

Detecting Depression using Natural Language Processing (NLP) and Machine Learning from Social Media and Clinical Notes: A Systematic Literature Review

Hadeel Taha Alsakkaf

Table of Contents

Abstract	3
Introduction	4
Background.....	5
The Global Mental Health Crisis	5
Limitations of Traditional Depression Detection.....	5
The Rise of Digital Mental Health & Passive Sensing.....	5
Problem Statement.....	6
Research Objectives	7
Research Scope.....	8
Literature Review.....	9
Research Methodology	15
Phase 1: Literature Review.....	15
Phase 2: Qualitative Research	16
Phase 3: Quantitative Research	16
Phase 4: Synthesis of Findings and Recommendations	16
Limitations	17
Expected Outcomes	17
Gantt chart	19
References	20

Abstract

In recent years, the rapid rise of artificial intelligence (AI) has brought both opportunities and responsibilities for integrating these advancements into society. Depression is still one of the greatest burdens on humanity, and early detection has the potential to significantly improve outcomes. Numerous studies have investigated detecting depression using AI, particularly through natural language processing (NLP) and machine learning (ML). Despite promising results in controlled research settings, significant challenges persist in translating these models to real-world clinical practice. Issues such as limited dataset sizes, overfitting, lack of external validation, and poor model explainability continue to delay progress. This systematic literature review examines recent studies in this field, with a focus on the progress achieved to date and the critical challenges and limitations that remain to be addressed for clinical adoption.

Keywords: Natural Language Process, Machine learning, Depression Detection, Systematic literature review.

Introduction

Depression is one of the most burdensome mental health disorders worldwide, affecting approximately 332 million people and significantly harming the quality of living, relationships, and professional function (World Health Organization, 2021). Regardless, depression is still underdiagnosed and untreated, partially due to barriers such as limited access to mental health professionals, reliance on self-reported screening tools that require active patient participation, or even stigma.

Recently, individuals have been using social media platforms and online forums to express emotions, seek advice, and form communities. These platforms have emerged as rich, large-scale sources of textual data for mental health research. However, manually analyzing such a large amount of data is impractical, requiring advanced computational tools such as Natural Language Processing (NLP), data mining, and machine learning (ML).

Depression detection using NLP has been documented in numerous studies, these studies have shown a promising accuracy in controlled research settings. However, a critical gap still exists between this controlled accuracy and the adaptability required for real-world clinical deployment. For instance, Vaci et al., (2020) found that even models trained on 526 annotated clinical papers did not perform well when confronted with complex, real-world medical entities. Similarly, Tadesse et al., (2019) reported that limited dataset size led to overfitting when detecting depression-related posts on Reddit, undermining generalizability.

These limitations introduce a fundamental challenge: models trained on social media data often learn correlational patterns rather than clinically meaningful indicators (Alghamdi et al., 2020). This can lead to false positives (e.g., mistaking grief for depression), false negatives (e.g., missing a depressed individual who posts positively), and low performance when applied to clinical data. Besides, the lack of external validation of NLP models is essential for clinical adoption.

Following, this research aims propose strategies to improve detection accuracy, evaluate existing methodologies, and labeling practices, and synthesize key limitations affecting real-world relevance, through identifying the factors that contribute to performance degradation and recommending future research directions.

Background

The Global Mental Health Crisis

The global mental health crisis is a widely recognized concept in clinical practice and professional discourse, as it profoundly affects individuals, families, and entire societies. This crisis manifests in acute psychiatric episodes, significant social disruption, and elevated suicide rates. Addressing this crisis requires urgent, scalable, and innovative approaches to mental health screening and intervention.

Limitations of Traditional Depression Detection

Usual depression detection techniques meet a major limitation that impedes early diagnosis. Many individuals rely on self-diagnosing, which is described as problematic, as individuals may under-report symptoms or over-report. Clinician's diagnosis introduces subjectivity, with assessments varying based on individual experience and bias, and Frequent visits to a professional are critical, as in a single clinic visit provides only a momentary snapshot. Additionally, mild or atypical depression usually does not get detected, which leads to missed cases and delayed intervention. Overall, these limitations underscore the urgent need for complementary, scalable, and objective detection approaches.

