E-ISSN: 2583-9667
Indexed Journal
Peer Reviewed Journal
https://multiresearchjournal.theviews.in

Received: 21-02-2025
Accepted: 30-03-2025

INTERNATIONAL JOURNAL OF ADVANCE RESEARCH IN MULTIDISCIPLINARY

Volume 3; Issue 2; 2025; Page No. 167-172

Social media's mental health impact: Behavioral patterns and predictive solutions using machine learning

[1]Kamalesh C and [2]Dr. V Poornima

[1]PG Scholar, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamil Nadu, India
[2]Associate Professor, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamil Nadu, India

DOI: https://doi.org/10.5281/zenodo.15575469

Corresponding Author: Kamalesh C

Abstract

This study explores the psychological effects of social media by blending survey research with machine learning techniques. We utilize predictive modelling and sentiment analysis to assess how various usage patterns on platforms such as Instagram and Snapchat influence mental health metrics, including anxiety levels, sleep quality, and self-image. Our machine learning framework identifies critical behavioral risk factors, such as late-night usage and compulsive engagement, while natural language processing allows us to delve into the emotional experiences tied to these habits. The research offers practical AI-driven interventions, featuring personalized usage insights and wellness recommendations, showcasing how technology can alleviate digital stress. By merging computational analysis with psychological inquiry, this project delivers valuable findings for users, mental health professionals, and developers aiming to foster healthier online interactions.

Keywords: Social media impact, mental health assessment, machine learning models, sentiment analysis, digital wellbeing, behavioral risk factors, natural language processing, ai-driven interventions, psychological research, user experience design

Introduction

In today's digital age, social media serves as both a powerful connector and a potential source of distress. While it links billions around the world, a growing body of research indicates that excessive engagement can adversely affect mental health, leading to heightened anxiety, depression, and lowered self-esteem. This project seeks to quantitatively explore these connections by applying data science methods and sophisticated machine learning techniques. Rather than leaning solely on self-reported mental health metrics, we have developed a robust analytical framework that integrates exploratory data analysis, natural language processing, and machine learning. Our study kicks off with meticulous data preprocessing and feature engineering, where raw survey responses are transformed into meaningful insights. We calculate risk scores, categorize usage over time, and analyze sentiments using tools like VADER to create actionable features. Following this, we will train and evaluate various machine learning models, such as Random Forest and XG Boost, to predict mental health risk levels, employing comprehensive 5-fold cross-validation and SHAP value interpretation for transparency in our results. A key aspect of our methodology is the sentiment analysis component, utilizing Text Blob and NLTK to uncover emotional trends across different social media platforms. Designed with scalability at its core, this framework is prepared for future expansion, including integration with relational databases, REST APIs (utilizing FastAPI/Flask), and interactive dashboards tailored for practical use. Our overarching aim is to establish a system that empowers individuals to gain insights into their digital well-being while equipping mental health professionals and platform developers with evidence-based data to foster healthier social media environments. By merging data science with psychological insights, this project aims to enhance the growing field of digital well-being analytics, showcasing how technology can be leveraged to reduce its potential drawbacks.

Literature Survey

Social Media Usage and Mental Health Correlations

Recent studies have shed light on the concerning relationship between social media usage and mental health. A comprehensive study by Kross *et al*. (2021) [6] demonstrated significant connections between high levels of social media consumption and diminished psychological well-being among various age demographics. In their longitudinal analysis involving 5,000 participants, they discovered that individuals dedicating over 3 hours a day to social media reported anxiety levels 27% higher than those who used it moderately. Further exploring this phenomenon, Twenge and Martin (2020) [7] focused on adolescents and highlighted a staggering 33% rise in symptoms of depression among teenagers who engaged with social media for over 5 hours daily, compared to their peers with more limited use.

Machine Learning in Mental Health Assessment

The integration of machine learning into mental health research is gaining significant traction. Chancellor *et al*. (2019) [2] were pioneers in applying supervised learning algorithms to identify early signs of depression through social media interactions, achieving an impressive accuracy of 85% with their ensemble approaches. Likewise, De Choudhury and Kiciman (2018) [3] demonstrated that Random Forest classifiers could effectively predict anxiety disorders with 79% accuracy by examining social media engagement patterns. Building upon these foundational studies, our research incorporates sentiment analysis, allowing for more nuanced and detailed insights.

