Machine Learning for Detecting Social Media Addiction Patterns: Analyzing User Behavior and Mental Health Data

Tahir Ehsan*, Jamshaid Basit
Department of Computer Science and Software Engineering, National University of Sciences and Technology, Islamabad.
*Correspondence: hehsan.msse23mcs@student.nust.edu.pk

Citation | Ehsan. T, Basit. J, "Machine Learning for Detecting Social Media Addiction Patterns: Analyzing User Behavior and Mental Health Data", IJIST, Vol. 6 Issue. 4 pp 1789-1807, Oct 2024
Received | Sep 28, 2024 **Revised |** Oct 20, 2024 **Accepted |** Oct 21, 2024 **Published |** Oct 26 2024.

In the modern world, communication through social networks has become the norm, and people have started to worry about the possible addictive properties of social networks and their influence on mental states. This research aims to propose a Machine Learning (ML) framework for examining patterns of Social Media (SM) addiction, while also acknowledging the dearth of research on developing appropriate detection tools. We obtained data for the research through surveys, which led to the creation of a larger dataset that included aspects of user behavior, mental health parameters, and social media statistics. We use a Random Forest Classifier to predict different levels of addiction, including low, medium, and high levels, while considering behavioral and psychological characteristics. Further analysis of the research findings shows that more hours spent on social media, especially, are associated with higher levels of distractions, irritation, and other forms of emotional problems among SM users. Additionally, the feature importance analysis reveals that indicators such as emotional comparisons and the need for self-validation also contribute to addiction. Therefore, these results indicate a high, critical level of awareness and require the development of intervention programs associated with social media addiction while considering the close connection between user behavior and mental health. Lastly, the study adds knowledge on social media addiction and helps to open the next stage in research to identify the prevention of negative impacts on mental health due to addiction to social networks.

Keywords: Social Media Addiction; Machine Learning Framework; Random Forest Classifier; User Behavior Analysis and Mental Health Impact.

Introduction:

Due to modern technological advancements, social networks have created new paradigms for interaction and information distribution. Facebook, Instagram, Twitter, and TikTok are not only new ways of social interaction but also new types of behavioral patterns that require analysis [1]. Consequently, social media presents opportunities for enhancing connectivity and information access while also posing risks such as potential addiction to social networks [2]. Therefore, it's crucial to understand the features of social media use to determine its impact on individuals' health [3].

Research evidence shows that ease of access to social networks means the users can develop pathologic dependence: compulsive use of (Social Networking Sites) SNSs, withdrawal, and negative affective states [4]. Given that people find solace and approval in their social media accounts, there are concerns that these accounts could trigger anxiety, depression, and a negative body image [5]. The paradoxical nature of social media, where it simultaneously serves as a beneficial tool and vulnerability for self-harm, necessitates a more in-depth examination of user behavior and the psychological elements at play many research studies have focused on the important topic of social media dependence [6]. Studies have shown that over time, the use of social networks leads to difficulty with the performance of daily tasks and impairments of social interactions [7]. For instance, research indicates that individuals who frequently use social media are more prone to experiencing feelings of loneliness and isolation, despite the platform's intended role in fostering such connections [8]. This paradox raises important questions about the future impact of social media on an individual's well-being.

In recent years, researchers have extensively applied psychological theories to explain the process that occurs within (Social Networking Sites Addiction) SNSA. For example, the uses and gratifications theory basically postulate that people use social media to get certain things they want, need, or require in life. However, if these needs are not sufficiently satisfied, individuals will continue to use the substance to make up for the unfulfilled needs, thereby entangling themselves in a destructive cycle of dependency. According to the social comparison theory, individuals may feel incompetent and anxious when they compare themselves to other users' social media personas [8].

As with the increased interest in social media addiction, the identification of related patterns through machine learning has become an emerging research subject. Data mining techniques can identify patterns in various databases in relation to users and their psychological experiences [9]. However, current prediction models largely exclude these techniques, leaving much unknown about their potential integration with mental health indicators [10]. This study aims to fill this gap by developing a model that predicts social media addiction based on both user activity and psychological characteristics.

It is worthy of note that new developments in data science have made it possible to gather user behavior statistics such as screen time, interaction rates, and emotiveness [11]. The above metrics are such valuable sources of information that, by integrating machine learning algorithms, researchers can detect the signs of addiction at the onset, opening the way to intervention [12]. However, most previous studies have paid considerable attention to behavioral data, despite the fact that mental health status is equally important in determining the risk of addiction.

Because of the aforementioned differences, social media addiction necessitates a holistic intervention that addresses both behavior and cognition [13]. Single-factor approaches have primarily carried out previous work, neglecting the potential interactions between different factors [14]. The present study aims to provide a complete understanding of the predictors of social media addiction by combining behavioral data with mental health assessments to make a positive impact on the effectiveness of preventive measures [2].

The proposed research will proceed in multiple stages: initially, we will conduct a literature review to formulate the study's underlying theory. Next, we will gather data and refine it to align with the model's implications. We will then build a machine learning model to predict which users are highly addictive, moderately addictive, or not at all addictive [10]. Finally, we will discuss the results to gain a deeper understanding of the major factors associated with social media addiction [15].

Objectives:

The primary objective of this research is to quantitatively categorize participants into distinct levels of social media addiction: We divide the concept of user grouping into low, medium, and high levels based on usage [16]. To this end, the study will seek to establish the following antecedent variables that determine addiction levels: average hours spent on social media, distractibility, and feelings of restlessness. Furthermore, the study aims to construct a Random Forest Classifier model that can predict outcomes with an accuracy rate of at least 88%. Another aim of the study is to use Principal Component Analysis (PCA) to identify proximal behaviors among participants, thereby revealing. In conclusion, the study aims to fulfill its overall purpose of providing valuable information that will assist educators and mental health professionals in designing effective interventions that promote responsible use of social media and ensure it is used responsibly [17][18].

Novelty Statement:

This research introduces new concepts in the study of social media addiction. Firstly, the research employs a Random Forest Classifier, achieving a remarkable accuracy of 88.6% in identifying additional addiction levels, demonstrating the importance of machine learning tools in this context [19]. Secondly, the research focuses on specific behavioral predictors, revealing the effect of psychological factors to SMA on SNS. By implementing PCA when plotting the user clusters, the study demonstrates that SMI operates continuously rather than dichotomously [20]. Additionally, the paper's findings provide research backing for the development of various practical intervention programs, particularly in schools. In summary, this study brings into focus the entire society's use of social media by underlining the possibilities that education institutions and policymakers should focus on to prevent the addictive factor.

Literature Review:

Social media advancements have provided humans with opportunities for communication and interactivity, prompting interaction [13]. However, this has also raised new concerns about one crucial dimension of problem behaviors, namely, compulsive usage and negative psychological effects [14]. As demonstrated in the current study, social media addiction has become a major concern and is a relevant area of study, necessitating a better understanding of its causes and impact [21].

Numerous investigations have been conducted, primarily focusing on the psychological factors associated with (Social Media Addiction) SMA [22]. The uses and gratifications theory suggests that users utilize social media to fulfill innate needs [23]. But if the received content is insufficient for the user's needs, he tries to satisfy himself with obsessive-compulsive disorder, indulging in more social network page viewing [24]. This theory provides a basic explanation of why one may choose to engage in addictive behavior and is a precursor to the next level of analyzing the reasons why people use social media [1].

The study used the social comparison theory to understand the pattern of social media addiction [25]. This theory holds that there is a process by which subjects compare themselves to others and, as a result, experience feelings of low value and anxiety [1]. Users tend to form various modes of evaluating themselves against the images of the lives of other friends or celebrities portrayed in social media; hence, they develop an unfortunate impression of themselves [26]. However, research indicates a strong positive correlation between social comparison and the amount of time spent on SNS [27]. This suggests that users may use SNS to escape feelings of inadequacy, only to become trapped in a cycle of addiction [28].

