A Systematic Review on Leveraging Machine Learning and Deep Learning for Early Mental health Depression detection on Social Media Platforms

[1]Alakananda K, [2]Ananth Prabhu G, [3]Melwin D Souza, [4]Vidya VV

[1,2,4]*Department of Computer Science and Engineering, Sahyadri College of Engineering and Management, Mangalore, India*

[3]*Department of Computer Science and Engineering, Moodlakatte Institute of Technology Kundapura, India*

KEYWORDS

RoBERTa ,DistilBT, Electra, GPT-3.5 Turbo 1106 model (proprietary model), LLaMA2-7B(open-source), BXCNN Model, BERT, XGBoost, and CNN

ABSTRACT:

The use of machine learning and deep learning models in detecting depression on social media has become a promising approach to address the challenges in traditional depression diagnosis methods. This systematic literature review examines the recent advancements in the application of these models to analyze social media data for early identification of depressive symptoms. The review highlights the advantages of leveraging user-generated content on social media platforms, as well as the potential of large language models and neural networks in achieving high accuracy in depression detection. However, the review also discusses the need to address challenges related to the interpretability and transparency of these models, and the importance of integrating them with clinical data and comprehensive mental health monitoring frameworks. The findings of this review provide valuable insights for researchers and healthcare professionals interested in utilizing innovative technologies to enhance the early detection and management of depression. Social media platforms have emerged as a valuable source of data for analyzing various mental health conditions, including depression. Researchers have explored the potential of leveraging textual content from social media posts to detect and monitor depressive disorders, as individuals often express their mental health struggles and experiences on these platforms. This paper presents a comprehensive literature review on the current state of research in machine learning models for mental health analysis on social media.

1. Introduction

Lakhs of people are suffering from mental illness due to unavailability of early treatment and services for depression detection. Thus, it is a very challenging task to recognize people who are suffering from mental health disorders and provide them treatments as early as possible. It is the major reason for anxiety disorder, bipolar disorder, sleeping disorder, depression and sometimes it may lead to self-harm and suicide.

The widespread use of social media platforms has provided researchers with a wealth of user-generated data that can be leveraged for early detection of mental health conditions, such as depression (Garg, 2023) [9]. The use of machine learning and deep learning models has emerged as a promising approach to analyze social media data and identify patterns indicative of depressive symptoms.

These advanced computational techniques offer several advantages over traditional depression diagnosis methods (Zhang et al., 2024) [10] (Shah et al., 2024) [11], including the ability to process large volumes of unstructured textual data, detect subtle linguistic and behavioral cues, and achieve high accuracy in predicting the onset of depressive episodes.

This systematic literature review examines the recent advancements in the application of machine learning and deep learning models for the early identification of depression on social media platforms.

2. Methods

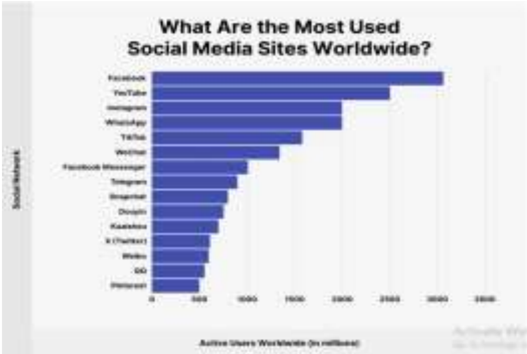


Figure 1: Worldwide active selected users of social media sites

Figure 1 describes the worldwide active users of social media sites [15]. Some of the most popular social media platforms are WhatsApp, Instagram, Facebook, YouTube, WeChat, TikTok, Facebook Messenger, Douyin, Telegram, Snapchat.

Conventional techniques for diagnosing depression, like clinical interviews and self-report questionnaires, are reliable but labor-intensive and require qualified specialists. Social networking sites like WhatsApp, Facebook, Instagram, and Twitter (X) have grown rapidly in the last ten years, opening new opportunities for timely and scalable depression screening. These platforms, which offer a multitude of user-generated material where people regularly share their thoughts and emotions, are intricately woven into the daily lives of a sizable section of the world's population.

Suicide is a huge global public health concern that is significantly influenced by depression. Early detection of suicidal thoughts and symptoms of depression can help avoid such consequences.

Because of their vast and varied user bases, publicly accessible information, and succinct posts that frequently capture individuals' emotions, Twitter and Reddit are especially well-suited for this role.

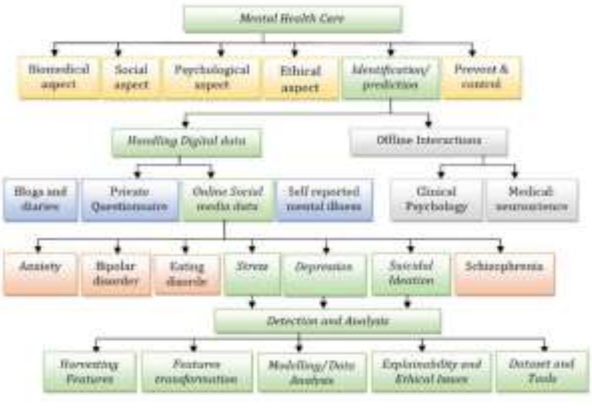


Figure 2: Taxonomy of mental health

Figure 2 describes the taxonomy of mental health [9] in order to automate mental health forecasts, this work presents a taxonomy for mental healthcare and examines it as an interdisciplinary field integrating computational linguistics and human–computer interaction. Biomedical (e.g., ECG, EHR), social (human behavior), psychological (mental health theories), and ethical (data security) factors are important components of mental healthcare. Due to stigma and restricted clinical access, traditional techniques of identifying mental health issues are becoming less common. As a result, digital mental health approaches that use internet data such as blogs, self-reports, and social media are becoming more popular [50]. In order to identify stress, depression, and suicide risk, the emphasis is on text-based data from websites such as Twitter, Reddit, and Sina Weibo. Because of their disparate semantics, multimodal sites (like YouTube and Instagram) are not included. Six more mental health issues are also examined in social NLP research, with the most extensively researched being stress, clinical depression, and suicide risk [51]. AI developments have sparked additional study and possible industrial uses.

Social Media Data Features:

Unstructured, user-generated content makes data curation for mental health diagnosis on social media difficult. While feature extraction research has helped identify neuropsychiatric diseases from self-reported text, classifier advancements have enhanced natural language understanding for mental state inference. Commonly used platforms include Facebook, Instagram, Reddit, and Twitter; research on these platforms focuses on social aspects and natural language processing (NLP) rather than multimodal data. Both cross-sectional and longitudinal research benefit from the use of feature extraction techniques to find prevailing trends [52, 53,54].

Learning-based models based on four feature classes—user-profile, linguistic, social, and multimedia features—are used in cross-sectional studies to infer mental states from social media data. Social NLP research focuses on text and image modalities, and textual features are extracted using either traditional (handcrafted linguistic and semantic features) or automated (vector representations in pre-trained models) methods.

In this literature review, methodology of this study includes a detailed description of the RoBERTa, DistilBERT, Electra, GPT-3.5 Turbo 1106 model (proprietary model), LLaMA2-7B(open-source), their fine-tuning process, and their use for depression detection from the users' social media data. Figure 1 illustrates the different methodology used for depression detection.

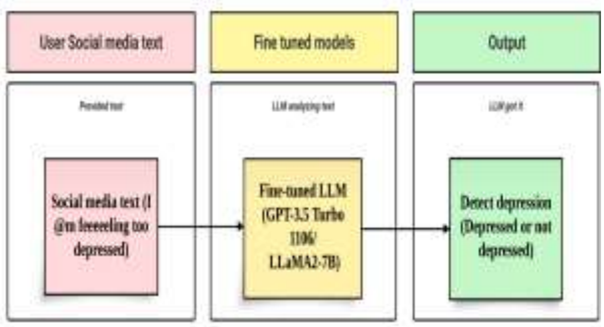


Figure 3: General Process

In this study, the improved GPT-3.5 Turbo 1106 (proprietary) and LLaMA2-7B (open-source) models are used to investigate depression detection. These algorithms use text data from social media to categorize posts as either depressive or non-depressive. Figure 1 shows the details of the model and methodology [27, 49].

