outcome more certain. According to their findings, Facebook's launch had a negative impact on mental health. As a result of Facebook's rollout, their index of poor mental health by NCHA which includes all relevant mental health factors rose by around 0.085 standard deviation units. According to a meta-analysis by Paul and Moser, this size is roughly 22 percentage of the effect of losing one's employment on mental health according to Paul and Moser . They also found out that Depression and anxiety-related diseases are the mental health issues that are driving the outcomes. "Forms of electronic communication" is how Merriam-Webster defined social media as. However nowadays social media is much more than just a form of communication medium as people use it for much more than just communicating with each other. With more use cases it's becoming a part of everyone's life day by day. In 2015, Facebook claimed that it had an estimated 1 billion active users, which means on any given day one in every seven individual visits Facebook. And its rising popularity also raises concerns about its consequences. In the research paper [22] Bashir and Bhat shows us that there has been evidence of connection between social media usage and obsessive behavior. For this paper Bashir and Bhat used other papers to prove his claim. They found that 40 percantage of British people became aggressive being unable to access social media. Which is not a good sign to start with. Moreover they found that because of social media addiction a new kind of syndrome has been found which is nothing but a social media addict sensing his phone vibrating even when it's not and it's called PVS. Along with PVS they found that depression, harassment, cyberbullying, sexting, anxiety are all negative aspects of social media. They found that young adults are most prone to anxiety and depression and excessive social media usage makes them have anxiety and it progresses to depression slowly. Davila (2012)[11] said that younger generations have fewer positive social connections due to significant depressive symptoms among them.

According to research work [14] and Grossman (2017),People with low self esteem tend to suffer from many clinical disorders like anxiety, depression and suicidal thoughts and it's also found that they are more likely to show a fake personality online. After picking out the adolescents who are involved in risky online behaviors and need external validations while also being influenced by social comparisons, it is found that FOMO can help them overcome many challenges that arise from the glorification of social media and social media websites such as depression and suicide. While we have seen mixed results from studies that involve social media and mental health we have also seen one consistent result among them which is it seems to heighten both positive and negative feelings. Many positive sides have also been described by many researchers like Vogel et al. (2017), who said that people tend to see idealized versions of themselves when they look at their own profile which definitely helps to enhance one's self esteem. Furthermore, social media platforms like Instagram and Snapchat allow young people to try out new social behaviors in a less risky context than in real life, and gaining external validation in the form of likes and favorable comments can assist to reduce loneliness and enhance self-esteem. Adolescents with more susceptible dispositions are especially prone to the harmful effects of these social networking platforms.

The study was based on a review which was conducted in 2020. According to research paper [15] there were reviewed mental health impacts on social media through some data and the article was stated in Persian and English language. They collected 501 items in total. The articles and texts were examined in three steps for purposive sampling such as initial search which focused on the social media and mental health related titles. The secondary screening gives a comprehensive study, categorization based on theoretical bounty, and a full examination of the articles. There were two sectors in the study one is inclusion criteria and exclusion criteria. The article was published between 2000 and 2020 years and was stated in in Persian and English languages where highlights how mental health is being affected for social media and also in exclusion criteria there elaborated the impacts of social media and its aspects, reviewing protocols, Inadequate data presented or a poor description of the methodology used and so on. Following the identification of relevant papers, the reference lists of those articles were reviewed for additional research. The remaining items were then thoroughly scrutinized. Finally, out of 501 papers analyzed and also in total 50 cases were extensively examined and featured in the study. The article selection structure is explained by a flowchart demonstrating the steps from unsorted search queries to the maximum number of articles selected. After the investigation there found two aspects of results - positive and negative. The widespread usage of social media has prompted studies exploring a connection between social media and mental health. There was substantial proof that social media usage is somehow connected to unfavorable

mental health outcomes. According to the research paper, there stated that social media usage occur anxiety, stressed, depression, sadness, insomnia, mentally unhappy, indicators, being suicidal, enhancing stress, cyber bullying, dissatisfaction of own image, and life dissatisfaction. On the other hand, from the findings, there were some beneficial benefits regarding the effects of social media on mental health such as recovering from depression and anxiety, maintaining relationships, making friendships with new people and meeting with new ones, increasing virtual communication, creating innovative things, developing mental health and so on.