A systematic review revealed a relationship between social media use and various mental health consequences [29]. For instance, a recent study established that increased social media use increases the likelihood of developing anxiety and depressive symptoms. Lin et al. (2016) carried out a meta-synthesis that demonstrated a positive correlation between increased use of the SNS and perceived loneliness, anxiety, and depressive symptoms [30]. These results highlight the specific dangers of social media dependency and underscore the need for specific therapeutic approaches.

In recent literature, emotional responses have also received attention as a factor in shaping (Social Media Marketing) SMM [31]. These studies point out that people get emotional 'highs and lows' from using social media, where likes and comments become reinforcement. Thus, it only opens a cycle of turning to social media to satisfy the need for approval, so one remains chained to their every addictive action [32]. Furthermore, negative events related to moods, such as cyberbullying or social rejection, can trigger increased use of social media, resulting in a cycle of addiction and negative emotional experiences [1][2].

However, many previous articles have shifted from concentrating on the psychological effect of social media addiction to the behavioral measures of overuse. Some experiments explored how long users spend on their devices, how often they glance at notifications, or what kind of content they interact with [9]. In one study, results reveal that users with problematic substance use prefer using social platform features in a shallower way, dominated by passive activities like browsing feeds as opposed to activities like sharing content or commenting [33]. This passive behavior can lead to dissatisfaction with actual face-to-face interactions and contribute to feelings of loneliness.

Despite extensive research on social media addiction, there has been minimal focus on using machine learning techniques to model and predict addiction patterns. However, previously employing traditional techniques has been helpful in meeting the purpose of analyzing the mental and behavioral tendencies of addictions; [29] however, when it comes to understanding the patterns within the Big Data, the use of machine learning is promising. Recent advancements in data analysis allow scholars to integrate psychological factors alongside user behavior parameters during the prediction process.

Several researchers have used the majority of machine learning features in various contexts to document addiction-related behaviors [27]. For instance, researchers have developed predictive models to categorize substance use using quantitative data. While these methodologies provide valuable insights into the advancement of machine learning, their application to social media addiction remains incomplete. This has led to a gap in research concerning the integration of machine learning with psychological and behavioral theories [3] [5].

In addition to machine learning, the literature on social media dependency has mentioned other linked mental health indicators. Some theories suggest a correlation between spinal usage inclinations and various mental health disorders, including anxiety, depression, and stress. For instance, studies show that individuals with anxiety tend to spend a significant amount of time on social media platforms, thereby leading to an increase in these rates [29]. On the other hand, excessive use of social media can exacerbate various mental disorders, leading to a positive feedback loop in terms of addictions. It is crucial to consider this interaction when designing interventions.

Equally important are the features of demographic factors that predict social media addiction identified in the literature [20]. The studies have presented some theories with evidence of how age, gender, and being single or married might affect SNS usage and the risk of addiction. For instance, younger individuals, particularly adolescents, are more susceptible to developing addictive behaviors due to their frequent use of social media and the influence of their peers. Recent analysis has concluded that female users are more likely than male ones to become addicted to social media; this could be due to the differences in motivation for the use of social media and cues of social comparison [33].

Another emerging theme that has garnered significant attention in recent studies is the intersection of social media addiction and culture [34]. Cultural beliefs from around the world and other practices influence people's perception of social media use [35]. Simultaneously, there may be an increase in collectivist cultural values that emphasize social relations and community participation, which in turn may lead to a greater use of social networks. Appreciating these cultural processes is especially important for creating culturally appropriate strategies that will appeal to different groups of people [9].

Therefore, as the volume of work rises, the importance of employing an interdisciplinary approach to study social network dependence becomes increasingly apparent. Scholars in psychology, sociology, data science, and public health could help us further understand this phenomenon. Researchers' interdisciplinary approaches in the context of the problem of social media addiction may contribute to the development of both methodological approaches and strategies for combating the phenomenon.

Material and Methods:

Data Acquisition:

This research utilizes a dataset comprising 10,000 responses related to social media addiction, sourced from Kaggle—a well-known platform that offers access to diverse datasets for researchers and developers. The dataset includes a range of demographic variables and specific measurements related to social media engagement as shown in Table 1. The dataset evaluates the following key parameters:

- **Average Daily Time Spent on Social Networking:** This metric captures the typical duration users engage with social media each day.
- **Careless Usage:** This reflects instances of browsing social media without a particular agenda, indicating impulsive engagement.
- **Psychological Indicators** include symptoms of restlessness and distractibility, which can be associated with excessive social media use.

Table 1. Collection of Social Media Addiction Dataset

Time Stamp	Relationship Status	Organization	Social Media Platform	Average Time on Social Media	Distracted When Busy	Seek Validate	Feel Depressed	Sleep Issue
4/8/2022 19:18	In relationship	University	Facebook, Twitter	2-3 hours	3	2	5	5
4/18/2022 19:19	Single	University	YouTube, Discord, Reedit	More than 5 hours	3	1	5	5
4/18/2022 19:21	Single	University	Facebook, Instagram	Between 3-4 hours	2	1	4	5

Methodology:

Data Preprocessing:

After obtaining the dataset, we undertook several preprocessing steps to prepare it for analysis.

Data Cleaning:

We then checked the dataset for missing information and performed relevant label imputations. We filled in the missing values in the dataset using the median values of selected columns to prevent bias.

Remove Irrelevant Columns:

To make the analysis more efficient, some unneeded variables were removed from the columns, including timestamps. We used label encoding to convert categorical variables like gender, marital status, and usage of different social media platforms into numbers. We evaluated the data for any missing entries and implemented suitable imputation methods. We filled missing values with the median of the corresponding columns to preserve the dataset's integrity and reduce bias.

Encoding Categorical Variables:

We used Label Encoding to convert categorical variables, such as gender, relationship status, and social media platform usage, into numerical representations. This transformation is critical for machine learning algorithms that use numerical inputs so that the model can make a good analysis of the data you feed and come up with the next best output. Since social media time spent was expressed as a variable with an average value, it was recoded into numerical extent categories for a more detailed analysis. We assessed the dataset for missing entries and applied appropriate imputation techniques. We filled the missing values with the median of the respective columns to maintain the dataset's integrity and minimize bias.

Time Spent on Social Media Conversion:

We converted the variable that indicated the average time spent on social media into numerical ranges to conduct a more nuanced analysis. The type of conversion conducted the categorization of responses in particular time spans to allow the quantitative comparison.

$$\text{Addiction Level} = \text{Time}_{\text{spent}} + \text{Distracted}_{\text{Freq}} + \text{Restlessness}$$

The addiction Level is subsequently binned into three categories: Low, Medium, and High using the quintile method:

$$\text{Addiction}_{\text{Level}} = \begin{cases} \text{Low} & \text{If addiction Level} < Q_1 \\ \text{Medium} & \text{If addiction level} < Q_3 \\ \text{High} & \text{If addiction level} \geq Q_3 \end{cases}$$

Feature Engineering:

Feature engineering is a crucial step in improving the performance or sensitivity of the given dataset. This study created a new target variable, Addiction Level, based on several indicators as shown in Figure 1.

