Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. ICMASD (2025)

Deep Learning and Natural Language Processing Techniques for Depression Detection in Social Media Texts: A Comprehensive Review

Neelam Kumari[1], Jagdeep Kaur[2]

[1]Research Scholar, Sant Baba Bhag Singh University, Jalandhar, India

[2]Professor, CSE, Sant Baba Bhag Singh University, Jalandhar, India

ndidhra1403@gmail.com, universitysbbsuuiet@gmail.com

Article History:

Received: 12-12-2024

Revised: 25-01-2025

Accepted: 05-02-2025

Abstract:

With the increased prevalence of social media, depression is becoming a prominent health issue worldwide and it has opened new opportunities for researchers to explore new techniques for depression detection in social media texts. According to the World Health Organization (WHO) depression may lead to various mental health diseases and suicide if not detected at an early stage. Deep learning and Natural Language Processing are becoming widely adopted techniques among researchers. This review provides a thorough examination of these techniques for depression detection in posts of users. These techniques are harnessed to identify linguistic markers related depression such as sentiment and emotional tone and capture temporal dependencies in texts. Transformer models represent the next level of deep learning techniques, enhanced with self-attention mechanism that enables the automatic analysis of text sequences over time, sematic feature extraction and the interpretation of context-sensitive language. Multimodal approaches using these techniques integrate textual and visual data to improve the accuracy of depression detection. Despite notable advancements, still there are many challenges to address such as data availability, privacy and ethical and model interpretability. Primary aim of this paper is to explore concepts such as evolution of techniques, deep learning and NLP for depression detection, dataset for testing and research gaps and future directions. We conducted a Systematic Literature Review (SLR) on research and review studies published in various conferences and peer-reviewed journals. At last, we provided the brief summary of key findings.

Keywords: Deep Learning, Natural Language Processing, Depression, Social Media, Mental Health.

1. Introduction

Social media become an essential platform for connecting, discovering and interacting with people globally. Facebook, Reddit, Twitter and Instagram are widely used platforms where users express their feelings, exchange thoughts and document their daily activities through text, blogs and emojis. But extensive use of social media leads to the depression in users all around the world, which is a prominent cause of mental illness if left undetected. A report from the WHO states that 1 in every 8 individuals worldwide suffers from depression, which can be treated if recognized at an earlier stage (Tahir, Khalid, Almutairi, Memon, & Khan, 2025).Depression not only contributes to mental health disorders but can also lead to suicide and impulsive behaviors such as self-harm or other dangerous actions

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. ICMASD (2025)

(Tahir et al., 2025). Suicide rates have been steadily rising, with a recent WHO report indicating that in 2021, approximately 14,000 individuals worldwide died by suicide (Vasha, Sharma, Esha, Al Nahian, & Polin, 2023). Early detection of mental health issues is crucial in preventing severe disorders and reducing the risk of suicide. Natural Language Processing (NLP) has emerged as a prominent field for sentiment analysis and emotional feature extraction, in conjunction with Machine Learning (ML) and Deep Learning (DL) models, to aid in the detection of depression.

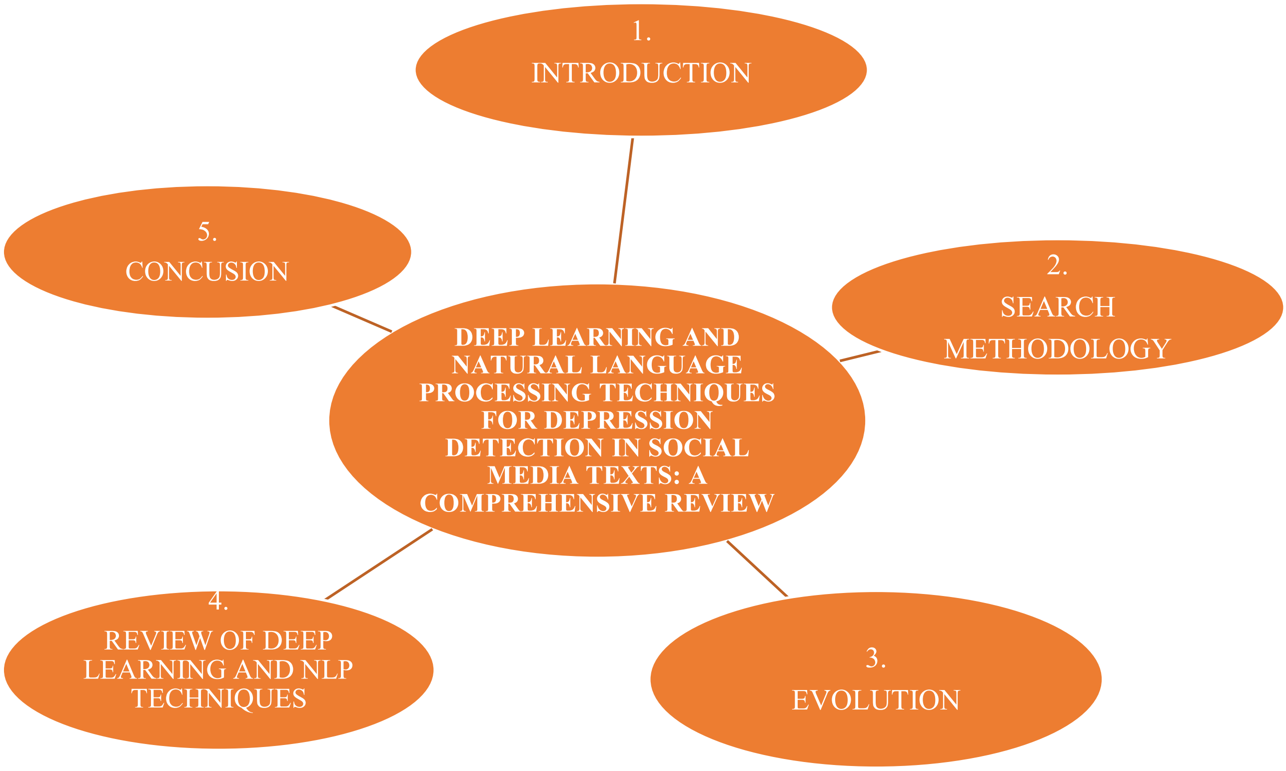

Figure 1: Outline of the paper Structure.

In early studies, machine learning classifiers such as Support Vector Machines (SVM), Naive Bayes (NB), Decision Trees (DT), Random Forests (RF), K-Nearest Neighbors (KNN), and Logistic Regression (LR) were predominantly employed for classifying user posts into two categories: depressed or non-depressed (Vasha et al., 2023; Islam et al., 2018). However, deep learning has significantly advanced the field of Natural Language Processing (NLP) by offering enhanced capabilities, particularly in handling large-scale data that includes text, images, audio, and video. Models like Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Long Short-Term Memory Networks (LSTMs) excel in automatically learning hierarchical data representations and can capture sequential and contextual information from text. More recently, transformer-based models have gained widespread attention due to their self-attention mechanisms, which efficiently propagate sequential information through multiple layers of data. Models such as BERT, XLNet, and RoBERTa are increasingly utilized for the automated classification of depressive versus non-depressive content in social media posts (Shetty, Singh, Hegde, Cenitta, & Dhruthi, 2025).

In this review, we aim to provide a comprehensive overview of the methodologies employed by researchers in recognizing efforts within the fields of Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning for the detection of depression and other mental health disorders on social media. We also learned about several datasets utilized in various research implementations. This paper strives to achieve the following objectives:

- To identify the historical evolution of the DL and NLP techniques for depression detection in social media texts
- Provide a comprehensive review of Deep Learning and NLP techniques. Also included studies employed transformer models for detect depression, dataset used for testing in research articles.
- Identifying the research gaps and future directions from existing studies.