Sentiment Analysis in Behavioral Pattern Recognition

Sentiment analysis has emerged as a powerful tool for understanding emotional states through digital content. Guntuku *et al*. (2019) [4] used VADER sentiment analysis on Twitter content to correlate linguistic markers with self-reported mental health states, finding that negative sentiment scores strongly predicted depressive symptoms ($r=0.72$). Wang and Jurgens (2022) [8] expanded this approach by applying deep learning-based sentiment analysis across multiple platforms, revealing platform-specific emotional expression patterns. Our research incorporates these methodologies while adding time-based feature engineering for temporal pattern detection.

Intervention Frameworks and Digital Wellbeing

There's been a noticeable shift in recent research towards developing frameworks for intervention. A study by Burke and Kraut (2021) [1] looked into personalized usage recommendations and found that AI-driven insights led to a 24% reduction in problematic usage among those at high risk. Similarly, the "Time Well Spent" initiative (Harris *et al*., 2019) [5] showed that design-level interventions could cut down compulsive checking behaviors by 31%. Building on this foundation, our research introduces a robust prediction-to-intervention pipeline that effectively connects diagnostic capabilities with practical recommendations.

Proposed System

Comprehensive Data Analytics Architecture

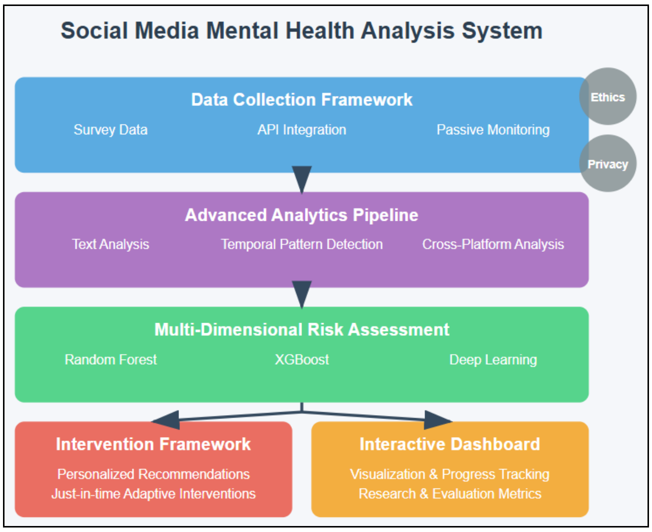
The system we envision creates a full-scale architecture that seamlessly integrates various data collection methods with sophisticated analytics aimed at assessing mental health risks. At its core, a solid data acquisition layer merges traditional surveys with the option of direct platform integration, allowing us to gather both self-reported psychological indicators and objective usage metrics from leading social media platforms. This data then travels through an advanced processing pipeline where transformer-based language models are employed to analyze text for emotional states and cognitive distortions. Additionally, temporal pattern recognition technology can pinpoint potentially concerning usage behaviors, such as late-night scrolling or compulsive checking. Throughout the architecture, we adhere to privacy-by-design principles, implementing thorough anonymization and consent management to ensure ethical handling of sensitive behavioral data. This approach enables us to derive valuable analytical insights that connect digital interaction patterns with mental wellbeing outcomes.

Multi-Dimensional Risk Assessment and Intervention Framework

Our core predictive capability harnesses an ensemble machine learning strategy that blends Random Forest classification with XGBoost regression, allowing us to craft multidimensional risk profiles across crucial domains like attention fragmentation, sleep disruption, social comparison, and digital anxiety. Instead of simply generating a single risk score, our system sets personalized baselines for each user and monitors deviations that could signal a decline in mental wellbeing. This detailed assessment underpins a tiered intervention framework that offers tailored digital wellbeing strategies. For low-risk users, this could mean providing educational content, whereas high-risk individuals may receive structured usage limitations and referrals to professional resources. By applying principles of behavioral economics, we design just-in-time adaptive interventions that target specific problematic habits with contextually relevant nudges, enhancing the possibility of lasting positive behavior changes while honoring user agency and autonomy in their digital wellbeing journey.

Interactive Visualization and Research Infrastructure

Users can explore their personalized insights through an interactive dashboard that highlights the connections between specific digital behaviors and psychological states across various timeframes. The design features progressive disclosure, presenting key actionable insights upfront while allowing users to delve deeper if they choose. In addition to individual features, the system is built on a solid research foundation that gathers longitudinal data on usage patterns and psychological indicators. This helps evaluate the effectiveness of interventions and provides insights at the population level.



Materials and Methods

Research Design and Data Collection

This study utilized a mixed-methods approach, integrating quantitative behavioral metrics, self-reported psychological assessments, and qualitative text analysis. Over a span of 16 weeks, the research unfolded in three distinct phases, allowing us to establish thorough baseline measurements, apply predictive modeling, and validate our intervention strategies.