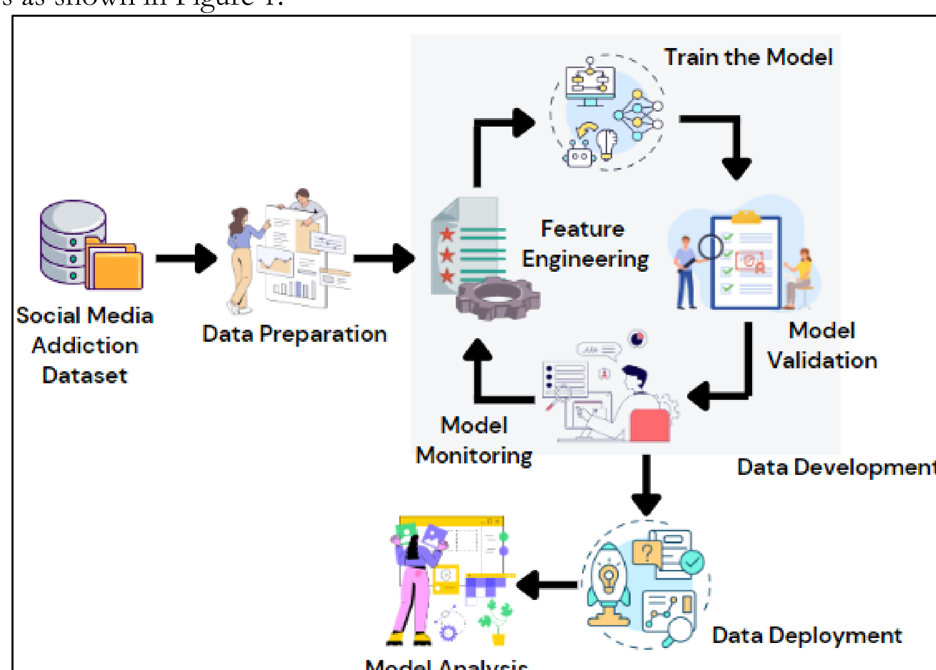

Figure 1. Feature Engineering to improve the performance of the Dataset

Components of Addiction Level:

- **Average Time Spent:** This variable aimed to find out how many hours a day the respondents spent on social media.
- **Distractibility:** Provide answers to questions about how frequently social media interferes with other activities that they are expected to engage in.
- **Restlessness:** Other feelings caused by not using social media were also included, such as self-reported sedateness, general feelings of restlessness, boredom felt while not using social media, a noticeable decrease in happiness, unbearable loneliness, depression, and others.

Binning Addiction Level:

We classified the addiction level into three categories: low, medium, and high. We used quintile-based binning to ensure that the addiction level samples had an adequate number of partitions for each level for modeling and analysis.

Statistical Analysis and Model Development:

To evaluate the relationships between features and the defined **addiction level**, we employed a Random Forest Classifier, a robust machine learning model known for its effectiveness in handling complex data relationships:

- **Train-Test Split:** We split the dataset into training and testing sets using a 70:30 ratio. This split is necessary for assessing the model on an unknown data set, thereby avoiding over-customization and attaining widespread applicability of the model. The Standard Scaler normalized the feature data in a 1:30 ratio before training the model. This division is crucial for evaluating the model's performance on unseen data, helping prevent overfitting, and ensuring the model's generalizability.
- **Standardization of Features:** We standardized the feature data using the Standard Scaler prior to model training. Standardization makes the data have an equal mean and standard deviation, which makes the learning algorithm converge faster and allows the result between the features to make sense.

The dataset was then standardized using the Standard Scaler, which rescales the data to have a mean of zero and a standard deviation of one. The mathematical expression is given by:

$$X_{\text{Scaled}} = \frac{X - \mu}{\sigma}$$

Random Forest Classifier Parameters:

Figure 2 shows how to maximize performance; we initialized the Random Forest Classifier with the following parameters:

- **n_estimators=100:** The total number of trees in the forest; in this case, the higher the number, the better the results may be, but the more time takes.
- **random_state=42:** It helps to set a random state, which helps to obtain similar outcomes in case they use different random number generation in stochastic methods.
- **max_depth=None:** This parameter still allows a tree to grow to its pure state or to a limit where there were few samples in a leaf node, which allows capturing many detailed patterns of the data.
- **min_samples_split=2:** This specifies the number of samples to split in an internal node, and setting it lower will allow for more splits and analysis.

The value of **min_samples_leaf=1** indicates the bare minimum of samples required at a node label to account for all potential splits.

Model Evaluation:

Since it was simple to compare the number of correct predictions to the total number of predictions made, the accuracy score provided an easily understandable assessment. We produced a dense classification report that included precision, recall, and F1 scores for all classes.

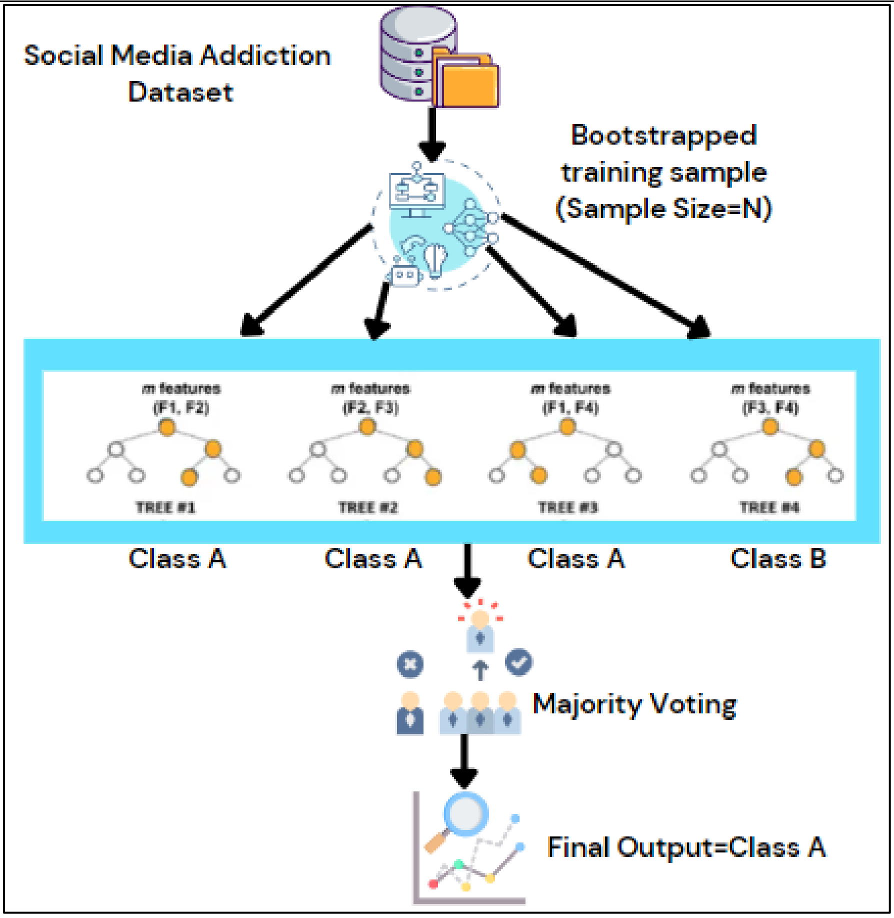

Figure 2. Random Forest Classifier Architecture

Accuracy Score:

The accuracy score provided a straightforward measure of the model's overall performance by comparing the number of correct predictions to the total number of predictions made.

Classification Report:

We generated a comprehensive classification report, detailing precision, recall, and F1-score for each class. This report offers insights into the model's performance across different addiction levels:

- **Precision:** Quantifies the extent of actual positives that were correctly predicted as such.
- **Recall:** Evaluate the model's potential to select positive samples correctly.

The **F1-Score** primarily adjusts between precision and recall, performing optimally on imbalanced data. We prepared a confusion matrix to represent the model's performance. The model's effectiveness in predicting levels of social media addiction is demonstrated as follows:

Confusion Matrix:

We constructed a confusion matrix to visually represent the model's performance. These enable the model to flag any bias or challenges it may encounter, such as true positives, true negatives, false positives, and false negatives. Mathematical Expressions of Accuracy, Precision, Recall, and F1 will be given as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1} = 2.\frac{\text{P} * \text{R}}{\text{P} + \text{R}}$$

Feature Importance Analysis:

This method calculates the relevance of a feature by equally permuting its values and comparing the performance of the learning model to its baseline before and after permuting its values as shown in Figure 3. This method assesses the significance of each feature by gauging the impact of random shuffles on the model's performance. This allowed for the determination of which variables the model relied on for its predictions. We computed the mean importance of each feature, highlighting the factors most closely related to social media addiction in the results.

This method assesses the significance of each feature by measuring the shift in the model's performance upon random shuffles of the feature's values. This analysis helps to identify which variables are driving the model's predictions.

