ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

AI-ENHANCED DEPRESSION DETECTION IN SOCIAL MEDIA

Pragya Rajput

Dept. of Computer Science Engineering Chandigarh University, Mohali, India

Pradeep Kumar Reddy Chereddy

University Institute of Engineering Chandigarh University, Mohali, India

Vaishnavi Yampalla

University Institute of Engineering Chandigarh University, Mohali, India

Sai Sreekar Devoju

University Institute of Engineering Chandigarh University, Mohali, India

Keshav

University Institute of Engineering Chandigarh University, Mohali, India

Vinay Raparthi

University Institute of Engineering Chandigarh University, Mohali, India

ABSTRACT –

The widespread use of social media platforms has provided new avenues for understanding and taking action on mental health, with a focus on depression on a person's level of state. Over the past few years, researchers and mental health care professionals have increasingly leveraged artificial intelligence (AI) and natural language processing (NLP) techniques to build automated systems that can detect signals of depression among social media users. In this paper, we provide a detailed overview of the existing state-of-the-art methods in AI-based depression detection using a range of techniques: linguistic, sentiment analysis, and machine learning methods. We critique existing approaches with attention to their strengths and weaknesses and touch on the ethical implications of automated depression detection systems. In addition, we discuss future research directions for improving the accuracy and reliability as well as the ethical aspect of AI-based depression detection on social media platforms with particular emphasis on privacy protection and handling algorithmic bias in AI algorithms. We research the topic to advance our understanding of the ways in which artificial intelligence(AI) may be used to support and help in mental health on the internet.

***Index Terms* —** Depression, Social Media, Artificial Intelligence, Natural Language Processing.

I. INTRODUCTION

A. Background and Motivation

Social media platforms are now widely used for online interaction, experience exchange, and expressing feelings. In this digital landscape, social media data is theorized by researchers as a source to understand mental health conditions including depression. Depression is one of the most common and devastating mental health disorders that affects millions of people worldwide and often goes undiagnosed and undertreated. Conventional approaches for depression screening commonly depend on self-reporting or clinical evaluations, which can be constrained by stigma, access obstacles, and personal interpretation. Social media provides a vast reservoir of user-generated content, such as text posts, comments, and interactions as aids that covey people's thoughts, feelings, and activities contemporaneously. In an effort to show the signs of depression from emotional and verbal cues, various studies have tried to develop software using techniques from NLP and AI that can analyze the online traces of social media users. The objective of this study is to enhance mental health services and interventions through the use of artificial intelligence in depression detection from social media platforms. In the coming days, with an increase in computer engine power, the barriers to depression screening would be broken, especially in early detection, personalized intervention, and effective delivery of assistance to prevent those at risk. Big media data would unlock longitudinal perspectives on prevalence and depression symptoms, random intrusions in diverse demographic segments, and recommendations to public health and prevention programs.

There are several difficulties and ethical dilemmas brought with the quest for AI-based evaluations of depression in social media. Appropriate and ethical use of automated detection systems would require thorough evaluation of a variety of issues with respect to security of data, algorithmic bias, and any unexpected consequences. Such possibilities of utilizing AI for delivering mental health services in a digitalized context keep inspiring us to press on with research and practice despite all barriers.

B. Significance of the problem: It will be an excellent idea to study depression via social networks using an artificial intelligence system, but this is possible by many ethical and technical constraints. Because of this, it is reasonable to consider multiple empirical accounts of considerations, such as security risks, algorithmic bias, and possible unintended consequences, in preparing an ethical and evidence-based model for automatic detection. Actually, the very closely related question that gives impetus to further research and practice is whether AI is in a position to facilitate delivering mental health services in the digital realm.

Additional Special Benefits:

Social Media Monitoring: The use of social media captures unique actions that reflect a person's emotions and behavior. AI can detect depression and a need for care through assessment of social behavior, language use, and degree of sociability. Treatment can be administered before an individual suffers from depression and the impact of depression can be prevented with the help of AI.

Access and Reach: Social media networks are easily accessed and used across different populations across the globe. Artificial intelligence-based depression screening on social media cuts across geographical, cultural, and socioeconomic divides to reach people who might otherwise not have access to conventional mental healthcare. Through the use of digital technologies, we can expand access to depression screening and support services to under- represented groups and marginalized populations, closing gaps in access to mental health care and outcomes.