According to Figure 3, social media text data is provided to fine-tuned RoBERTa, DistilBERT, Electra, GPT-3.5-Trurbo 1106 [26] and fine-tuned LLaMA2 7B [27] models that in turn recognize the provided text as a depressive or non-depressive post. A detailed description of each of these models is provided below:

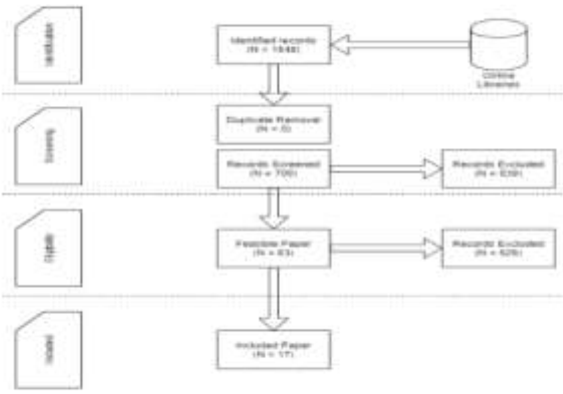
Mental health care	Identification/ Prediction	Online Social media data	focuses on detecting expressions of anger, fear, joy, and sadness, scoring these emotions	Identify Depressed/ Not depressed

Figure 4: General Extraction Process

In this figure 4: first column indicates mental healthcare indicate human-computer interaction to automate the predictions, 2nd and 3rd column indicates identification and prediction of mental health depression using social media platforms, it focuses on detecting emotional expression like anger, fear, joy, and sadness, scoring these emotions and detect the person is having depression or not. It deals with the text data analysis and classification from social media platform data (Twitter, Reddit, Sina Weibo). Also, among 9 mental health problems, stress, tension, clinical depression and suicidal risk detection are most widely studied areas on social media platforms [28].

Figure 5: Research Process PRISMA framework

In this model [114], the PRISMA framework is used in this literature review to arrange the research process in an organized manner. Using the query ["depression" OR "depressed"] AND ["mental" OR "disorder"] AND ["detection"] AND ["social media"], articles were obtained from five databases: ScienceDirect, ACM Digital Library, IEEE Xplore, Springer Link, and Emerald Insight. Articles written in languages other than English, those released prior to 2015, duplicate research, and unrelated depression-related subjects were also excluded. Abstracts were examined for their applicability to the study's goals, and eligible records were further reviewed for full-text availability. Included were only studies that addressed the predetermined study question [114]. Table 1 provides a summary of the



outcomes of this procedure.

Table 1: Difficulties in text-based detection

No	Issue	No of papers	Study identified
1	No means to verify the data validity	1	10
2	Data unable to provide sample characteristics	2	10,19

3	Ethical concerns	6	10, 6, 17, 18, 20, 21
4	Insufficient data	4	11, 12, 18, 19
5	Stigma and/or lack of awareness	3	14, 20, 21
6	Severity of error	2	6, 17
7	Oversimplification	1	21

These are the challenges faced in doing text-based approach for depression detection

Table 1 demonstrates that the most prevalent ethical issues in the chosen studies are privacy protection and data accessibility. Data pertaining to mental health is regarded as extremely sensitive even though it is publicly available.

Table 1 identifies ethical issues as a significant obstacle, especially with relation to data accessibility and privacy for sensitive mental health information. Due to the difficulty of confirming users' depression state, API data may be relied upon, perhaps resulting in subpar sample quality. The research process is complicated by the interconnectedness of many difficulties [114].

Out of the original 1548 data, 17 articles on text-based depression detection were chosen for analysis in this systematic review. The stigma associated with mental health, data scarcity, and ethical considerations are among the main problems noted. Classifiers, Support Vector Machines, and Probabilistic Classifiers are popular approaches; the BiLSTM + Attention method produces the best results. Furthermore, an experiment using a refined BERT-based model that included text summarization for lengthy sequences produced better results, although more improvement is required. In this paper architecture diagram, they have collected different papers and compared with different algorithms and what are the challenges faced for detecting depression through social media text-based approaches [114].

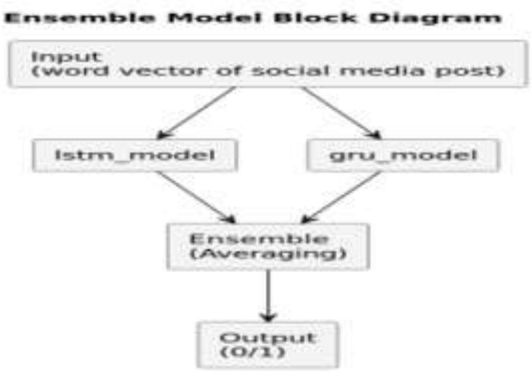


Figure 6: Ensemble architecture

In this model describe the architecture of ensemble model of multilingual languages.

This study looks at social media depression detection in eight Indian languages with limited resources. In order to integrate the advantages of the LSTM and GRU models, it suggests an ensemble model. By overcoming data constraints, a translation strategy achieves competitive accuracy and precision, and the ensemble model performs well [48].

Deep learning methods, particularly transformer models like BERT and GPT, have demonstrated great promise in the analysis of social media data for the identification of depression.

These models perform better than more conventional machine learning techniques like logistic regression and support vector machines [55, 56], which necessitate intensive feature engineering, since they can catch subtle linguistic patterns.

In order to identify depressive symptoms, the study intends to create a large dataset of tweets tagged with depression markers and apply both transformer-based and conventional machine learning models. The relative usefulness and efficacy of these models in automated depression screening will be ascertained through a comparative analysis.

The study consists of a performance comparison, model creation and training, and a review of previous research. The study intends to solve the growing mental health crisis in the digital age by developing precise, scalable methods for mental health monitoring and early intervention by utilizing the power of transformers and real-time social media data. Social media's worldwide reach, and continuous data flow offer a previously unheard-of chance to identify mental health problems and guarantee prompt assistance for people in need.

3. Related Works

Mental health is a crucial aspect of overall well-being, and depression is one of the most prevalent mental health disorders worldwide. Given the heavy toll of depression and its widespread prevalence, there is a growing interest in developing computational methods to detect depression using data from social media platforms.

Muzafar et al.2023[1] highlight that the wealth of personal expression shared on ubiquitous social media platforms presents an opportunity to develop machine learning models capable of detecting patterns indicative of depression and facilitating early intervention.

The paper employs AI techniques, specifically machine learning and natural language processing, to analyze social media data (Mansoor & Ansari, 2024) [2]. This reflects a broader trend in mental health research towards utilizing AI for tasks such as risk assessment, diagnosis, and personalized treatment recommendations. The authors' multilingual, multi-platform approach is particularly noteworthy, addressing a key limitation of previous studies that often focus on single platforms or languages.

Mansoor and Ansari developed a multi-modal deep learning model that achieved 89.3% accuracy in detecting mental health crises, demonstrating a predictive lead time of 7.2 days prior to expert identification. Linguistic patterns and behavioral changes were key indicators in this model's prediction [2].

Social Media and Mental Health

Social media's explosive growth and broad use have completely changed how people communicate, express themselves, and share their experiences. Because social media data provide important insights into users' emotional states and behavioral patterns, this digital transition has opened up new avenues for mental health research and interventions [116, 117]. As digital extensions of people's social life, platforms like X, Facebook, and Reddit—which have billions of active users—have become essential to contemporary culture. They make it possible to share ideas, emotions, and experiences in real time. These networks' vast amounts of user-generated material offer a wealth of data for researching and tracking mental health issues, such as depression [118,119].

Users' emotional and mental states are frequently reflected in social media material [120,121]. People who are depressed may exhibit more negative mood, hopelessness, and less social interaction

[118,119]. Examining these trends over time facilitates prompt interventions and aids in the detection of early indicators of mental health problems [118,119].