Nowadays, adolescents cannot imagine life without technology and the internet. The research paper [16] is conducted based on the impacts on social media on adolescent on mental health. Basically , this article's review is to help instructors understand how social media affects adolescents both socially and mentally in their classes. Teachers must grasp both social media and its potential mental health consequences, since studies show that over 90 percentage of kids who have been cyber bullied do not tell their teachers or parents because they have a fear that their parents and teachers would not understand them or won't do anything . This study's literature evaluation examines the advantages and disadvantages of teen social media use outside and inside the classroom. Recent research has found that how a teenager uses social media, rather than how they use social media in general, has an impact on their mental health. This literature review of the study highlighted on giving information for the teachers so that they can make them aware of the issues that social media brings to the classroom due to teachers' limited training on the effects of social media on mental health. Also they explained the reasons why adolescents utilize social media, what kind of way they use to form and to maintain virtual connections vs offline interactions, and an explanation of cyber bullying will all be discussed. Teachers will hear opposing points of view in order to understand the benefits as well as the risks of social media. Teachers will also be given an overview of accessible resources to use in their classes. According to the analysis of data, it's shown that due to spending too much time on social media, it causes a great impact on an adolescent's mental health. And there was a finding that there needs to be more qualitative data in future regarding adolescents' daily activities on social media and also, they should be asked to track offline activities in addition to regular social media use.

Excessive use of social media is seen to be detrimental to one's mental health. In the research paper [17] explored the effects of social media on mental health through some previous research papers. In addition, there is no specific methodology or dataset in this study but there are discussions on different articles regarding the topic. In this study, it's shown that according to a national poll of young adults in the United States (US) there found that people who use 7 to 11 social media platforms there had much greater chances of suffering depression and anxiety symptoms but the people who use 0 to 2 social media platforms, they have less chances of suffering these. Social media has become a life schedule. Social media has been connected to anxiety and obsessive behavior in several studies. When they couldn't access their email or social media accounts, 45 percent of British adults felt frightened or uncomfortable. Another study of hundred Facebook users of a university examined that the people who use extreme social media have low selfesteem. Also, between 2010 and 2016, they published a meta-analysis of 62 studies and according to that meta-analysis they found that trait narcissism is linked to social media usage. The connection between social media use, self-harm, and possibly suicide is concerning. It's concerning that adolescents can access obnoxious online content that encourages self-harm and suicide. This material aims to "normalize" self-harm and suicide, which may promote to adolescents imitating the behaviours they observe. Furthermore, cyber bullying through online or social networking sites has detrimental mental health implications. For the victim, this can be exceedingly humiliating, resulting in a loss of self-esteem and confidence. Depression, anxiety, insomnia, self-harm, and feelings of loneliness are all possible symptoms. Gonzales and Hancock, then again, exhibited the positive advantages of Facebook on confidence, reinforcing the "Hyperpersonal Model," which expresses that particular self-show helps confidence. As per the objective mindfulness hypothesis, any experience that makes oneself become the item (as opposed to the subject) of mindfulness which will prompt a decreased impression of oneself. If utilized responsibly, social media may be quite beneficial. Government, business, trade, education, and information technology have all been invaded by social media. Social media has infiltrated government, business, trade, education, and IT. The deleterious impacts of social media on young people could have broad consequences. This subject necessitates continuing worldwide exploration to not only detect negative consequences, but also to look into prevention and therapy solutions. Technology is always here for us. So, all of us should figure out how to cohabit and optimize its utilization. Social media has infiltrated government, business, trade, education, and IT. The deleterious impacts of social media on young people could have broad consequences. This subject necessitates continuing worldwide exploration to not only detect negative consequences, but also to look into prevention and therapy solutions. Technology is always here for us. So, all of us should figure out how to cohabit and optimize its utilization.

This research work [19] presented the relationship between social media and self-esteem and mental health was extensively discussed. Songs lyrics and videos were used to show how social media is promoting eating disorders , identity building and manipulations , and cyber bullying. The effects of social media on self-identity, behavioral and psychological maladjustment , and behavioral and psychological maladjustment were used in this study to counteract the traits of physiological distress and negative emotions caused by social media . It also covers topics such as the impact of social media on children's mental health and so on. This study shows that, despite the fact that online social environments can benefit and support people in many ways, they can also be harmful.

