Mental Wellbeing Assessment Through Social Media and Machine Learning – A Web-Based Tool

V. Aravind Reddy (Dept of ECE) IARE,

Dr. S China Venkateshwarlu ,Professor,(Dept of ECE) IARE ,

Dr. V Siva Nagaraju ,Professor ,(Dept of ECE) IARE ,

Abstract :–

Mental health disorders such as depression and suicidal ideation affect hundreds of millions globally, posing serious public health challenges. With the rise of social media platforms, user-generated content has become a valuable source for detecting psychological distress through computational methods.This study presents a comprehensive roadmap for predicting mental health conditions using social media data, focusing on machine learning and deep learning approaches. We examine key components of the predictive pipeline, including data collection, feature extraction, and classification algorithms.

Recent research efforts are reviewed to highlight effective techniques for identifying at-risk individuals based on linguistic, behavioral, and interactional patterns online. The paper also discusses the development of automated detection systems, ethical considerations, and future research directions. These advances support scalable, real-time mental health monitoring and offer potential to augment traditional screening and intervention strategies.

***Key Words*:** Mental healthcare, Mental disorder prediction, Social media analysis, Big social data, Machine learning, Deep learning

1. INTRODUCTION

Mental health disorders, including depression, anxiety, and suicidal ideation, represent a growing global concern, affecting nearly one billion individuals worldwide. The World Health Organization highlights that mental illnesses are among the leading causes of disability, yet they often remain underdiagnosed and undertreated due to limited access to mental health services, social stigma, and insufficient screening tools. Traditional diagnostic methods typically rely on self-reported symptoms, clinical interviews, and structured assessments, which can be resource-intensive and lack scalability for population-wide monitoring.The widespread use of social media platforms offers a novel and powerful medium for mental health surveillance. Individuals frequently express their emotions, experiences, and behavioral patterns through posts, comments, and interactions on platforms like Twitter, Facebook, Instagram, and Reddit. These digital footprints contain rich linguistic, temporal, and visual cues that can reflect a user's psychological state.

Recent advances in machine learning and natural language processing (NLP) enable the analysis of such large-scale, unstructured data to detect mental health indicators with increasing accuracy.This study aims to explore how social media data can be harnessed for mental health prediction using machine learning techniques. We present a structured overview of current methodologies, including data collection strategies, feature engineering, and classification algorithms. Furthermore, we review recent developments, discuss implementation challenges, and propose directions for future research.

By integrating artificial intelligence with digital behavior analysis, this approach has the potential to complement traditional mental health assessments, enable early intervention, and improve outcomes through scalable, real-time surveillance systems.

Table 1. Common mental disorders definition

Mental Disorder	Definition
Depression	There is a difference between depression and mood swings or short-lived emotional reactions to daily experiments. A mental state causing painful symptoms adversely disrupts normal activities (e.g., sleeping). For at least two weeks, the person experiences depressive moods (sad, irritable, empty) or a lack of interest in activities for most of the day or the week.
Anxiety	Several behavioral disturbances are associated with anxiety disorders, including excessive fear and worry. Severe symptoms cause significant impairment in functioning cause considerable distress. Anxiety disorders come in many forms, such as social anxiety, generalized anxiety, panic, etc.
Bipolar Disorder	An alternating pattern of depression and manic symptoms is associated with bipolar disorder. An individual experiencing a depressive episode may feel sad, irritable, empty, or lose interest in daily activities. Emotions of euphoria or irritability, excessive energy, and increased talkativeness can all be signs of manic depression. Increased self-esteem, decreased sleep need, disorientation, and reckless behavior may also be signs of manic depression.
Post-Traumatic Stress Disorder (PTSD)	In PTSD, persistent mental and emotional stress can occur after an injury or severe psychological shock, characterized by sleep disturbances, constant vivid memories, and dulled response to others and the outside world. People who re-experience symptoms may have difficulties with their everyday routines and experience significant impairment in their performance.
Seasonal Affective Disorder (SAD)	In most cases, SAD occurs in the fall/winter and enters remission in the spring/summer, although in some cases, it may happen in the summer and remit in the autumn and winter. A majority of the cause-and-effect mechanisms of SAD have not yet been discovered. However, several hypotheses have been posed regarding the disease, and they promise to deliver new information to scientists.
Schizophrenia	A schizophrenic disorder is characterized by episodes of psychosis that occur continuously or recur continuously. Disorganized thinking, hallucinations, and delusions are some of the significant symptoms.