Machine Learning for Depression Detection

Machine learning is being used more and more in mental health research to identify depression in texts from social media. To find depressed signs, supervised algorithms such as SVM, Decision Tree, Naive Bayes, and Random Forest examine text data from websites like Facebook and Twitter [122–126]. For the investigation of depression, these models record visual information, affective patterns, and language clues. Studies utilizing classifiers like KNN, Adaboost, and logistic regression have shown great accuracy, proving the effectiveness of sentiment analysis and behavioral pattern analysis [14,125]. Nevertheless, difficulties include limited model interpretability and laborious manual feature engineering [120,121]. Notwithstanding this, machine learning provides scalable, real-time insights into mental health, and developments in deep learning hold promise for increased precision and effectiveness in the diagnosis of depression [127]

Deep learning for mental health

Deep learning has been used to identify mental health problems in social media and clinical data [20]. Neuroimaging EEG, and EHR are examples of clinical data types that are utilized to diagnose disorders such depression, schizophrenia, and ADHD [21–25]. CNNs and RNNs are two examples of techniques that have been used for feature extraction and predictive modeling. For example, EEG has been used to diagnose depression, and RNNs have been used to analyze EHRs in order to predict depressive episodes [11,26]. In order to detect depression, deep learning models have also examined unstructured clinical notes [12].

Heterogeneous Subgraph Networks:

Chen et al.2024 [3] employed prompt learning to improve the interpretability and user interaction modeling of their depression detection method. This approach significantly outperformed existing techniques.

The BXCNN model, a novel approach to depression detection, integrates BERT, XGBoost, and CNN to effectively extract features and address data imbalance,

resulting in improved accuracy for predicting depressive symptoms (Xi et al., 2024) [4].

Multi-Modal Feature Fusion: This method made use of cross-attention mechanisms and achieved an accuracy of 94.95%, which demonstrated a strong potential for detecting hair species by combining various types of data (Li & Xiao, 2024) [5]. Though the potential of social media data in mental health monitoring is manifested in these advancements, ethical issues including privacy and the risk of stigmatization are still key challenges to overcome in future studies (Mansoor & Ansari, 2024) [2].

This approach, posts, comments, and multimedia data could also be analyzed using machine learning and deep learning models to identify depressive symptoms. Methods such as decision trees and convolutional neural networks improve early detection and assist expert decision making in the field of mental health. A recent literature review shows that depression is a worldwide problem and the use of machine learning (ML) and deep learning (DL) techniques allows for the automatic detection of depressive symptoms on social media (Wadzani et al. 2024) [6].

In this paper focuses advanced machine learning techniques, such as LSTM (long short-term memory), CNNs (convolution neural networks). Using these methods, it analyzes the language and temporal context of tweets and analyzes psychological states of individual (Kartik et. al 2024) [7].

This paper contributes to this field by proposing a systematic approach to classify depression levels based on social media activity. In summary, the paper discusses depression detection through social media, with a focus on machine learning and NLP techniques to enhance early detection and intervention strategies (Dhanashree, Patil. el al (2024)) [8].

The study involves automated depression detection methos like social media audio, video, posts. The research survey uses existing ML methods, analyzes over 140 related articles, and processes 80 for comparative analysis. It also discusses various systems for depression detection and future directions and identified open challenges. In this it highlights the early detection of depression symptoms and different techniques used to assist for diagnosis and treatment [12].

The literature review focuses on analyzing social media text data, particularly post-Covid, it uses ML techniques to detect and extract emotions related to depression. It highlights the challenges involved in pre-processing the text data and applying algorithms to uncover hidden emotional layers [13].

In this paper utilized machine learning techniques like Naïve Bayes and Random Forest on a Kaggle dataset to see the comparison of their accuracy in predicting three emotions associated with depression. The results intend to lead to the right counseling interventions, that may prevent exacerbation in mental health problems [14].

In this paper, we point out the trend to use social media data like Twitter for depression detection with machine learning algorithms. It draws parallels between the rise of social media and the increased rates of depression. However, limitations of this paper is diverse feature extraction from user activities to enhance accuracy and F-measure scores. While the approach shows promise, challenges remain in ensuring comprehensive data representation and addressing privacy concerns associated with analyzing personal social media content for mental health predictions [14].

In this paper it uses various machine learning models, such as logistic regression, random forest, naïve bayes also advanced deep learning models like RoBERTa and DeBERTa which excel in capturing contextual nuances in language (Bokolo & Liu, 2024) [15]. The paper evaluates machine learning and transformer models for depression detection in social media, highlighting that transformer models like RoBERTa and DeBERTa outperform traditional algorithms, while logistic regression excels in simpler datasets, contributing to mental health monitoring advancements [15].

In this paper it identified convolution neural networks and recurrent neural networks in detecting depressive symptoms from various datasets, such as Twitter & Reddit (Gadzama et al., 2024) [6]. It identifies 28 machine learning articles focused on depression detection on social media, using technology like decision tree, KNN, and SVM. It emphasizes the importance of deep learning model for improved diagnosis and prediction [6].

Emotional intensity analysis has been utilized to quantify expressions of various emotions, providing insights into users' mental states (Myee et al., 2024) [17].

In this paper (Myee et al., 2024) [17], focuses on depression detection in social media post through emotional intensity using AI & ML models like RoBERTa, DistilBERT, and Electra, emphasizing their effectiveness in identifying mental health issues by analyzing emotional expressions in textual data [17].

BERT and its variations performed very well in detecting signs of depression. They achieved impressive scores of 99.29% for macro F-measure and 99.56% for micro F-measure, showing their strong capability in identifying depression-related cues (Balcı & Essiz, 2024) [18].

The BERT model showed very high accuracy with 99.29% macro and 99.56% micro-F-measure values. It performed better than traditional algorithms like SVM, Naive Bayes, and Random Forest, which had difficulty with sentiment analysis in unbalanced datasets. This demonstrates BERT's effectiveness in detecting depression [18].

Support Vector Machine (SVM): Achieved 71.14% accuracy using metadata, indicating its potential in specific applications (Begum & Sait, 2024) [19].

This shows that SVM could be useful in certain applications

In this paper various ML models were used, SVM, Logistic Regression, Random Forest & KNN. SVM has the highest accuracy of 71.14% using Glove embedding. Other models' performance was not specified [19].

The recent advancements in the application of machine learning and deep learning models to analyze social media data have shown remarkable progress in the early identification of depressive symptoms.

A study by (Hassantabar et al., 2021) [20] in this study used neural networks & DL models for detecting mental health disorders, such as depression detection, using both clinical and social media data. These models were able to achieve higher levels of accuracy.

A study by (Bokolo & Liu, 2024) (Bokolo & Liu, 2023) [21], here it used 2 models RoBERTa and DeBERTa, these models

good in contextual nuances of languages and achieving highest accuracy for depression detection tasks. The model RoBERTa achieved an accuracy of 0.981%. Depression detection from tweets data. This model identified as best model for early detection of mental health depression detection from social media platforms, while compared to DeBERTa, [21].

In this paper it uses a machine learning model Naïve Bayes, SVM to classify depression level from social media post. The system achieved 74% accuracy in classifying the depression level based on sentiment analysis according to their post. It is used leveraging NLP and beck depression inventory II (BDI-II) [37].

Alqazaaz et al. [38] here it compares the performance of machine learning model with deep learning LSTM networks for the depression detection through twitter data. LSTM is achieving better results.

The social media platforms like Facebook and Reddit have lakhs of active users. These platforms enable the users to share their thoughts, feelings, and experience. The huge, generated content on these sites serves as a data source for monitoring and understating the mental health conditions, like depression [39,40].

In [41], this paper describes a model for depression analysis by developing correlations between depressive indicators and textual indicators.

The author (S. AlSagri & Ykhlef, 2020) [42] examining the potential of machine learning models to identify depression among twitter users. Here the users diagnosed with depression will be posting negative emotional sentiments. And it used SVM classifier, which achieved 70% accuracy.

The systematic review by the author (Liu et al., 2022) [43] used various machine learning models, which is used for text mining techniques. This research states that machine learning models are achieved high accuracy and reliable model for early detection of depressive symptoms.

The author (Safa et al., 2023) [44] focuses on assessment for future developments in mental health, also states the importance of mental health state languages. In this paper describes the feature extraction processes provide a valuable framework for ongoing research.

Moreover, the author (Nasrullah & Jalali, 2022) [45] and (Gupta et al., 2022) [46], focused on the importance of preprocessing textual data & the psychological analysis of languages. These studies state linguistic features and detect depressed and non-depressed individuals. The (Ahmed et al., 2022) [47], author describes the challenges and opportunities presented by social media data for depression detection.

In this paper it proposes a model LSTM & GRU for depression detection across 8 low-resource Indian languages, future work focuses to improv the dataset and different model and advanced Architecture [48].