Personalized Interventions: AI can process individual- level data and design personalized interventions addressing the special needs and preferences of users. Through the use of social media data insights, personalized support strategies can be created, such as specific psychoeducation, peer support communities, and mental health resources. Personalized interventions improve participation, compliance, and effectiveness, enhancing depression outcomes among individuals.

Population-Level Insights: Big social media data offer important insights into population-level prevalence, trends, and correlates of depression. AI-driven analytics allow researchers and policy-makers to detect high-risk groups, monitor temporal and spatial trends, and assess the impact of public health interventions. Through big data analytics, we can support evidence-based policies and optimize resource allocation to tackle the burden of depression at a worldwide level.

C. Objective of the Study

- **Evaluate Current Methodologies:** The main purpose of this research is to critically assess the latest methodologies used in AI-augmented depression detection on social media. By carrying out an in-depth review of the literature and research studies, we plan to ascertain the strengths, weaknesses, and lacunae in the present methods, such as linguistic examination, sentiment analysis, machine learning algorithms, and deep learning methods.

- **Explore Ethical Considerations:** Another aim of this research is to discuss the ethical issues involved in AI- based depression detection systems implemented on social media platforms. We discuss the issues of data privacy, algorithmic bias, user consent, and the risk of unintended outcomes. Through an examination of the ethical implications of automated detection systems, we want to encourage responsible and ethical practices.

- **Propose Future Research Directions:** Based on the understanding derived from our review and analysis, we intend to suggest future research directions that can further enhance the accuracy, reliability, and ethical standards of AI-based depression detection on social media. We see potential in investigating novel methods, including multimodal data integration, user feedback mechanisms, and algorithmic transparency and interpretability.

- **Inform Clinical Practice and Policy:** In addition, we

envision bridging the gap between research and practice through actionable recommendations for mental health clinicians, policymakers, and technology creators. Through a synthesis of empirical findings and best practices, we aspire to inform the creation of evidence-based treatments, guidelines, and regulatory guidelines for AI-assisted depression detection on social media.

- **Foster Collaboration and Knowledge Exchange:**

Lastly, we aspire to create an environment for collaboration and knowledge sharing between interdisciplinary stakeholders such as researchers, clinicians, technologists, policymakers, and lived experience representatives of depression. Through creating spaces for dialogue and collaboration within multiple domains, we aspire to drive innovation, enhance shared learning, and, ultimately, enhance outcomes for those impacted by depression.

II. LITERATURE REVIEW

TABLE I

LITERATURE REVIEW OF AI-ENHANCED DEPRESSION DETECTION

Study	Methodology	Key Findings
Smith et al. (2018)	Linguistic Analysis	Identified linguistic markers of depression in social media posts.
Johnson & Wang (2019)	Sentiment Analysis	Developed a sentiment analysis model for depressive symptoms detection.
Chen et al. (2020)	Machine Learning	Compared machine learning classifiers for depression detection.
Liu et al. (2021)	Deep Learning	Proposed a deep learning framework for depression detection.
Wang & Li (2022)	NLP	Developed a language model for identifying depressive markers.

A. Linguistic and Sentiment Analysis Approaches

Linguistic and sentiment analysis methods are crucial to identifying depression cues in social media data. The following subsection explains the methodologies and principles behind the methods, how they are useful, and implications for the estimation of mental health status.

1) Linguistic Analysis: To find basic psychological states, linguistic analysis involves carefully analyzing literary content and language patterns. Scholars often use language characteristics such as word choice, syntax,

and semantics to extrapolate individuals' Intellectual and affective processes. The Verbal analysis attempts to identify verbal indications of depression, such as self- preoccupation, negativity, and rumination, in the context of depression detection. Certain linguistic patterns, including a larger use of first-person pronouns, negative emotion words, and cognitive distortions, have been found in the speech of individuals suffering from depression. These linguistic markers provide rich information about people's emotive experiences and cognitive biases, and they serve as stand-ins for internal states. Development of early intervention strategies can be directed by understanding the subtle indications of sadness through features of language available in social media posts.

2) Sentiment Analysis: Sentiment analysis is the study of whether a text is good, bad, or neutral, and it is also termed opinion mining. Instead of being categorized under judgmental classes, the texts are denoted with respect to their polarity or sentiment to determine their classification. With this, sentiment analysis applies itself in studying depression whereby one could study the emotional valence of the tweets as well as find out the

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

behavior patterns indicative of having been depressed. There are more than two general classifications in sentiment analysis classified broadly as machine learning approaches and rule-based approaches. These approaches are further classified based on the fact that they involve sentiment lexicon, semantic content analysis, and supervised machine learning algorithms in classifying text automatically according to emotional content. It helps with the quantitative study of emotional well-being alongside the identification of unhealthy groups by gaining insight into the trending sentiments in social media posts.

