Interaction Between Social Media Habits & Mental Health of Students

Raghav Vaidya, Aashima Malhotra, Sarvjeet Herald

Abstract— Social media's effects on users' mental health have been a topic of discussion since the medium's advent. Although both the benefits and harms of social media on mental health have been debated extensively, it is generally agreed upon that the younger generation is far more susceptible than the rest, given their higher impressionability. The emergence of COVID-19 saw increased utilization of social media platforms, as people practiced social distancing and self-isolation and generally had more time to themselves. Students developed social media habits that continue to this day and affect their lifestyles. In many instances, they showcased symptoms of depression, anxiety, decreased/delayed sleep, and other related disorders, which are disproportionately underreported, especially in countries such as India, where mental health is given less importance. Hence it becomes imperative to analyze the extent of side effects of social media usage on students to mitigate the risks among students in the era of technology. Borrowed from standardized scales, we create a new survey questionnaire to understand depressive traits and social media habits among students aged 13-24. We identified that primarily, negative thoughts and social media usage during the morning hours leads to depression. Our Machine Learning model implemented for predictive analysis was 91.89% accurate. Further research is required to understand the potential association between social media use and positive outcomes among students.

Index Terms— Social Media, Mental Health, Data Science, Predictive Analysis

1 Introduction

The importance of mental health has been ignored by all societies across the world, regardless of their economic standing. According to the 2020 National Survey on Drug Use and Health, mental disorders affected more than 20 percent of American adults [1]. Suicide alone was responsible for 44,834 deaths in the United States in 2020, making it the 11th leading cause of death [2]. Even in a developed country like the US, stigma continues to play an inimical part in affecting patient perception and in a few instances creates moral injury, either initiating or aggravating mental health illness [3]. In 2022, WHO published the "World Mental Health Report: Transforming Mental Health for All" formally acknowledging the escalation of mental health disorders post the pandemic [4]. The report also highlights that one in eight people suffer from mental disorders, which is highly concerning.

Countries like India have only begun to acknowledge the problem, which till very recently was considered taboo and meant to be kept under wraps. The World Health Organization (WHO) has constantly cautioned India of its severely understaffed mental health workforce. "The National Mental Health Survey 2016" found that nearly 15 percent of India's adult population required active mental health interventions (which amounts to more than the population of Mexico) [5]. The fact that the number of suicides has increased by nearly 26% from around 130,00 in 2017 to more than 160,000 in 2021 is an alarming pointer to the severity of the problem [6]. Unfortunately, it is always assumed that if a person lives a healthy and productive lifestyle, they are unlikely to suffer from mental illnesses, unless from trauma. Lack of awareness and sensitivity to mental health issues compounds the problem manifold.

The "Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic" study, published in Lancet public health, has estimated an increase of about 35 percent in the prevalence of anxiety and depression in India during the COVID-19 pandemic forcing the Government of India to take formal cognizance of the issue [7].

Worryingly, this spike in statistics was noted more in the relatively younger age group including teenagers and juveniles. Severe restrictions on movements and social interaction on account of the severe Lockdown imposed by the government for a prolonged period and with offices and schools shutting out people, most of this age group was left isolated devoid of their regular social framework making them more susceptible to anxiety and depression. With more time on their hands, much of this age group turned to social media to fill the void in their lives.

The transition to work-from-home and study-from-home further minimized the need for physical socialization. Additionally, the practices of social distancing, self-isolation, and quarantine further helped establish the new norm, in which people subsequently felt more alone than ever [8]. Many in this age group have not been able to give up their dependence/addiction to social media despite returning to normalcy.

The damaging impact of social media is particularly noticeable in the youth as a young person is more likely to have greater social media accounts than an older person [9]. One reason for this is that more popular social media networks such as Instagram, and Snapchat are geared towards serving teens and young adults [10], as evidenced by their advertising campaigns and minimal user interface. Not only does the younger generation form the larger part of the user base, but they are also highly impressionable [11]. Recent studies have inextricably linked this increase in social media usage to the increase in mental health disorders. Increased distraction

• *Raghav Vaidya conducted this research under the supervision of Aashima Malhotra and Dr. Sarvjeet Herald at the RoboGenius. India*

from academics [12], disrupted sleep [13], exposure to online bullying, harassment [14], rumor spreading, peer pressure, and the "fear of missing out", abbreviated popularly as FOMO [15] are being noticed in varying degrees in most of this age group. While most psychologists acknowledge the problem and its severity, very few formal research studies have been conducted to study the detrimental effects of social media on the mental health of the younger population across the world.

Especially under-examined are the effects of social media on the youth of India, who have one of the highest rates of suicide among people aged 15 and 24 [16]. This pattern continues to grow unabated post the pandemic. It is important to note, that given that India is home to the largest youngest population in the world, the statistics can be alarming in terms of the size of this fatality. There are some social pressures too that are making students use social media more than usual. Given the sheer numbers, Indians face higher levels of competition to make it to better universities and more lucrative jobs, as compared to their counterparts in the more developed or less populated nations. The pressure to excel in studies is primordial in almost all modern Indian households, and education is viewed as the most important aspect of the life of a student. This parental insistence to succeed and the complete repugnance to failure adds to the existing social pressure pushing students towards an external outlet where they are not likely to be judged or where they can release their stress by being anonymous. They find solace in social media networks as it offers a social network without societal pressures and supervision.