Machine learning algorithms are being used more and more in mental health research to identify depression in texts from social media. Research has shown that supervised algorithms can detect minor emotional patterns and linguistic clues that indicate depressive symptoms in user-generated material by analyzing data from social media sites like Facebook and Twitter [57].

Machine learning (ML) approaches for identifying depression in social media data have been studied recently [128]. Through text analysis, studies have demonstrated the potential of supervised algorithms such as SVMs, Decision Trees, Naïve Bayes, and Random Forest in processing massive amounts of social media data for mental health monitoring [60]. Additionally, some researchers have used behavioral pattern analysis and sentiment analysis, showing great accuracy in predicting depression levels. For instance, Obagbuwa et al [58]. employed classifiers like KNN and Adaboost to identify depressed symptoms in tweets, whereas Angskun et al [59]. used logistic regression and SVMs.

Despite their efficacy, standard machine learning techniques necessitate manual feature engineering, which takes a lot of effort and could miss subtle linguistic indicators of sadness. Furthermore, these models frequently lack interpretability, which makes it more difficult to comprehend how depression indicators are found.

However, machine learning approaches have made considerable progress in mental health research by providing effective and scalable tools for early depression identification. By increasing the precision and breadth of social media-based depression analysis, emerging deep learning algorithms hold the potential to overcome current constraints.

In this paper, by correcting asymmetrical data distribution through a one-shot judgment strategy, this work tackles bias in the diagnosis of depression from social media posts [16]. With a 98.05% classification accuracy, an ensemble model that combines SVM and KNN with noisy label correction is suggested to differentiate between suicidal thoughts and depression symptoms. The model guarantees objective and comprehensible data classification [16].

In this they state that [49], three methods developed by Team6 to forecast chatbot response quality in DSTC 11 Track 4 using LLMs such as ChatGPT and Llama 2 are presented in this study. Dynamic few-shot examples were used to achieve improvements over the baseline, and an ablation research revealed discrepancies between Llama 2 and ChatGPT's use of these instances [49].

This study examines 37 papers on machine learning-based social media suicidal ideation detection, emphasizing the technology's potential to stop suicide and self-harm. It offers academics a thorough resource for classifying suicidal texts [93].

This paper state that 18 emotions associated with stress are identified in this work, with a focus on anger, fear, and melancholy. Complement Naïve Bayes and Bi-LSTM with GloVe embeddings are used to identify stress-inducing emotions in text. With accuracy rates of 88.40% and 89%, respectively, the suggested techniques surpass current algorithms [103].

This study investigates the use of deep learning methods, such as CNNs, for facial expression-based emotion recognition in videos [124]. The FER2013 dataset was used in experiments to detect depression, with 97% training accuracy and 57.4% testing accuracy [124].

PERFORMANCE METRICS

In the performance metrics, social media textual analysis these are the model used to analyse textual data for depressive language pattern and classification it used NLP techniques, (transformer-based models like BERT & RoBERTa, TF-IDF, n-grams) it is achieving high accuracy in identifying depressive sentiments from social media data. Pandey & Kumar, 2024) [22] (Agoylo et al., 2024) [23].

In this study (Lu & Koehn, 2024) [24] it used LLMs – Large language models like LLM2Vec for textual analysis and classification for depression detection from social media data by analysing linguistic cues. Also, it analyses Multimodal data integration, including text, audio, and videos. This approach allows more comprehensive analysis for mental health depression detection of a patient, nuances that may be missed by text-only models (Lu & Koehn, 2024) [24].

Haq et al., 2024 [25] the author used advanced machine learning techniques, such as AdaBoost Classifier (ABC), SVC – Support vector classifier. These models are used to detect the depression levels based on textual data like questionnaire PHQ-9. By using these model state that it shown high accuracy in classifying depression and early detection of depression symptoms.

As evident from studies in the past, social media platforms have strong association with feelings expressed by users [29–31]. About 8 out of 10 people tend to disclose their suicidal tendencies on social media [20]. Mental health prediction from social media [32] facilitates suicidal risk assessments [33] and early detection of suicidal tendencies by using emotion spectrum from social media user's historical timeline [34] due to the presence of Papageno effect [35]. Such path-breaking developments intensifies faith in developing learning-based mechanisms to capture mental health levels using language.

4. Results & Discussion

Table1: different models and their description

Algorithm	Description
Support Vector Machine (SVM)	The most popular methods used in the set of selected articles are the usage of classifiers, support vector machine (SVM), and probabilistic approach (Bayesian, Hierarchical Dirichlet Process, etc.)
RNN on the early detection of depression	difficulties finding proper context from a word in a long sentence
BiLSTM + Attention model	textual data best result is obtained by using BiLSTM + Attention model
Recurrent Neural Network Model (RNN)	difficulties finding proper context from a word in a long sentence. on the early detection of depression
Random Forest	94.9%
Bernoulli Naive Bayes	90.1%
Logistic Regression	97.0%
RoBERTa	98.0%
DeBERTa	98.0%
DistilBERT	97.0%
SquuezeBERT	95.0%

NAÏVE BAYES DEPRESSION PREDICTION	88%
GPT-3.5	GPT-3.5 Turbo 1106 has fine-tuning capabilities, and we can fine-tune it on different datasets to perform different tasks
Turbo 1106	Turbo 1106 is equipped with an appropriate and advanced understanding of the language
LLaMA2-7B	offers robust fine-tuning capabilities, allowing it to be adapted for specific tasks. s like sentiment analysis, language translation, and more.
Multinomial Naive Bayes	accuracy of 83%
Decision Tree (max_depth = 5)	accuracy of 91%
K-Nearest Neighbours (K = 7)	gave the lowest accuracy of 67%
Long Short-Term Memory(LSTM)	accuracy of 92%
Universal Sentence Encoder (USE)	performing best with an accuracy of 98%.
Sequential Model	accuracy of 87%

With 98.0% accuracy and F1 scores, transformer models such as RoBERTa and DeBERTa scored noticeably better on the Sentiment 140 Dataset than conventional machine learning models. With an accuracy and F1-score of 97.0%, logistic regression outperformed transformers among conventional models. In contrast, Bernoulli Naive Bayes performed somewhat well across all metrics, whereas Random Forest had high precision with somewhat worse recall.

As indicated in Table 1, a unique technique for assessing depression levels has been created based on the Beck Depression Inventory-II (BDI-II) questionnaire [61]. A person is classified as non-depressed if their score falls between 1 and 55%, which encompasses normal, mild, and borderline depression. A depressed state is indicated by a score more than 55%.

Table 2: text-based approach for depression detection algorithms.

No	Method	Study identifiers
1	Scikit-Learn Toolkit machine classification	10
2	Classifiers	11, 14, 19, 23
3	Support Vector Machine	13, 14, 15, 22
4	Random Forest	14,15
5	Probabilistic classifier	14, 15, 22, 24
6	Clustering Approach	15, 24
7	Association rule mining	19

8	Logistic	22
9	Gaussian Process	22
10	Term Frequency – Inverse Document Frequency	25
11	BiLSTM + Attention	26

The given Table 2 is indicating the most effective text-based approach for early depression detection.

Table 3: multilingual language models

Paper	Language	Accuracy	Precision	Recall	Model
[62]	Bengali	0.86	0.85	0.86	BiGRU
[63]	Bengali	0.757	-	-	GRU
[64]	Bengali	90.3	0.694	0.547	BiGRU
[65]	Gujarati			0.84	Linear SVC
[66]	Hindi			0.78	Linear SVC
[66]	Hindi	0.8321	0.8323	0.8321	CNN-BiLSTM
[67]	Hindi	0.7143			BERT
[68]	Malayalam	0.7633	0.7607	0.7633	GRU
[69]	Tamil		0.744	0.645	BERT
[70]	Malayalam		0.647	0.69	BERT
[70]	Malayalam	0.8428	0.6303	0.6304	LSTM
[70]	Tamil	0.5293	0.4232	0.5272	LSTM
[70]	Kannada	0.5745	0.5062	0.5455	LSTM

The table 3 describe multilingual language, this study uses precision, recall, accuracy, and F1 score to assess sentiment analysis models across Indian languages. Performance was improved by using a translation-based dataset building strategy; languages such as Hindi, Gujarati, and Kannada achieved F1 scores above 0.90. The model successfully analyzed sentiment in a variety of languages, demonstrating a high degree of linguistic sensitivity. Reliance on data quality and linguistic variances in the translation methodology are among the drawbacks, which could affect generalizability. Furthermore, the LSTM and GRU components of the ensemble model impart certain constraints. In spite of this, the findings show that multilingual sentiment analysis has advanced significantly.