The paper is structured into the following sections: Section 2 outlines the search methodology and criteria for selecting study material for a thorough review. Section 3 examines the evolution of techniques such as Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), transformation models, and other ensemble models, as well as the proposed methods for identifying depressive posts. Section 4 provides an in-depth review of the literature, narrating the use of Natural Language Processing and Deep Learning techniques for detecting depression in social media data. Finally, we conclude with a summary of the key findings. Figure 1 presents an overview of the paper's structure.

2. Search Methodology

In this section, we outline the search methodology employed to conduct the systematic literature review (SLR). We focused on papers published from 2018 to January 2025, ensuring a comprehensive search process aligned with the core theme of the review. To identify relevant studies, we leveraged prominent scholarly databases, including Google Scholar, Web of Science, ResearchGate, IEEE Xplore, ScienceDirect, and the ACM Digital Library. Boolean operators, specifically "AND" and "OR," were used to refine the search. We focused on titles and abstracts, utilizing search phrases and keywords such as "depression" OR "mental health disorders" OR "mental illness" AND "social media" OR "social networking sites" AND "Natural language processing" OR "NLP" AND "Deep learning" OR "DL." PRISMA methodology used for inclusion and exclusion process of selection criteria is shown in Figure 2.

2.1 Selection Criteria

Articles published in reputed journals and conferences are used for this study. We identified 102 records based on the search strategy. After removing 10 duplicate records, remained 92 records out of which 78 full text articles are screened for eligibility. 28 primary studies and 4 survey studies are included in synthesis of the review.

2.1.1 Inclusion Criteria:

a. Articles published from 2018 to January 2025
b. Studies published in reputed journals
c. Articles written in English language only
d. Studies focussing on the centric theme of review

e. Research employed NLP and DL techniques for depression detection
f. Evaluations of studies utilizing only social media platforms for depression detection.

Table 1: List of Studies screened and included in synthesis.

Year	**No. of papers Screened**	**No of papers included**
2018	4	1
2019	9	2
2020	6	2
2021	8	3
2022	12	4
2023	18	5
2024	17	12
2025	4	3

2.1.2 Exclusion Criteria:

a. Studies irrelevant to the centric theme
b. Papers not written in English.
c. Articles not use social media for depression detection
d. Articles based on topics other than depression such as hate speech and cyberbullying
e. Low quality research papers

We defined the scope of our review based on the details of the selected articles, including the publication years, authors, publishers, methodologies employed and key findings. The selection of papers is presented in a year-wise format in Table 1

3. Evolution

In this section we will try to get overview of progressive growth of ML, DL used with NLP to detect depression from social media platforms like Reddit, Facebook, Twitter, and Instagram. This section will pinpoints the progressive growth of conventional approaches to modern techniques such as Deep Learning and transformer models together with NLP.

3.1 Traditional Machine Learning techniques and ensemble models with NLP (2015-2024)

Machine learning classifiers gained significant popularity starting in the mid-2010s for depression detection in social media data (Hassan et al., 2017; McManus et al., 2015). The advancement of natural language processing (NLP) techniques in combination with machine learning (ML) methods drew the attention of researchers, as emotional feature extraction notably improved the performance of these classifiers. Commonly used supervised ML algorithms include: Support Vector Machines (SVM), Decision Trees (DT), K-Nearest Neighbors (KNN), Random Forests (RF), Naive Bayes (NB) (Al

Asad, Mahmud Pranto, & Afreen, 2019; Almouzini, Khemakhem, & Alageel, 2019; Govindasamy & Palanichamy, 2021; Islam et al., 2018; Pande et al., 2024; Michael M. Tadesse, Lin, Xu, & Yang, 2019; Vasha et al., 2023), Gradient Boosting (GB) (Chiong et al., 2021), and XGBoost (Kumar et al., 2022). Frequently utilized NLP techniques with ML models include Linguistic Inquiry and Word Count (LIWC), Latent Dirichlet Allocation (LDA), Term Frequency-Inverse Document Frequency (TF-IDF), and N-Grams. Some studies have also incorporated fastText embeddings as a word representation technique, combined with XGBoost (Ghosal & Jain, 2023). In late 2020, ensemble models demonstrated superior performance compared to traditional models in certain studies (Shah et al., 2020).

However, the performance of ML models can be significantly impacted by large, highly imbalanced, and poorly designed datasets used for training and testing. Additionally, the lack of context in textual data often limits the models' ability to identify the fine-grained symptoms of depression in social media content.

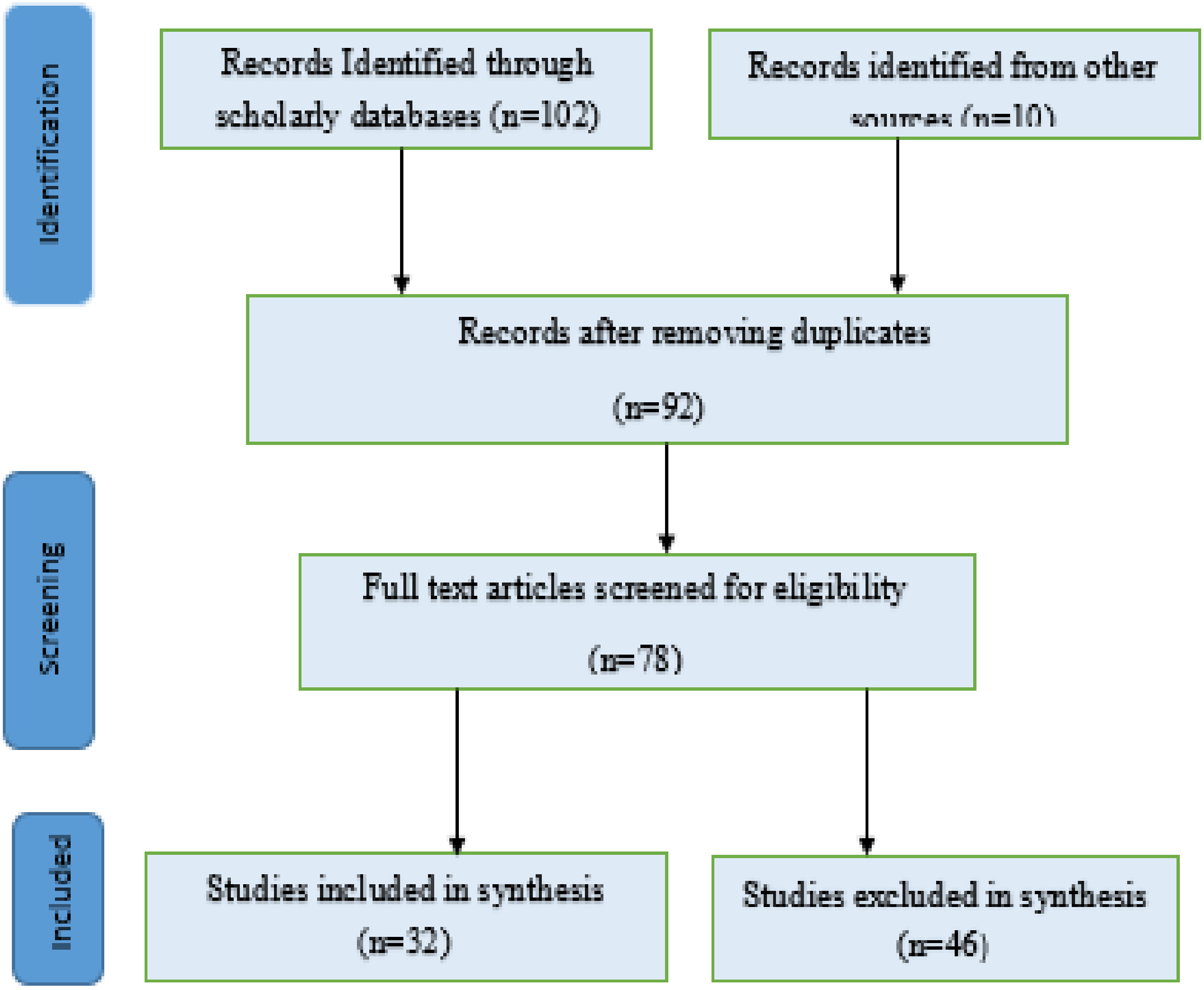

Figure 2. PRISMA flowchart for selection process of systematic review of Deep Learning and NLP techniques for depression detection in social media.

