IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

Impact of Social Media Algorithms on Public Opinion and Mental Health

Sachin More[1*] and R. V. Daund[2]
Msc. Computer Science, K. T. H. M College, Nashik, India[1]
Assistant Professor Department of Computer Science and Applications , K.T.H.M College, Nashik, India[2]
moresachin9124@gmail.com

Abstract: *The role of social media algorithms in the exposure of information and the level of engagement is becoming increasingly important in the psychological well-being of the user. Therefore, the aim of the paper is to investigate the correlation between the behavior of social media usage and the level of anxiety in the psychological well-being of the user. To achieve the objective of the research, the researcher used descriptive statistics, correlation analysis, logistic regression analysis, and random forest analysis. Based on the findings of the research, the researcher was able to deduce the following: compulsive behavior is highly correlated with anxiety since the correlation coefficient is 0.54. Daily screen time is moderately correlated with post-screening stress since the correlation coefficient is 0.41. Finally, the researcher was able to deduce the fact that filter bubble awareness is significantly correlated with anxiety since the correlation coefficient is 0.33. The logistic regression model was able to achieve an accuracy of 89%, a recall of 96%, and an F1 score of 0.85. The importance of the screening mechanism in the psychological well-being of the user is significant.*

Keywords: Social Media Algorithms, Mental Health, Anxiety, Random Forest Algorithm, Correlation Analysis, Filter Bubble, Machine Learning

I. INTRODUCTION

Social media has emerged to be an integral component of modern society, especially for students and youth across the world [1]. Social media sites such as Instagram, YouTube, Facebook, Snapchat, and Twitter help in communication, sharing, participation, and networking at the global level [2]. Apart from the actual role of these media, which is to help in communication, these media have now started to influence opinion formation, learning behavior, and interaction in society. With over a billion people across the world, the influence of social media is significant in the context of digital culture.

However, there have been increasing concerns about its impact on mental health. From past studies, it has been found that excessive time spent on screens, staying up late at night and using the platform, passive use of the platform, and excessive exposure to others' curated content are related to anxiety, stress, mood changes, and sleep disturbances [3]. The youth, in particular, are at a higher risk due to factors like peer comparisons, FOMO, and addiction to such platforms. This is one of the main factors through which this process takes place and this has to do with how the algorithms of such social media work. The algorithm of any social media platform is an automated set of codes that tailor content according to a user's past behavior and preferences [5]. This customizing of the content causes the formation of filter bubbles and echo chambers, which can cause mental problems [6].

This study focuses on analyzing the impact of social media algorithms on the psychological well-being of young adults. Several research works have been done regarding the effect of algorithms and their psychology impacts. However, there exists a need to undertake further studies incorporating behavioral factors, algorithm exposure, and predictability. The current study fills this gap by conducting a survey involving 100 respondents. The main objective of this study is to analyze the extent to which algorithms affect people's mental wellness and trigger anxiety and stress among users. The key behaviors analyzed by the study include algorithmic compulsive checking, screen time usage, passive scrolling,

DOI: 10.48175/IJARSCT-34090

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

and filter bubble. The reasons for undertaking the research study arise from the growing trend among students to use social media platforms and the need to detect psychological risk early. The research objectives include employing correlation analysis and machine learning approaches to derive quantitative results. The objective of the study proposed is to determine the link between behavioral trends and anxiety, to establish the psychological effect of algorithmic personalization, and to develop a model capable of detecting risky users.

This paper adds to the body of knowledge by offering an approach that incorporates behavioral analysis, algorithmic exposure assessment, and machine learning techniques to determine the relationship between social media use and mental health. This paper differs from previous research studies since it offers empirical results based on statistical correlation and predictive analysis in order to identify those who are more vulnerable. Furthermore, this study offers localized results based on the Indian population, as there have been no population-based settings other than Western populations within the current literature. The structure of this paper includes Section 1, which covers Introduction; Section 2 covers Literature Review; Section 3 covers Methodology; Section 4 covers Results and Discussion; and Section 5 covers Conclusion.