The Rise of Digital Mental Health & Passive Sensing

The growth of social media offers a new way to detect depression, with language offering a powerful natural signal. people generate a large amount of textual data from social media, journals, and therapy transcripts. This amount of data provides a valuable insight that can be analyzed and studied. Studying these researchers found a clear pattern in psycholinguistics, people with depression use first-person pronouns like "I," "me," and "my," which shows self-focused thinking. They also use more negative words such as "sad," "lonely," and "hopeless," along with absolute terms like "always," "never," and "everything" this reflect all-or-nothing thinking. Additionally, they use less positive words like "happy" or "love" Frequent use of words like "because" "think," and "know" may indicate hesitation. Overall, the language becomes more negative, and uncertain. These patterns make text a useful tool for detecting depression early and monitoring it over time.

Problem Statement

Depression is described as one of the biggest burdens affecting a human life, with approximately 332 million people affected worldwide. Depression can affect the quality of everyday life, relationships with family and friends, and even work and school life (WHO, 2021). With social media platforms, individuals can express, seek advice, or form communities. As a result, these online platforms have become a major source of data, providing valuable insights for investigation. For this amount of data, it will require having an advanced collecting tool for extracting, such as natural language process (NLP), data mining, and machine learning (ML).

Detecting depression using NLP and ML has been studied in numerous research. Sourcing data from social media platforms such as Reddit (Tadesse et al., 2019) and from medical records (Vaci et al., 2020). Nonetheless, there is still a critical gap between the high accuracy achieved in controlled research settings and the adaptability required for real-world clinical deployment. According to Veci et al. (2020), even with data from 526 annotated papers, the model did not perform as expected with complicated medical entities. Similarly, detecting depression via Reddit (Tadesse et al., 2019) was taken back by the limited data size, which led to overfitting. This proves the high need for a larger dataset to improve testing and overall model performance. Moreover, several studies rely on a narrow range of NLP techniques where they usually apply one approach such as dictionary-based methods (e.g., LIWC) or basic machine learning classifiers. To the contrary, some studies adopt complex deep learning architectures, while neglecting more interpretable linguistic methods. Deep learning models achieves a high accuracy by capturing contextual relationships but operate as "black boxes," which leads to limited understanding of why the choice is made.

These interconnected challenges small dataset sizes, single-platform reliance, poor labeling methodologies, reliance on a single NLP approach, black-box opacity. This collectively explain why real-world clinical practice has not yet used NLP-based depression detection. Addressing these limitations is essential for bridging the gap between research accuracy and clinical applicability.

Research Objectives

1. To address the gap between social media platform and real-world detecting technique (Controlled accuracy vs. ecological validity)
 - a. To acknowledge the difference between the real-life diagnosing technique and how can it affect detecting depression.
 - b. To propose and discuss methods to improve the accuracy of detecting techniques.
2. To establish a comprehensive evaluation framework that prioritizes clinical relevance.
 - a. To explain how NLP/ML can assist in the medical field.
 - b. To discuss the need for a clinically approved model.
3. To systematically identify and select relevant studies on NLP/ML based depression detection.
 - a. To define the search strategy.
4. To critically analyze the methodologies employed in existing studies.
 - a. To critically discuss the overreliance on single platforms.
 - b. To investigate the lack of external validation.
 - c. To look over the choices of NLP/ML features.
5. To evaluate the datasets used in existing research.
 - a. To label methodology and its limitations.
 - b. To discuss Dataset size and overfitting concerns.

6. To identify and synthesize the key limitations affecting real-world applicability.
 - a. To identify the key challenges and limitations of current NLP/ML models
 - b. To understand the underlying causes of performance degradation when models are applied outside controlled research settings.

Research Scope

Analysis of Existing Challenges: emphasizing and understanding the issues that affect real-life application of NLP/ML models for depression detection.

Highlighting that a small dataset, reliance on single-platform data (e.g., Reddit only), poor explainability of deep learning models, would lead to overfitting. Small datasets provide insufficient variation, causing models to memorize noise instead of learning generalizable patterns (Tadesse et al., 2019).

Evaluation of Current Solutions: Exploring existing solutions from cross-domain validation techniques, multi-modal integration (combining text with clinical or behavioral data), longitudinal modeling, clinician-in-the-loop evaluation to explainable AI frameworks (e.g., SHAP, LIME). Looking into ethical guidelines for mental health in AI. This evaluation seeks to assess the efficacy and limitations of these solutions in addressing the identified challenges.

Synthesis of Findings and Recommendations: Integrating insights from the analysis and evaluations to produce a structured synthesis of the current state of knowledge. This synthesis aims to offer practical recommendations for researchers and clinicians to enhance the trustworthiness, generalizability, and clinical relevance of NLP-based depression detection systems.