3.2 Emergence of Deep Learning in Natural Language Processing (NLP) (2018-Present)

The advent of Deep Learning (DL) has significantly surpassed the limitations of traditional machine learning (ML), particularly in Natural Language Processing (NLP). Unlike ML, DL models can autonomously detect complex patterns and features from raw data, making them exceptionally efficient at handling large datasets with highly contextualized information. Deep learning techniques began to gain prominence around 2011 due to their remarkable performance in various domains, including NLP.

One of the pivotal milestones in NLP was the use of Long Short-Term Memory (LSTM) networks for textual data (Coppersmith, Leary, Crutchley, & Fine, 2018). LSTM was soon followed by advancements like Bidirectional LSTM (BiLSTM) (Shah et al., 2020), Convolutional Neural Networks (CNN) combined with LSTM (García-Noguez, Tovar-Arriaga, Paredes-García, Ramos-Arreguín, & Aceves-Fernández, 2023; Kumnunt & Sornil, 2020; Merayo, Ayuso-Lanchares, & González-Sanguino, 2025), and Recurrent Neural Networks (RNN) (Amanat et al., 2022; Phaninder et al., 2024; Kanahuati-Ceballos et al., 2024). These models represent the state-of-the-art in deep learning for NLP tasks, offering superior performance over conventional techniques.

Moreover, DL classifiers often integrate NLP pre-processing techniques such as word embeddings, tokenization, stemming, and lemmatization for feature extraction. These methods serve to enhance the representation of textual data, facilitating better model learning and improving the accuracy of predictions.

3.3 Transformer models with self-attention mechanism (2021-Present)

BERT, introduced in late 2018, gained significant attention starting in 2020 and has since become one of the most widely used classifiers by researchers. Transformer models, like BERT, have exceeded expectations in performance for tasks such as depression detection, not only in text but also in images, audio, and video, outperforming traditional benchmarks. These deep learning models utilize a self-attention mechanism to capture linguistic features and excel at handling contextual dependencies. Unlike sequential deep learning models, transformers process input sequences in parallel, enabling faster and more efficient computation. BERT, as a fine-tuned model, is applied to a variety of NLP tasks, and has variants such as RoBERTa, DistilBERT, and DeBERTa, which are optimized for different types of data. (Alhamed et al., 2024; Ramirez-Cifuentes et al., 2021; Zhang, Yang et al., 2023).

Multimodal techniques and explainable ensemble models are emerging research fields for detecting depression in social media text. The progressive growth of techniques is illustrated in figure 3.

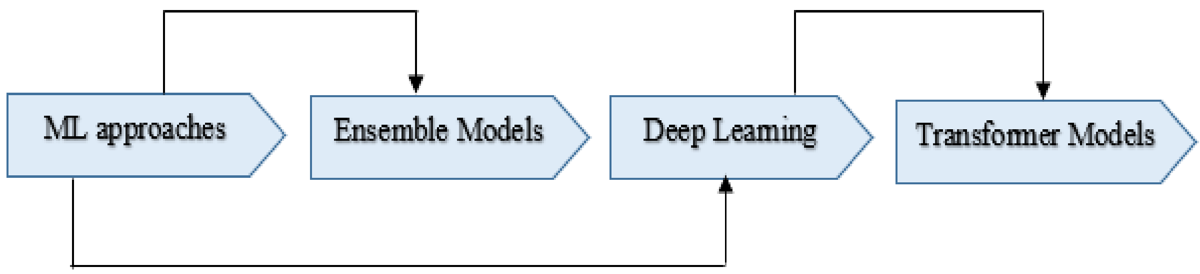

Figure 3. Evolution of techniques for depression detections.

4. Review of Deep Learning Techniques and NLP

This section provides an overview of the deep learning techniques and associated transformation methods employed for detecting depression from social media. Natural Language Processing (NLP) techniques have evolved alongside deep learning models to uncover hidden patterns and extract linguistic features, enabling the identification of depressive signs within textual data on social media platforms. The objective of this review is to gain comprehensive insights into the current machine learning (ML), deep learning (DL), and NLP techniques used for depression detection. The studies are categorized into three subsections: first, we discuss the foundational studies focused on depression detection; second, we present a selection of survey studies; and finally, we include a few studies beyond depression to provide a broader understanding of the topic. Tables 2, 3, and 4 offer a summarized view of the literature review conducted within each of these three subsections.

4.1 Review of Primary Studies for Depression Detection

This study employed a combination of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for the early detection of suicide risk based on Reddit users' posts. **Tadesse et al. (2019)** also compared the proposed model with other machine learning classifiers integrated with NLP techniques for detecting suicidal ideation. The dataset from Reddit consisted of both suicide-related and non-suicidal posts. Statistical features such as TF-IDF and Bag-of-Words (BOW) were input into the machine learning classifiers. However, the proposed model, utilizing Word2Vec embeddings, outperformed the others, achieving an accuracy of 93.8%, an F1-score of 93.4%, recall of 94.1%, and precision of 93.2%(Michael Mesfin Tadesse, Lin, Xu, & Yang, 2019).

Aswathy et al. (2019) proposed a CNN+LSTM model for depression detection using Twitter sentiment, leveraging a public dataset containing both depressive and non-depressive tweets. They also employed Support Vector Machine (SVM) for a comparative analysis of the results. The proposed model outperformed the SVM, achieving higher F1-score, recall, and precision, with values of 97%, 96%, and 97%, respectively (K S, P C, & Murali, 2019).

A hybrid model utilizing the deep learning classifier BiLSTM with various word embedding features was proposed by **Shah et al. (2020)** for early depression detection. This model was tested on the public CLEF eRisk 2017 dataset and demonstrated that the Word2Vec+Meta feature outperformed other word embedding techniques proposed in the study (Shah et al., 2020) .

A study conducted in 2020 proposed an approach for depression detection on Thai social media, where **Kumnunt et al. (2020)** employed a CNN-LSTM classifier with various activation functions and Word2Vec embeddings. They also compared the performance of SVM and Logistic Regression (LR) using TF-IDF on a dataset created from social media messages addressing depression-related issues. The three-branch CNN-LSTM model with the ReLU activation function outperformed the state-of-the-art models (Kumnunt & Sornil, 2020).