The research work [18] also suggests using PRISMA criteria to learn more about the impact social media has on our mental health. This study shows that social networking sites can be more impactful on mental health. Social media has a massive connection with anxiety and depression. Quantitative examinations were included in the Study. Various databases and different search strategies were used to audit the articles. Utilizing PICO we did an efficient audit and after checking for clearness of the contents they were used for the survey. The determination of checked articles was finished utilizing the PRISMA Flow chart. The pertinent articles were considered and the information was separated. The extraction of the information was finished by one researcher and one more analyst looked at the examinations to check whether they were predictable with the point and consideration measures. Cochrane's rules were utilized and only the quality appraisal of the chosen studies was chosen. From the data set, an underlying 884 investigations were recognized which include the utilization of virtual entertainment. After applying the incorporation standards and the PRISMA rules the absolute number of studies that remained for the study was 16.

In the study [20] Descriptive research design was considered. Primary data was collected from 90 MBA students through a survey form and secondary data was collected from various websites, magazines and journals. A structured set of questions distributed to the participants which collected information from demographic questions in the first section while the other section asked Fred's choice questions. Which included topics such as Patterns and time spent on the website on their social media usage , their daily activity and engagements with other people on the websites , they also included health related questionnaires such as if they faced any eating or sleeping disorder and also if they felt stressed and anxiety. A reliability test was done for this research which gave us a score of 0.713 which means it's reliable. And the study confirmed that there was a relationship between feeling anxious and being active on social networking sites which further affirms that more usage of social media is affecting the mental health of the students. The research paper [21] proposes using Self-presentation theory by Goffman, which implies that people may choose how they show themselves in a range of situations. Furthermore, Communicating using computers provides advantages over real life communication in terms of allowing users to present themselves better, according to the "hyperpersonal model" of behavior by Gonzales and Hancock. As a result, one could argue that the digital self's performance is a meticulously choreographed self presentation. The fact that some social media platforms are termed "Facebook, YouTube, or MySpace" rather than "WeBook, OurTube, or Our Space," according to Fuchs , demonstrates their self-centered orientation. According to Marwick , marketing and advertising methods have been used to relationships and social behavior through social media. It presented the relationship among social media, social comparison and mental health. This study shows that social media can be more addictive than cigarettes or alcohol according to the Royal society for public health. Social media has a massive connection with depression and anxiety in 19-32 year olds, said Primaketal .This study shows people have 338 online friends on average which used to be 10-12 friends. Moreover, according to Gonzales and Hancock people can optimize their self- presentation on social media, hence it has advantages over face to face communication. However, nursing practice is required for the people who may be negatively affected by social media.

Every knowledge synthesis, including this scoping review, has limitations that must be stated. During the course of this literature review, it was revealed that in the settings of "young generation" and "mental health," variable terms are commonly employed equally. While the most popular versions were utilized, it's possible that some relevant content was overlooked, which could be a suggestion for the further research in this topic.

V. Data Collection

For the proposed model one of the biggest challenges was to Obtain the input data. When obtaining a set of input data confirming the validity of the questionnaire was important. As the Data had to accurately Show individuals Social media usage pattern along with an accurate mental health data for training the model. For that reason when it comes to primary data collection it was very hard to obtain ,but on the other hand there were some datasets available online which satisfied the needs for the proposed model. The dataset was collected from participants across Bangladesh aged between 15-40 years by researchers from University of Asia Pacific Dhaka. And the dataset consists of 793 participants.

A. Recruitment and Values :

The dataset shows replies from 793 adults to a self-made questionnaire with five sections that was distributed between February 4, 2021, and March 18, 2021 using the Google survey tool (Google form). The responses were gathered in the dataset in order to construct a paradigm for the association between the usage of social networking sites and three different types of psychological distress, including depression, anxiety, and loneliness. This data set consists of 669 users and 122 non-users between the ages of 15 and 40. The study asked questions about socio-demographic information, social networking sites usage patterns, and an evaluation of mental health problems. Numbers of their socio-economic standing, respondents were recruited from around the country. Therefore, this information can be used by government organizations and medical professionals to address the mental health problems associated with social networking sites use.