Existing Block Diagram

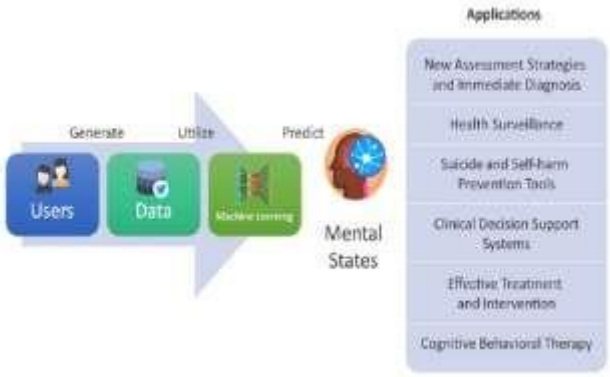


Figure-1

Proposed Block Diagram

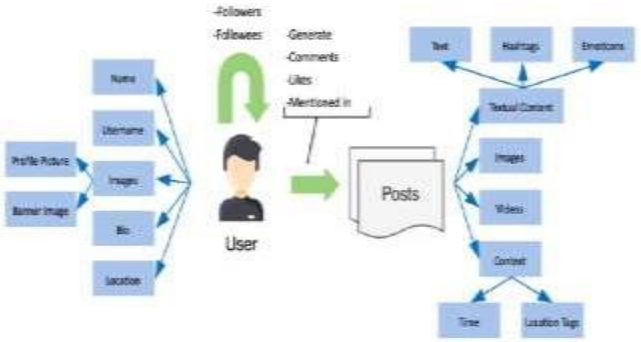


Figure- 2

2. Related Work

2.1 Mental Health and Digital Expression

Previous studies have demonstrated that mental health symptoms can be reflected in online behavior and language. Individuals experiencing depression or anxiety often show distinct patterns in social media usage—such as reduced social engagement, changes in posting frequency, or the use of negative emotional language.

2.2 Social Media as a Mental Health Data Source

Social platforms like Twitter, Reddit, Facebook, and Instagram provide rich, user-generated content. Researchers have used keyword tracking, self-declared diagnoses, forum memberships, and annotated posts to collect data. Reddit and Twitter, in particular, are popular due to their openness and ease of data access via APIs.

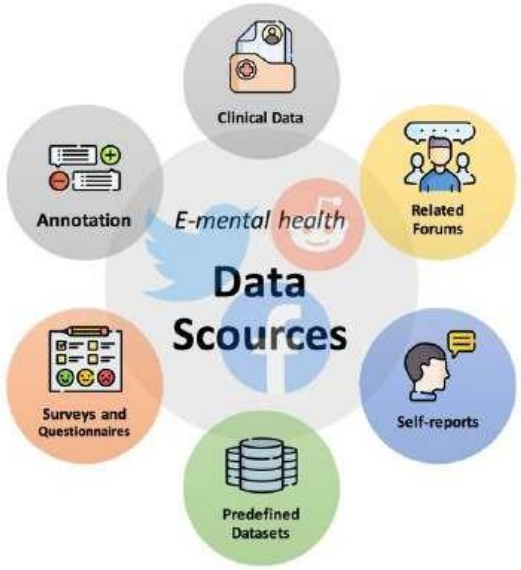


Figure 3. The data source for e-mental health research

2.3 Machine Learning in Mental Health Prediction

Various machine learning models have been employed for predicting mental health states from social media data:

- **Supervised learning** (e.g., SVM, Random Forests, Logistic Regression) has been widely used for binary classification (e.g., depressed vs. not depressed).
- **Unsupervised learning** (e.g., clustering) is applied for discovering latent behavioral groups.
- **Deep learning**, particularly CNNs and RNNs (e.g., LSTM, GRU), has shown promise in capturing complex temporal and linguistic features.
- **Transfer learning** and **reinforcement learning** are emerging trends to improve prediction with limited labeled data.

3. Methodology

The methodology for predicting mental health using social media involves a structured pipeline that integrates data acquisition, preprocessing, feature extraction, and machine learning modeling. The process begins with sourcing relevant social media content, followed by a series of analytical techniques aimed at inferring users' psychological states.

3.1 Data Collection

Data for mental health prediction is primarily sourced from social media platforms such as Twitter, Reddit, Facebook, Instagram, and Sina Weibo. Researchers employ two primary strategies: survey-based data collection and public data harvesting. In the first approach, participants voluntarily complete psychometric surveys such as CES-D, PHQ-9, BDI, SPS, and SWLS, often alongside sharing their social media account information via platforms like OurDataHelps or MTurk. This method provides clinically reliable ground truth but is resource-intensive and limited in scale.

The second strategy leverages self-declared mental health disclosures from public platforms using regular expressions and mental illness-related keywords (e.g., "I was diagnosed with depression"). APIs are used to retrieve posts, after which filters are applied to ensure relevance and to eliminate negations or third-party references. Additional sources include mental health forums such as Reddit's subreddits, which offer valuable insights into individuals' lived experiences with mental disorders.