Wongkoblap et al. (2021) proposed two novel predictive models for depression detection from Twitter, designed to address the anaphora resolution problem and emphasize anaphoric interpretations. Anaphora resolution refers to the process of determining the person previously mentioned in social media text messages. Initially, they proposed the MIL-SocNet model and later introduced an extended version called MILA-SocNet, which addresses the anaphora resolution problem. LIWC was used for information interpretation. MILA-SocNet outperformed other deep learning models in performance (Wongkoblap, Vadillo, & Curcin, 2021).

Mendes et al. (2022) conducted a study to detect 21 signs of depression in Brazilian undergraduate students by collecting data from Facebook pages where they express their emotions. A total of 783 posts were collected, with manual annotations assigned to address lengthy posts containing long texts. They employed SVM, XGBoost, and BERT, along with various NLP techniques, for each depression sign. The deep learning techniques outperformed the state of art (Mendes, Caseli, Luis, & Sp, 2022).

Amanat et al. (2022) authors proposed a machine learning model that utilized Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), comparing it with SVM, One-hot + SVM, Naive Bayes (NB), and DBN + One-hot for detecting depression from the texts and emoticons used by social media users. The CNN with TF-IDF model achieved an accuracy of 99.0%. A performance comparison of the proposed model was conducted against other models (Amanat et al., 2022).

Gupta et al. (2022) employed machine learning classifiers to detect depression on social media sites and analyzed the behavior and activity patterns of the depressed group. They used two public datasets, Sentiment_140 and Sentiment_tweets3, for training and testing. The study highlighted the high activity rates during both day and night. The proposed classifiers included Decision Tree (DT), k-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), and Long Short-Term Memory (LSTM). LSTM outperformed the other models in detecting depression from Twitter (Gupta, Goel, Singh, Prasad, & Ullah, 2022).

To detect depressed users on Twitter, **Garcia-Noguez et al**. **(2023)** incorporated temporal analysis, and through this innovative approach, they were able to distinguish between signs of Major Depressive Disorder (MDD) and Persistent Depressive Disorder (PDD) based on DSM-V criteria. They selected the CLPsych 2015 dataset, which had been used in previous studies. The results were evaluated using CNN with GloVe and XGBoost with TF-IDF classifiers. CNN with GloVe outperformed the baseline model, achieving an accuracy of 88.61% (García-Noguez et al., 2023).

XLNet was proposed by **Apriliani et al. (2023)** for depression detection from Twitter texts. They collected data for their dataset using DASS-42 questionnaires filled out by the respondents. The DASS-42 is a self-report questionnaire that assesses 42 negative emotional symptoms related to depression, anxiety, and stress. XLNet is an autoregressive model that learns bidirectional context and features a two-stream self-attention architecture. The performance of XLNet was evaluated in three different scenarios, and the model performed well with an average accuracy rate of 93.3% (Apriliani & Maharani, 2023).

A Kumar et al. (2023) employed a BERT and BiLSTM pipeline for depression detection in Arabic social media, using two Arabic language datasets. The performance of these models was compared with that of RF, KNN, Naive Bayes (NB), Gradient Boosting (GB), and Decision Tree (DT) classifiers with TF-IDF. Additionally, a fine-tuned RoBERTa model was proposed for detecting signs of depression (depressive, non-depressive, and severe depressive) in an English language dataset. The authors utilized three datasets from previous studies, two in Arabic and one in English. RoBERTa outperformed the baseline models, achieving an F1 score of 1.00 and 0.82 on the two Arabic datasets, and 0.60 on the English dataset (A. Kumar, Kumari, & Pradhan, 2023).

Bokolo et al. (2023) employed multiple deep learning models for depression detection in user tweets, including RoBERTa, squeezeBARTa, DistilBERT, and DeBERTa. These models were evaluated for depression detection and compared to traditional machine learning classifiers such as Logistic Regression (LR), Bernoulli Naive Bayes (NB), Random Forest (RF), and the TF-IDF technique of NLP, based on various performance metrics. This study repurposed and customized the Sentiment140 dataset, which had been used in previous experimental studies. All deep learning models outperformed the state-of-the-art machine learning models, with RoBERTa achieving the highest performance, recording accuracy, precision, recall, and F1 scores of 98.0%, 98.0%, 99.0%, and 98%, respectively (Bokolo & Liu, 2023).

An ensemble model proposed by **Adarsh et al. (2023)** combined SVM and KNN for depression detection from Reddit. They selected data from two subreddits, r/SuicideWatch and r/Depression, to create the dataset. During the pre-processing phase, Neural Machine Translation (NMT) and one-shot decision approaches were used to remove noisy data and address the imbalanced data issues across different age groups of users, respectively. Afterward, TF-IDF was applied for text mining. Additionally, they employed Local Interpretable Model-Agnostic Explanations (LIME) for explainable artificial intelligence to identify depression and suicidal ideation. The results of the proposed ensemble model were compared with those of CNN and other machine learning classifiers, including SVM, Decision Tree (DT), Random Forest (RF), and XGBoost. The proposed model outperformed the others, achieving a higher accuracy of 84.77% (Adarsh, Arun Kumar, Lavanya, & Gangadharan, 2023).

S. Khan et al. (2024) proposed four hybrid models combining supervised and unsupervised techniques for detecting signs of depression from Twitter posts. BERT, TF-IDF, and the SpaCy package were used as unsupervised NLP techniques for feature extraction, while Artificial Neural Networks (ANN), Logistic Regression (LR), and linear Support Vector Machine (SVM) were employed as supervised classifiers. Model 2 achieved the highest accuracy of 0.994. These models were tested on the Public Sentiment dataset sourced from Kaggle (Khan & Alqahtani, 2024).

Thekkekara et al. (2024) proposed a deep learning model utilizing an attention mechanism on a CNN-BiLSTM (CBA) architecture for depression detection on social media. They conducted a comparative analysis of CBA, CNN, LSTM, and BiLSTM models using the public CLEF 2017 dataset. The results showed that the CBA model outperformed the state-of-the-art models based on AUC-ROC and MCC performance metrics (Philip Thekkekara, Yongchareon, & Liesaputra, 2024).

Dalal et al. **(2024)** proposed a CNN model with multiple channels, called MCNN-IA, which incorporates an individual attention layer capable of capturing global features from the local features of various context levels. They also experimented with CNN, BiLSTM, and a simple MCNN without the individual attention layer. The experimental results of MCNN-IA, using the imbalanced CLEF eRisk 2018 dataset, demonstrated superior performance, achieving accuracy, recall, and F1 scores of 91%, 76.50%, and 70.51%, respectively, compared to the MCNN model without the attention layer (Dalal, Jain, & Dave, 2024).