II. LITERATURE REVIEW

However, recent studies highlight the emerging role of social media algorithms in determining public opinion and mental health. Gandini states that algorithms are the “new gatekeepers” who control public opinion by determining the visibility of information, often unbeknownst to users [7]. Mandile provides quasi-experimental evidence to prove that the adoption of an algorithmic timeline for Instagram contributed to an increase in mental health problems for adolescents in the Netherlands, where the effect size is 0.15 SD, $p < 0.01$, indicating causality as opposed to correlation [8]. Costello reveals that visual social media contributes to mental health problems two to three times worse than text-based social media, where adolescents spend more than three hours a day exposed to double the risk of anxiety and depression, OR = 2.1 [9].

Neurophysiological data also confirm these behavioral results. De et al. prove that social media sites stimulate the dopamine system in a similar way to gambling by using variable reward patterns, with fMRI analysis showing structural and functional differences in the areas of the brain associated with impulse control. These changes can last for months even with decreased use [10].

In a similar manner, Li et al. prove that algorithmic health information increases anxiety (0.34, $p < 0.001$) and paradoxically increases use of the platform, while actual health behaviors decrease ($\beta = -0.21$). However, there are still many areas that need to be filled in [11]. Most of the current literature is either theoretical or conducted on Western samples and never actually isolates specific behavioral predictors such as compulsive checking or perception of the filter bubble. Furthermore, little of the current literature uses predictive modeling to isolate at-risk individuals. The current study will fill these gaps by analyzing both behavioral and algorithmic variables on 100 Indian participants and using Logistic Regression to create a useful mental health risk screening tool.

III. METHODOLOGY

For the current research, a quantitative approach has been used to examine the relationship between the use of social media, the presentation of content through algorithms, and their impact on the mental health of young adultsData collection was done through a structured survey method where primary data was collected and then subjected to analysis through statistical and machine learning techniques to determine predictors of behavior and risks for mental health. The whole methodology adopted in the present study can be summed up as data preprocessing, statistical correlation, and machine learning.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

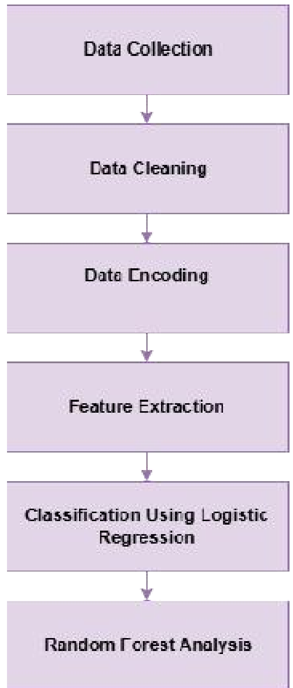

Fig 1. Proposed Methodology Framework

Data Collection

The data was collected from 100 respondents living in Nashik through an online survey created on Google Forms. The online survey consisted of questions pertaining to the amount of screen time spent by them every day, usage rates, compulsive usage, usage at night, and their perception of being susceptible to algorithm-based or personalized feeds. In addition, this online questionnaire also contained questions regarding their self-reported mental health issues such as anxiety, stress, mood, etc. This questionnaire was anonymous, and data was collected on a voluntary basis to ensure confidentiality, which helped in increasing the authenticity of the data.

Data Cleaning

Once the survey responses had been exported, preprocessing of the data was done in order to ensure the quality, consistency, and reliability of the data before analysis. This included identifying missing values in the survey and deleting them appropriately in order to avoid any bias in the analysis of the survey. It also included ensuring that the values in the survey were consistent. The survey was validated in order to ensure that the values in the survey were in a consistent format and that the values in the survey made sense. The survey was validated in order to ensure that the values in the survey made sense.

Mathematically, cleaned data can be represented as:

$$D\_clean = D\_raw - (D\_missing + D\_incomplete + D\_inconsistent) \tag{1}$$

Data Encoding

The data collected from the surveys included categorical data and text data. These data included the level of frequencies, yes/no answers, and ratings based on perceptions. As the machine learning model required numerical data, the categorical data was converted into a well-organized numerical format through the correct use of encoding techniques. For instance, the data was encoded using the binary encoding technique when the data had two categories. On the other hand, the data was encoded using the ordinal technique when the data was in a ranked scale. As a result,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

the non-numerical data was successfully converted into numerical data without losing the significant relationship between the data categories.