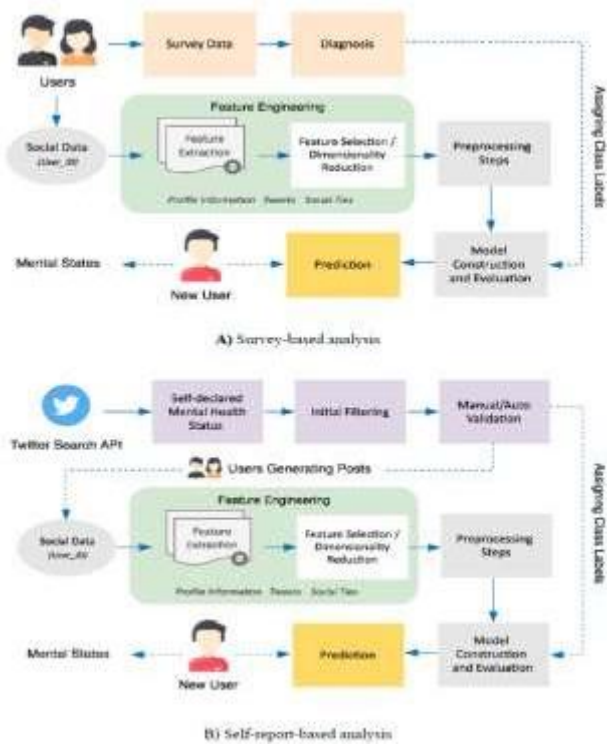
3.2 Preprocessing

After data collection, the raw content is cleaned and prepared for analysis. Preprocessing steps include removing URLs, emojis, usernames, and special characters, converting emojis to ASCII format, anonymizing sensitive data, and normalizing text through lowercasing and stop-word removal. Language detection is also applied to retain only English content, ensuring consistency in linguistic analysis.

3.3 Feature Extraction

Feature engineering involves extracting meaningful indicators from the data to train predictive models. These include:

- **Textual features**: Linguistic Inquiry and Word Count (LIWC) metrics, sentiment analysis (using tools like SentiStrength, OpinionFinder, and VADER), and character/word n-grams with TF-IDF scores.
- **Topic modeling**: Latent Dirichlet Allocation (LDA) is used to uncover thematic content associated with psychological conditions.
- **Visual features**: Image-based signals such as hue, saturation, and brightness (especially from Instagram), processed using CNNs and APIs like Imagga.
- **Temporal and behavioral metrics**: Posting frequency, time-of-day activity, and user interaction patterns provide auxiliary cues.



A) Survey-based analysis

B) Self-report-based analysis

3.4 Model Development

Various machine learning models are employed to classify or predict mental health conditions:

- **Supervised learning** models include Logistic Regression, Random Forest, Decision Trees, and Support Vector Machines (SVM), trained on labeled datasets where output categories (e.g., depressed, not depressed) are known.
- **Deep learning** approaches utilize architectures like LSTM, CNN, and hybrid CNN-LSTM for modeling complex language sequences and image data.
- **Transfer learning** uses pre-trained models (e.g., BERT for text, VGGNet for images) to improve performance in scenarios with limited labeled data.

The feature vector derived from social media content is treated as the independent variable, while the corresponding mental health state serves as the dependent variable. Dimensionality reduction methods such as PCA, gain ratio, and forward greedy stepwise are applied to improve model efficiency and reduce overfitting.

3.5 Evaluation

Model performance is evaluated using cross-validation techniques and metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. For ordinal prediction tasks (e.g., depression severity levels), additional metrics like mean squared error (MSE), mean absolute error (MAE), and correlation coefficients are reported. Confusion matrices are used to identify the types of errors made by the model and inform further refinement

4. CONCLUSION

This project introduced a web-based platform that employs machine learning and natural language processing to monitor mental wellbeing through social media content. By analyzing user posts for emotional cues and behavioral patterns, the system offers early insights into potential mental health concerns such as depression and anxiety.The results demonstrate the feasibility of using digital data for mental health surveillance. However, ethical considerations such as data privacy and algorithmic fairness must be addressed. With future enhancements, this can serve as a valuable supplementary tool for mental health monitoring and awareness**.**

REFERENCES

- Benton, A., Mitchell, M., & Hovy, D. (2017). Multi-task learning for mental health using social media text. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 152–162.
- Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: A critical review. *npj Digital Medicine, 3*(1), 1–11. https://doi.org/10.1038/s41746-020-0233-7
- Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: An integrative review. *Current Opinion in Behavioral Sciences*, *18*, 43–49. https://doi.org/10.1016/j.cobeha.2017.07.005
- Resnik, P., Armstrong, W., Claudino, L., Nguyen, T., Nguyen, V. A., & Boyd-Graber, J. (2015). Beyond LDA: Exploring supervised topic modeling for depression-related language in Twitter. *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 99–107.
- Tsakalidis, A., Ilias, L., & Liakata, M. (2022). Meta-learning for mental health prediction. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1–13.
- World Health Organization. (2022). *Mental health and COVID-19: Early evidence of the pandemic's impact*. https://www.who.int/publications/i/item/9789240047934