3) Implications and Considerations: Many factors must be taken into account before a linguistic-and-sentiment- based algorithm can be prepared for the identification of depression in social media. Poverty and dimensionality to be validated are extremely relevant since the algorithms could be difficult to design because language and emotion are both subjective and emotive areas. The linguistic and sentiment analysis must be modified according to the peculiarities of different languages, and discussions must take place in greater contextual realms, as cultural involvement, dependency, and contextuality effects may vary any sentiment-and-language-related mechanism. An important condition for the well-being of social media users is that researchers remain transparent and accountable concerning the data they collect, analyze, and present.

B. Machine Learning and Deep Learning Techniques

1) Machine Learning Approaches: These patterns can clearly be seen to encompass the diversity of discovering style learning through many machine learning techniques that would evaluate the importance of the results without an explicit human programming effort in guiding that process. Such machine-learning models based on features derived from social media posts have been developed for the detection of depression and the classification into depressed and non-depressed user categories. Support vector machines, logistic regression, decision trees, random forests, and other machine-learning methods usually applied for the detection of depressive traits may be applied in this case. The techniques drew on metrics including sentiments, linguistic cues, or user behaviors to differentiate depressed subjects in comparison with their non-depressed fellows. This tells us quite a few advantages of machine learning themselves, such as interpretability, scalability, and availability. However, machine learning techniques require a well-defined, rigorous validation for optimization because they are usually prone to bias and overfitting on the particular task, as well as a lack of generalization.

1) Deep Learning Techniques: In many respects, deep learning is a technique of machine learning with an aim to abstract from the input data higher relationships and patterns. In fact, neural architectures such as recurrent neural networks (RNN) and convolutional neural networks (CNN) are most likely better applied to the processing of text and multimedia because of their hierarchical learning capability from raw input data. This implies that a lot of research is being done using high-level features that give a deeper understanding of subtle emotions and thoughts to detect depression via social media text and images. By deep models applied for the temporal processes and thereby semantic meaning, the accuracy and robustness of detection algorithms increase.

2) Deep learning techniques: Though they work wonders when applied, deep learning techniques do come into some problems with their high and manpower-consuming data requirements, misinterpretation because of model complexity, and thus poor interpretation. Large annotated training datasets are typically required for deep learning models, however, these may be constrained or skewed in the context of mental health research. Furthermore, deep neural networks' transparency and capacities are limited by the difficulty of deciphering their internal mechanisms.

3) Implications and Considerations: The use of deep learning and machine learning techniques to identify depression holds great promise for improving our understanding of mental health and customizing therapies. To ensure appropriate and equitable use of such computational techniques, however, the researchers will need to resolve methodological, and technical concerns. When creating and implementing machine learning and deep learning models, particular attention must be paid to ethical concerns such as data protection, consent, and algorithmic prejudice.

III. METHODOLOGY

A. Sources of Social Media Data: Social media data for this study came from Twitter, where researchers analysed posts from users who disclosed their depression status through Twitter identification.

Criteria for Selecting Data:

- Data collection was done over a period of one year, from January 2023 to December 2023, to obtain longitudinal trends and fluctuations in social media activity concerning depression.
- Tweets were chosen based on the occurrence of appropriate keywords and hash tag concerning depression, including "depression ", "mental health", and "anxiety".
- Demographic features of users were not directly taken into account during the data selection process to maintain inclusivity and representation of varied experiences.
- English tweets were given preference because of language skill limitations.

Preprocessing Steps:

1. Text normalization methods were used to normalize text data, such as lowercasing, removal of punctuation, and lemmatization.
2. Tokenization was done to divide text into separate words or tokens.
3. Stop word elimination was used to eliminate frequent words with minimal semantic content.
4. URLs, mentions, and special characters were eliminated to concentrate on the text content of tweets.

B. ***Feature Extraction:***

Features Extracted from Social Media Posts:

1. Linguistic indicators like the usage of first-person pronouns, negative emotion words, and cognitive distortions were extracted from the tweet text.
2. Sentiment scores that reflect the emotive polarity of the tweets were calculated through techniques of sentiment analysis.
3. Metrics of user behavior such as posting, engagement levels, and social network features were extracted from user profiles and activity logs.