Three different models proposed by **Kanahuati-Ceballos and Valdivia (2024)** utilized RF, RNN with LSTM and RNN with BiLSTM classifiers for detecting depressive comments in social media posts. They combined two datasets, sourced from Kaggle, one from Twitter and the other from Reddit, for dataset selection. The models were compared based on common evaluation metrics, including accuracy, recall, precision, F1 score, specificity, and sensitivity. Additionally, they employed the Optuna optimization framework on LSTM to fine-tune hyperparameters and identify the optimal combinations for detecting depressive content. Two performance metrics were evaluated: the confusion matrix and the ROC curve, to assess the optimization of the tuned LSTM model. They found that the tuned model achieved higher precision (Kanahuati-Ceballos & Valdivia, 2024)**.**

Gul et al. (2024) proposed a novel system called DORIS for depression detection in social media, leveraging medical knowledge of depression symptoms and emotional features. They employed large language models (LLMs) with transformer architecture on the public SWDD dataset. The DSM-V diagnostic criteria were used to assess the depression scale, and baseline machine learning models combined with NLP techniques were employed for comparative analysis. The evaluation, based on precision, recall, F1 score, AUROC, and AUPRC metrics, showed that their model outperformed the Mood2Content model (Gül, Lebret, & Aberer, 2024)**.**

The next study focuses on detecting the early stages of depression and the transition from healthy to unhealthy mental conditions in Twitter users. Authors **Alhamed et al. (2024)** collected data from November 2018 to February 2019 using specific keywords to gather self-reported tweets from depressive users. They then identified tweets before and after diagnosis from the same users to track changes in mental state. The dataset was used to evaluate the performance of the following models: SVM, Random Forest, BERT, RoBERTa, MentalBERT, and large language models (LLMs) such as GPT-3, GPT-3.5, Bard, and Alpaca. The results showed that BERT, MentalBERT, and RoBERTa outperformed the LLMs (Alhamed et al., 2024)**.**

Zhang et al. (2024) proposed a hierarchical transformer model (HTN) for depression detection from Sina Weibo, a popular social media platform in China. This model was used to generate semantic representations of users' posts. Data collection was performed using two methods: one involved directly searching for the keyword “depression” on Weibo, while the second focused on searching other keywords related to the broader topic of depression. The model incorporated three user-level encoders: one for word embedding, another for post word embedding, and a third for semantic aggregation. The proposed model utilized a Feed forward Neural Network (FNN) and LSTM, both with and without attention mechanisms, to generate user-level semantic representations of depression. Additionally, other deep learning classifiers, including CNN, GRU, bidirectional GRU, and LSTM (with and without attention), were employed for comparison with the proposed model, using different sampling strategies: no sampling, random sampling, and retrieval sampling. HTN outperformed the baseline models, achieving an accuracy of 95% (Z. Zhang et al., 2024).

The study presented by **B. P. Kumar et al. (2024)** proposes an approach using BERT and RNNs to uncover insights through NLP techniques such as semantic embedding, sentiment analysis, and syntactic features from online debates, clinical notes, and social media. The proposed system collected Twitter posts via the API and pre-processed the textual data by removing personal identifiers, stop words, and applying stemming. After pre-processing, LSTM was used for part-of-speech (POS) tagging and grammatical analysis of the texts. This study aims to facilitate the early detection of depressive symptoms in users enrolled in the system via the Twitter API. The code provided achieved 96% accuracy on the tested data (B. P. Kumar et al., 2024).

Table 2: Review of Deep Learning approaches for depression detection in Primary Studies.

References/ Authors	**ML/DL Methods**	**NLP**	**Outperformed**
Tadesse et al. (2019)	SVM, RF, NB, XGBoost LSTM-CNN	TF-IDF,BoW, Statistics and Word2Ve	CNN-LSTM with Word2Vec
Aswathy et al. (2019)	SVM LSTM-CNN	Word2Ve	LSTM-CNN
Shah et al. (2020)	BiLSTM	Word2Ve,FastText,TrainableEmbed, GloVe and Meta feature	Word2Vec+Meta
Kumnunt et al. (2020)	LR, SVM, LSTM, CNN-LSTM , Three-branch CNN-LSTM	TF-IDF, Word2Ve	Three-branch CNN-LSTM with Relu activation Function
Wongkoblap et al. (2021)	MIL-SocNet, MILA-SocNet, Deep Learning	Word2Ve, Topic modelling, User2Vec, LIWC	MILA-SocNet
Mendes et al. (2022)	SVM, XGBoost, LR BERT , RoBERTa, MentalRoberta and BERTimbau	Engineering Features, Embedding features,	BERT , RoBERTa, MentalRoberta, BERTimbau, LR(engineered)
Amanat et al. (2022)	SVM, One-hot+ SVM, NB , LSTM ,RNN,CNN and DBN+One-hot	TF-IDF	CNN+TF-IDF
Gupta et al. (2022)	DT, KNN, SVM, LR and LSTM	TF-IDF	LSTM
Garcia-Noguez et al. (2023)	CNN and XGBoost	GloVe , TF-IDF	CNN with GloVe

Apriliani et al. (2023)	XLNet	Tokenization, stemming	XLNet
A Kumar et al. (2023**)**	RF, KNN, NB, GB and DT, BERT and BiLSTM pipeline	TF-IDF	RoBERTa
Bokolo et al. (2023)	LR, Bernoulli NB, RFRoBERTa, squeezeBARTa, DistiBERT and DeBERTa	TF-IDF	RoBERTa
Adarsh et al. (2023)	SVM-KNN ensemble SVM, DT, RF, XGBoost, CNN	TF-IDF	SVM-KNN
S. Khan et al. (2024)	ANN, LR, Linear SVM, BERT	TF-IDF and SpaCy Package	LR + TF-IDF
Thekkekara et al. (2024)	CNN, LSTM, BiLSTM CNN-BiLSTM (CBA)	Tokenization	CNN-BiLSTM (CBA)
Dalal et al (2024)	CNN, BiLSTM and MCNN(IA)	GloVe	MCNN (IA)
Kanahuati-Ceballos &Valdivia (2024)	RNN-LSTM, RNN-BiLSTM, optimized Fine-Tuned LSTM	Tokenization, lemmatization	optimized Fine-Tuned LSTM
Gul et al. (2024)	DORIS, XGBoost, BERT, RoBERTa, Mood2Content, MentalLLaMA	TF-IDF and FastText	DORIS
Alhamed et al. (2024)	SVM, Random Forests, BERT, RoBERTa, MentalBERT, LLMs: GPT-3, GPT-3.5, Bard, Alpaca and HAN	TF-IDF	RoBERTa and MentalBERT
Zhang et al. (2024)	CNN, GRU, bidirectional GRU, LSTM with and without attention and HTN	Tokenization	HTN

B. P. Kumar et al. (2024)	BERT, RNN and LSTM	Tokenization, stemming and POS embedding	BERT, RNN

4.2 Review of Survey studies

We added some secondary studies in this review to analyse the growth of deep learning and natural language processing techniques in last few years.

William et al. (2021) conducted a systematic review of deep learning models for depression detection from social media textual posts. They found a limited amount of research focused on textual data for depression detection. A total of 1,548 records were retrieved from well-known databases, and 709 were screened, with 83 full papers considered for eligibility. Ultimately, 17 studies were included in the synthesis. The review identified that LSTM was initially used alongside machine learning algorithms such as SVM and Naive Bayes (NB). Later, BiLSTM and attention-based LSTM models emerged, incorporating more advanced features. Ensemble models combining RNN, CNN with LSTM, and BiLSTM enhanced feature extraction, as these models utilized Feed forward Neural Networks for bidirectional sequential interpretation of words and sentences, even with large datasets. The review also highlighted that the BERT transformer model has become widely used due to its self-attention mechanism, which is fine-tuned for automatic feature detection (William & Suhartono, 2021).

A systematic literature review (SLR) conducted by **Salas-Zarate et al. (2022)** screened 192 studies published between 2016 and mid-2021. Out of these, 34 full-text primary studies were selected for review to address their research questions related to depression detection in social media. The study aimed to gain insights into machine learning techniques, linguistic feature extraction methods, computing tools, and statistical tools used in this field. As a result, they found that SVM was the most commonly used classifier, word embedding was the predominant feature extraction method, Python was the most widely used computing tool, and cross-validation was the most frequently employed statistical tool (Salas-Zárate et al., 2022).