India can no longer afford to ignore the magnitude of the side effects of social media and its impact on the mental health of its young populace. The main aim of this research is to establish a definitive association between social media usage and mental health decline among the youth. We plan on correlating the social media usage trends of the students with their mental health profiles, to analyze the nature of this correlation. The research attempts to identify the age groups that are more susceptible to the negative effects of social media. Additionally, it attempts to identify the nature of the problem and side effects vis a vis the specific social media and technology they engage. It concludes by proposing measures to restrict this harm without impinging on the personal freedom of this age group.

In a structural overview, we initially studied the literature germane to the aspects of mental health and how they are treated in the broader spectrum of healthcare. After recognizing the common factors affecting mental health, we chose to focus on social media which has seen a sharp increase in its usage particularly in the younger population over the years and even more so post-COVID. We chose to use students as our focal group, as they are the largest consumers of social media. Subsequently, we decided to conduct a survey using various standardized scales (to identify depressive traits and social media habits) modified to fit our specificity, to get a general notion of the student's mental well-being and their correlation with social media consumption. Finally, a detailed analysis was conducted to identify patterns and describe prescriptive steps to allay the negative effects of social media.

2 Existing Works

Good mental health can be defined as a state of well-being that allows individuals to cope with the normal stresses of life and function productively. Life stresses can come in many forms, for example, being unhappy about something, being overburdened, having excess responsibilities, having prolonged hours of activity, having poor prioritization skills, having unclear decision-making, experiencing high-risk and life-threatening situations, and facing discrimination or harassment. Although people are expected to overcome such life challenges and still function productively, many are unable to do so in present times. The mental disorders include but are not limited to anxiety, depression, bipolar disorder, post-traumatic stress disorder, schizophrenia, eating disorder, disruptive behavior, dissocial disorder, and neurodevelopmental disorder.

The approach of using data mining and data science techniques in psychiatry has opened a new area of research in the detection, diagnosis, and classification of psychiatric disorders. Before we employ our analysis, it is important to establish how far the field of health analytics has come and how the modern healthcare industry is adapting to and utilizing the vast amounts of ever-increasing data being collected. The comprehensive review titled "A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining" describes the use of data mining in healthcare and how it is becoming increasingly popular [17]. It also conveys that a lot of recent modern healthcare analyses, especially in the mental health sector, depend on data accumulated from social media services. Therefore, it follows that social media is an important factor affecting the mental health conditions of the present-day population.

Another review titled "Data Mining Algorithms and Techniques in Mental Health: A Systematic Review" focuses on common mental disorders such as dementia, Alzheimer's, schizophrenia, and depression, and how the past research discussed in it employed various data mining techniques to predict the associated risk factors [18]. The review concluded that the use of data mining techniques and analytics applied to diseases such as depression, schizophrenia, etc. can be of considerable aid to clinical decisions, diagnosis predictions, and improvement of patients' quality of life. On top of that, another review called "Machine learning in mental health: A systematic scoping review of methods and applications" (aimed to synthesize the literature on machine learning and big data analytics for mental health) concluded that the application of systematic data processing techniques and machine learning show a range of benefits across the areas of diagnosis, treatment, support, research and clinical administration of mental health [19]. Thus, it is evident that the application of data analytics in the mental health sector can yield a lot of useful insights.

Although mental illness seems to be uniformly ubiquitous, the fact remains that it affects separate demographics differently. Gender seemed like a practical separator to identify and differentiate the groups at higher risk. One study titled "Gender and Mental Health" establishes that men and women do suffer from different kinds of mental health problems [20]. It concluded that while depression and anxiety were more common in women, men exhibit higher rates of substance abuse and antisocial behavior. However, as the gender gap blurs globally, it can be assumed that the gap in types of mental health issues faced by men and women is also vanishing. Another separator, age, is also a major factor that extensively plays a role in mental health upkeep. A major study conducted in 2012 at the University of Pittsburgh revealed that the circumstances and events in the surrounding environment have a stronger hold on an adolescent's brain, as compared to an adult's brain. This means that a teen is more susceptible to all kinds of emotions, be it happiness, stress, or depression. Stress can also be an important attribute of depression and a study conducted in 2011 at the University of California, Los Angeles aimed to distinguish between the stress perceived by adolescents from that of adults. It suggests that stress affects adolescents more than adults, hence consequently adolescents see higher rates of depression.