There are some outcomes that displays the results of the dataset initially, which are as follows :

- The data can help researchers pinpoint the reasons behind the poor mental health of young social networking site users from Bangladesh.
- This study which is supported by data can be utilized to develop prevention strategies for mental health issues like depression, anxiety, and loneliness.

- Responses from younger people who fit within the intended age range are included in this data collection (15–40). The data set shows how the use of social networking sites has harmed the mental health of Bangladesh's young population.

B. Dataset Description :

Due to social media, new ways of connecting and communicating with people have emerged. There is a clear time-killing component in it. A vital question for us is whether anything in our daily lives is positive or negative. It seems logical, in my opinion, to examine how it correlates to our mental health. This cross-sectional study was designed, created, and executed with a specific goal in mind. which is to exactly find that out, wheather social media could be affecting our meantal health or not. The survey questionnaire was created using the UCLA Loneliness Scale-8 (UCLA-8), Patient Health Questionnaire-9 (PHQ-9) and the 7-item Generalized Anxiety Disorder (GAD-7) Scale to investigate three mental illnesses. It was challenging to determine whether use of Social networking sites had an impact on the mental health of the young Bangladeshi population based solely on the current dataset. Due to the continuous COVID-19 lockdowns, more people than ever were actively using social media. Therefore, regular Social networking site usage throughout the COVID-19 era may have resulted in more mental health issues.

C. Ethics Statement :

This study protocol (No. UAP/Pharm/2021/01004) was approved by the Department of Pharmacy, University of Asia Pacific's Committee for Advanced Studies. All study participants provided electronic informed consent, which was obtained. Legal guardians of underage participants in the study were also asked for their informed consent.

VI. METHODOLOGY

A. Workflow

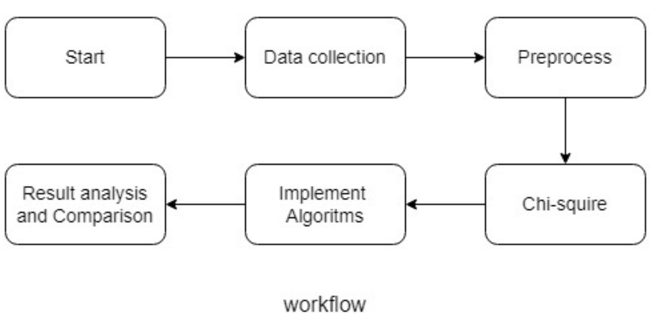
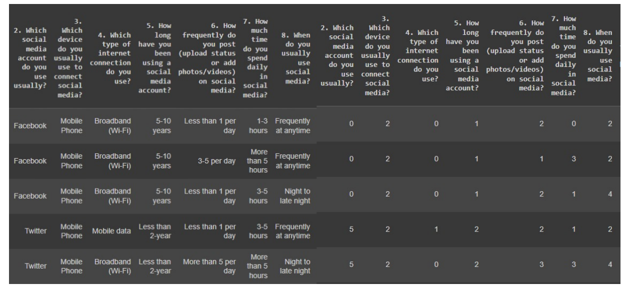


Fig. 1. Workflow

B. Data Pre-Processing :

Processing the data manually for 793 datasets consisting of 32 columns with categorical data is very hard. So Scikit-learn, a free machine learning library for python was used to process the data. From the library,the LabelEncoder class was used to encode categorical data into integer values for training purposes. We can see the dataset before and after pre-processing in figure [2].