The next paper provides an insightful overview of the diverse applications of AI and sentiment analysis (SA) for emotion detection, particularly with the increasing use of social media, digital marketing, and e-commerce platforms. This study, presented by **Chutia & Baruah (2024)**, focuses on extracting emotional features from various industrial sectors, including social media, healthcare, education, human resources, and more. They reviewed studies published between 2013 and 2023, summarizing over 330 papers to support the research objectives of the study. This review explored articles published in scholarly journals, covering techniques, methods, deep learning models, data collection, and evaluation techniques (Chutia & Baruah, 2024).

Tahir et al. (2025) evaluated deep learning and machine learning techniques used for depression detection from social media. A comprehensive review was conducted to achieve objectives such as the evolution of techniques for depression detection, classification methods, datasets used, and identification of open research areas for future exploration. The review included 86 full-text articles out of 144 records identified from academic databases (Tahir et al., 2025)

Table 3. Review of secondary studies (Review papers)

Ref.	Paper Title	No. of records screened for review	Studies included for synthesis of the review
William et al. (2021)	Text-Based Depression Detection on Social Media Posts: A Systematic Literature Review	709	17
Salas-Zarate et.al. (2022)	Detecting Depression Signs on Social Media: A Systematic Literature Review	192	34
Chutia & Baruah (2024)	A review on emotion detection by using deep learning techniques	330	-----
Tahir et al. (2025)	Depression Detection in Social Media : A Comprehensive Review of Machine Learning and Deep Learning Techniques	144	86

4.3 Review of Primary Studies beyond Depression: Exploring Mental Health

We also include some studies that are relevant to this review to explore broader dimensions of depression. We know severe depression may leads to variety of mental health conditions that may be the reason of suicide. So multi-class and multi-task are the mostly used approaches for early detection of depression.

Coppersmith et al. (2018) evaluated a study for suicide risk screening in social media posts. They obtained their data of the users who attempted suicide from OurDataHepls.org, data was provided by the loved ones after they committed suicide. They also used data originally described in other studies which depicts the self-stated users who openly discussed on suicides attempts on social media. They combined the data of these two sources and found total 547 users attempted suicide after 6 months they mentioned in their social media posts. Most of the users were females of 18-24 age groups. They employed bidirectional LSTM to capture more contextual and linguistic representations of the words. GloVe embedding was already employed by the original authors (Coppersmith et al., 2018).

Next study is proposed by **Ramirez–Cifuentes et al. (2021)** to analyse abusive substances and mental health problems in Reddit texts. They created a two multiclass datasets. In first Dataset 4 classes were extracted from the users writings: Suicide (SUI), Depression (DEP), Alcoholism (ALC) and Eating Disorder (ED) , second dataset consists binary classification with Mental Conditions (MCON) in general and Control cases (CON).They employed multiple baseline models Word2vec, DistilBERT and fine-tuned pre-learned GloVe, leveraging embedded features. But proposed variations of enhanced representations outperformed the baseline models (Ramirez-Cifuentes et al., 2021).

Yang et al. (2024) presented a novel approach using variations of MentalLLaMA model. For testing they created their own dataset named Interpretable Mental Health Instruction (IMHI) sourced from 10 previous datasets and contained 105K data samples. It was first ever dataset with multi class and multi-tasking features. Results evaluation was compared with Pre-learned transformer models such as BERT, LLaMAs, RoBERTa, MentalBERT and MentalRoBERTa. Fine-tuned MentalLLaMA-chat-13B outperformed for correct classifications (Yang et al., 2024).

Large language models (LLMs) based approach proposed by **Alhamed et al. (2024)** provided the evidences of suicidal risks that are pre-annotated in user level posts of Reddit. They assigned 4 labels to users' texts for suicide risk assessment from a Subreddit dataset r/SuicideWatch. They prompted the Meta Llama 2 7B open source chat LLM with questions for retrieval of evidences on the behalf of 4 labels: No Risk, Low risk, moderate and high risk. Meta Llama 2 7B with lexicons provided the higher precision of 0.96 (Alhamed, Ive, & Specia, 2024)

Approach used in this study proposed a VADER tool for semantic and syntactic analyses of social media data to identify the changes in behaviour of users. Again this study utilized "Suicide watch" and "depression" Subreddit along with Twitter dataset obtained from Github for testing of proposed models. After Pre-processing feature extraction was done using BERT

Table 4. Exploring Mental Health: Studies beyond depression.

References	**Beyond Depression Studies**	**Approaches**
Coppersmith et al. (2018)	Screening for Suicide Risk	GloVe, LSTM
Ramirez–Cifuentes et al. (2021**)**	Detection of Substance Abuse and Mental Health Issues	Word2vec, DistilBERT and fine-tuned pre-learned GloVe
Yang et al. (2024)	Interpretable Mental Health Analysis	BERT, LLaMAs, RoBERTa, MentalBERT and MentalRoBERTa
Shukla and Sing (2024)	Early Prediction of Behaviour Changes	VADER, BERT, MUSE, LSTM, BiLSTM, GRU, BiGRU and XGBoost
Shetty et al. (2025)	Emotion detection for unravel mental health insights	XLNet, RoBERTa and ELECTRA
Merayo et al. (2025)	Polarity and Stigma of Mental Health Disclosures	SVM, RF, BERT, LSTM and hybrid CNN-LSTM

through two experiments, where stacking classifier with base RF, DT, GB and Meta XGBoost along with BERT and VADER achieved high accuracy of 94, 95 % on Reddit and twitter dataset respectively. In second experiment various embedded models using deep learning were employed including BERT, MUSE, LSTM, BiLSTM with attention GRU, and BiGRU. Proposed model with embedding of

VADER with all of individual models outperformed with BERT, MUSE and XGBoost for Reddit dataset with higher accuracy of 98%, 96%, and 95%, respectively and 97%, 95%, and 94% on X (Twitter) dataset (Shukla & Singh, 2024).

One more similar type of study with ensemble approach was introduced by **Shetty et al. (2025)** for automatic detection of mental health disorders from social media. Three Fine-Tuned models XLNet, RoBERTa and ELECTRA employed on dataset obtained from Reddit social media posts with 15 distinct labels for tracking mental health of the users. After evaluating predictions of individual models, they employed another voting ensemble approach for merging the predictions of all three independent models and this approach achieved higher accuracy of 0.780 (Shetty et al., 2025).

This study provided a novel dataset sourced directly from Instagram for addressing the mental health conditions of the Spanish Instagram followers. **Merayo et al. (2025)** collected the fresh publication from September to December 2020 and excluded males for removing biasness. Only 20 high profile female followers were identified using similar type of comments on depression and mental health issues. They used three labels for this type of comments that is negative, positive and neutral. SVM, RF, BERT, LSTM, hybrid CNN-LSTM were employed on corpus and BERT outperformed the State of art models with 96% accuracy (Merayo et al., 2025).

Review studies help the researchers for detailed examination of primary research for new disclosures following the existing studies. This review enlightened us to about the techniques and various approaches employed by the researchers to find depressive content in the texts of users of social media. We learned that Twitter, Reddit and Facebook are the prominent platforms where users spent their time and share every moment of their life. So maximum studies are performed research on the data of these sites for depression detection and for other mental health analysis.

5. Conclusion

A detailed examination of primary research in the field of depression detection from social media has illuminated several key aspects. 1. The evolution of techniques, incorporating feature embeddings for the detection of mental health disorders, including depression. 2. The application of multimodal, multi-class, and shared-task techniques to address the previously overlooked insights related to depression in social media content. 3. The challenges to be addressed, such as privacy and ethical concerns surrounding user data collection. 4 The use of advanced natural language processing features, such as XLNet, GloVe, and Word2Vec, for pre-processing, which has enhanced the potential of multimodal approaches. 5. The availability of open-source platforms like Kaggle, GitHub, and CLPsych, which provide valuable datasets.