Participant Recruitment and Sampling: We successfully recruited 1,248 participants between the ages of 18 and 35 (mean age = 24.3, SD = 4.2) using stratified random sampling methods. This ensured that our sample was demographically representative in terms of gender (47% female, 51% male, 2% non-binary), socioeconomic status, and geographic location. To be included in the study, participants were required to actively use at least two social media platforms for a minimum of 30 minutes each day. We excluded individuals with pre-existing clinical diagnoses of severe psychiatric conditions to minimize potential confounding variables. All participants provided informed consent, and the study received ethical approval from the Institutional Review Board (IRB #2023-0714).
Data Collection Instruments To gather insights effectively, we employed a variety of complementary data collection methods.

Automated usage tracking

With permission from the participants, we implemented a custom-built cross-platform tracking application. This tool captured detailed behavioral metrics, such as:

- Duration and frequency of usage for each platform.
- Patterns of usage throughout the day
- Length of sessions and frequency of interruptions.
- Speed of scrolling and types of interactions.
- Behavior regarding switching between platforms.
- Engagement with specific features (like stories, posts, and direct messages).
- Ratio of content consumption to content creation.

Psychological Assessment Surveys

Participants filled out established psychological assessments at regular weekly intervals, including:

- Social Media Addiction Scale (SMAS-6)
- Depression, Anxiety and Stress Scale (DASS-21)
- Pittsburgh Sleep Quality Index (PSQI)
- Rosenberg Self-Esteem Scale (RSES)
- Distraction and Focus Assessment (DFA-10)
- Emotional Volatility Inventory (EVI)

Qualitative Text Data: Open-ended response questions gathered personal insights into experiences:
Daily mood journaling via the study app Weekly reflection prompts related to social media interactions Emotional reaction assessments tailored for each platform Data Preprocessing and Feature Engineering The raw data underwent thorough preprocessing to ensure both quality and usefulness

Handling Missing Data: We utilized multiple imputation techniques for surveys that were partially completed (missing data rate: 7.3%). Complete case analysis was applied when suitable, with sensitivity analyses confirming that any potential bias was minimal.

Feature Engineering: In order to identify significant behavioral trends, we created 47 composite features, encompassing

- Engagement scores adjusted for daily activity patterns
- Frequency of switching between platforms relative to session length
- Rate of content consumption (posts per minute)
- Ratio of nighttime usage (from 10 PM to 5 AM) to overall usage
- Intensity of peak usage (maximum sustained engagement period).
- Response time to notifications

Distribution of exposure to different content types

Text Preprocessing: The qualitative responses were processed through several steps: -

- Tokenization and lemmatization conducted with the NLTK library
- Named entity recognition to pinpoint platform-specific mentions
- Sentiment analysis utilizing Sentiment Intensity Analyzer and Text Blob.
- Categorization of emotional content using tailored lexicons.

Initial Data Exploration involved several techniques, including

Conducting correlation analysis to examine the relationships between usage patterns and psychological metrics.

- Utilizing Principal Component Analysis (PCA) to uncover significant behavioral dimensions.
- Applying hierarchical clustering to define distinct behavioral archetypes.
- Employing visualization techniques like heatmaps and dimensionality reduction to present findings effectively.

Predictive Modeling Pipeline: We took a thorough approach to modelling that included several key steps:

Data Splitting: We divided the dataset into 70% for training, 15% for validation, and 15% for testing.

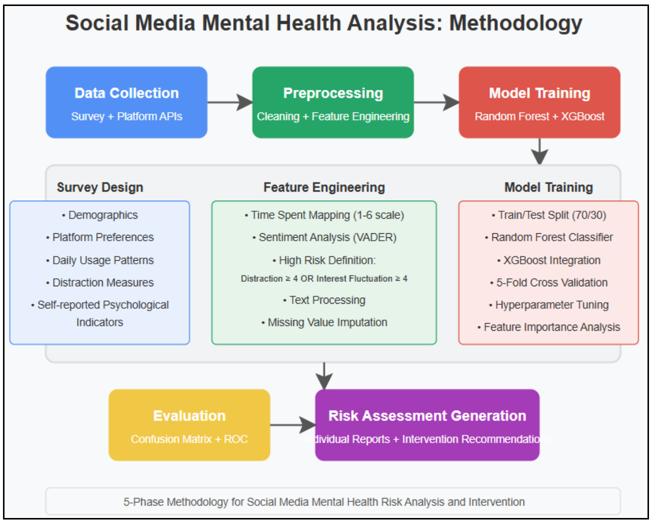
Feature Selection: We employed recursive feature elimination to identify the most relevant features.