Literature Review

Teferra et al. (2024), published on JMIR publications, discusses a comprehensive literature review on the use of Natural Language Processing (NLP) techniques for screening and detecting depression from textual data, sourcing most of their data from social media posts, clinical interviews, and textual mental health records. These datasets require specific NLP techniques such as sentiment analysis, linguistic markers, word embedding, and advanced models like CNNs, RNNs, and large language models. This finding contributes to a powerful toolkit for detecting not only depression but also is a helpful tool to early detect mental illness disorders. In addition, the authors discuss the ethical concerns of the individual, lack of textual understanding, and bias and fairness issues. Overall, this paper concludes with how powerful and potential NLP is in improving mental illness detection, but further work is needed to address ethical, technical, and cross-cultural challenges.

DeSouza, Robin, Gumus, and Yeung in 2021 discuss in this paper the use of natural language processing (NLP) to detect late-life depression (LLD) compared to younger adults using speech. Therefore, studying a speech dataset will need computational linguistics with statistical, machine learning, and deep learning models to take unstructured, free-form data (e.g., a voice recording or writing sample) and produce structured, quantitative acoustic and linguistic outputs. Nevertheless, using NLP in LLD will have its unique challenges in comprehending the identification of speech alterations, seeing older adults as it is more possible for them to have additional medical comorbid conditions that may affect their way of speech. Results say that NLP can recognize the speech changes for depressed individuals. However, the review may not have included all pertinent studies because it is narrative rather than systematic. It is challenging to distinguish depression-specific alterations in older persons, as they frequently have medical comorbidities and take various medications that impact speech. Concerns about misdiagnosis are raised by the overlap between speech abnormalities and dementia. The authors note that more studies with larger sample sizes and closely matched control groups are required.

In 2019 Vaci et al, published Natural language processing for structuring clinical text data on depression using UK-CRIS. This research paper demonstrates how Natural Language Processing (NLP) can elevate the medical field in diagnosing individuals by developing an algorithm that extracts clinical information from unstructured Electronic Health Records (EHRs) of patients with depression. The goal of the study is to support the Clinical Record Interactive Search (CRIS), which provides real-world data. The authors describe their methodology in two steps as follows: 1) Clinical experts identify the relevant information and then develop the training and testing datasets used for building statistical models. 2) They developed and refined the statistical models through active learning procedures to iteratively enhance performance while minimizing the annotation workload. This study demonstrates the success of the NLP models, which outperform human annotators in writing prescriptions. However, models can fail to achieve satisfactory results on more complex medical entities. Therefore, they introduce the model with a new large dataset into the training to adapt the model to a new pattern of textual data. Nevertheless, adding more information disrupted the weights of the categories, so the model underperformed at recalling subsequent information. In summary, the study shows how well-developed NLP models are in extracting complex medical information; the records show high accuracy results regarding drug-related information, but do not give the same accuracy on the auxiliary variables.

Alghamdi, Mahmoud, Abraham, Alanazi, and Hernández (2023) present a study in Natural Language Processing (NLP) and a machine learning (ML) approach to predict depression symptoms from Arabic text posted online. This study has been done before on English and Japanese, but not in Arabic. The authors' goal was to address the lack of research in the Arabic language in mental health prediction. Therefore, they developed both a lexicon-based approach and a machine learning-based approach, where they collected 20,000 posts from "Nafsany," which is an online platform that covers substantive subjects regarding diverse medical circumstances. Thus, they labeled the collected data into categories and further annotated them by psychologists using the "Diagnostic and Statistical Manual of Mental Disorders" (DSM-5) and "Patient Health Questionnaire" (PHQ-9). For the lexicon-based approach, they created the ArabDep Lexicon to collect depression-related terms using category-based and annotation-based methods. As for the

machine-learning approach, six classifiers (AdaBoost, Decision Trees, KNN, Random Forest, SGD, and SVM) were used with many extraction methods such as n-grams and word embedding. Ultimately, they achieved promising performance with an accuracy that exceeds 80%, a recall of 82%, and a precision of 79% for the lexicon-based approach, and with an accuracy that exceeds 73%, a recall of 72%, and a precision of 74% for the machine-learning-based approach. Although the study shows reasonable results, further study is needed.