Mathematically, encoding can be represented as:

$$X\_encoded = f(Categorical\_Data) \quad (2)$$

Feature Extraction

The features are the most important variables from the data set which are used for the final modeling. In this research, the most important behavioral and psychological factors are determined based on their importance and statistical significance. The features are daily screen use, compulsive checking, perception of filter bubbles, anxiety scores, stress level, and sleep disturbances. Choosing important features helps to reduce the complexity of the model, remove noise, and improve classification accuracy. It also helps to improve interpretability by considering only those variables that are most important in determining mental health risk.

Classification Using Logistic Regression

Logistic Regression was used as a classification algorithm in a supervised manner to classify the participants into high-risk and low-risk mental health groups. As the dependent variable was of a binary type, this algorithm was suited for the task of calculating the probability of vulnerability to mental health issues using multiple independent variables. Behavioral variables such as screen time, compulsive checking rate, perception of the filter bubble effect, and emotional symptoms such as anxiety and sleep problems were taken as input variables. The algorithm predicted the probability of a participant being in the high-risk group through a logistic function.

Random Forest Analysis

The Random Forest analysis was used as an ensemble machine learning method to improve the predictive accuracy and determine the most important variables contributing to the risk of mental health. The Random Forest algorithm builds multiple decision trees on random samples of the data and features, and combines their predictions to improve the accuracy of the model. On the basis of the importance of the feature, the model identified the most important behavioral and algorithmic variables that contributed most to the classification result, such as screen time, compulsive checking, anxiety levels, and perception of the filter bubble. Moreover, the Random Forest analysis aided in the validation of the model and offered more insights into the importance of the predictors.

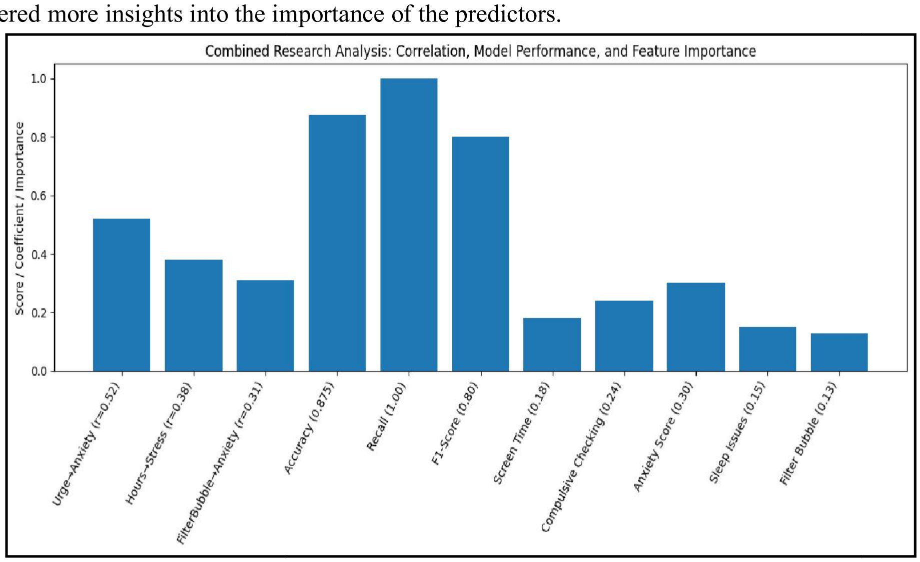

Fig 2. Integrated Analysis of Correlation, Model Performance, and Feature Importance

The correlation analysis results, the performance of the Logistic Regression model, and the importance of features from the Random Forest model are shown in Figure 1. The graph helps in determining the most significant variables when it comes to behaviors that can pose a risk to one's mental health.