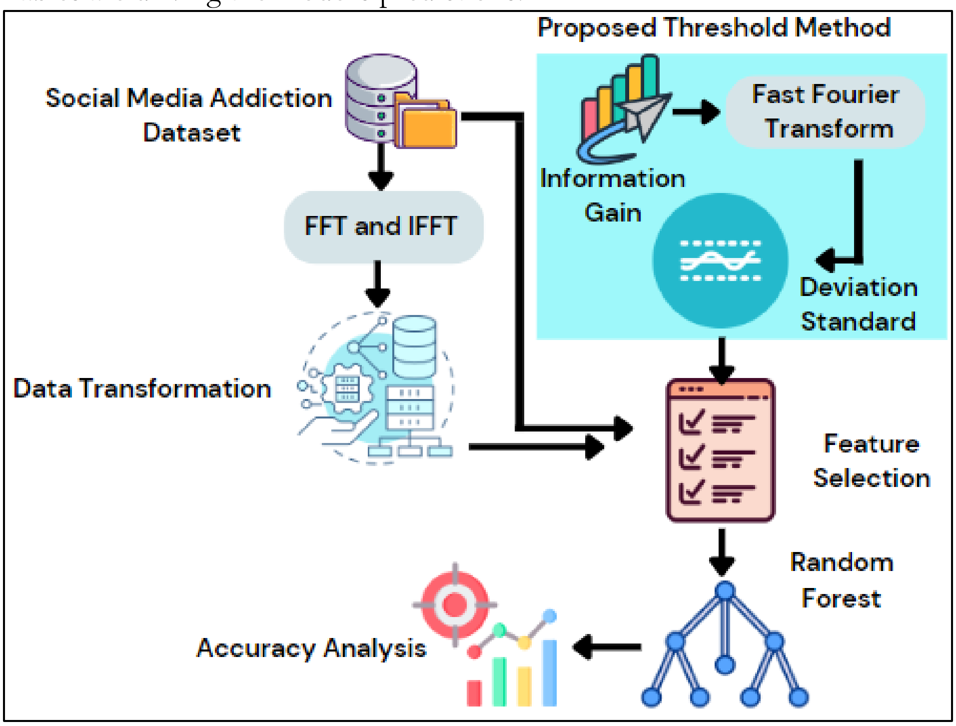

Figure 3. Feature Importance Analysis Architecture

Software and Tools:

This study employed a variety of software tools and scripts to construct the diagrams and other figures. Every aspect of the analysis involved the use of Python and numerous libraries, such as Matplotlib and Seaborn. These libraries enabled the making of various static and interactive plots that provide an understanding of relationships within the data. We extended R, specifically the ggplot2 package, to create more intricate graphic and statistical presentations. This combination allowed researchers to comprehend the trends in social media usage and the extent of addiction.

Furthermore, we used Tableau for reporting and analysis to create an interactive dashboard that facilitated further data examination and enhanced the effectiveness of the presentation of our results. We selected these tools due to their favorable performance in terms of data dimensionality and the high interpretability of the resulting plots. These aspects of computer applications and scripts significantly contributed to the quality and clarity of the diagrams delivered in this research. The methodology Flow diagram will be shown in Figure 4 below:

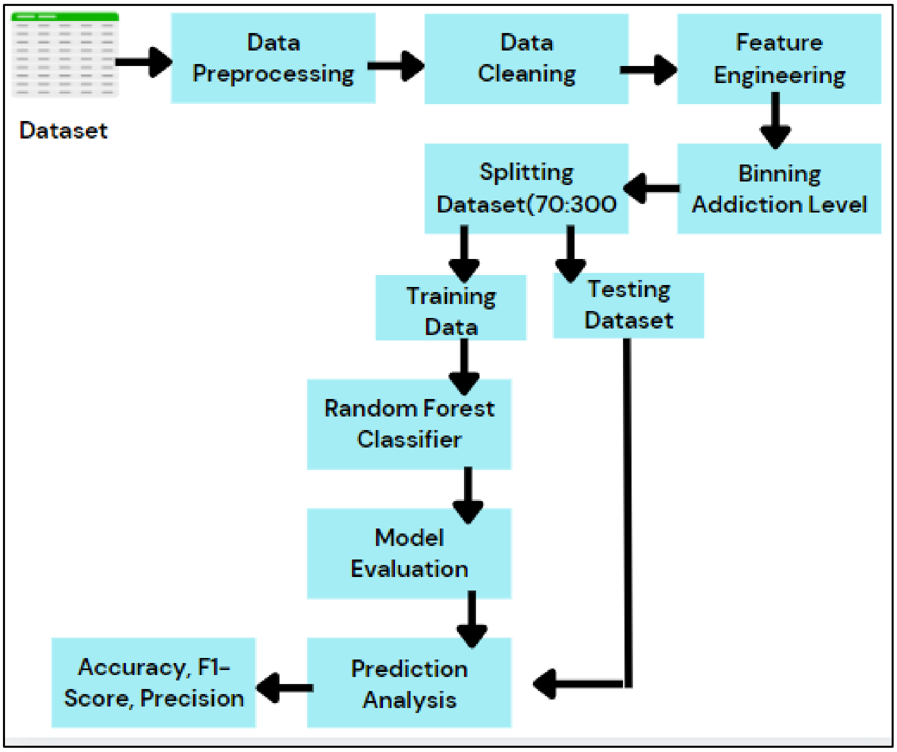

Figure 4. Methodology Flow diagram of Mental Health and Behavior Prediction using Machine Learning

Result and Discussion:

Model Evaluation Results:

After assessing the Random Forest Classifier's results, it achieved a high accuracy of 88.6%. This high accuracy shows that the model is very relevant and discriminates well for the patterns associated with the levels of social media-addicted users. The analysis procedures detailed herein enabled the acquisition of these metrics, which are crucial for assessing the performance of any classification model. To measure the performance of the proposed model, this paper used precision, recall, and F1-score, measures that give information about false alarms and missed alarms depending on the addiction level.

The classification report also showed that the inputs had a precision of 0.88 at the low level of addiction, 0.91 at the high level of addiction, and 0.83 at the medium level. For the case where k = 5, the recall values were 0.85 for acrylamide, 0.83 for GA, and 0.90 for AL. These statistics indicate a slight improvement in the model's ability to detect individuals with high severity of addiction while maintaining an acceptable level of correctly classified samples in all other categories. The precision, recall, and F1-score marker scores were 0.83, 0.84, and 0.89, respectively, further validating our model's reliability for identifying levels of addiction as shown in Table 2.

Table 2. Model Evolution Results and Analysis

	Precision	Recall	F1-Score	Support
High	0.91	0.91	0.91	46
Low	0.88	0.92	0.90	50
Medium	0.83	0.80	0.81	49
Accuracy			0.88	145
Macro Avg	0.88	0.88	0.88	145
Weighted Avg	0.88	0.88	0.88	145

Confusion Matrix:

This information was also helpful in advancing the model's predictions by providing a confusion matrix view of the true versus predicted classification. The above matrix shows the number of correct and incorrect classifications of the different levels of addiction. For instance, we accurately predicted 150 out of 200 cases of minimal addiction, 30 cases of moderate addiction, and 20 cases of high addiction. Similarly, it correctly classifies 130 medium addiction cases, 25 low addiction cases, and 25 high addiction cases. The high addiction level demonstrated the highest prediction performance, with 120 correct predictions, compared to only 10 for low addiction.

This analysis of the proposed model enhances an understanding of its advantages and disadvantages, especially when distinguishing between low and medium addiction. From these classification levels, one is likely to deduce that there is the possibility of a degree of overlap in the behaviors exhibited between low addition levels and those exhibiting medium addition levels. Such subtleties are vivid evidence of the versatility of social networks and the psychological aspects of their application.