Table 4: Accessible datasets for different Language of Mental Health

Year	Dataset	Psychological Outcome	Paper
2015	CLPsych [71]	Suicide Risk	[71–74, 75, 76, 76, 77]
2017	MDDL [39]	Depression	[78–80, 81]
2017	RSDD [84]	Depression	[82, 81, 83, 84]

2018	SMHD [85]	Mental Health	[85.86]
2018	eRISK[88]	Stress	[87, 80, 88]
2018	Pirina [90]	Depression	[89, 90]
2018	Ji [91]	Suicidal id.	[92, 52]
2019	Sina Weibo [95]	Suicide Risk	[94, 95]
2019	Dreaddit [96]	Stress	[97, 98]
2019	SRAR [99]	Suicide Risk	[99]
2019	Aladaug [100]	Suicidal Id.	[92, 100]
2020	UMD-RD [101]	Suicide Risk	[101, 102, 99]
2020	GoEmotion [112]	Emotion	[97, 104]
2021	SDCNL [105]	Suicide/Depression	[105]
2022	CAMS [106]	Mental Health	[106, 107]
2022	RHMD [108]	Mental Health	[108]
2022	Kayalvizhi [111]	Depression	[109, 110]

The scientific community has previously seen the use of publicly accessible datasets like the Language of Mental Health [113], Reddit Self-reported Depression Diagnosis [84], CLPsych shared task [71], and early risk prediction on the Internet (eRISK) from CLEF Forum [88]. As was previously mentioned, very few datasets are in the public domain, while many are replicable or accessible upon request. We encounter around a dozen datasets each year that use social media data to predict mental health. Due to their limited availability, we have chosen to use either the most widely used and replicable dataset or those that can be obtained upon request or through a signed contract. Table 4 contains a list of replicable datasets. These are the different dataset used for detecting mental health depression, suicidal risk and stress of a person.

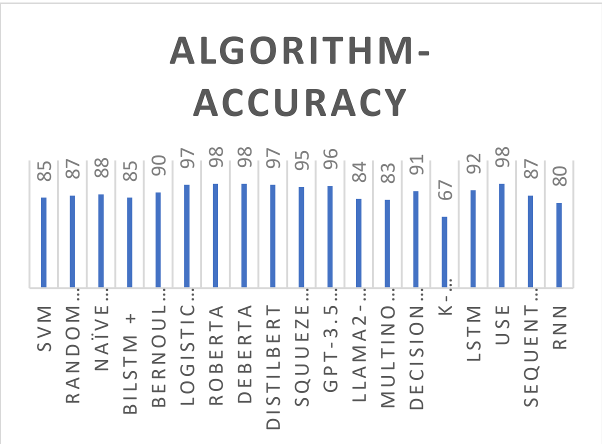


Figure 7: Comparative analysis of different ML model accuracies.

The accuracy comparison of the several techniques used to train the dataset is shown in Figure 7, SVM,RF,Naïve bayes, BiLSTM +, Bernoulli Naive Bayes, Logistic Regression, RoBERTa, DeBERTa, SquuezeBERT, NAÏVE BAYES, GPT-3.5 Turbo 1106, LLaMA2-7B, Multinomial Naive

Bayes, Decision Tree, K-Nearest Neighbours, Long Short-Term Memory, Universal Sentence Encoder, Sequential Model these are the different algorithm is used in the different papers. Comparing to all model these three model RoBERTa, DeBERTa, Universal Sentence Encoder attains the highest accuracy among them. Bernoulli Naive Bayes, SquuezeBERT, GPT-3.5 Turbo 1106 perform competitively as well, coming very close to RoBERTa. The accuracy of the remaining three algorithms, however, is noticeably lower.

The evaluation methodology used accuracy, to compare three NLP models: RoBERTa, DistilBERT, and Electra. With the best recall (0.93) and F1-score (0.91), RoBERTa showed excellent true positive detection and struck a compromise between recall and precision. With an accuracy and F1-score of 0.86, DistilBERT had strong overall performance; however, its precision was lower at 0.77. There is a trade-off between detecting real positives and preventing false positives, as evidenced by Electra's somewhat worse precision and recall despite its F1-score of 0.84. The comparison shows each model's advantages and disadvantages in terms of textual depression detection.

Using Deep Learning models like the Universal Sentence Encoder, LSTM, and Sequential Model in conjunction with Machine Learning models like Multinomial Naive Bayes, Decision Tree, K-Nearest Neighbor, and SVM, RoBERTa, DeBERTa, SquuezeBERT, BiLSTM +, Bernoulli Naive Bayes, Logistic Regression, RF, this study investigates text classification for depression diagnosis. Future research will concentrate on implementing cutting-edge techniques to improve robustness and reliability, given the rising popularity of sophisticated models like Transformers. Only two datasets were used for training because there were not many labeled datasets available [115]. In order to enhance model performance and evaluation, future initiatives will give priority to obtaining larger datasets.

5. Conclusion

The systematic literature review examined the recent advancements in the application of machine learning and deep learning models for the early identification of depressive symptoms using social media data. The review highlighted both the advantages, and the challenges associated with these innovative approaches, providing valuable insights for researchers and healthcare professionals interested in leveraging these technologies to enhance the early detection and management of depression.

The scarcity of easily accessible web resources makes gathering a different language dataset extremely difficult. In order to solve this, the dataset size was constrained by manually selecting data from social media with a focus on two particular domains. The model's performance is impacted by this constraint. The different language also has special difficulties, like frequent spelling mistakes, where even small vowel changes can completely affect the meaning. Accurate interpretation is made more difficult when English or other language words are written in script. In order to increase accuracy and prediction performance, future developments will try to address these issues.

References

1. Muzafar, D., Khan, F Y., & Qayoom, M. (2023, January 1). Machine Learning Algorithms for Depression Detection and Their Comparison. Cornell University. https://doi.org/10.48550/arxiv.2301.0322
2. Masab, A., Mansoor., Kashif, Ansari. (2024). Early Detection of Mental Health Crises through AI-Powered Social Media Analysis: A Prospective Observational Study. doi: 10.1101/2024.08.12.24311872

3. Chen, Chen., Wei-Chiang, Hong., Fenghuan, Li., Haopeng, Chen., Y., H., Lin. (2024). 2. Heterogeneous Subgraph Network with Prompt Learning for Interpretable Depression Detection on Social Media. doi: 10.48550/arxiv.2407.09019