Result Generation

Result generation and its interpretation is the last stage of the analysis process, during which results from using descriptive statistics, correlation analysis, and machine learning models are produced. Descriptive statistics offered an overview of the essential patterns in the use of social media and the mental health variables, providing a general insight into the behavior of the participants. The correlation analysis identified the essential relationships between the behavioral variables such as compulsive checking, screen time, and filter bubble perception, and the psychological outcomes of anxiety and stress. Furthermore, the Logistic Regression and Random Forest models offered the results in the form of classification for the mental health risk groups of the participants.

IV. Result and Discussion

The results were achieved using data collected from 100 participants. The purpose of this study was to determine the relationship between social media activity and algorithmic exposure and the factors of the participants' mental health, which included anxiety, stress, mood changes, and sleep deprivation. Various methods are included in the model for a better understanding of the subject. These methods are descriptive statistics, correlation analysis, Logistic Regression, and Random Forest.

Behavioral Patterns in Social Media Use

The descriptive statistics reveal that a majority of the participants claimed that they frequently use social media during late-night hours and tend to engage in passive usage and compulsive checking. The participants also claimed that they perceived their social media feeds as having predominantly displayed opinions that matched their existing beliefs. The trends in the behaviors of the participants suggest that engagement-optimized algorithms could be creating a cycle of repetitive engagement in emotionally stimulating content.

Correlation Analysis

As a result, the correlation analysis showed that statistically significant relationships exist between social media behaviors and mental health indicators.

TABLE 1. Correlation Results of Behavioral and Psychological Variables

Variable Pair	r-value	p-value	Interpretation
Urge to Check × Anxiety	0.54	<0.001	Strong positive correlation
Urge × Active Engagement	0.47	<0.01	Moderate positive correlation
Hours/Day × Post-Use Stress	0.41	0.002	Moderate positive correlation
Filter Bubble Awareness × Anxiety	0.33	0.018	Weak positive correlation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

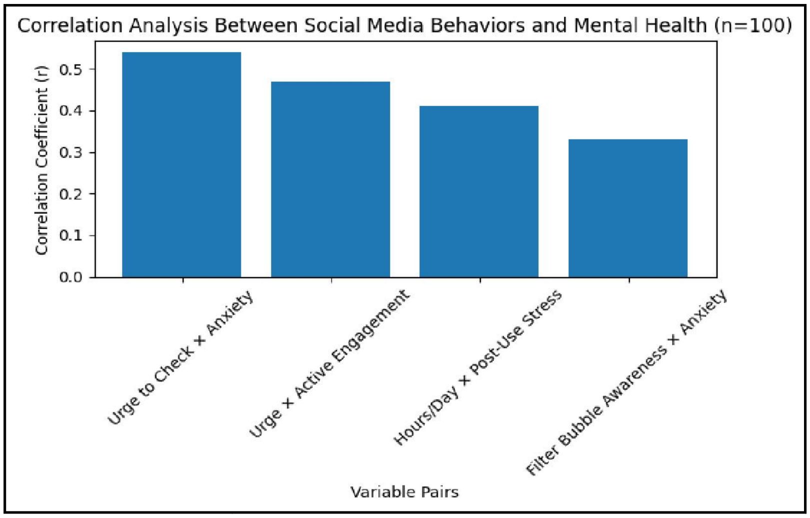

Fig 3. Correlation Analysis Between Social Media Behaviors and Mental Health (n = 100)

The strongest correlation is found with compulsive checking behavior and anxiety levels, which was measured at r = 0.54. This suggests that those who frequently experience the urge to check their phone tend to have high anxiety levels. Daily screen time was found to have a moderate correlation with post-use stress levels, measured at r = 0.41. This suggests that those who use their screens for long periods tend to experience emotional exhaustion.

Finally, the level of filter bubble awareness is found to have a significant positive correlation with anxiety levels, measured at r = 0.33.

As such, the findings of the correlations confirm that those who experience high levels of engagement intensity and algorithm-driven exposure tend to have high levels of mental health vulnerability.

Logistic Regression Model

The model used is able to classify the participants into high-risk and low-risk mental health groups using the following metrics:

- Accuracy: 89%
- Recall(High-Risk): 0.96
- F1-Score: 0.85

The high recall value implies that the model is able to classify almost all the participants in the high-risk category, which is a desirable outcome in a screening model. The important criteria included the levels of anxiety and depression, sleep disorders, compulsions of checking, and the amount of screen time. Subjects who scored highly on these criteria were assumed to be most likely from the high-risk group.