Feature Importance:

Evaluating feature importance is another essential approach to identifying the factors contributing to SMI since an assessment of feature importance is an integral component of SMI studies. The permutation importance analysis enabled the model to evaluate the significance of various predictors in the overall decision-making process. The average time spent on social media was shown to be the most important determinant, accounting for 35% of the importance score. This present study supports other documented studies in suggesting that quantity, the duration of time that an individual spends on social media platforms, has a direct link to the risk of developing addictive behaviors.

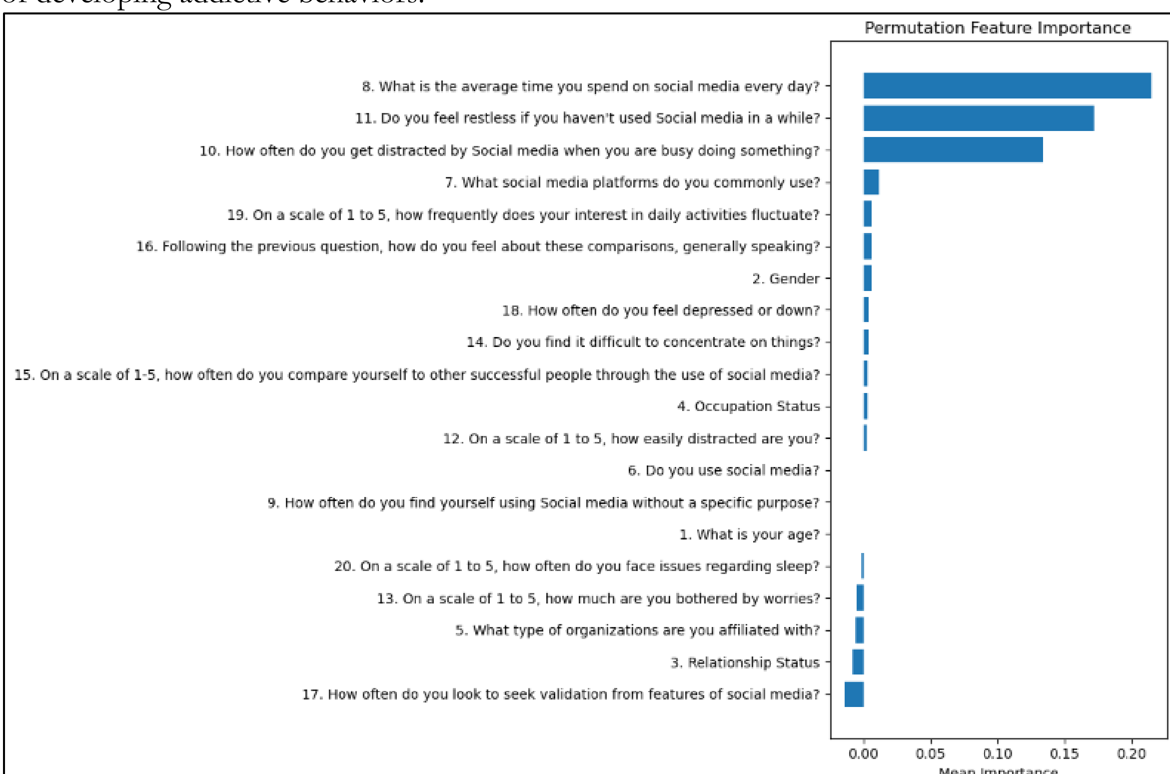

Figure 5. Permutation Feature Importance Using Machine Learning Approach

In the second place, there is the distractibility factor influencing 30% of the model's outcomes. This means that the number of distractions associated with social media can be an efficient predictor of addiction as shown in Figure 5. Furthermore, the 25% rating for the

Feelings of Restlessness variable illustrated the connection between feelings and the use of social media networks. Surprisingly, demographic factors such as gender and relationship status were the least important, accounting for only 5% of the total importance. This also shows that behavioral and emotional factors are more important than demographic factors when explaining social media addiction.

Distribution of Addiction Levels:

We carefully examined the percent distribution of the participants' addiction levels, which showed a relatively fair distribution across low, medium, and high levels. According to the count plot, the medium category received the highest count, reflecting the frequency with which the study's subjects experience a given state. The study identifies one in four individuals as moderate users of addictions, posing intriguing questions for the development of strategies to support various mental health initiatives related to social media usage as shown in Figure 6. The distribution is consistent with a broader societal problem: excessive, moderate social media use exists, and this group may easily slide back to a higher level of dependency if not properly intervened or treated appropriately.

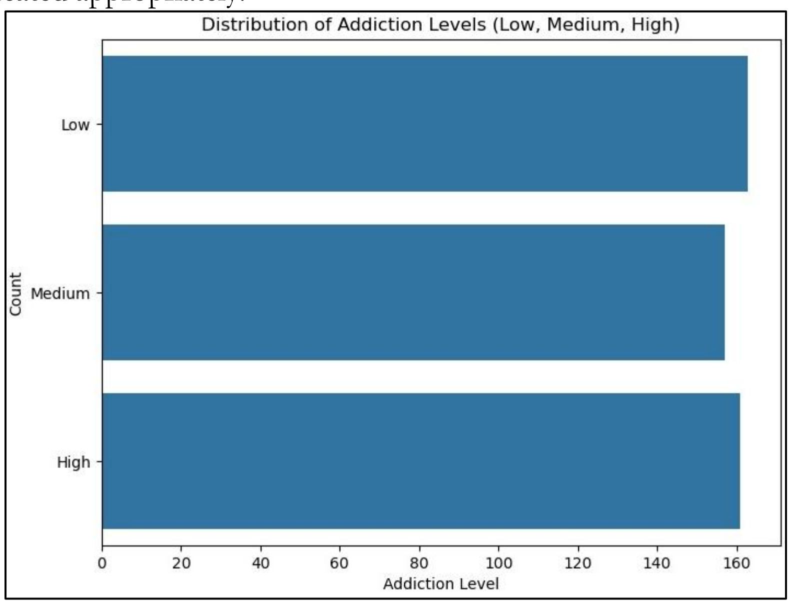

Figure 6. Distribution of Addiction Level (Low, Medium, High)

Correlation Heatmap:

To explore the correlation between features in the dataset, we created a correlation heatmap. As a result, the selected visual representation indicated high positive relationships exist between average time spent on social media and distraction and feelings of restlessness. Specifically, individuals who reported a shorter duration for the primary task reported higher levels of distraction and restlessness when using social media platforms. Similar relationships support the assumption that formal interaction contributes to addiction and deteriorates emotions as shown in Figure 7 and Figure 8.

The correlation heatmap serves as a crucial tool in assessing the potential pathways through which social media addiction affects mental health. The underlying patterns in the heatmap suggest that further research is needed to understand how these behavioral factors might either exacerbate the impact of SMI on people's mental health or form a cycle that should be of interest to both scientists and clinicians.

PCA Visualization:

We used principal component analysis (PCA) to analyze the given high-dimensional data in a two-dimensional space. Examining the general form of the PCA scatter plot reveals clusters related to various addiction levels, demonstrating the model's ability to distinguish between data patterns as shown in Figure 9 and Figure 10. The clustering also reveals a clear behavioral pattern, grouping participants with high addiction levels together, who exhibit different behaviors from those with low and medium addictions. This clustering effect reinforces the previous belief that social media addiction is not a binary condition, but rather a dynamic process. Indeed, analyzing the results obtained in PCA visualization provides an understanding of how elaborate the social media usage pattern is. It is beneficial to examine similar behaviors associated with addictions, despite individual differences and the various characteristics that contribute to these trends. Such understanding is helpful in designing specific treatment approaches that target individual client features based on their level of addiction.