A study on Detection of Depression-Related Posts in Reddit Social Media Forum done by Tadesse, Lin, Xu, and Yang published in 2023. The study showed in this paper presents a Natural Language model (NLP) and Machine learning (ML) approach to find depression words from Reddit social media posts. They aimed to identify linguistic markers of depression and improve detection performance through proper feature selection and feature combinations. With data of 1,293 depression indicative posts and 548 standard posts from Reddit, which were labeled under the subreddit categories. The authors followed these three methodologies: 1) pre-processing the data, which includes removing URLs, punctuation, and stop words, and applying stemming to reduce words to their root forms. 2) extracted three types of features: N-grams (unigrams and bigrams using TF-IDF), LIWC (68 psycholinguistic categories), and LDA topic modeling (70 topics) 3) evaluated five classifiers: Logistic Regression, Support Vector Machine (SVM), Random Forest, AdaBoost, and Multilayer Perceptron (MLP). This resulted in 80% accuracy and 0.080 F1-score. However, the best overall performance was achieved by combining LIWC, LDA, and bigram features with the MLP classifier, reaching 91% accuracy and 0.93 F1-score. Finally, the authors state that the reasonable results are great for future work.

Uddin, Dysthe, Følstad and Brandtzaeg (2021) in Deep learning for prediction of depressive symptoms in a large textual dataset present a research paper suggests Deep Learning methods to identify depression/depressive symptoms from textual data using Recurrent Neural Network (RNN) based on Long Short-Term Memory (LSTM). Their aim is to develop an efficient system to identify texts with self-perceived symptoms of depression using symptom-based features rather than traditional word frequency techniques. 11,807 to 21,470 text entries classified as depression or non-depression by medical and psychological experts were gathered

from a Norwegian public internet information channel (ung.no). The authors used these three approaches: 1) used a one-hot encoding process to extract strong features based on 189 unique phrases that reflect potential symptoms of depression as described by experts. 2) demonstrated superior clustering compared to conventional TF-IDF and one-hot techniques by using Linear Discriminant Analysis (LDA) to visualize and verify the robustness of the suggested features. 3) used the text data to train an LSTM-based RNN with an attention layer to simulate time-sequential emotional experiences. The authors concluded that their symptom-based feature approach with LSTM-RNN is highly effective for automatic depression detection in text after using Local Interpretable Model-Agnostic Explanations (LIME) to provide explainable AI decisions, demonstrating the significance of features in predicting depression.

Study by Zhang, Schoene, Ji and Ananiadou (2022) in Natural language processing applied to mental illness detection: a narrative review analyzed 399 research articles from over 10,000 records. They focused on NLP methods to detect mental illness from multiple data sources such as social media posts, electronic health records, clinical interviews, and online forums. The review reveals an ascending trend in NLP-driven mental illness detection research. The authors state that classical machine learning methods account for 59% of approaches, typically employing features such as Part-of-Speech tags, Bag-of-Words, Linguistic Inquiry and Word Count (LIWC), and n-grams. However, deep learning methods have shown superior performance. These methods are categorized into CNN-based approaches (17%), RNN-based methods (36%), transformer-based methods (17%), and hybrid models (30%). There are still several challenges, such as data quality and volume, which remain significant obstacles, as most supervised learning approaches require extensive human annotation, which is time-consuming and expensive, and model instability due to class imbalance, noisy labels, and diverse writing styles across different data sources. The review concludes with presenting promising results and suggesting interpretable models, novel deep learning paradigms, and multi-modal approaches to support clinician adoption of mental illness detection applications.

The paper Depression detection from social network data using machine learning techniques, published in 2018 by Islam, Kabir, Ahmed, et al., presents a machine

learning approach to detect depression in Facebook by analyzing their comments and posts. applying multiple psycholinguistic features and machine learning classifiers to automatically detect depressive content. from 7,145 comments, 58% classified as depression-indicative and 42% as non-depressive. The researchers extracted features from these comments using LIWC (Linguistic Inquiry and Word Count) software, and four machine learning classifiers were tested: Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Decision Tree, and Ensemble methods. The researchers used 10-fold cross-validation and evaluated performance using precision, recall, and F-measure metrics. The results show the decision tree with the best performance, achieving the highest recall and F-measure scores across different feature combinations. This study proves that machine learning techniques can effectively detect depression by analyzing linguistic and emotional patterns in social media posts.

In Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model by Tejaswini, Babu, Sahoo (2024) they proposed FCL (Fasttext Convolution Neural Network with Long Short-Term Memory) model, a hybrid deep learning architecture designed to detect depression from social media text with high accuracy. The technique combines three key components: fasttext embeddings for superior text representation, convolutional neural networks (CNNs) to extract global features, and long short-term memory (LSTM) networks to capture local features with long-term dependencies. They collected real-world data from Twitter and Reddit using NLP. They performed comprehensive preprocessing steps, including tokenization, lowercasing, stop word removal, stemming, and lemmatization. The FCL model outperformed existing techniques across both datasets. On the Reddit dataset, FCL achieved 87% accuracy compared to 84.6% for CNN with word2vec and 77.7% for LSTM-RNN with word2vec. On the Twitter dataset, FCL achieved 88% accuracy, outperforming CNN with word2vec at 84.8% and LSTM-RNN at 85.8%. The proposed method showed improvements of 2.4-14.2% in accuracy over advanced baseline methods. This approach proves that combining methods can upgrade the depression detection research field.

Alsagri and Ykhlef 2020 study paper in Machine Learning-Based Approach for Depression Detection in Twitter Using Content and Activity Features proposes

extracting Twitter data to detect depression using machine learning techniques. This study follows a quantitative approach to train and test various machine learning classifiers. The data preparation involved multiple preprocessing steps such as tokenization, normalization, and stemming. The paper evaluates three main classification algorithms: Support Vector Machine (SVM), Naive Bayes (NB), and Decision Tree (DT). Alsagri and Ykhlef developed a Depression Detection using Activity and Content Features (DDACF) classification model that combines features from both user account activities and tweet content analysis. The experimental results demonstrated that enriching the model with discriminant features yields better results. The SVM-linear classifier achieved the best performance with an accuracy of 82.5%, precision of 0.7391, recall of 0.85, and F-measure of 0.79. When combining categorical account measures with synonyms, SVM-linear significantly outperformed other classifiers. The authors concluded that their main contribution lies in merging both tweet text analysis with account behavioral trends to monitor changes in users' mental and psychological states. The study can serve as a foundational step toward building complete social media-based platforms for analyzing and predicting mental health issues.

Research Methodology

A systematic literature search was conducted using Semantic Scholar, PubMed, and Google Scholar to identify studies on depression detection using Natural Language Processing (NLP). Search keywords included "depression screening," "depression detection," "natural language processing," and "machine learning." Studies were included if they discussed the application of NLP or machine learning techniques for depression screening or detection, with no exclusion based on data source (e.g., social media, clinical records, or online forums).

Identified studies were screened for relevance based on their titles and abstracts, followed by full-text review for eligible articles. Data were extracted from each included study and synthesized to identify common themes, with a specific focus on two areas: the challenges reported in developing NLP/ML models for depression detection, and the existing solutions proposed or evaluated to address those challenges. These themes were analyzed to assess the effectiveness and limitations of current approaches.

This methodology enabled a structured understanding of the current status of NLP-based or ML-based depression detection, highlighting the primary challenges faced by researchers and the strategies developed to overcome them.

Phase 1: Literature Review

The first phase involves a systematic literature review to identify and synthesize existing research on NLP-based or ML-based depression detection. This review covers scholarly articles and conference papers that discuss the application of NLP techniques or ML techniques for depression screening or detection. The purpose is to map the current landscape of the field, identifying the main challenges reported in existing studies and the solutions proposed to address them.

Data Sources: Semantic Scholar, PubMed, Google Scholar, IEEE Xplore, and the ACM Digital Library.

Selection Criteria: Peer-reviewed publications from the last ten years, focusing on depression detection, natural language processing, and machine learning approaches to mental health text analysis.

Analysis Technique: Thematic analysis to identify recurring themes, challenges, and solutions across the literature.

Phase 2: Qualitative Research

The second phase involves a qualitative analysis of the included studies to extract and synthesize the experiences, challenges, and solutions reported by researchers in the field. Rather than conducting primary interviews, this phase analyzes qualitative data already present in the literature, including authors' discussions of limitations, methodological challenges, and proposed solutions.

Unit of Analysis: Authors stated challenges, limitations, methodological justifications, and recommendations for future work.

Analysis Technique: Qualitative content analysis to extract recurring themes related to challenges in NLP-based depression detection and the effectiveness of existing solutions.