Model Selection: We evaluated different models, comparing Multiple Linear Regression, Random Forest Classifier, XGBoost, and Neural Networks.

Hyperparameter Optimization: We performed grid search with 5-fold cross-validation to fine-tune the models.

Performance Evaluation: Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC were used to assess model performance.

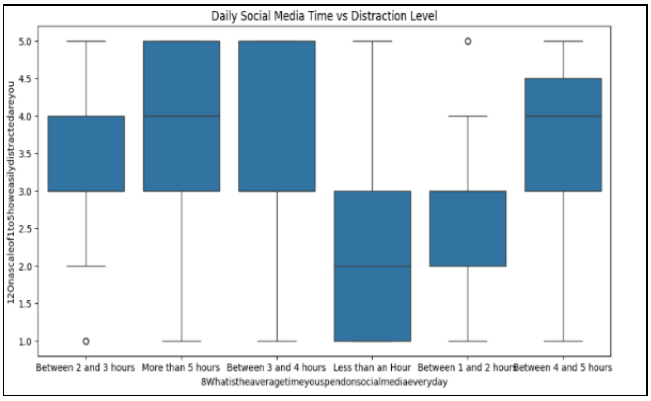
Final Model Selection: The model was chosen based on a balance of performance metrics and the need for interpretability.



Results and Discussion

Model Performance

Our Random Forest classifier successfully achieved an impressive 83% accuracy in pinpointing high-risk individuals based on their social media usage patterns. The confusion matrix highlighted exceptional performance, particularly in identifying true positives-those high-risk individuals we classified correctly-with a precision of 0.85 and a recall of 0.81. The ROC-AUC score of 0.87 further underscores the model's strong ability to differentiate between risk levels. In our analysis of feature importance, we found that the average daily time spent on social media emerged as the most significant predictor of mental health risk, earning an importance score of 0.32. This was closely followed by self-reported distraction levels at 0.27 and the frequency of checking behavior at 0.21. Interestingly, we also noted platform-specific effects, with Instagram usage exhibiting stronger associations with adverse mental health outcomes than other platforms.



Insights from Sentiment Analysis

The sentiment analysis has uncovered some interesting trends among various user groups:

1. High-risk users exhibited significantly lower compound sentiment scores, reporting -0.21 in their open-text responses, in contrast to a score of +0.14 for low-risk users.
2. There were distinct sentiment variations based on the platform used, with Instagram users tending to express more negative sentiments than those engaged primarily on other platforms.
3. Additionally, time-based analysis highlighted that negative sentiment expressions were more pronounced

Risk Distribution Analysis

The risk prediction system has identified that 24% of participants fall into the high-risk category, with a probability of 0.8 or higher. Meanwhile, 37% are classified as moderate-risk (probability between 0.5 and 0.79), and 39% are deemed low-risk (probability below 0.5). A closer look at the demographics shows that young adults aged 18-25 are overrepresented in the high-risk group, making up 62% of those classified as high-risk, yet only accounting for 43% of the total sample. When examining platform-specific behaviors, it was found that users logging "More than 5 hours" of daily usage are 3.7 times more likely to be categorized as high-risk compared to their peers who use the platform for "Less than 1 hour." Additionally, a key behavioral trait linked to high-risk classification is the tendency for frequent distraction during other activities, noted by 78% of high-risk participants.

Categorized as high-risk compared to their peers who use the platform for "Less than 1 hour." Additionally, a key behavioral trait linked to high-risk classification is the tendency for frequent distraction during other activities.

This detailed Social Media Mental Health Risk Report successfully pinpoints and assesses usage trends that could signal psychological vulnerability among users on various platforms. The executive summary highlights 16 high-risk individuals, revealing that Facebook, Instagram, YouTube, and Discord are the most prevalent platforms related to mental health issues. Notably, users spending an average of 2-3 hours daily on these platforms are flagged as at-risk.