Exploring the lifestyle choices of adolescents gives further insight into the causes that lead to depression and anxiety in them. One apparent characteristic that most of the modern youth share is their increased usage of social media. Social media has a huge impact on society today and it is being used to openly discuss political, religious, and other contemporary topics. Its increase in usage is also commensurate to increased self-harm and suicide, as discussed in the study titled "Effect of Social Media Use on Mental Health during Lockdown in India" [21]. The study also explores the role of COVID-19 in exacerbating the already substandard mental health condition of Indians. The usage of social media increased substantially during the lockdown placed by the Indian government at the pandemic's onset, further negatively impacting the mental health of people, the repercussions of which are still observable to this day. The prolonged impact on mental health and the subsequent increase in social media usage creates an opportunity to develop a new framework that could guide the development, implementation, and assessment of mental health interventions, especially in adolescents - the most at-risk group.

3 Methodology

Measuring patient-reported outcomes have increased significantly in recent years. There are many methods available for measuring similar mental health traits of individuals. Three popular scales used to correctly identify depressive symptoms in the student participants are discussed in this section. Borrowing from these, we further discuss our approach to the newly designed questionnaire and the motivations behind it. As data collection was imperative to our overall analysis, this probing survey helped in collecting data from 122 student participants about their mental health state and social media habits.

3.1 Beck Depression Inventory (BDI)

The Beck Depression Inventory was derived from clinical observations about the attitudes and symptoms displayed by depressed psychiatric patients including among students [22,23]. The observations were systematically reduced to 21 symptoms and attitudes which could be rated from 0 to 3 in terms of intensity. Importantly, the items were chosen to assess the intensity of depression and were not selected to reflect any developmental theory of depression. The symptoms and attitudes were (1) mood, (2) pessimism, (3) sense of failure, (4) lack of satisfaction, (5) guilt feelings, (6) sense of punishment, (7) self-dislike, (8) self-accusation, (9) suicidal wishes, (10) crying, (11) irritability, (12) social withdrawal, (13) indecisiveness, (14) distortion of body image, (15) work inhibition, (16) sleep disturbance, (17) fatigability, (18) loss of appetite, (19) weight loss, (20) somatic preoccupation, and (21) loss of libido. The final score received indicates the level of depression as per Table 1.

Total Score	Levels of Depression
1-10	These ups and downs are considered normal
11-16	Mild mood disturbance
17-20	Borderline clinical depression
21-30	Moderate depression
31-40	Severe depression over
40	Extreme depression

Table 1: BDI score and corresponding level of depression

3.2 Patient Health Questionnaire (PHQ-9)

The Patient Health Questionnaire is a self-administered variety of the PRIME-MD diagnostic instrument for common mental disorders [24]. The PHQ-9 consists of nine questions and scores each of the nine DSM-IV criteria based on the frequency of symptoms. It has been used as an initial screening tool as well as a follow-up instrument to monitor treatment response. Its sensitivity and specificity were reported as 89% and 77%, respectively, with a positive predictive value (PPV) of 15% and a negative predictive value (NPV) of 99%. The screen generally takes less than 5 min to administer, is available in several languages, and is currently freely accessible online through Pfizer. One obvious criticism of the tool has been a potential conflict of interest, given that increased screening theoretically may lead to false-positive diagnoses for MDD and unnecessary treatment with antidepressants. The PHQ-9 is the depression module, which scores each of the 9 DSM-IV criteria as "0" (not at all) to "3" (nearly every day). The final score received indicates the severity of the depression given in Table 2.

Total Score	Depression Severity
0-4	No or Minimal depression
5-9	Mild depression
10-14	Moderate depression
15-19	Moderately severe depression
20-27	Severe depression

Table 2. PHQ-9 score and corresponding depression severity

3.3 Center for Epidemiologic Studies Depression Scale

The Center for Epidemiological Studies-Depression (CES-D), initially published by Radloff in 1977, is a 20-item measure that asks caregivers to rate how often over the past week they

ISSN 2229-5518

experienced symptoms associated with depression, such as restless sleep, poor appetite, and feeling lonely [25,26]. Responses range from 0 to 3 for each item (0 = Rarely or None of the Time, 1 = Some or Little of the Time, 2 = Moderately or Much of the time, 3 = Most or Almost All the Time). Totals span from 0 to 60, with high scores indicating greater depressive symptoms.

In scoring the CES-D, a value of 0, 1, 2, or 3 is assigned to a response depending on whether the item is worded positively or negatively. For items 1-3, 5-7, 9-11, 13-15, and 17-20 the scoring is:

- Rarely or none of the time (less than one day) = 0
- Some or a little of the time (1-2 days) = 1
- Occasionally or a moderate amount of time (3-4 days) = 2
- Most or all of the time (5-7 days) = 3

Items 4, 8, 12, and 16 are scored in reverse as follows:

- Most or all the time (5-7 days) = 0
- Occasionally or a moderate amount of time (3-4 days) = 1
- Some or a little of the time (1-2 days) = 2
- Rarely or none of the time (less than 1 day) = 3

The total score is calculated by finding the sum of 20 items. Scores range from 0-60. A score equal to or above 16 indicates a person is at risk for clinical depression.

3.4 New Proposed Survey

Using the already established methods discussed above, we borrowed important elements from BDI, PHQ-9, and CES-D. As the use of digital tools is growing among students, we developed questions about social media usage to ascertain the social media habits of the participants and to further develop a regression model to correlate social media habits with depressive traits.