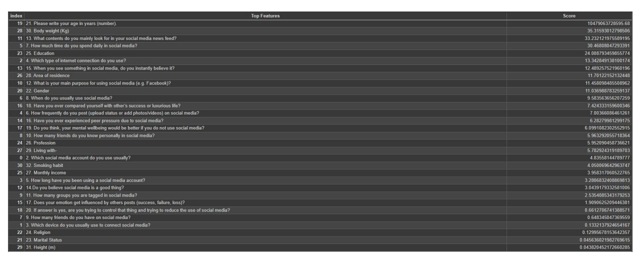


2. Which social media account do you use usually?	3. Which device do you usually use to connect social media?	4. Which type of internet connection do you use?	5. How long have you been using a social media account?	6. How frequently do you post (upload status or add photos/videos) on social media?	7. How much time do you spend daily in social media?	8. When do you usually use social media?	2. Which social media account do you use usually?	3. Which device do you usually use to connect social media?	4. Which type of internet connection do you use?	5. How long have you been using a social media account?	6. How frequently do you post (upload status or add photos/videos) on social media?	7. How much time do you spend daily in social media?	8. When do you usually use social media?
Facebook	Mobile Phone	Broadband (Wi-Fi)	5-10 years	Less than 1 per day	1-3 hours	Frequently at anytime	0	2	0	1	2	0	2
Facebook	Mobile Phone	Broadband (Wi-Fi)	5-10 years	3-5 per day	More than 5 hours	Frequently at anytime	0	2	0	1	1	3	2
Facebook	Mobile Phone	Broadband (Wi-Fi)	5-10 years	Less than 1 per day	3-5 hours	Night to late night	0	2	0	1	2	1	4
Twitter	Mobile Phone	Mobile data	Less than 2-year	Less than 1 per day	3-5 hours	Frequently at anytime	5	2	1	2	2	1	2
Twitter	Mobile Phone	Broadband (Wi-Fi)	Less than 2-year	More than 5 per day	More than 5 hours	Night to late night	5	2	0	2	3	3	4

Fig. 2. Data before and after pre-processing

C. Chi-Square

For algorithm implementation, at first the Chi-Square test was carried out. Utilizing this function eliminates the characteristics that are most likely to be unrelated to class and measure dependence between random variables rather than their independence from class. Using this method ensures we don't have any irrelevant data that doesn't have much effect on the target column of the proposed algorithms. To calculate chi square values once again the Scikit-learn library was used. Values greater than 1 were used to minimize overfitting in the dataset.Chi-Square values were calculated for all 3 of our mental health conditions separately. To minimize overfitting for all three cases. Here from the chi square table the less important factors were dropped from the table and used and then used in the algorithms.



index	Top Features	Score
19	21. Please write your age in years (number).	10479063728595.68
28	30. Body weight (Kg)	35.31593012798506
11	13. What contents do you mainly look for in your social media news feed?	33.232121975589195
5	7. How much time do you spend daily in social media?	30.46808047293391
23	25. Education	24.008793459855774
2	4. Which type of internet connection do you use?	13.342849138100174
13	15. When you see something in social media, do you instantly believe it?	12.489257521960196
26	28. Area of residence	11.70122152132448
10	12. What is your main purpose for using social media (e.g. Facebook)?	11.458098485508962
20	22. Gender	11.036988783259137
6	8. When do you usually use social media?	9.583563656207259
16	18. Have you ever compared yourself with other's success or luxurious life?	7.424333159600346
4	6. How frequently do you post (upload status or add photos/videos) on social media?	7.0036608646261
14	16. Have you ever experienced peer pressure due to social media?	6.28279981299175
17	19. Do you think, your mental wellbeing would be better if you do not use social media?	6.0991082302552915
8	10. How many friends do you know personally in social media?	5.963292055718364
24	26. Profession	5.952090458736621
27	29. Living with-	5.782924319189703
0	2. Which social media account do you use usually?	4.83558144789777
30	32. Smoking habit	4.050069642963747
25	27. Monthly income	3.958317060522765
3	5. How long have you been using a social media account?	3.2806832408869813
12	14.Do you believe social media is a good thing?	3.0439179332581006
9	11. How many groups you are tagged in social media?	2.5354085343179253
15	17. Does your emotion get influenced by others posts (success, failure, loss)?	1.9090625209446381
18	20. If answer is yes, are you trying to control that thing and trying to reduce the use of social media?	0.6612706741388571
7	9. How many friends do you have on social media?	0.648345047369559
1	3. Which device do you usually use to connect social media?	0.1332137924654167
22	24. Religion	0.12995678153642357
21	23. Marital Status	0.045636021982769615
29	31. Height (m)	0.043820452172660285

Fig. 3. Chi-Squire Score

D. Training and testing

Our study aims to determine a persons mental state, which means whether or not a person is more likely to experience sadness, anxiety, or loneliness based on the frequency and duration of their social media use. Our data can have multiple outcomes in each of the three cases of mental illness. Therefore, based on 32 questions about social media usage patterns, which were thought to be some of the significant domains that might affect people's mental health, we utilized seven algorithms to predict depression. The relevant fields were collected and used from the research conducted by university of asia pacific as according to the researchers this information

was intended to be used to address the mental health problems associated with SNS use.We therefore ran the algorithms based on those reasons to determine the accuracy, precision, recall, and f-measure of our predictive model.