Phase 3: Quantitative Research

This phase complements the qualitative insights with quantitative data extracted from the included studies, assessing the frequency of identified challenges and evaluating the reported effectiveness of existing solutions.

Data Extracted: Dataset size reported performance metrics (accuracy, precision, recall, F1, AUC), NLP/ML techniques employed, and validation methods used.

Analysis Approach: Descriptive statistics to summarize the distribution of reported metrics across studies, including range, mean, and median values. Comparative analysis will examine whether performance differs by data source (social media vs. clinical), model type (interpretable vs. black box), or dataset size.

Phase 4: Synthesis of Findings and Recommendations

Based on insights gathered from the previous phases, a comprehensive synthesis of findings will be developed. This phase integrates the identified challenges and existing solutions from the literature into a structured set of recommendations for future research and practice in NLP-based or ML-based depression detection.

Synthesis Development: Integration of identified challenges reported limitations, and effective strategies into a coherent set of findings organized by the six research objectives.

Recommendations Formulation: Actionable recommendations for researchers, clinicians, and policymakers based on gaps identified in the literature.

Evaluation Criteria: Recommendations will be assessed for clarity, applicability, comprehensiveness, and alignment with solutions that have shown promise in existing studies.

Limitations

This research methodology acknowledges potential limitations that may affect the findings and conclusions of this systematic literature review. Efforts will be made to reduce these limitations as much as possible.

Expected Outcomes

This systematic literature review aimed to produce several key results to check research objectives. These outcomes will contribute to the academic understanding of NLP-based depression detection and provide practical guidance for future research and development.

Comprehensive Analysis of Challenges in NLP and ML Based Depression Detection: A detailed examination of the current challenges and barriers to effective depression detection using NLP/ML techniques across various data sources (social media, clinical records, online forums). This analysis is expected to highlight the multifaceted nature of the problem, including issues related to: dataset size and quality, Labeling methodology limitations, overfitting and generalizability concerns, lack of external validation, model explainability and transparency, and absence of clinical approval pathways.

Evaluation of Existing Solutions: An assessment of current methodologies, NLP techniques, ML techniques and validation approaches used in the literature. This evaluation is anticipated to uncover: Which NLP techniques (e.g., traditional ML, deep learning, transformers, linguistic features) have shown the most promise, the effectiveness of different validation practices (e.g., cross-validation, external

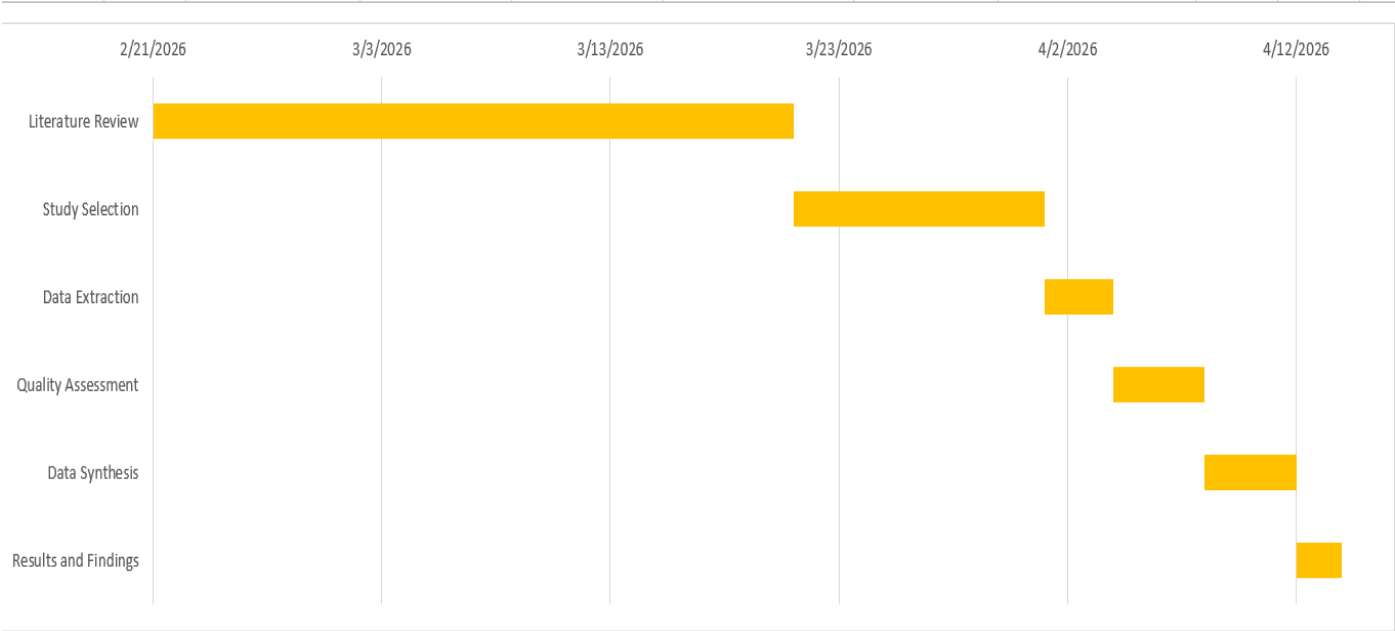
validation), the strengths and weaknesses of interpretable vs. black-box approaches, gaps in existing solutions, particularly regarding clinical applicability.