4. Haocheng, Xi., Haoyang, Zhong., Yutong, Wang. (2024). 3. BXCNN: A novel depression detection model. Applied and Computational Engineering, doi: 10.54254/2755-2721/68/20241448.
5. Shengjie, Li., Yinhao, Xiao. (2024). 4. A Depression Detection Method Based on Multi-Modal Feature Fusion Using Cross-Attention. arXiv.org,doi: 10.48550/arxiv.2407.12825
6. Wadzani, Aduwamai, Gadzama., Danlami, Gabi., Musa, Sule, Argungu., H., U., Suru. (2024). 5. The use of machine learning and deep learning models in detecting depression on social media: A systematic literature review. Personalized Medicine in Psychiatry, doi: 10.1016/j.pmip.2024.100125.
7. Kartik, Singh., Utkarsh, Gupta., P., Renukadevi. (2024). Psychological Analysis using Social Media Tweets. doi: 10.1109/icaaic60222.2024.10574935.
8. Prof., Dhanashree, Patil. (2024). Depression detection by analyzing social media post of user: a review. Indian Scientific Journal Of Research In Engineering And Management, 08(04):1-5. doi: 10.55041/ijsrem32471.
9. Garg, M. (2023). Mental Health Analysis in Social Media Posts: A Survey. In M. Garg, Archives of Computational Methods in Engineering (Vol. 30, Issue 3, p. 1819). Springer Science+Business Media. https://doi.org/10.1007/s11831-022-09863-z
10. Zhang, W., Xie, J., Zhang, Z., & Liu, X. (2024). Depression Detection Using Digital Traces on Social Media: A Knowledge-aware Deep Learning Approach. In W. Zhang, J. Xie, Z. Zhang, & X. Liu, Journal of Management Information Systems (Vol. 41, Issue 2, p. 546). Taylor & Francis. https://doi.org/10.1080/07421222.2024.2340822
11. Shah, S. M., Gillani, S. A., Baig, M., Saleem, M. A., & Siddiqui, M. (2024). Advancing Depression Detection on Social Media Platforms Through Fine-Tuned Large Language Models. In S. M. Shah, S. A. Gillani, M. Baig, M. A. Saleem, & M. Siddiqui, arXiv (Cornell University). Cornell University. https://doi.org/10.48550/arxiv.2409.14794
12. Uma, Yadav., Ashish, Sharma., Dipti, D., Patil. (2022). Review of automated depression detection: Social posts, audio and video, open challenges and future direction. Concurrency and Computation: Practice and Experience, 35(1) doi: 10.1002/cpe.7407
13. Baskaran, Jayapradha., T., Velmurugan. (2024). Analysing Depression in Social Media Data Using Machine Learning Techniques. 1043-1049. doi: 10.1109/icesc60852.2024.10689804
14. Prof., Saba, Anjum, Patel., Kalakshi, Jadhav., Sayali, Ligade., Vishal, Mahajan., Keshav, Anant. (2024). Depression Prediction using Machine Learning Algorithms. International Journal of Advanced Research in Science, Communication and Technology, doi: 10.48175/ijarsct-18279
15. Biodoumoye, George, Bokolo., Qingzhong, Liu. (2024). Advanced Comparative Analysis of Machine Learning and Transformer Models for Depression and Suicide Detection in Social Media Texts. Electronics, 13(20):3980-3980. doi: 10.3390/electronics13203980
16. Adarsh, V., Kumar, P. A., Lavanya, V., & Gangadharan, G. R. (2023). Fair and explainable depression detection in social media. Information Processing & Management, 60(1), 103168.
17. M., Kiran, Myee., R., Deepthi, Crestose, Rebekah., T, K, Deepa., G., Divya, Zion., K, Lokesh. (2024). Detection of Depression in Social Media Posts using Emotional Intensity Analysis. Engineering, Technology & Applied Science Research, 14(5):16207-16211. doi: 10.48084/etasr.7461
18. Emirhan, Balcı., Esra, Sarac, Essiz. (2024). Automated Depression Detection from Tweets: a Comparison of NLP Techniques. 1-5. doi: 10.1109/idap64064.2024.10711029

19. Shaik, Rasheeda, Begum., Saad, Yunus, Sait. (2024). Machine Learning-Based Depression Detection via Twitter Metadata. 1-6. doi: 10.1109/iceeict61591.2024.10718376
20. Hassantabar, S., Zhang, J., Yin, H., & Jha, N. K. (2021). MHDeep: Mental Health Disorder Detection System based on Body-Area and Deep Neural Networks. In S. Hassantabar, J. Zhang, H. Yin, & N. K. Jha, arXiv (Cornell University). Cornell University. https://doi.org/10.48550/arxiv.2102.10435
21. B., Bokolo., Qingzhong, Liu. (2023). Deep Learning-Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques. Electronics, doi: 10.3390/electronics12214396
22. Abhishek, Pandey., Sanjay, Kumar. (2024). Mental Health and Stress Prediction Using NLP and Transformer-Based Techniques. 61-66. doi: 10.1109/iswta62130.2024.10651843
23. Jose, C., Agoylo., Kim, N., Subang., Jorton, A., Tagud. (2024). Natural Language Processing (NLP)-Based Detection of Depressive Comments and Tweets: A Text Classification Approach. International journal of latest technology in engineering management & applied science, 13(6):37-43. doi: 10.51583/ijltemas.2024.130606
24. Alan, Lu., P., Koehn. (2024). Depression Detection with Novel Large Language Model Methods. International journal of computer science and information technology, 4(1):262-269. doi: 10.62051/ijcsit.v4n1.32
25. H., B., U., Haq., Muhammad, Nauman, Irshad., Muhammad, Daniyal, Baig. (2024). Unlocking Minds: An Adaptive Machine Learning Approach for Early Detection of Depression. doi: 10.56578/hf020205
26. OpenAI. Gpt-3.5 turbo: Model documentation. https://platform. openai.com/docs/models/gpt-3-5-turbo, 2024. Accessed: 2024 08-17.
27. Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Alma hairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhar gava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288, 2023.
28. Chancellor S, De Choudhury M (2020) Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digit Med 3(1):1–11
29. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC (2017) Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci 18:43–49
30. Chancellor S, De Choudhury M (2020) Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digit Med 3(1):1–11
31. Luxton DD, June JD, Fairall JM (2012) Social media and suicide: a public health perspective. Am J Public Health 102(S2):195–200
32. Golden RN, Weiland C, Peterson F (2009) The truth about illness and disease. Infobase Publishing, New York
33. De Choudhury M (2013) Role of social media in tackling chal lenges in mental health. In: Proceedings of the 2nd international workshop on socially-aware multimedia, 2013, pp 49–52
34. Shing H-C, Resnik P, Oard DW (2020) A prioritization model for suicidality risk assessment. In: Proceedings of the 58th annual meeting of the Association for Computational Linguistics, 2020, pp 8124–8137
35. Sawhney R, Joshi H, Flek L, Shah R (2021) Phase: learning emo tional phase-aware representations for suicide ideation detection on social media. In: Proceedings of the 16th conference of the European Chapter of the Association for Computational Linguis tics: main volume, 2021, pp 2415–2428
36. Niederkrotenthaler T (2017) Papageno effect: its progress in media research and contextualization with findings on harmful media effects. In: Media and suicide: international perspectives on research, theory, and policy. Routledge, London, pp 133–158

37. N. A. Asad, M. A. Mahmud Pranto, S. Afreen and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 13-17, doi: 10.1109/SPICSCON48833.2019.9065101.
38. Ali Alqazzaz, Mohammad Tabrez Quasim, Mohammed Mujib Alshahrani, Ibrahim Alrashdi, and Mohammad Ayoub Khan. A deep learning model to analyse social-cyber psychological problems in youth. Comput. Syst. Sci. Eng., 46(1):551–562, 2023
39. Coppersmith, G.; Dredze, M.; Harman, C. Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Quantifying Mental Health Signals in Twitter. 2014. Available online: https://aclanthology.org/W14-3207.pdf (accessed on 15 June 2024).
40. Reece, A.G.; Danforth, C.M. Instagram photos reveal predictive markers of depression. EPJ Data Sci. 2017, 6, 15. [CrossRef]
41. Jain, V.; Chandel, D.; Garg, P.; Vishwakarma, D.K. Depression and Impaired Mental Health Analysis from Social Media Platforms Using Predictive Modelling Techniques. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020. Available online: https://ieeexplore.ieee.org/abstract/ document/9243334 (accessed on 8 May 2021).
42. S. AlSagri, H. & Ykhlef, M. (2020). Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features. [PDF]
43. Liu, D., Lin Feng, X., Ahmed, F., Shahid, M., & Guo, J. (2022). Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review. ncbi.nlm.nih.gov
44. Safa, R., A. Edalatpanah, S., & Sorourkhah, A. (2023). Predicting mental health using social media: A roadmap for future development. [PDF]
45. Nasrullah, S. & Jalali, A. (2022). Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model. ncbi.nlm.nih.gov
46. Gupta, S., Goel, L., Singh, A., Prasad, A., & Aman Ullah, M. (2022). Psychological Analysis for Depression Detection from Social Networking Sites. ncbi.nlm.nih.gov
47. Ahmed, A., Aziz, S., T. Toro, C., Alzubaidi, M., Irshaidat, S., Abu Serhan, H., A. Abd-alrazaq, A., & Househ, M. (2022). Machine learning models to detect anxiety and depression through social media: A scoping review. ncbi.nlm.nih.gov
48. Rajderkar, V., & Bhat, A. (2024). Multilingual Depression Detection in Online Social Media Across Eight Indian Languages. 2024 3rd International Conference for Innovation in Technology (INOCON), 1-6.
49. Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.
50. Chancellor S, Birnbaum ML, Caine ED, Silenzio VM, De Choudhury M (2019) A taxonomy of ethical tensions in inferring mental health states from social media. In: Proceedings of the conference on fairness, accountability, and transparency, 2019, pp 79–88
51. Chancellor S, De Choudhury M (2020) Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digit Med 3(1):1–11
52. Shah FM, Haque F, Nur RU, Al Jahan S, Mamud Z (2020) A hybridized feature extraction approach to suicidal ideation detection from social media post. In: 2020 IEEE Region 10 symposium (TENSYMP), 2020. IEEE, pp 985–9
53. Stankevich M, Isakov V, Devyatkin D, Smirnov I (2018) Feature engineering for depression detection in social media. In: ICPRAM, 2018, pp 426–431
54. Hussain J, Satti FA, Afzal M, Khan WA, Bilal HSM, Ansaar MZ, Ahmad HF, Hur T, Bang J, Kim J-I et al (2020) Exploring the dominant features of social media for depression detection. J Inf Sci 46(6):739–759