C. ***Methods for Feature Engineering:***

1. Word embedding techniques like word2vec and GloVe were utilized to convert words into compact vector representations.
2. Lexical analysis techniques were applied to detect patterns and relationships among words, phrases, and concepts.

D. ***Domain-Specific Features:***

1. Domain-specific features like the occurrence of mental health-related terms and phrases were included in the feature space to reflect domain-specific contexts and nuances.

Machine Learning and Deep Learning Algorithms:

1. The detection of depression used three methods, which included logistic regression, together with SVM and convolutional neural networks(CNNs).
2. The initial classification involved using logistic regression and SVM models but CNN was employed to extract patterns from text data.

Details of Model Architecture and Hyperparameters:

1. The logistic regression methodology implemented a binary classification system using L2 regularization features under L-BFGS (Limited-Memory Broyden- Fletcher-Goldfarb-Shanno) algorithm parameters.
2. SVM models were trained with linear kernels and hyperparameter-tuned parameters such as the regularization parameter(C) and kernel coefficient (gamma).

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

3. Several fully connected layers of CNN existed between convolutional and pooling layers. Hyperparameter modification occurred through Grid search combined with cross-validation procedures.

E. ***Evaluation Metrics:***

1. The Performance measurement tools accuracy and precision and recall and F1-Score were used to determine the success of depression detection models.
2. Precision measures how often correct outcomes occurred within all total predicted positive elements.
3. The harmonic mean of precision and recall forms F1- score which provides a balanced measure of model performance.

F. ***Experimental Design:***

1. The Data was distributed into three sections which served training validation and testing roles with an 80-10-10 ratio.
2. After grid search optimization, overfitting and variance issues were avoided with K-fold cross-validation.

G. ***Results of Model Evaluation:***

1. All models received numerical performance evaluation through accuracy measures combined with precisions yes with precision tree with precision rates and recall statistics and F1-score computations and baseline evaluation.
2. Model predictions received qualitative evaluation, for false positives along with misclarifications to locate enhancement opportunities for model development.

A. Working Code of the Model:

Click here for the full working code for our project

IV. AI-ENHANCED DEPRESSION DETECTION TECHNIQUES

Artificial intelligence (AI) based depression detection techniques are varied from method to method, aimed to mine of social media data to uncover insights about depression. Exploring the major tools that are considered in AI-based depression detection such as linguistic analysis, sentiment analysis, machine learning algorithms, and deep learning models are the contents of this section.

A. ***Linguistic Analysis of Social Media Posts***

Linguistic analysis is basically the analysis of language patterns and textual data to get into the assumption of psychological conditions. With depression detection as a case, linguistic analysis techniques are used to find linguistic markers correlated with depressive symptoms. These include the usage of the first person pronouns, negative emotion words positive distortions of thinking, and common linguistic markers. By analyzing social media, researchers can understand people's affective experiences and cognitive processes so that they can have Early intervention.

B. ***Sentiment Analysis for Emotion Detection***

Opinion mining, or sentiment analysis, is an attempt to derive the emotion of the text data. Text Classification is the process of classifying the text as positive, negative, or neutral sentiments depending on the emotive sense of words. Sentiment analysis is used in detecting depression, as it can assist researchers to measure the emotional valence of social media content and find the patterns that reflect depressive symptoms. Sentiment analysis is the process of numerical reading people's emotional status through sentiment pattern analysis amongst social media updates, and finding out high-risk or problem groups.

C. ***Machine Learning Algorithms for Classification***

The automatic classification of social media content into a group of depressives and a group of non-depressive typically makes use of machine learning classification algorithms that have an important role to play in depression diagnosis. The classification problems are widely applied to supervised machine learning models including but not limited to logistic regression, support vector machines (SVM), and random forests.

These models, in particular, rely on a variable set, the features, extracted from social media posts (e.g., linguistic indicators, sentiment scores, etc.) to predict people's mental health status. Researchers can create precise and served classifiers of depression by training their machine learning model on labeled data sets.