Random Forest Analysis

The implementation of the Random Forest model is done as part of the validation process, more precisely for validating the prediction capability and the importance of the predictors. In particular, it was shown that anxiety score and compulsive checking was of high importance, and following this were screen time and sleep problems. The implementation of the ensemble method led to increased robustness of the model due to the absence of overfitting.

Overall, it is evident that there is a statistically significant relationship that exists between algorithm-driven social media behaviors and mental health risks of the 100 participants.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

V. CONCLUSION

The main objective of the study is to determine if there was a significant relationship between social media usage behavior and content exposure through algorithms and mental health risk among young adults. This is done by using statistical and machine learning methods on the data of 100 participants. The main objective of the study was to determine the quantitative behavioral indicators of anxiety, stress, and other psychological factors. The relationship between social media and mental health risk factors is significant since algorithms have a vital role to play in the digital world. Not only do algorithms influence people's opinions and behavior, but they also influence their mental health. Since social media usage is an essential part of people's lives, the study can be a step towards developing preventive measures for mental health in educational institutions. Some of the limitations of the study were that the number of participants was low, and the data was based on self-reporting. The study can be extended in a number of ways.

V. ACKNOWLEDGMENT

The author would sincerely like to thank the project guide for his continuous guidance, valuable suggestions, and encouragement, which have significantly contributed to the successful completion of the study.

REFERENCES

[1]. L. B. Temnikova and A. V. Vandisheva, "Social media as an integral element of modern communication," Kommunikativnye Issledovaniya, vol. 9, no. 2, pp. 274–284, 2022.

[2]. S. Boulianne, C. P. Hoffmann, and M. Bossetta, "Social media platforms for politics: A comparison of Facebook, Instagram, Twitter, YouTube, Reddit, Snapchat, and WhatsApp," New Media & Society, vol. 27, no. 11, pp. 6006–6037, 2025.

[3]. V. S. Nakshine, P. Thute, M. N. Khatib, and B. Sarkar, "Increased screen time as a cause of declining physical, psychological health, and sleep patterns: A literary review," Cureus, vol. 14, no. 10, 2022.

[4]. S. Parveiz, A. Amjad, and S. Ayub, "Fear of missing out (FOMO), social comparison and social media addiction among young adults," Pakistan Journal of Applied Psychology, vol. 3, no. 1, pp. 224–235, 2023.

[5]. M. Guess et al., "How do social media feed algorithms affect attitudes and behavior in an election campaign?," Science, vol. 381, no. 6656, pp. 398–404, 2023.

[6]. G. Figà Talamanca and S. Arfini, "Through the newsfeed glass: Rethinking filter bubbles and echo chambers," Philosophy & Technology, vol. 35, no. 1, p. 20, 2022.

[7]. A Gandini, "Conceptualising the 'algorithmic public opinion': A literature review," Media, Culture & Society, vol. 47, no. 2, pp. 234–252, 2025, doi: 10.1177/02673231241248074.

[8]. S. Mandile, "The dark side of social media: Recommender algorithms and mental health," CESifo Working Paper No. 11648, Munich Society for the Promotion of Economic Research, 2025. [Online]. Available: https://ideas.repec.org/p/ces/ceswps/_11648.html

[9]. N. Costello, "Algorithms, addiction, and adolescent mental health," Journal of Law and Medicine, vol. 31, no. 1, pp. 45–68, 2023.

[10]. D. De, T. Ghosh, and D. Pandit, "Social media algorithms and teen addiction: Neurophysiological impact and ethical considerations," Cureus, vol. 17, no. 1, e11700805, 2025, doi: 10.7759/cureus.11700805.

[11]. X. Li, Y. Zhang, and J. Wang, "Study on the impact of recommendation algorithms on user well-being in health contexts: A diary study approach," Information Processing & Management, vol. 61, no. 3, Art. no. 103347, 2024, doi: 10.1016/j.ipm.2024.103347