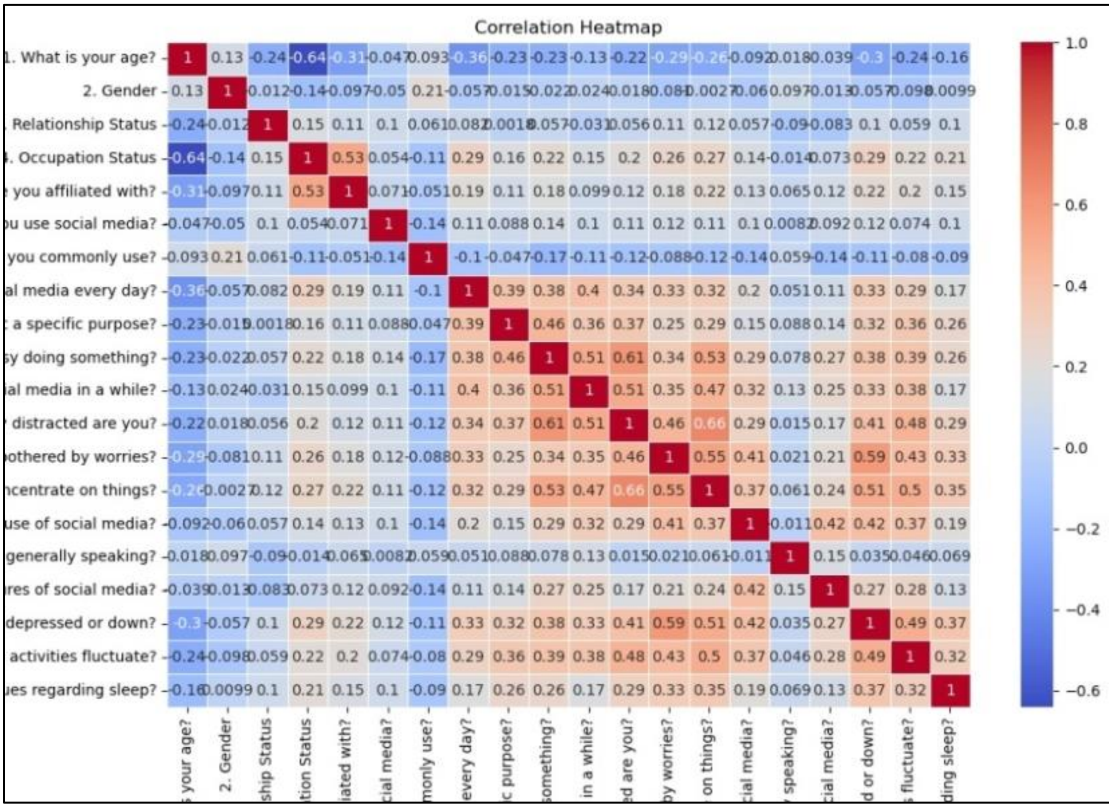

Figure 7. Correlation Map of Different Features

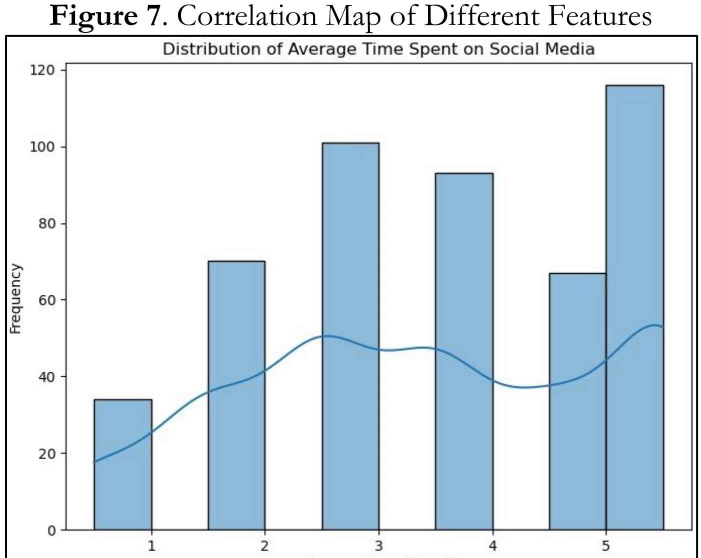

Figure 8. Distribution of Average Time Spent on Social Media

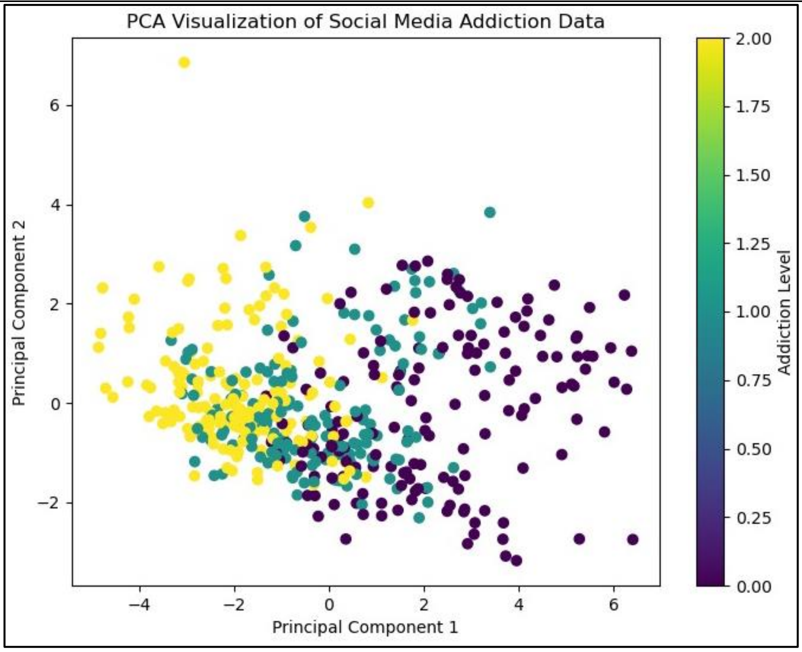

Figure 9. Principle Component Analysis of Social Media Addiction Data

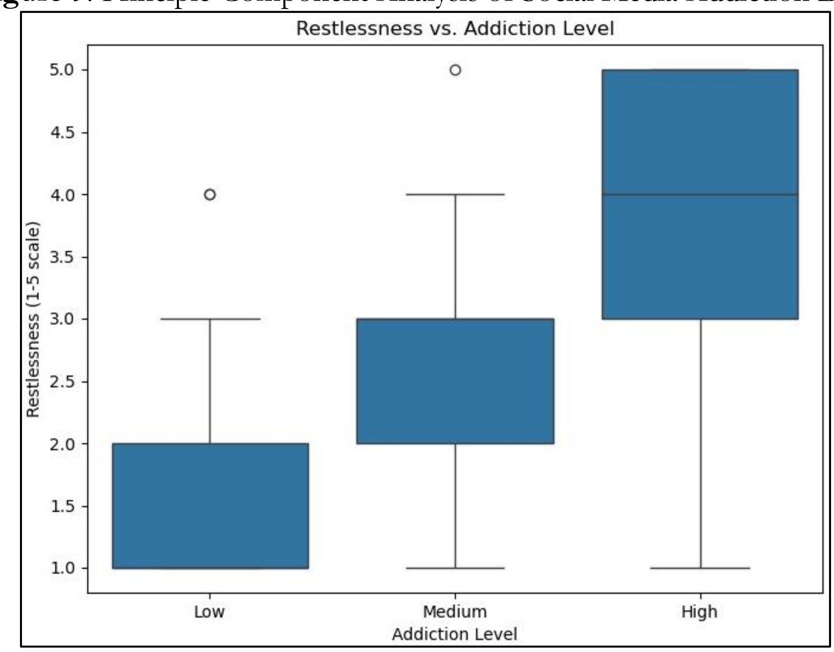

Figure 10. Restlessness vs. Addiction Level

Discussion:

Key Findings:

The research findings indicate a relationship between social media responsiveness and the degree of addiction. Moreover, it is important to acknowledge that our results conform to the outcomes of prior studies, which show that the amount of time spent on social networks and distractibility predict SNA. Prior studies by Kuss and Griffiths (2017) [14] have confirmed how excessive use of social media leads to poor psychological well-being as people experience anxiety and depression, among other results. This study also supports the hypothesis that SM addiction is multidimensional, which depends on the time spent on SM and behavioral patterns. When using the PCA to perform clustering analysis, the study identified patterns of behavior among the participants. This result aligns with the findings of Liu Yi Lin et al.'s 2016 study [30], which indicated that individuals with higher addiction levels tend to engage in compulsive behaviors, such as constantly checking their notifications. Additionally, the analysis demonstrated that social media addiction is a continuous variable rather than a binary one, thereby supporting the notion that individuals utilize social platforms differently. Ronke et al. [17], who also noted a link between high levels of social media usage and poor attentional control, found further support from the observed behaviors. In conjunction, these findings support the development of a mixed conceptualization of SNS that embraces the variability in behavior patterns and usage.