Since students were spending significant time on social media every day, it was important to understand how it affects their mental health and well-being. Hence, we designed a new 24-question questionnaire (including two supplementary questions related to age and gender), to provide insights on how to tackle the various problems faced by the students. The core questionnaire is given in Table 3.

The student respondents were asked their age based on these four age groups: 1) 13-15 years, 2) 16-18 years, 3) 19-21 years, and 4) 22-24 years. The people in the first two age groups were students who were exploring their career options and were aspiring to undertake an undergraduate degree in the future. The people in the third age group were those who were in the process of completing their undergraduate degree. The people in the fourth group were those who had established a need for a post-graduate degree and were working towards attaining it. The survey also asked for the respondent's gender, since the parameters evaluating the state of mental health and well-being of males and females may vary. Variations in ethnicity may have also impacted the results since this survey was only limited to the Indian Nationals studying in India.

S. No.	Questions	Options
1	How much time do you usually spend on social media in a day?	0-30 minutes; 1-2 hours; 2-4 hours; 4-8 hours; 8+ hours
2	What all social media platforms do you regularly use?	Instagram; Snapchat; Twitter; Facebook; Others
3	When do you usually use social media?	Morning; Afternoon; Evening; Night
4	Do you check your social media before going to bed?	Yes; No
5	Do you check your social media after waking up from the bed?	
6	What all do you use social media for?	Keeping in touch with friends and family; Event planning;Buying and selling; Inspiration; News; Dating; Making new Friends; Education
7	Do you feel that any of these conditions affect you and are enhanced by social media usage?	Depression; Anxiety; Insecurity; Peer pressure; FOMO; Anti-social; Self-conscious about life or body; Hatred
8	Do you feel trouble falling asleep or sleeping?	Strongly disagree; Disagree; Somewhat disagree; Neither agree nor disagree; Somewhat agree; Agree; Strongly agree
9	Do you face trouble concentrating on things like school/college work?	
10	Do you often find yourself lost in self-critical thoughts, depressed, irritable or hopeless?	
11	Do problems start to feel bigger when you think about them for long periods of time?	
12	Are you able to relax and sit still for short periods of time?	
13	Do you avoid social events because of overwhelming feelings or fear of how people see you?	
14	Do you feel bad about yourself – or feel that you are a failure, or that you have let yourself or your family down?	
15	Has there been any decline in your self-confidence and self-esteem level recently?	
16	Do you feel confident about where you are in life and in where you will be in the future?	
17	Do you have thoughts that you would be better off dead, or of hurting yourself in some way?	
18	Have you experienced any major changes, loss, or trauma recently?	
19	If you are struggling with something emotions how likely are you to talk to your parents about it?	Very Likely; Likely; Unlikely; Very Unlikely; Not at all
20	Do you find it difficult to pinpoint the source of anxious thoughts?	Strongly disagree; Disagree; Somewhat disagree; Neither agree nor disagree; Somewhat agree; Agree; Strongly agree
21	Has social media affected your relationship with loved ones in a negative way?	Yes; No; Not Sure
22	Do you consider yourself addicted to social media platforms?	Yes; No

Table 3. Survey Questionnaire

The scoring system we adopted was like that of the combination of BDI, PHQ-9, and CES-D. The scoring system is described below based on whether the question is worded positively or negatively: Scoring for questions 8 to 11, 13 to 15, 17, 18, 20:

- Strongly disagree = 0
- Disagree or somewhat disagree = 1
- Neither agree nor disagree =2
- Somewhat agree or agree = 3
- Strongly agree = 4

Scoring for questions 12 and 16, the scores are reversed as follows:

- Strongly disagree = 4
- Disagree or somewhat disagree = 3
- Neither agree nor disagree =2
- Somewhat agree or agree = 1
- Strongly agree = 0

Scoring for questions 19, the scoring is as follows:

- Very likely = 0
- Likely = 1
- Unlikely = 2
- Very unlikely = 3
- Not at all = 4

Total Score	Levels of Depression
1-10	These ups and downs are considered Normal
11-16	Mild mood disturbance
17-20	Borderline clinical depression
21-30	Moderate depression
31-40	Severe depression
40 +	Extreme depression

Table 4. Total score and corresponding level of depression

The total score is calculated by finding the sum of all 20 questions. The total score ranges between 0-60 points. A score equal to or above 16 indicates a person is at risk of developing

ISSN 2229-5518

clinical depression. Final scoring criteria is given in Table 4.

4 Result and Discussion

Data were collected through the questionnaire created using the survey administration software Google Forms. The survey was circulated to school and university students to examine their mental health and social media habits. Through this survey, we were able to obtain responses from 122 students aged between ages 13 and 24 years. Data analysis was done using Google spreadsheets, CODAP (Common Online Data Analysis Platform), and Google Forms Python.

77% of the responses came from postgraduate students; 12.3% of the responses were received from undergraduate students and the remaining responses were from school students aged between 13 and 18 years. The spread of moderate to extreme depression is seen in Fig. 1. All the participating students were of Indian origin and studying in India. 73% of the respondents identified as males and the rest were females.