We use these two factors - accuracy and f-measure, to determine which algorithms perform best for our system since the higher the accuracy and f-measure, the better the system will be. Precision — defined as the total number of true positives across all predictions of depressed states, is a crucial consideration. This means that the people from predicted depressed results are actually depressed. which means if we get a higher precision we can at least get a higher correct prediction of actually depressed people. Recall is another important part of the prediction system, The recall is determined as the proportion of Positive samples that were correctly identified as Positive compared to all Positive samples. Finally, using precision and recall, f-measure is calculated.

E. Result Analysis and comparison

Algorithms	Accuracy	f-measure
Decision tree	34%	30%
Random forest	36%	27%
SVM	38%	28%
ADA boost	28%	25%
KNN	33%	19%
Naive bayes	24%	14%
Voting Classifier	35%	26%

TABLE I
ALGORITHM AND CLASSIFICATION REPORT FOR DEPRESSION

Algorithms	Accuracy	f-measure
Decision tree	57%	35%
Random forest	65%	29%
SVM	65%	44%
ADA boost	60%	32%
KNN	65%	26%
Naive bayes	21%	12%
Voting Classifier	68%	30%

TABLE II
ALGORITHM AND CLASSIFICATION REPORT FOR LONELINESS

Algorithms	Accuracy	f-measure
Decision tree	42%	32%
Random forest	49%	38%
SVM	48%	42%
ADA boost	49%	41%
KNN	42%	31%
Naive bayes	17%	14%
Voting Classifier	53%	31%

TABLE III
ALGORITHM AND CLASSIFICATION REPORT FOR ANXIETY

As our research includes depression , anxiety and loneliness we used all 7 algorithms for all 3 mental conditions separately. And we already discussed how important Accuracy and f measure is to determine how good the system is. From figure I , II and III we can see a clear picture of all the algorithms and figure out which algorithms are best suited for us in each case. A system with highest accuracy and highest f-measure points are the ones that are good. So in case of depression we can see that Decision tree, random forest and SVM gives us the best result with Accuracy of 34% , 36% and 38% respectively And their respective f-measure scores were 30%, 27% , 28%. Lastly we used a Voting classifier in an attempt to get better results. which gave us an output of 35% accuracy and 26% f-measure score which is not much better than the other algorithms we used. So, in case of depression we can see that SVM gives us the best result. In case of loneliness we can see that random forest, SVM and KNN gives us the best result with all three of them having 65% accuracy respectively And their respective f-measure scores were 29%, 44% and 26%. However upon using voting classifier we got an accuracy of 68% and f-measure score of 30%.

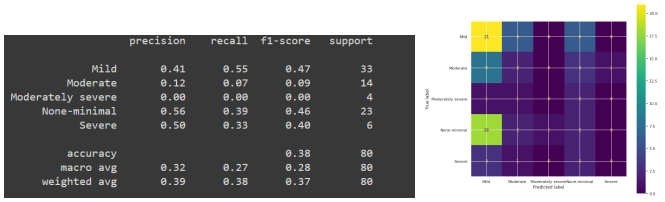


Fig. 4. SVM classification report for depression

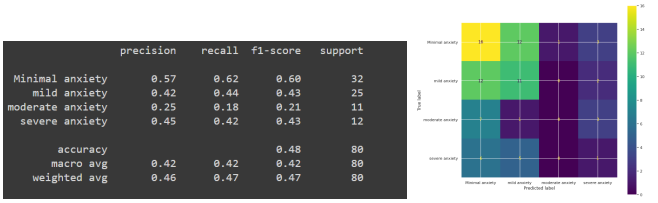


Fig. 5. SVM classification report for anxiety

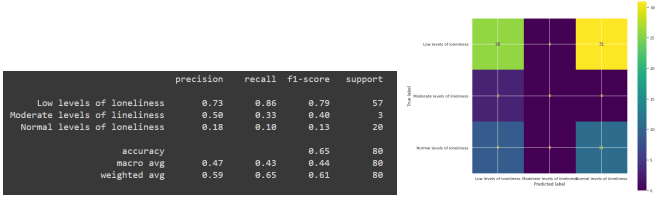


Fig. 6. SVM classification report for loneliness

At first glance we might think that the voting classifier is the best algorithm to choose for loneliness. However SVM had the best f-measure score of 44% which is also an important factor. So considering both accuracy and f-measure , SVM was the best model to use for loneliness. Then, in case of Anxiety we can see that random forest, SVM and ADA boost gives us the best result with an accuracy of 49% , 48% and 49% respectively And their f-measure scores were 38%, 42% and 41% respectively. Even though We got slightly better accuracy at 53% upon using the voting classifier. However in case of anxiety, SVM and ADABoost are still better algorithms to use because of their significantly higher f-measure scores.