55. Obagbuwa, I.C.; Danster, S.; Chibaya, O.C. Supervised machine learning models for depression sentiment analysis. Front. Artif. Intell. 2023, 6, 123. [CrossRef] [PubMed]
56. Lyu, H. Application of machine learning on depression prediction and analysis. Appl. Comput. Eng. 2023, 5, 712–719. [CrossRef]
57. Govindasamy, K.; Palanichamy, N. Depression Detection Using Machine Learning Techniques on Twitter Data. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 6–8 May 2021. [CrossRef]
58. Obagbuwa, I.C.; Danster, S.; Chibaya, O.C. Supervised machine learning models for depression sentiment analysis. Front. Artif. Intell. 2023, 6, 123. [CrossRef] [PubMed]
59. Angskun, J.; Tipprasert, S.; Angskun, T. Big data analytics on social networks for real-time depression detection. J. Big Data 2022, 9, 69. [CrossRef] [PubMed]
60. Ashraf, A.; Gunawan, T.S.; Riza, B.S.; Haryanto, E.V.; Janin, Z.A. On the review of image and video-based depression detection using machine learning. Indones. J. Electr. Eng. Comput. Sci. 2020, 19, 1677. [CrossRef]
61. E. Alberta, "Review of the Beck Depression Inventory 2nd Edition (BDI-II), Canada.," CommunityUniversity Partnership for the Study of Children, Youth, and Families.
62. D.B. Victor, J. Kawsher, M.S. Labib, S. Latif, "Machine learning techniques for depression analysis on social media-case study on Bengali community," 2020 4th International Conference on Electronics, Com- munication and Aerospace Technology (ICECA), p. 1118-1126, 2020, IEEE.
63. M.K. Kabir, M. Islam, A.N.B. Kabir, A. Haque, M.K. Rhaman, "Detec- tion of Depression Severity Using Bengali Social Media Posts on Mental Health: Study Using Natural Language Processing Techniques," JMlR Formative Research, vol. 6, no. 9, p. e36118, 2022, JMIR Publications Tnc., Toronto, Canada.
64. A.H. Uddin, D. Bapery, A.S.M. Arif, "Depression analysis of bangla social media data using gated recurrent neural network," 2019 Isl International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), p. 1-6, 2019, IEEE.
65. L. Gohil, D. Patel, "Multilabel classification for emotion analysis of multilingual tweets," Int J Innov Technol Explor Eng, vol. 9, no. I, p. 4453-4457, 2019.
66. T.T. Sasidhar, B. Premjith, K.P. Soman, "Emotion detection in hinglish (hindi+ english) code-mixed social media text," Procedia Computer Science, vol. 171, p. 1346-1352, 2020, Elsevier.
67. A. Wadhawan, A. Aggarwal, "Towards emotion recognition in Hindi- English code-mixed data: A transformer based approach," arXiv preprint arXiv:2102.09943, 2021.
68. S. Thara, P. Poornachandran, "Social media text analytics ofMalayalam- English code-mixed using deep learning," Journal of big Data, vol. 9, no. l, p. 45, 2022, Springer.
69. A. Hande, S.U. Hegde, R. Priyadharshini, R. Ponnusamy, P.K. Ku- maresan, S. Thavareesan, B.R. Chakravarthi, "Benchmarking multi-task learning for sentiment analysis and offensive language identification in under-resourced dravidian languages," arXiv preprint arXiv:2108.03867, 2021.
70. P.H.V. Pavan Kumar, B. Premjith, J.P. Sanjanasri, K.P. Soman, "Deep Learning Based Sentiment Analysis for Malayalam, Tamil and Kannada Languages," 202l.
71. Coppersmith G, Dredze M, Harman C, Hollingshead K, Mitchell M (2015) CLPsych 2015 shared task: depression and PTSD on Twitter. In: Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, 2015, pp 31–39
72. Milne DN, Pink G, Hachey B, Calvo RA (2016) CLPsych 2016 shared task: triaging content in online peer-support forums. In: Proceedings of the third workshop on computational linguistics and clinical psychology, 2016, pp 118–127

73. Preotiuc-Pietro D, Sap M, Schwartz HA, Ungar LH (2015) Mental illness detection at the world well-being project for the CLPsych 2015 shared task. In: CLPsych@ HLT-NAACL, 2015, pp 40–45
74. 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. In: Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, 2015, pp 99–107
75. Matero M, Idnani A, Son Y, Giorgi S, Vu H, Zamani M, Limbachiya P, Guntuku SC, Schwartz HA (2019) Suicide risk assessment with multi-level dual-context language and BERT. In: Proceedings of the sixth workshop on computational linguistics and clinical psychology, 2019, pp 39–44
76. Orabi AH, Buddhitha P, Orabi MH, Inkpen D (2018) Deep learning for depression detection of Twitter users. In: Proceedings of the ffth workshop on computational linguistics and clinical psychology: from keyboard to clinic, 2018, pp 88–97
77. Jamil Z (2017) Monitoring tweets for depression to detect at-risk users. PhD Thesis, University of Ottawa
78. Lin C, Hu P, Su H, Li S, Mei J, Zhou J, Leung H (2020) SenseMood: depression detection on social media. In: Proceedings of the 2020 international conference on multimedia retrieval, 2020, pp 407–411
79. Gui T, Zhu L, Zhang Q, Peng M, Zhou X, Ding K, Chen Z (2019) Cooperative multimodal approach to depression detection in Twitter. In: Proceedings of the AAAI conference on artifcial intelligence, 2019, vol 33, pp 110–117
80. 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. In: IJCAI, 2017, pp 3838–3844
81. Zogan H, Razzak I, Jameel S, Xu G (2021) DepressionNet: a novel summarization boosted deep framework for depression detection on social media. arXiv preprint. arXiv:2105.10878
82. Song H, You J, Chung J-W, Park JC (2018) Feature attention network: interpretable depression detection from social media. In: PACLIC, 2018
83. Cong Q, Feng Z, Li F, Xiang Y, Rao G, Tao C (2018) XABiLSTM: a deep learning approach for depression detection in imbalanced data. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM), 2018. IEEE, pp 1624–1627
84. Yates A, Cohan A, Goharian N (2017) Depression and self-harm risk assessment in online forums. arXiv preprint. arXiv:1709. 01848
85. Cohan A, Desmet B, Yates A, Soldaini L, MacAvaney S, Goharian N (2018) SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. arXiv preprint. arXiv:1806.05258
86. Gamaarachchige PK, Inkpen D (2019) Multi-task, multi-channel, multi-input learning for mental illness detection using social media text. In: Proceedings of the tenth international workshop on health text mining and information analysis (LOUHI 2019), 2019, pp 54–64
87. Almeida H, Briand A, Meurs M-J (2017) Detecting early risk of depression from social media user-generated content. In: CLEF (working notes), 2017
88. Losada DE, Crestani F, Parapar J (2018) Overview of eRISK: early risk prediction on the Internet. In: International conference of the Cross-Language Evaluation Forum for European Languages, 2018. Springer, pp 343–361
89. Tadesse MM, Lin H, Xu B, Yang L (2019) Detection of depression-related posts in Reddit social media forum. IEEE Access 7:44883–44893