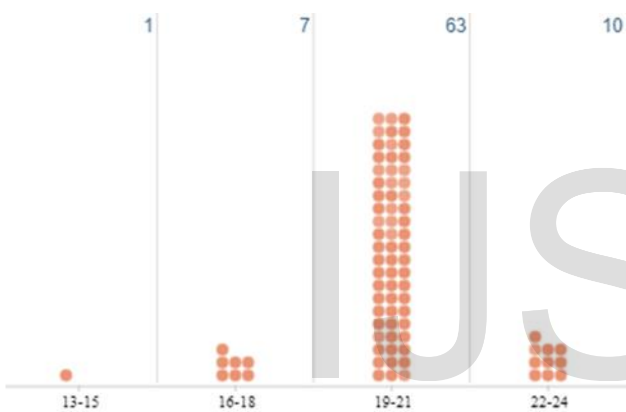

Fig 1. Moderate to extreme depression among age groups

The average student spent 2.4 hours every day on social media. It was also identified that of the students surveyed: around 46% spent between 2-4 hours on social media, around 27% spent 1-2 hours, a little less than 15% spend 4-8 hours daily, a little more than 8% spent 0-30 minutes and just over 4% spent more than 8 hours on social media. Fig. 2 shows the time spent on social media by students. Fig. 3 further shows the spread of major levels of depression, with those using more than 2 hours of social media experiencing the most extreme, moderate and severe forms of depression.

Instagram was the most popular social media network, with nearly 75% of the participants using it regularly. 47.5% were using the second most popular social media app, Snapchat. Facebook and Twitter were utilized by around 27% and 10% of the participants on a regular basis. 48.4% of the respondents also used social media networks apart from the ones mentioned earlier. Fig. 4 shows the usage of various social media networks.

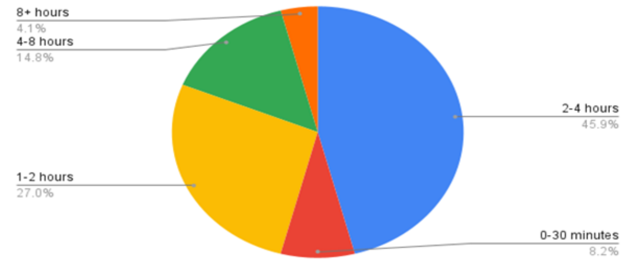

Fig 2. Time spent on social media

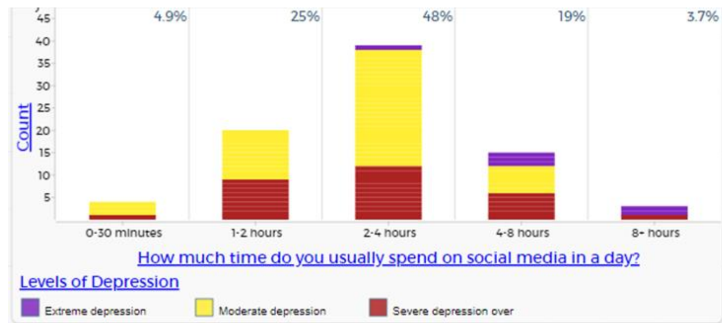

Fig 3. Time spent on Social Media by Students

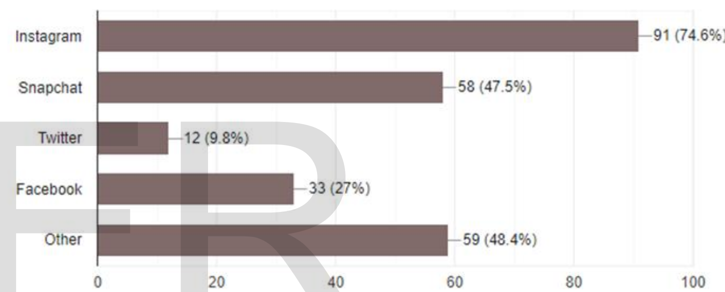

Fig 4. Use of popular social media network

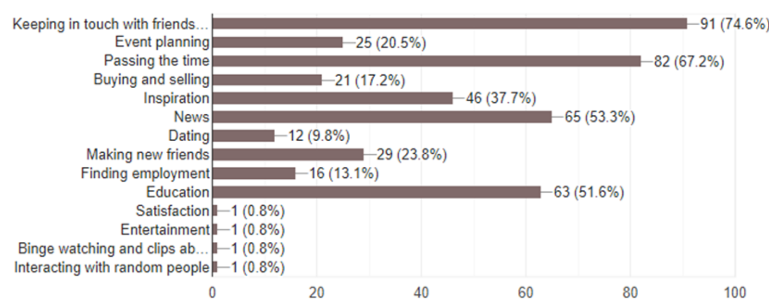

Fig 5. Activities done using social media

91 of the 122 participants used social media mainly for keeping in touch with friends and family, while 82 used it just to pass the time. 65 got their news from social media and 63 used it for educational purposes. The fifth most popular activity on social media was looking for new ideas and inspiration, with 46 participants agreeing to do so. Rest of the usage statistics are given in Fig. 5.