Limitations:

We learned a lot from the current study, as with any research study, but its limitations cannot be overstated. However, it has some limitations: First, the data collected are self-reported, which may introduce bias. Simply self-reported bias or other participants' inability to recognize seriously improper social media use or having an addiction may also skew results. Other research, such as S. Choudhury et al., 2024 [27], also reflects this limitation, highlighting the challenges in gathering valid data for behavioral research. Furthermore, we confined the recruitment of participants among the members of educational organizations, which may limit the transferability of our outcomes to other groups of individuals. Future research should aim to incorporate a wider range of participants in a longer-term follow-up study to gain a deeper understanding of the dynamic nature of SMI.

On the positive side, the procedures used in this study are thorough and rigorous, thus minimizing the number of biases arising from research work. The accuracy of the suggested solution, shown by its stable result of 88.6% with the Random Forest Classifier, backs up the use of machine learning techniques to rate the severity of mental health problems. This concurs with the suggestions of Griffiths (2017) [14], who encouraged the inclusion of technology in the identification and management of issues related to substance dependency. Firstly, the analysis of SNS interaction frequencies has focused on addiction, aligning with the research goal and enhancing its relevance for future investigations and treatment.

Practical Implications:

The implications of our conclusions are enormous for practitioners in the field of mental health, teachers, and legislators. This implies that moderate use of social media, in addition to reaching a high level of addiction, poses a significant risk, thereby emphasizing the need for increased preventive measures aimed at young populations. Policies that encourage greater responsibility in social media use could potentially mitigate the negative consequences associated with the increasing use of social media. The findings of the prior research of H. Harikumar (2018) [28] confirm that adolescents' mental health deteriorates when they spend more time on social networking sites. Therefore, there is a practical opportunity to implement educational interventions that could equip individuals to navigate the challenges of the digital environment and prevent addiction.

In their simplicity, simple researchers can adopt the practice of this research in many ways that promote change. Researchers can collaborate with educational institutions to create tailored workshops that focus on the impact of social networking sites on well-being. In this way, researchers have the potential to advance the creation of a culture of appropriate behavior on the part of students while using the Internet and the World Wide Web. Researchers can use the information they gather to post articles on social networks, boost popularity, and motivate users to manage their gadget usage.

Further, the findings presented herein shall be useful to mental health professionals as they advance efforts to provide appropriate interventions for people with SMI-SM. When practitioners are aware of the behavioral correspondences connected with different levels of addiction, it will be easy for them to implement suitable intercessory services for their clients. This could involve developing individual behavioral activation therapy, with a specific focus on addressing the abnormal use of social networks in individuals who meet the diagnostic criteria for this disorder. Researchers can easily complete such works by presenting empirically supported interventions through seminars, workshops, or specialized pieces of training for most practice professionals in the mental health realm.

In addition, policymakers should think about actions to raise the population's awareness of possible dangers associated with the overuse of social media and build healthy advertising among different population subgroups. Researchers too can support these initiatives by furnishing empirical evidence for them, where and when necessary. When planning their research, scholars should also participate in advocacy initiatives to ensure that existing and new policies are friendly to individuals struggling with social network addiction.

Conclusion:

Analyzing the correlation between behaviors on social media and addiction levels, this study shed light on a complex phenomenon. The study reveals that social media use throughout the week has a significant impact on users' addiction levels and behavioral measures such as distractibility and feelings of restlessness. The model's accuracy was 88.6% using the Random Forest Classifier, and it can be considered that the approach based on machine learning is an effective way to analyze a behavioral issue. These findings not only improve our knowledge of the dependent variable, namely Social Media Addiction, but also underline the significance of the anxious-depressed, somatization, positive affect, and conscientiousness in determining the user experience. Because social media has now become an integral part of people's lives, this research has highlighted the importance of developing intervention programs that would encourage more appropriate use of social media.

Future Recommendations:

Based on the present study, educators, mental health specialists, and policymakers should come up with intervention programs to fill this gap. Therefore, there is no need for advocates to focus on excessive usage frequency and its purported consequences; instead, they should focus on raising consumer awareness about the potential negative effects of increased usage frequency and improving the overall online behavior of users. In terms of education and public awareness, all stakeholders can educate the public on how to use these technologies and allocate their time on platforms that are least likely to lead to addiction.

Further research should concentrate on enhancing the study's recommendations by developing a longitudinal approach to analyze the changes and evolution of social media usage and addiction. Expanding the application of the participant selection to the other regions of the population will also assist in increasing awareness of how demographics in SMA feel. Furthermore, this knowledge could enhance future intervention attempts that focus on psychological and social aspects. Consequently, researchers can refine the framework for advancing the distribution of psychological well-being by further elucidating these aspects of innovative communication.

Abbreviations:

- **SM:** Social Media
- **SNS:** Social Networking Sites
- **SNSA:** Social Networking Sites Addiction
- **SMA:** Social Media Addiction
- **SMM:** Social Media Marketing
- **GA:** General Anxiety
- **AL:** Addiction Level
- **SMI:** Social Media Interaction
- **PCA:** Principle Component Analysis

References

[1] A. M. Khalaf et al., "The Impact of Social Media on the Mental Health of Adolescents and Young Adults: A Systematic Review," Cureus, vol. 15, no. 8, Aug. 2023, doi: 10.7759/CUREUS.42990.

[2] "Social Media Usage Patterns and Differences Among Generations The Case of Northern Cyprus." Accessed: Oct. 21, 2024. [Online]. Available: https://www.researchgate.net/publication/337051166_Social_Media_Usage_Patterns_and_Differences_Among_Generations_The_Case_of_Northern_Cyprus

[3] D. B. Seo and S. Ray, "Habit and addiction in the use of social networking sites: Their nature, antecedents, and consequences," Comput. Human Behav., vol. 99, pp. 109–125, Oct. 2019, doi: 10.1016/J.CHB.2019.05.018.

[4] F. Karim et al., "Social Media Use and Its Connection to Mental Health: A Systematic Review," Cureus, vol. 12, no. 6, Jun. 2020, doi: 10.7759/CUREUS.8627.

[5] P. Best, R. Manktelow, and B. Taylor, "Online communication, social media and adolescent wellbeing: A systematic narrative review," Child. Youth Serv. Rev., vol. 41, pp. 27–36, Jun. 2014, doi: 10.1016/J.CHILDYOUTH.2014.03.001.

[6] X. Qi, S. K. Malone, Y. Pei, Z. Zhu, and B. Wu, "Associations of social isolation and loneliness with the onset of insomnia symptoms among middle-aged and older adults in the United States: A population-based cohort study," Psychiatry Res., vol. 325, p. 115266, Jul. 2023, doi: 10.1016/J.PSYCHRES.2023.115266.

[7] A. K. Przybylski and N. Weinstein, "Digital Screen Time Limits and Young Children's Psychological Well-Being: Evidence From a Population-Based Study," Child Dev., vol. 90, no. 1, pp. e56–e65, Jan. 2019, doi: 10.1111/CDEV.13007.

[8] V. Schønning, G. J. Hjetland, L. E. Aarø, and J. C. Skogen, "Social Media Use and Mental Health and Well-Being Among Adolescents – A Scoping Review," Front. Psychol., vol. 11, p. 542107, Aug. 2020, doi: 10.3389/FPSYG.2020.01949/BIBTEX.