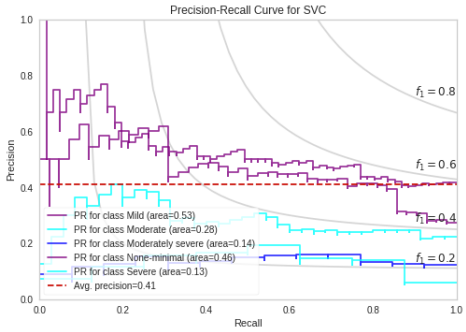


Fig. 7. Precision recall curve of SVM for depression

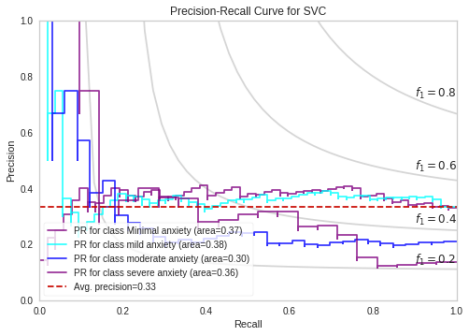


Fig. 8. Precision recall curve of SVM for Anxiety

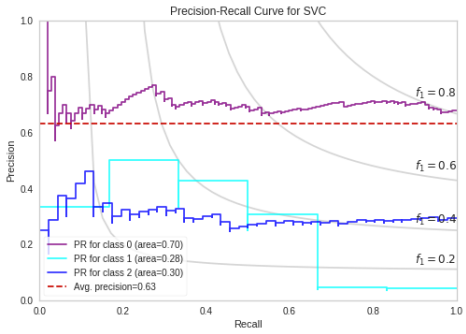


Fig. 9. Precision recall curve of SVM for Loneliness

However, from the figures we can notice that we have very poor accuracy precision and recall score . And from the classification report we can notice some of the cases have 0 accuracy and f1 score which indicates that our dataset might be imbalanced. Imbalanced dataset means our dataset might have one class which makes up a large proportion of the dataset. Which can lead us to a bad result like in our case. We tried to use RandomizedSearchCV, a hyperparameter tuning method to try and combat this issue however using RandomizedSearchCV seemed to further worsen our results. So we decided to not use it. Synthetic Class samples creating techniques such as SMOTE was also considered to try and balance the data distribution but as our work is on a very sensitive topic such as mental health, a synthetic data would means we can not actually measure mental health conditions reliably even though it might have given us a higher accuracy score in benchmarks. However an imbalanced data might not be the only problem here. As we only have data from 793 people, such a small number of data might have also contributed to us having a lower accuracy score.

Lastly, by comparing the above mentioned data and figures, we can conclude that SVM consistently has one of the greatest accuracy and f-measure scores across all three systems even with our limited dataset, and is therefore delivering the best results for our system across all three mental health situations from our dataset.

VII. Future work

There is still a lot of space for improvement in our research. In future we need to collect a lot more data for a much better accuracy score and f-measure score which we couldn't achieve this time because of the limited data sets that we have. Our work in the future will be to design a more accurate algorithm and use it to make a website or an app in which users can get a checkup of their mental health situation and get an accurate result on their mental health.

VIII. Conclusion

The purpose of the study is to look into how social media affects problems with mental health. The relationship between social media use and bad mental health determines the topic's significance. Because mental health concerns can impact anyone at any time in their life, social media use has harmful effects on all generations. According to the study, the majority of respondents use a variety of social networking sites.Because of the excessive usage of these sites so much throughout the day, they frequently experience mental health problems. Our study concentrated on three significant mental health conditions like anxiety, depression, and loneliness among all these mental health difficulties.To accurately measure these three key mental health issues—anxiety, sadness, and loneliness—we used seven different models. The most significant result of this investigation is that the SVM algorithm gives us the best result among all algorithms that have been used in this research.