The daily usage of social media by students saw an increasing trend as seen in Fig. 6. While only 27.9% used it in the morning, the usage gradually increased to 35.2% using it in the afternoon, then to 60.7% using it in the evening, and finally to 72.1% using it during nighttime. Also, people who checked their social media apps just before they slept experienced higher levels of depression (Fig. 7). It was also revealed that 81.48 % of students who are suffering from the major three levels of depression (extreme, severe, and moderate) are actively using social media during the night.

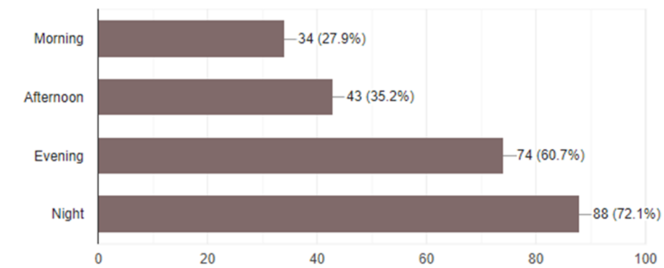

Fig 6. Social media usage during the day

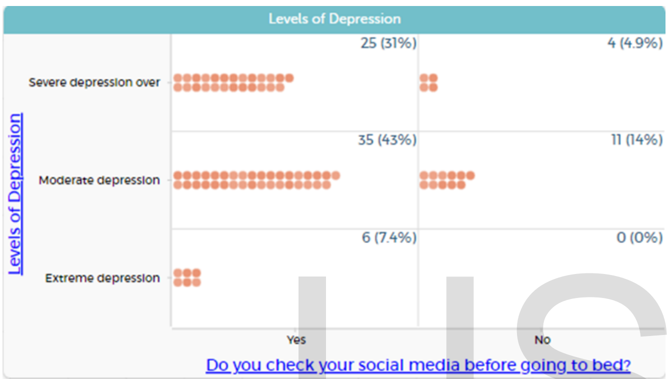

Fig 7. Social media use before sleep & level of depression

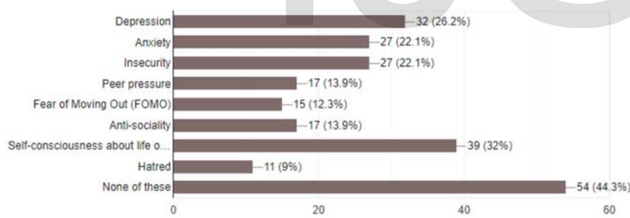

Fig. 8. Perceived effects of social media usage

For the perceived harmful effects of social media, the majority (55.7%) believed that social media is having some negative impacts on their lives with self-consciousness, depression, anxiety, and insecurity being the most reported effects as seen in Fig. 8.

As seen in Fig. 9, 46 participants experienced moderate depression, making it the most common form of depression. Next, 26 suffered from severe depression, while 18 and 10 suffered from mild mood disturbances and borderline clinical depression respectively. Only 13 people had a score that is considered normal, while 6 individuals' responses suggested cases of extreme depression.

The average score of a participating student was 23.98, implying that students who actively use social media likely suffer from a moderate form of depression and may require help. Additionally, 50% of the students participating in the survey also had trouble focusing on their studies or work. Finally, a considerable amount (44.3%) of students perceived themselves as 'social media addicts', as shown in Fig. 10.

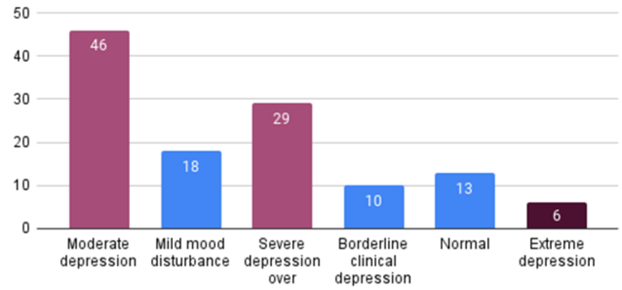

Fig. 9 Spread of forms of depression

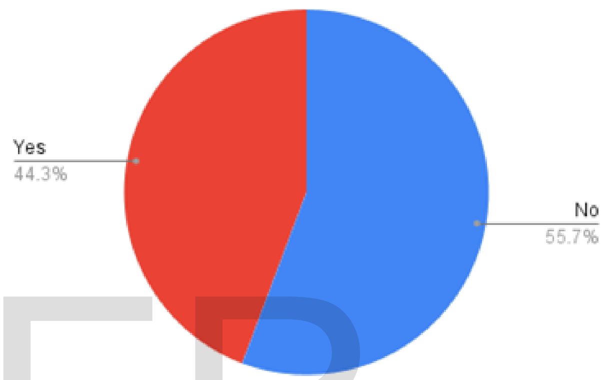

Fig 10. Social Media Addiction

The student responses were converted into a CSV file and then loaded into the Spyder IDE as a Pandas DataFrame to perform data analysis and to build and test the accuracy of the prediction model for depression identification. The text choices were converted into distinct serial numbers for further manipulation. The scikit-learn library was used to create an A.I., a test prediction system, and to produce a classification report.