Future research should aim to fill the gaps identified in earlier studies, such as the development of multimodal or ensemble models with feature embeddings using the latest NLP techniques, the creation of multilingual models for cross-cultural platforms, and the discovery of new corpora by incorporating multiple datasets, the exploration of depression detection on platforms beyond Reddit, Twitter, and Facebook. Furthermore, research should be conducted with user consent to mitigate bias and ensure accurate results, while also addressing privacy and ethical concerns during such studies.

This review study provides the overview of techniques and datasets used for depression detection in social media. Future research can be carried out to fulfil the research gaps following the future directions based on exiting studies.

Abbreviations

SVM - Support Vector Machine

NB - Naïve Bayes

LR - Logistic Regression

DT- Decision Tree

ANN - Artificial Neural Network

KNN - K- Nearest Neighbors

GB - Gradient Bosting

XGBoost - Extreme Gradient Bosting

RNN - Recurrent Neural Network

CNN - Convolutional Neural Network

XLNet - Extra-Long Network

LSTM - Long Short-Term Memory

BiLSTM - Bidirectional Long Short-Term Memory

BERT - Bidirectional Encoder Representations from Transformers

RoBERTa - Robustly Optimized Bidirectional Encoder Representations from Transformers

DistiBERT - Distilled Bidirectional Encoder Representations from Transformers

DeBERT - Decoding enhanced Bidirectional Encoder Representations from Transformers

MentalBERT - Mental health Bidirectional Encoder Representations from Transformers

LIME - Local Interpretable Model-Agnostic Explanations

TF- IFD - Term Frequency – Inverts document frequency

Word2Ve - Word to Vector

GloVe - Global Vector

ReLu - Rectified Linear Unit

MIL-SocNet - Multiple Instance Learning and Social Network

MILA-SocNet - Multiple Instance Learning with Anaphora Resolution and Social Network

DASS-42 - Depression, Anxiety, and Stress Scale - 42 Items.

DSM-V - Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

DORIS - **D**iagn**O**stic C**R**iteria-Guided Mood H**IS**tory-Aware

AUC-ROC - Area under Curve-Receiver Operating Characteristic

MCNN-IA - Multiple channel Convolutional Neural Network – Individual Attention

FNN - Feed forward Neural Network

LLMs - Large Language Models

MentalLLaMA - Mental Language Learning Model for Autism

DASS-42 - Depression, Anxiety, and Stress Scale - 42 Items.

DSM-V - Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

DORIS - **D**iagn**O**stic C**R**iteria-Guided Mood H**IS**tory-Aware

SWDD - Social Web Depression Detection

GRU - Gated Recurrent Unit

HTN - Hierarchal Transformer Network

IMHI - Interpretable Mental Health Instruction

VADER - Valence Aware Dictionary and Sentiment Reasoner

MUSE - Multilingual Unsupervised and Supervised Embedding

ELECTRA - Efficiently Learning and Encoder that classifies Token Replacements Accurately

---

References

[1] Adarsh, V., Arun Kumar, P., Lavanya, V., & Gangadharan, G. R. (2023). Fair and Explainable Depression Detection in Social Media. *Information Processing and Management*, *60*(1), 103168. https://doi.org/10.1016/j.ipm.2022.103168

[2] Al Asad, N., Mahmud Pranto, M. A., & Afreen, S. (2019). 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON) : 28th-30th November, FARS Hotel and Resorts, Dhaka, Bangladesh. *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, 13–17.

[3] Alhamed, F., Ive, J., & Specia, L. (2024). *Classifying Social Media Users before and after Depression Diagnosis via Their Language Usage: A Dataset and Study*. 3250–3260. Retrieved from https://aclanthology.org/2024.lrec-main.289

[4] AlHamed, F., Ive, J., & Specia, L. (2024). Using Large Language Models (LLMs) to Extract Evidence from Pre-Annotated Social Media Data. *CLPsych 2024 - 9th Workshop on Computational Linguistics and Clinical Psychology, Proceedings of the Workshop*, (CLPsych), 232–237.

[5] Almouzini, S., Khemakhem, M., & Alageel, A. (2019). Detecting Arabic Depressed Users from Twitter Data. *Procedia Computer Science*, *163*, 257–265. https://doi.org/10.1016/j.procs.2019.12.107

[6] Amanat, A., Rizwan, M., Javed, A. R., Abdelhaq, M., Alsaqour, R., Pandya, S., & Uddin, M. (2022). Deep Learning for Depression Detection from Textual Data. *Electronics (Switzerland)*,

11(5), 1–13. https://doi.org/10.3390/electronics11050676

[7] Apriliani, F., & Maharani, W. (2023). Depression Detection on Social Media Twitter Using Xlnet Method. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *8*(1), 172–180. https://doi.org/10.29100/jipi.v8i1.3345

[8] Bokolo, B. G., & Liu, Q. (2023). Deep Learning-Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques. *Electronics (Switzerland)*, *12*(21). https://doi.org/10.3390/electronics12214396

[9] Chiong, R., Budhi, G. S., Dhakal, S., & Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. *Computers in Biology and Medicine*, *135*, 104499. https://doi.org/10.1016/j.compbiomed.2021.104499

[10] Chutia, T., & Baruah, N. (2024). A review on emotion detection by using deep learning techniques. In *Artificial Intelligence Review* (Vol. 57). Springer Netherlands. https://doi.org/10.1007/s10462-024-10831-1

[11] Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural Language Processing of Social Media as Screening for Suicide Risk. *Biomedical Informatics Insights*, *10*, 117822261879286. https://doi.org/10.1177/1178222618792860

[12] Dalal, S., Jain, S., & Dave, M. (2024). Convolution Neural Network Having Multiple Channels with Own Attention Layer for Depression Detection from Social Data. *New Generation Computing*, *42*(1), 135–155. https://doi.org/10.1007/s00354-023-00237-y

[13] García-Noguez, L. R., Tovar-Arriaga, S., Paredes-García, W. J., Ramos-Arreguín, J. M., & Aceves-Fernandez, M. A. (2023). Automatic classification of depressive users on Twitter including temporal analysis. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *12*(1), 1–13. https://doi.org/10.1007/s13721-023-00434-1

[14] Ghosal, S., & Jain, A. (2023). Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier. *Procedia Computer Science*, *218*, 1631–1639. https://doi.org/10.1016/j.procs.2023.01.141

[15] Govindasamy, K. A. L., & Palanichamy, N. (2021). Depression detection using machine learning techniques on twitter data. *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, (Iciccs), 960–966. https://doi.org/10.1109/ICICCS51141.2021.9432203

[16] Gül, İ., Lebret, R., & Aberer, K. (2024). *Stance Detection on Social Media with Fine-Tuned Large Language Models*. 1–16. Retrieved from http://arxiv.org/abs/2404.12171

[17] Gupta, S., Goel, L., Singh, A., Prasad, A., & Ullah, and M. A. (2022). Psychological Analysis for Depression Detection from Social Networking Sites. *Computational Intelligence and Neuroscience*, *2022*(2022), 28–31.