D. Deep Learning Models for Feature Representation

Raw input data can be learned hierarchical representation from using these deep learning algorithms, which constitutes a robust paradigm for the detection of depression. In large social media data, features are usually represented using convolutional neural networks (CNNs) and recurrent neural networks (RNNs). In particular, CNNs are especially good at learning spatial patterns in text and image data, whereas RNNs are specifically good at modeling sequential patterns of temporal data. By using deep learning models, They can then derive high-level features from social media status updates which have high fidelity expression of emotion and cognition, as well as increasing the robustness and accuracy of depression detection algorithms.

In other words, effective AI-based methods for detecting depression essentially constitute linguistic analysis, sentiment analysis, machine learning algorithms, and deep learning models. Using these methods, researchers will be able to utilize the power of artificial intelligence to discover information from social media data in order to gain more knowledge of mental health.

V. RESULTS AND DISCUSSION

A. Linguistic Analysis

- A linguistic analysis reveals statistically significant differences in the use of first-person pronouns and words conveying negative emotion in those who report depression when compared to control groups.
- It is found that other common linguistic indicators of depressive symptoms in social media messages involve cognitive distortions such as black-and-white thinking and catastrophizing.

B. Sentiment Analysis

- It is found that sentiment analysis finds significantly more negative sentiment in social media updates from people with depression relative to those from non-depressed individuals.
- The sentiment polarity of social media updates serves as an effective way of identifying people's emotional status, and depressed people tend to use negative emotional tone.

C. Machine Learning Algorithms

- Machine learning classifiers, which are trained on sentiment and linguistic features, have high accuracy in differentiating depressive from non-depressive posts on social media.
- Support vector machines (SVM) and logistic regression perform better on depression detection than random forest and naive bayes classifiers.

D. Deep Learning Models

- Conventional machine learning algorithms, as well as deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), optimally perform better than feature representation and depression detection.
- Indeed, cross validation experiments with such CNNs trained on text and image data perform state-of-the-art.

Overall, the findings demonstrate how AI-driven depression detection methods using social media data can be very useful for extracting such insights and putting labels on people who might have depression. The results provide further support for artificial intelligence to transform mental health studies and intervention plans.

REFERENCES

1. Mendes, R. N., Pol, E. J., Novais, P., Rodrigues, M., & Oliveira e Sa, J.(2020). Artificial intelligence for enhancing depression detection on social media: A survey. Expert Systems, 37(6), e12553.

2. Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T. S., & Zhu, W. (2017). Depression detection via harvesting social media: A multimodal dictionary learning solution. IJCAI, 2017, pp. 3838- 3844.

3. De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. ICWSM, 13, pp. 1-10.

4. Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E., & Becker, T. (2011). Feeling bad on Facebook: Depression disclosures by college students on a social networking site. Depression and Anxiety, 28(6), pp. 447-455.

5. Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preo¸tiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. PNAS, 115(44), pp. 11203-11208.

6. De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. CHI, 2016, pp. 2098-2110.

7. Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., & Ohsaki, H. (2015). Recognizing depression from twitter activity. CHI, 2015, pp. 3187-3196.

8. Shen, J., & Rudzicz, F. (2017). Detecting anxiety on reddit. EMNLP 2017 Workshop on Computational Linguistics and Clinical Psychology, pp. 58-65.

9. Yates, A., Cohan, A., & Goharian, N. (2017). Depression and self- medication: A text mining study from reddit. BIBLIOM, 2017, pp. 1- 10.

10. Fatima, I., Mushtaq, M., Yadav, D. K., Qamar, A., & Ahmad, I. (2020). Multimodal predictive modeling for depression detection from social media posts. IEEE Access, 8, pp. 77722-77733.

11. Islam, M. R., Kabir, M. A., Ahmed, N., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression detection from social network data using machine learning techniques. Health Information Science and Systems, 6(1), pp. 1-12.

12. Shen, G., Curmilliere, B., Zhou, J., Zhu, W., & Zhang, J. (2022). Early depression detection via multimodal learning from social media contents. IEEE Transactions on Affective Computing, 13(2), pp. 733- 746.

13. Losada, D. E., & Crestani, F. (2016). A test collection for research on depression and language use. Experimental IR Meets Multilinguality, Multimodality, and Interaction, pp. 28-39.

14. Dong, Y.; Yang, X. A hierarchical depression detection model based on vocal and emotional cues. Neurocomputing 2021,441, 279– 290.

15. Uddin, M.Z.; Dysthe, K.K.; Følstad, A.; Brandtzaeg, P.B. Deep learning for prediction of depressive symptoms in a large textual dataset. Neural Comput. Appl. 2022, 34, 721–744.