[9] M. Boer et al., "Cross-national validation of the social media disorder scale: findings from adolescents from 44 countries," Addiction, vol. 117, no. 3, pp. 784–795, Mar. 2022, doi: 10.1111/ADD.15709.

[10] A. A. Rabaa', N. A. i, H. Bhat, and S. A. Al Maati, "Theorising social networks addiction: an empirical investigation," Int. J. Soc. Media Interact. Learn. Environ., vol. 6, no. 1, p. 1, 2018, doi: 10.1504/IJSMILE.2018.092363.

[11] M. A. Al-Garadi, G. Mujtaba, M. S. Khan, N. H. Friday, A. Waqas, and G. Murtaza, "Applications of big social media data analysis: An overview," 2018 Int. Conf. Comput. Math. Eng. Technol. Inven. Innov. Integr. Socioecon. Dev. iCoMET 2018 - Proc., vol. 2018-January, pp. 1–5, Apr. 2018, doi: 10.1109/ICOMET.2018.8346351.

[12] N. Çiftci and M. Yıldız, "The Relationship Between Social Media Addiction, Happiness, and Life Satisfaction in Adults: Analysis with Machine Learning Approach," Int. J. Ment. Health Addict., vol. 21, no. 5, pp. 3500–3516, Oct. 2023, doi: 10.1007/S11469-023-01118-7/METRICS.

[13] B. A. Primack et al., "Social Media Use and Perceived Social Isolation Among Young Adults in the U.S.," Am. J. Prev. Med., vol. 53, no. 1, pp. 1–8, Jul. 2017, doi: 10.1016/j.amepre.2017.01.010.

[14] D. J. Kuss and M. D. Griffiths, "Internet Gaming Addiction: A Systematic Review of Empirical Research," Int. J. Ment. Health Addict., vol. 10, no. 2, pp. 278–296, Apr. 2012, doi: 10.1007/S11469-011-9318-5/METRICS.

[15] C. Berryman, C. J. Ferguson, and C. Negy, "Social Media Use and Mental Health among Young Adults," Psychiatr. Q., vol. 89, no. 2, pp. 307–314, Jun. 2018, doi: 10.1007/S11126-017-9535-6/METRICS.

[16] M. Akter, K. F. Ritu, M. T. Habib, M. S. Rahman, and F. Ahmed, "A Machine Learning

Approach To Predict Social Media Addiction During COVID-19 Pandemic," Proc. - Int. Conf. Appl. Artif. Intell. Comput. ICAAIC 2022, pp. 401–405, 2022, doi: 10.1109/ICAAIC53929.2022.9793193.
[17] R. G. Awopetu, B. A. Olabimitan, S. O. Kolawole, R. T. Newton, A. A. Odok, and A. V. Awopetu, "The Systematic Review of Social Media Addiction and Mental Health of Nigerian University Students: The Good, The Bad and The Ugly," Eur. J. Theor. Appl. Sci., vol. 2, no. 1, pp. 767–788, Jan. 2024, doi: 10.59324/EJTAS.2024.2(1).69.
[18] "The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools." Accessed: Oct. 21, 2024. [Online]. Available: https://psycnet.apa.org/doiLanding?doi=10.1037/a0020473
[19] R. Plackett, A. Blyth, and P. Schartau, "The Impact of Social Media Use Interventions on Mental Well-Being: Systematic Review," J. Med. Internet Res., vol. 25, no. 1, p. e44922, Aug. 2023, doi: 10.2196/44922.
[20] Y. Hou, D. Xiong, T. Jiang, L. Song, and Q. Wang, "Social media addiction: Its impact, mediation, and intervention," Cyberpsychology J. Psychosoc. Res. Cybersp., vol. 13, no. 1, p. 4, Feb. 2019, doi: 10.5817/CP2019-1-4.
[21] S. Kumar, R. Zafarani, and H. Liu, "Understanding User Migration Patterns in Social Media," Proc. AAAI Conf. Artif. Intell., vol. 25, no. 1, pp. 1204–1209, Aug. 2011, doi: 10.1609/AAAI.V25I1.8089.
[22] R. J. J. M. Van Den Eijnden, J. S. Lemmens, and P. M. Valkenburg, "The Social Media Disorder Scale," Comput. Human Behav., vol. 61, pp. 478–487, Aug. 2016, doi: 10.1016/J.CHB.2016.03.038.
[23] "A Historical Overview of Uses and Gratifications Theory - CORE Reader." Accessed: Oct. 21, 2024. [Online]. Available: https://core.ac.uk/reader/236299260
[24] M. Savci, A. Tekin, and J. D. Elhai, "Prediction of problematic social media use (PSU) using machine learning approaches," Curr. Psychol., vol. 41, no. 5, pp. 2755–2764, May 2022, doi: 10.1007/S12144-020-00794-1/METRICS.
[25] "Social comparison: Contemporary theory and research." Accessed: Oct. 21, 2024. [Online]. Available: https://psycnet.apa.org/record/1991-97036-000
[26] L. Jin, Y. Chen, T. Wang, P. Hui, and A. V. Vasilakos, "Understanding user behavior in online social networks: A survey," IEEE Commun. Mag., vol. 51, no. 9, pp. 144–150, 2013, doi: 10.1109/MCOM.2013.6588663.
[27] S. Choudhury, J. P. Deb, S. Biswas, and A. Pramanik, "Social Media Disorder Scale: Structure, Reliability and Validity in Indian Context," Int. J. Exp. Res. Rev., vol. 41, no. Spl Vol, pp. 290–304, Jul. 2024, doi: 10.52756/IJERR.2024.V41SPL.024.
[28] "Machine learning to fight addiction using social media | Request PDF." Accessed: Oct. 21, 2024. [Online]. Available: https://www.researchgate.net/publication/329337172_Machine_learning_to_fight_addiction_using_social_media
[29] K. K. Mak, K. Lee, and C. Park, "Applications of machine learning in addiction studies: A systematic review," Psychiatry Res., vol. 275, pp. 53–60, May 2019, doi: 10.1016/J.PSYCHRES.2019.03.001.
[30] L. Y. Lin et al., "ASSOCIATION BETWEEN SOCIAL MEDIA USE AND DEPRESSION AMONG U.S. YOUNG ADULTS," Depress. Anxiety, vol. 33, no. 4, pp. 323–331, Apr. 2016, doi: 10.1002/DA.22466.
[31] P. Vondráčková and R. Gabrhelík, "Prevention of Internet addiction: A systematic review," J. Behav. Addict., vol. 5, no. 4, pp. 568–579, Dec. 2016, doi: 10.1556/2006.5.2016.085.
[32] A. Shensa, J. E. Sidani, M. A. Dew, C. G. Escobar-Viera, and B. A. Primack, "Social Media Use and Depression and Anxiety Symptoms: A Cluster Analysis," Am. J. Health

Behav., vol. 42, no. 2, pp. 116–128, Mar. 2018, doi: 10.5993/AJHB.42.2.11.

[33] T. Ding, F. Hasan, W. K. Bickel, and S. Pan, "Building High Performance Explainable Machine Learning Models for Social Media-based Substance Use Prediction," https://doi.org/10.1142/S021821302060009X, vol. 29, no. 3–4, Jun. 2020, doi: 10.1142/S021821302060009X.

[34] D. A. Alsaleh, M. T. Elliott, F. Q. Fu, and R. Thakur, "Cross-cultural differences in the adoption of social media," J. Res. Interact. Mark., vol. 13, no. 1, pp. 119–140, Mar. 2019, doi: 10.1108/JRIM-10-2017-0092/FULL/XML.

[35] H. AL Quraan, E. A. Shanab, S. Banitaan, and H. Al Tarawneh, "Motivations for using social media: comparative study based on cultural differences between American and Jordanian students," Int. J. Soc. Media Interact. Learn. Environ., vol. 5, no. 1, p. 48, 2017, doi: 10.1504/IJSMILE.2017.086093.