Finally,we can't deny that social media has many advantages. Government, business, trade, education, and information technology are just a few of the sectors where social media is prevalent and has a foothold.One of the most convenient places to access information and stay up to date on everything around and far away is here. All of the benefits and drawbacks that exist in this world should be constantly ingrained in the minds of the people in our nation. So, in order to prevent being impacted by social media's negative aspects, we must be aware of how we use social media platforms for ourselves effectively so that mental health can't be affected.

REFERENCES

[1] Julia , B Elke, R. (2019). Relationships Between Addictive Facebook Use, Depressiveness, Insomnia, and Positive Mental Health in an Inpatient Sample: A German Longitudinal Study.

[2] Kalpana, S. , Suprakash, C. , Joyti, P. Sana, D. (2019) Social media and mental health challenges.

[3] Royal Society for Public Health (RSPH) Young Health Movement. (2017). STATUSOFMIND- Social media and young people's mental health and wellbeing. Royal Society for Public Health, May, 32.

[4] McBride, D. L. (2011). Risks and benefits of social media for children and adolescents. Journal of Pediatric Nursing, 26(5),498–499.

[5] Warrender, D. and Milne, R.(2020).Social media, social comparison and mental health.

[6] Safayet, H. (2021) Nature and Aftermath of Cyberbullying with Female University Students in Bangladesh. DOI: 10.9790/0837- 2610074553

[7] Deepa, R. V Krishna, P. (2020) Impact of social media on mental health of students.

[8] Osman, U. , Ajlina, k. , Mahdi, B. , Ateka, B. , Jacob, S. Ubydul, H.(2022) Social Media Use and Mental Health: A Global Analysis

[9] Wenhong, C. Kye-Hyong Lee (2013). Sharing, Liking, Commenting, and Distressed? The Pathway Between Facebook Interaction and Psychological Distress.

[10] Luca, B. , Ro'ee, L. Alexey M (2021). Social media and mental health.

[11] Davila, J., Hershenberg, R., Feinstein, B. A., Gorman, K., Bhatia, V., Starr, L. R. (2012). Frequency and quality of social networking among young adults: Associations with depressive symptoms, rumi-nation, and co-rumination. Psychology of Popular Media Culture, 1(2), 72–86

[12] Grossman, M. (2017). Study of social media users: The relationship between online deception, machiavellian personality, self-esteem, social desirability. Retrieved from ProQuest Dissertations The-ses Global. (1946736580)

[13] Vogel, E.A., Rose, J.P., Roberts, L.R., Eckles, K. (2014). Social Comparison, Social Media, and Self-Esteem. Psychology of Popular Media Culture, 3(4), 206-222. doi:10.1037/ppm0000047

[14] Keigler Isabeau, J(2020). Social Media and Its Effect on SelfEsteem and Mental Health. http://dx.doi.org/10.26153/tsw/8588

[15] Hassan, E. S. Farin, T. (2020) Investigating the role of social media on mental health. DOI 10.1108/MHSI-06-2020-0039

[16] Nicole, J. C. (2020) The Impact of Social Media on Adolescent Mental Health.

[17] Kalpana, S. , Suprakash, C. , Jyoti, P. Sana, D. (2020). Social media and mental health challenges.

[18] Chantal, S. (2020) Examining the Impact of Social Media on Youth Self-Perceived mental health.

[19] Manoj, K. S. , Nisha J. Maya, S. (2020) Influence of social media on mental health: a systematic review. DOI: 10.1097/YCO.0000000000000631

[20] Deepa, M. Dr.V.Krishna, P. (2020) Impact of social media on mental health of students.

[21] WARRENDER, D. and MILNE, R. (2020) Social media, social comparison and mental

[22] Hilal, B Shabir, B. (2017) Effects of Social Media on Mental Health: A Review

[23] Islam, M. R., Tushar, M. I., Jannath, S., Moona, A. A., Akter, S., Islam, S. M. A. (2021). Data set concerning the use of social networking sites and mental health problems among the young generation in Bangladesh. Data in Brief, 39, 107593. https://doi.org/10.1016/j.dib.2021.107593