The SelectKBest (chi2) class from Scikit-learn was used for extracting the best features of the dataset. The SelectKBest method chooses the features according to the k highest score. By modifying the 'score_func' parameter we can apply the function for both classification and regression data. Selecting the best features is an important process while preparing datasets as it helps to eliminate less important parts of the data, reduce training time and improve accuracy score. Table 5 below shows the features in increasing order of their contributing nature to identify whether someone has depression. Although using social media at the night could lead to depression, it is the least contributing factor for calculating the depression score. Instead, negative thoughts whilst using social media and the amount of time spent on social media are the top contributing factors. Hence social media consumption and the chances of developing depression are strongly related. Age difference does not seem to contribute to the score calculation as people of all ages who spent more time on social media, could all generate negative thoughts and hence were likely to develop depression. Gender, though once an important metric for identifying depression, does not seem to play a major role. Both men and women have similar levels of depression, although their experiences may vary.

Post identifying the best contributing features, the dataset was split into the ratio of 70% training data and 30% testing data using the train_test_split method of Scikit-learn. Logistic Regression was used for training, testing, and predicting. Logistic Regression is a Machine Learning algorithm that is used for classification problems: it is a predictive analysis algorithm based on the concept of probability. It assigns observations to a discrete set of classes. Some classic examples of classification problems are identifying spam emails, identifying fraudulent online transactions, and finding whether a tumor is malignant or benign. Logistic regression transforms its output using the logistic sigmoid function to return a probability value.

Features	Score
social_media_night	0.728059
gender	2.16146
social_media_morning	4.69543
social_media_negative	9.82865
social_media_time	10.1503
social_media_addiction	16.5816
total_score	434.853

Table 5. Most contributing features

```
Accuracy:  0.918918918918919
Recall:  0.875
Precision: 0.96875
CL Report:               precision    recall  f1-score   support

           0       1.00      1.00      1.00         1
           1       1.00      1.00      1.00         4
           2       1.00      0.25      0.40         4
           3       0.81      1.00      0.90        13
           4       1.00      1.00      1.00         4
           5       1.00      1.00      1.00        11

   micro avg       0.92      0.92      0.92        37
   macro avg       0.97      0.88      0.88        37
weighted avg       0.93      0.92      0.90        37
```

Table 6. Accuracy of Mental Health Prediction System

A classification report was used to measure the quality of predictions from a classification algorithm and to check how many predictions are true and how many are false. The report shows the main classification metrics: precision, recall, and f1-score on a per-class basis as seen in Table 6. The metrics are calculated by using true and false positives, and true and false negatives. Positive and negative in this case are generic names for the predicted classes. There are four ways to check if the predictions are right or wrong: a) TN / True Negative: when a case was negative and predicted negative, b) TP / True Positive: when a case was positive and predicted positive, c) FN / False Negative: when a case was positive but predicted negative, and d) FP / False Positive: when a case was negative but predicted positive.

Precision is the ability of a classifier not to label an instance positive that is actually negative. For each class, it is defined as the ratio of true positives to the sum of true and false positives. Recall is the ability of a classifier to find all positive instances. For each class, it is defined as the ratio of true positives to the sum of true positives and false negatives. The F1 score is a weighted harmonic mean of Precision and Recall such that the best score is 1.0 and the worst is 0.0. The accuracy of our prediction A.I. is 91.89%, the Recall 87.5%, and the Precision of our model is 96.87%.

5 CONCLUSION AND FUTURE WORK

We identified leading factors contributing to mental illnesses. Social media usage is very strongly related to signs of depression. Also, if a person perceives themselves as addicted to social media, they have a very high chance of displaying depressive traits. Negative thoughts during social media usage also contribute strongly to developing depression. Surprisingly social media usage during the morning hours is more of a contributing factor than usage during the night, likely because morning depression is a common part of depression and people may want to distract themselves by using social media instead of starting their days.

We were also able to identify steps to reduce the probability of the formation of depression. According to our research, reducing time spent on social media to 30 minutes a day can significantly reduce the chances of forming a mental illness. Also, checking social media just before sleeping should be avoided, as a person who does so is 4 times as likely to form depression than one who does not. Checking social media in the morning is also a major contributor and should be avoided. It was also identified that people aged between 19-21 and 22-24 are more susceptible to form depression than people aged 16-18, hence people in these groups should be monitored more actively for signs of depression.

Young professionals also face a similar predicament as the students in our research. With increasing displacement to urban centers, the younger generation of India is going through a social transition. Post the pandemic, those in jobs are feeling the pressure of working harder to retain their positions as younger and more skilled Indians are joining the workforce every year, giving employers options to choose from. A lot of these professionals again turn to social media to help divert their attention from their work and personal stress. Social media is addictive as it gives them the high of being in control over their lives and what makes them happy. They can navigate and control their life on social media, a feature they find missing in their own lives where they feel controlled all the time. Social media gives them the outlet to say what they want to say unfettered, and this is addictive. Hence further investigations can be made into the relationship between young professionals' mental health and their social media usage.