90. Pirina I, Çöltekin Ç (2018) Identifying depression on Reddit: the efect of training data. In: Proceedings of the 2018 EMNLP workshop SMM4H: the 3rd social media mining for health applications workshop and shared task, 2018, pp 9–12
91. Ji S, Yu CP, Fung S-F, Pan S, Long G (2018) Supervised learning for suicidal ideation detection in online user content. Complexity. https://doi.org/10.1155/2018/61572
92. Tadesse MM, Lin H, Xu B, Yang L (2020) Detection of suicide ideation in social media forums using deep learning. Algorithms 13(1):7
93. Abdulsalam, A., & Alhothali, A. (2024). Suicidal ideation detection on social media: A review of machine learning methods. Social Network Analysis and Mining, 14(1), 188.
94. Cao L, Zhang H, Feng L (2020) Building and using personal knowledge graph to improve suicidal ideation detection on social media. IEEE Trans Multimed. https://doi.org/10.1109/tmm.2020.3046867
95. Cao L, Zhang H, Feng L, Wei Z, Wang X, Li N, He X (2019) Latent suicide risk detection on microblog via suicide-oriented word embeddings and layered attention. arXiv preprint. arXiv: 1910.12038
96. Turcan E, McKeown K (2019) Dreaddit: a Reddit dataset for stress analysis in social media. arXiv preprint. arXiv:1911.00133
97. Turcan E, Muresan S, McKeown K (2021) Emotion-infused models for explainable psychological stress detection. In: Proceedings of the 2021 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies, 2021, pp 2895–2909
98. Harrigian K, Aguirre C, Dredze M (2020) On the state of social media data for mental health research. arXiv preprint. arXiv: 2011.05233
99. Gaur M, Alambo A, Sain JP, Kursuncu U, Thirunarayan K, Kavuluru R, Sheth A, Welton R, Pathak J (2019) Knowledgeaware assessment of severity of suicide risk for early intervention. In: The World Wide Web conference, 2019, pp 514–525
100. Aladağ AE, Muderrisoglu S, Akbas NB, Zahmacioglu O, Bingol HO (2018) Detecting suicidal ideation on forums: proofof-concept study. J Med Internet Res 20(6):21
101. Shing H-C, Resnik P, Oard DW (2020) A prioritization model for suicidality risk assessment. In: Proceedings of the 58th annual meeting of the Association for Computational Linguistics, 2020, pp 8124–8137
102. Matero M, Idnani A, Son Y, Giorgi S, Vu H, Zamani M, Limbachiya P, Guntuku SC, Schwartz HA (2019) Suicide risk assessment with multi-level dual-context language and BERT. In: Proceedings of the sixth workshop on computational linguistics and clinical psychology, 2019, pp 39–44
103. Pendke, K., Digalwar, M., & Goel, C. (2023, August). Identification and Analysis of Emotions from the Text for Stress Detection. In Proceedings of the 2023 Fifteenth International Conference on Contemporary Computing (pp. 623-630).
104. Burkhardt H, Pullmann M, Hull T, Aren P, Cohen T (2022) Comparing emotion feature extraction approaches for predicting depression and anxiety. In: Proceedings of the eighth workshop on computational linguistics and clinical psychology, 2022, pp 105–115
105. Haque A, Reddi V, Giallanza T (2021) Deep learning for suicide and depression identifcation with unsupervised label correction. arXiv preprint. arXiv:2102.09427
106. Garg M, Saxena C, Krishnan V, Joshi R, Saha S, Mago V, Dorr BJ (2022) CAMS: an annotated corpus for causal analysis of mental health issues in social media posts. arXiv preprint. arXiv:2207.04674
107. Saxena C, Garg M, Ansari G (2022) Explainable causal analysis of mental health on social media data. In: Proceedings of ICONIP, 2022

108. Naseem U, Khushi M, Kim J, Dunn AG (2022) RHMD: a realworld dataset for health mention classifcation on Reddit. IEEE Trans Comput Soc Syst. https://doi.org/10.1109/TCSS.2022.3186883
109. Tavchioski I, Koloski B, Škrlj B, Pollak S (2022) E8-IJS@ LTEDI-ACL2022-BERT, AutoML and knowledge-graph backed detection of depression. In: Proceedings of the second workshop on language technology for equality, diversity and inclusion, 2022, pp 251–257
110. Sivamanikandan S, Santhosh V, Sanjaykumar N, Durairaj T et al (2022) scubeMSEC@ LT-EDI-ACL2022: detection of depression using transformer models. In: Proceedings of the second workshop on language technology for equality, diversity and inclusion, 2022, pp 212–217
111. Kayalvizhi S, Thenmozhi D (2022) Data set creation and empirical analysis for detecting signs of depression from social media postings. arXiv preprint. arXiv:2202.03047
112. Demszky D, Movshovitz-Attias D, Ko J, Cowen A, Nemade G, Ravi S (2020) GoEmotions: a dataset of fne-grained emotions. arXiv preprint. arXiv:2005.00547
113. Gkotsis G, Oellrich A, Hubbard T, Dobson R, Liakata M, Velupillai S, Dutta R (2016) The language of mental health problems in social media. In: Proceedings of the third workshop on computational linguistics and clinical psychology, 2016, pp 63–73
114. William, David, and Derwin Suhartono. "Text-based depression detection on social media posts: A systematic literature review." Procedia Computer Science 179 (2021): 582-589.
115. Ghosh, Smita, et al. "Depression detection using machine and deep learning models to assess mental health of social media users." Machine Learning Techniques and Data Science Trends (2022).
116. Di Cara, N.H.; Maggio, V.; Davis, O.S.P.; Haworth, C.M.A. Methodologies for Monitoring Mental Health on Twitter: Systematic Review. J. Med. Internet Res. 2023, 25, e42734. [CrossRef] [PubMed
117. Chatterjee, M.; Modak, S.; Sarkar, D. Mental Health Predictions Through Online Social Media Analytics. In Cognitive Cardiac Rehabilitation Using IoT and AI Tools; IGI Global: Hershey, PA, USA, 2023; pp. 44–66. [CrossRef]
118. Coppersmith, G.; Dredze, M.; Harman, C. Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Quantifying Mental Health Signals in Twitter. 2014. Available online: https://aclanthology.org/W14-3207.pdf (accessed on 15 June 2024).
119. Reece, A.G.; Danforth, C.M. Instagram photos reveal predictive markers of depression. EPJ Data Sci. 2017, 6, 15. [CrossRef]
120. Guntuku, S.C.; Yaden, D.B.; Kern, M.L.; Ungar, L.H.; Eichstaedt, J.C. Detecting depression and mental illness on social media: An integrative review. Curr. Opin. Behav. Sci. 2017, 18, 43–49. [CrossRef]
121. Shen, J.; Rudzicz, F. Detecting Anxiety on Reddit. In Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology, Vancouver, BC, Canada, 3 August 2017; pp. 58–65. Available online: https://aclanthology.org/W17-3107.pdf (accessed on 19 June 2024).
122. Govindasamy, K.; Palanichamy, N. Depression Detection Using Machine Learning Techniques on Twitter Data. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 6–8 May 2021. [CrossRef]
123. Jain, V.; Chandel, D.; Garg, P.; Vishwakarma, D.K. Depression and Impaired Mental Health Analysis from Social Media Platforms Using Predictive Modelling Techniques. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020. Available online: https://ieeexplore.ieee.org/abstract/ document/9243334 (accessed on 8 May 2021).
124. Gunawan, T. S., Ashraf, A., Riza, B. S., Haryanto, E. V., Rosnelly, R., Kartiwi, M., & Janin, Z. (2020). Development of video-based emotion recognition using deep learning with Google Colab. TELKOMNIKA (Telecommunication Computing Electronics and Control), 18(5), 2463-2471.

125. Angskun, J.; Tipprasert, S.; Angskun, T. Big data analytics on social networks for real-time depression detection. J. Big Data 2022, 9, 69. [CrossRef] [PubMed]
126. M. D. Souza, V. Kumara, R. D. Salins, J. J. A Celin, S. Adiga and S. Shedthi, "Advanced Deep Learning Model for Breast Cancer Detection via Thermographic Imaging," *2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, Mangalore, India, 2024, pp. 428-433, doi:10.1109/DISCOVER62353.2024.10750727
127. Souza, M. D., Prabhu, A. G., & Kumara, V. (2019). A comprehensive review on advances in deep learning and machine learning for early breast cancer detection. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, *10*(5), 350-359.
128. P. M. Manjunath, Gurucharan and M. Dsouza, Shwetha, "IoT Based Agricultural Robot for Monitoring Plant Health and Environment", *Journal of Emerging Technologies and Innovative Research vol.* 6, no. 2, pp. 551-554, Feb 2019