[18] Hassan, A. U., Hussain, J., Hussain, M., Sadiq, M., & Lee, S. (2017). Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression. *International Conference on Information and Communication Technology Convergence: ICT Convergence Technologies Leading the Fourth Industrial Revolution, ICTC 2017*, *2017-Decem*, 138–140. https://doi.org/10.1109/ICTC.2017.8190959

[19] Islam, M. R., Kabir, M. A., Ahmed, A., 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), 1–12. https://doi.org/10.1007/s13755-018-0046-0

[20] K S, A., P C, R., & Murali, R. (2019). Deep Learning Approach for Detection of Depression in Twitter. *ICSEE2019*, (July 2019).

[21] Kanahuati-Ceballos, M., & Valdivia, L. J. (2024). Detection of depressive comments on social media using RNN, LSTM, and random forest: comparison and optimization. *Social Network Analysis and Mining*, *14*(1). https://doi.org/10.1007/s13278-024-01206-z

[22] Khan, S., & Alqahtani, S. (2024). Hybrid machine learning models to detect signs of depression. *Multimedia Tools and Applications*, *83*(13), 38819–38837. https://doi.org/10.1007/s11042-023-16221-z

[23] Kumar, A., Kumari, J., & Pradhan, J. (2023). Explainable Deep Learning for Mental Health Detection from English and Arabic Social Media Posts. *ACM Transactions on Asian and Low-Resource Language Information Processing*. https://doi.org/10.1145/3632949

[24] Kumar, B. P., Kala, C. N., Durga, G. L., JAswanth, P., & Jasmine, S. (2024). DEEP LEARING FOR DEPRESSION DETECTION FROM TEXTUAL DATA UNVEILING INSIGHTS THROUGH NATURAL LANGUAGE PROCESSING. *Journal of Nonlinear Analysis and Optimization*, *15*(1), 37–48.

[25] Kumar, P., Samanta, P., Dutta, S., Chatterjee, M., & Sarkar, D. (2022). Feature Based Depression Detection from Twitter Data Using Machine Learning Techniques. *Journal of Scientific Research*, *66*(02), 220–228. https://doi.org/10.37398/jsr.2022.660229

[26] Kumnunt, B., & Sornil, O. (2020). Detection of depression in Thai social media messages using deep learning. *DeLTA 2020 - Proceedings of the 1st International Conference on Deep Learning Theory and Applications*, 111–118. SciTePress. https://doi.org/10.5220/0009970501110118

[27] McManus, K., Mallory, E. K., Goldfeder, R. L., Haynes, W. A., & Tatum, J. D. (2015). Mining Twitter Data to Improve Detection of Schizophrenia. *AMIA Joint Summits on Translational Science Proceedings. AMIA Joint Summits on Translational Science*, *2015*, 122–126. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26306253%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4525233

[28] Mendes, A. R., Caseli, H. D. M., Luis, R. W., & Sp, C. (2022). *Identifying Fine-grained Depression Signs in Social Media Posts*. 8594–8604.

[29] Merayo, N., Ayuso-Lanchares, A., & González-Sanguino, C. (2025). Machine Learning Algorithms to Address the Polarity and Stigma of Mental Health Disclosures on Instagram. *Expert Systems*, *42*(2), 1–17. https://doi.org/10.1111/exsy.13832

[30] Pande, S. D., Hasane Ahammad, S. K., Gurav, M. N., Faragallah, O. S., Eid, M. M. A., & Rashed, A. N. Z. (2024). Depression detection based on social networking sites using data mining. *Multimedia Tools and Applications*, *83*(9), 25951–25967. https://doi.org/10.1007/s11042-023-16564-7

[31] Philip Thekkekara, J., Yongchareon, S., & Liesaputra, V. (2024). An attention-based CNN-BiLSTM model for depression detection on social media text. *Expert Systems with Applications*, *249*(PC), 123834. https://doi.org/10.1016/j.eswa.2024.123834

[32] Ramirez-Cifuentes, D., Largeron, C., Tissier, J., Baeza-Yates, R., & Freire, A. (2021). Enhanced Word Embedding Variations for the Detection of Substance Abuse and Mental Health Issues on Social Media Writings. *IEEE Access*, *9*, 130449–130471. https://doi.org/10.1109/ACCESS.2021.3112102

[33] Salas-Zárate, R., Alor-Hernández, G., Salas-Zárate, M. D. P., Paredes-Valverde, M. A., Bustos-López, M., & Sánchez-Cervantes, J. L. (2022). Detecting Depression Signs on Social Media: A Systematic Literature Review. *Healthcare (Switzerland)*, *10*(2), 1–22. https://doi.org/10.3390/healthcare10020291

[34] Shah, F. M., Ahmed, F., Saha Joy, S. K., Ahmed, S., Sadek, S., Shil, R., & Kabir, M. H. (2020). Early Depression Detection from Social Network Using Deep Learning Techniques. *2020 IEEE Region 10 Symposium, TENSYMP 2020*, *1*(June), 823–826. https://doi.org/10.1109/TENSYMP50017.2020.9231008

[35] Shetty, N. P., Singh, Y., Hegde, V., Cenitta, D., & Dhruthi, K. (2025). Exploring emotional patterns in social media through NLP models to unravel mental health insights. *Healthcare*

Technology Letters, (October 2024), 1–17. https://doi.org/10.1049/htl2.12096

[36] Tadesse, Michael M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. *IEEE Access*, *7*, 44883–44893. https://doi.org/10.1109/ACCESS.2019.2909180

[37] Tadesse, Michael Mesfin, Lin, H., Xu, B., & Yang, L. (2019). Detection of suicide ideation in social media forums using deep learning. *Algorithms*, *13*(1), 1–19. https://doi.org/10.3390/a13010007

[38] Tahir, W. B. I. N., Khalid, S., Almutairi, S., Memon, S. A. L. I., & Khan, J. (2025). Depression Detection in Social Media : A Comprehensive Review of Machine Learning and Deep Learning Techniques. *IEEE Access*, *PP*, 1. https://doi.org/10.1109/ACCESS.2025.3530862

[39] Vasha, Z. N., Sharma, B., Esha, I. J., Al Nahian, J., & Polin, J. A. (2023). Depression detection in social media comments data using machine learning algorithms. *Bulletin of Electrical Engineering and Informatics*, *12*(2), 987–996. https://doi.org/10.11591/eei.v12i2.4182

[40] William, D., & Suhartono, D. (2021). Text-based Depression Detection on Social Media Posts: A Systematic Literature Review. *Procedia Computer Science*, *179*(2019), 582–589. https://doi.org/10.1016/j.procs.2021.01.043

[41] Wongkoblap, A., Vadillo, M. A., & Curcin, V. (2021). Deep learning with anaphora resolution for the detection of tweeters with depression: Algorithm development and validation study. *JMIR Mental Health*, *8*(8). https://doi.org/10.2196/19824

[42] Yang, K., Zhang, T., Kuang, Z., Xie, Q., Huang, J., & Ananiadou, S. (2024). MentaLLaMA: Interpretable Mental Health Analysis on Social Media with Large Language Models. *WWW 2024 - Proceedings of the ACM Web Conference*, 4489–4500. https://doi.org/10.1145/3589334.3648137

[43] Zhang, T., Yang, K., Ji, S., & Ananiadou, S. (2023). Emotion fusion for mental illness detection from social media: A survey. *Information Fusion*, *92*(December 2022), 231–246. https://doi.org/10.1016/j.inffus.2022.11.031

[44] Zhang, Z., Zhu, J., Guo, Z., Zhang, Y., Li, Z., & Hu, B. (2024). Natural Language Processing for Depression Prediction on Sina Weibo: Method Study and Analysis. *JMIR Mental Health*, *11*, e58259–e58259. https://doi.org/10.2196/58259