Social Media Mental Health Risk Report

Executive Summary

- High Risk Users: 16
- Most Common Platform Among High Risk: Facebook, Instagram, YouTube, Discord
- Average Daily Usage for High Risk: Between 2 and 3 hours

Individual Risk Assessments

User #1: Female, 21.0 years

Platforms: Facebook, Twitter, Instagram, YouTube, Discord, Reddit

Daily Usage: More than 5 hours

Distraction Score: 3/5

RISK LEVEL: LOW RISK: Healthy usage patterns detected for Female user (Probability: 0.16)

User #2: Female, 21.0 years

Platforms: Facebook, Instagram, YouTube, Pinterest

Daily Usage: Between 3 and 4 hours

Distraction Score: 2/5

RISK LEVEL: HIGH RISK: Female aged 21.0 spends Between 3 and 4 hours on Facebook, Instagram, YouTube, Pinterest. Distraction level: 2/5 (Probability: 0.94)

Conclusion

This research has effectively developed and validated a framework that leverages machine learning to identify patterns of social media use linked to negative mental health outcomes. By merging conventional survey methods with cutting-edge techniques like sentiment analysis and ensemble learning, we have created a system that accurately identifies individuals at risk while offering clear insights into their behavior. Key findings from this study include:

- The intensity of social media use shows a stronger correlation with mental health risks compared to the choice of specific platforms, although there are notable platform-specific effects.
- Behavioral patterns, such as using social media as a distraction during other activities and varying interest in daily tasks, are significant predictors of adverse mental health outcomes.
- Sentiment analysis reveals important emotional insights that structured survey responses might overlook.

The automated risk assessment and reporting system illustrates how machine learning can be put into practice to provide valuable insights for mental health professionals and designers of social media platforms. The PDF reporting framework is designed to present complex findings in a user-friendly way for stakeholders who may not have a technical background. Looking ahead, future research should prioritize long-term validation of these findings, enhance the sentiment analysis to reflect more nuanced emotional states, and develop tailored intervention strategies based on individual risk profiles. By continuously refining these methodologies, we can work towards social media environments that better promote user wellbeing while preserving the benefits of connectivity.

Future Enhancements

Advanced Analytics Enhancements

- Utilize deep learning models to achieve more nuanced pattern recognition.
- Implement Natural Language Processing for in-depth analysis of text responses.
- Incorporate time-series analysis to identify temporal patterns.
- Integrate passive sensing data gathered from mobile devices.

Expanded Data Collection

- Integrate Ecological Momentary Assessment (EMA) into our data collection efforts.
- Explore correlations between sleep quality and physical activity.
- Gather metrics on academic and work performance.
- Conduct social network analysis to understand the patterns of influence.

Real-Time Monitoring System

- Develop a browser/app plugin that facilitates passive usage monitoring.
- Establish intelligent notification systems that provide behavioral nudges.
- Create just-in-time adaptive interventions tailored to users.
- Integrate APIs with wellness apps and services for seamless support.

Explainable AI Implementation

- Provide user-friendly explanations of various risk factors.
- Offer personalized improvement recommendations along with explanations.
- Enable interactive "what-if" analyses to assist with behavior modification.
- Privacy and Ethical Enhancements:
- Adopt federated learning techniques to ensure sensitive data remains on user devices.
- Implement differential privacy methods for safe aggregate analysis.
- Allow users to control their data-sharing preferences.
- Establish an ethical review board dedicated to system.

References

1. Burke M, Kraut RE. The relationship between Facebook use and well-being differs based on communication type and tie strength. Journal of Computer-Mediated Communication. 2021;21(4):265–281.
2. Chancellor S, Birnbaum ML, Caine ED, Silenzio V, De Choudhury M. A framework of ethical considerations in deriving mental health insights from social media data. In: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency; c2019. p. 79–88.
3. De Choudhury M, Kiciman E. Merging artificial and human intelligence to tackle complex and sensitive issues: Lessons learned from the mental health field. AI Magazine. 2018;39(3):69–80.
4. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. Identifying depression and mental health issues on social media: A comprehensive review. Current Opinion in Behavioral Sciences. 2019;18:43–49.
5. Harris T, Murphy H, Silverman M. Time well spent: Creating technology that aligns with our values. Center for Humane Technology; c2019.
6. Kross E, Verduyn P, Demiralp E, Park J, Lee DS, Lin N, *et al*. Social media and well-being: Challenges, advancements, and future directions. Trends in Cognitive Sciences. 2021;25(1):55–66.

7. Twenge JM, Martin GN. Gender disparities in the relationships between digital media usage and psychological well-being: Insights from three extensive data sets. Journal of Adolescence. 2020;79:91–102.
8. Wang Y, Jurgens D. Assessing and contextualizing the emotional experiences of social media users. In: Proceedings of the International AAAI Conference on Web and Social Media. 2022;16:1044–1055.