Focus on the potential association between social media use and positive outcomes seems to be rarer in the current literature. Further research is required to identify the traits of social media that have a more positive effect than a negative one. The effect of the social media content itself needs to be evaluated, and differentiation between the content that encourages and the content that inhibits positive values/feelings/behaviors needs to be done. Additionally, more data acquisition and exploration of the Indian population is required, as the current data publicly available is limited. Further investigations on the policies that balance a healthy lifestyle in the era of technology, especially among teenagers, need to be done. Data-driven techniques and data science are going to play a very important role in the future in identifying a balanced lifestyle that will maximize the happiness index of not only India but the world.

REFERENCES

[1] Key substance use and mental health indicators in the United States: Results from the 2020 National Survey on Drug Use and Health", Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, 2021.

[2] F. Ahmad and R. Anderson, "The Leading Causes of Death in the US for 2020", JAMA, vol. 325, no. 18, 2021.

[3] M. Brondani, R. Alan and L. Donnelly, "Stigma of addiction and mental illness in healthcare: The case of patients' experiences in dental settings", PLOS ONE, vol. 12, no. 5, 2017.

[4] "World mental health report: transforming mental health for all", Geneva: World Health Organization, 2022.

[5] Gururaj G et al., "National Mental Health Survey of India, 2015-16: Summary", Bengaluru: National Institute of Mental Health and Neuro Sciences, NIMHANS Publication No. 128, 2016.

[6] "Accidental Deaths & Suicides in India 2021", New Delhi: National Crime Records Bureau (Ministry of Home Affairs), Government of India, 2022.

[7] D. Santomauro et al., "Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic", The Lancet, vol. 398, no. 10312, 2021.

[8] R. Yavich, N. Davidovitch and Z. Frenkel, "Social Media and Loneliness - Forever connected?", Higher Education Studies, vol. 9, no. 2, 2019.

[9] A. Perrin, "Social Networking Usage: 2005-2015.", Pew Research Center, October 2015.

[10] A. Kusá, Z. Záziková, "Influence of the social networking website snapchat on the generation Z", European Journal of Science and Theology, vol. 12, no. 5, 2016.

[11] J. Krosnick and D. Alwin, "Aging and susceptibility to attitude change.", Journal of Personality and Social Psychology, vol. 57, no. 3, 1989.

[12] W. Lau, "Effects of social media usage and social media multitasking on the academic performance of university students", Computers in Human Behavior, vol. 68, 2017.

[13] J. Twenge, Z. Krizan and G. Hisler, "Decreases in self-reported sleep duration among U.S. adolescents 2009–2015 and association with new media screen time", Sleep Medicine, vol. 39, 2017.

[14] W. Craig et al., "Social Media Use and Cyber-Bullying: A Cross-National Analysis of Young People in 42 Countries", Journal of Adolescent Health, vol. 66, no. 6, 2020.

[15] Roberts and M. David, "The Social Media Party: Fear of Missing Out (FoMO), Social Media Intensity, Connection, and Well-Being", International Journal of Human–Computer Interaction, vol. 36, no. 4, 2019.

[16] S. Gupta and D. Basera, "Youth Suicide in India: A Critical Review and Implication for the National Suicide Prevention Policy", OMEGA - Journal of Death and Dying, 2021.

[17] M. Islam, M. Hasan, X. Wang, H. Germack and M. Noor-E-Alam, "A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining", Healthcare, vol. 6, no. 2, 2018.

[18] S. Alonso, I. de la Torre-Díez, S. Hamrioui, M. López-Coronado, D. Barreno, L. Nozaleda and M. Franco, "Data Mining Algorithms and Techniques in Mental Health: A Systematic Review", Journal of Medical Systems, vol. 42, no. 9, 2018.

[19] A. Shatte, D. Hutchinson and S. Teague, "Machine learning in mental health: a scoping review of methods and applications", Psychological Medicine, vol. 49, no. 09, 2019.

[20] S. Rosenfield and D. Mouzon, "Gender and Mental Health", Handbooks of Sociology and Social Research, 2012.

[21] S. Swarnam, "Effect of Social Media Use on Mental Health during Lockdown in India", arXiv.org, 2022.

[22] T. Beck, C. Ward, M. Mendelson, J. Mock, and J. Erbaugh, "An inventory for measuring depression", Arch. Gen. Psychiatry 4, 1961.

[23] Lasa, Lourdes, et al. "The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis." Journal of affective disorders 57.1-3 (2000): 261-265.

[24] K. Kroenke, R. Spitzer and J. Williams, "The PHQ-9", Journal of General Internal Medicine, vol. 16, no. 9, 2001.

[25] L. Radloff, "The CES-D Scale", Applied Psychological Measurement, vol. 1, no. 3, 1977.

[26] Khader, Yousef Saleh. "Mental Health and Psychosocial Concerns and Provision of Services for Adolescent Syrian Refugees in Jordan." Handbook of Healthcare in the Arab World (2021): 2967-2989.