Synthesis of Key Limitations Affecting Real-World Applicability: A synthesized understanding of the persistent limitations that prevent current NLP/ML models from being deployed in real-world clinical settings. This synthesis will identify the most common and critical limitations reported across studies, the underlying causes of performance degradation when models are applied outside controlled research settings, and the relationship between methodological choices and reported performance.

Contribution to Academic Discourse: The enrichment of academic discourse on NLP-based mental health detection. By providing a well-researched and critically analyzed body of work, this review aims to: Stimulate further discussion on the gap between research accuracy and clinical applicability, bridge the divide between computational approaches and clinical requirements, and encourage more rigorous validation practices in future research.

Foundation for Future Research: Establishment of a solid foundation for future research on NLP/ML based depression detection. This study is expected to Identify new research questions and areas for further investigation and provide a benchmark for evaluating the quality of future studies.

Gantt chart



	Literature Review	Study Selection	Data Extraction	Quality Assessment	Data Synthesis	Results and Findings
Start Date	2/21/2026	3/21/2026	4/1/2026	4/4/2026	4/8/2026	4/12/2026
Duration	28	11	3	4	4	2

References

1. Teferra BG, Rueda A, Pang H, Valenzano R, Samavi R, Krishnan S, Bhat V
Screening for Depression Using Natural Language Processing: Literature Review
Interact J Med Res 2024;13:e55067. <https://www.i-jmr.org/2024/1/e55067/>
2. DeSouza DD, Robin J, Gumus M and Yeung A (2021) Natural Language Processing as an
Emerging Tool to Detect Late-Life Depression. Front. Psychiatry 12:719125. doi:
10.3389/fpsyt.2021.719125.
<https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2021.719125/full>
3. Vaci N, Liu Q, Kormilitzin A, De Crescenzo F, Kurtulmus A, Harvey J, et al. Natural
language processing for structuring clinical text data on depression using UK-
CRIS. *Evidence Based Mental Health*. 2020;23:.. <https://doi.org/10.1136/ebmental-2019-300134>
4. N. S. Alghamdi, H. A. Hosni Mahmoud, A. Abraham, S. A. Alanazi and L. García-
Hernández, "Predicting Depression Symptoms in an Arabic Psychological Forum,"
in *IEEE Access*, vol. 8, pp. 57317-57334, 2020, doi: 10.1109/ACCESS.2020.2981834.
keywords: {Depression;Feature extraction;Natural language processing;Social
networking (online);Mental health;Mental disorders;Supervised learning;semi-supervised
learning;machine learning;predictive models;depression;lexicon;text analysis}
<https://ieeexplore.ieee.org/document/9040556/authors#authors>
5. M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of Depression-Related Posts in
Reddit Social Media Forum," in *IEEE Access*, vol. 7, pp. 44883-44893, 2019, doi:
10.1109/ACCESS.2019.2909180. keywords: {Feature extraction;Task
analysis;Linguistics;Twitter;Natural language processing;Support vector
machines;Natural language processing;machine learning;Reddit;social
networks;depression}, <https://ieeexplore.ieee.org/document/8681445/authors#authors>
6. Uddin, M.Z., Dysthe, K.K., Følstad, A. *et al.* Deep learning for prediction of depressive
symptoms in a large textual dataset. *Neural Comput & Applic* **34**, 721–744 (2022).
<https://doi.org/10.1007/s00521-021-06426-4>
7. Zhang, T., Schoene, A.M., Ji, S. *et al.* Natural language processing applied to mental
illness detection: a narrative review. *npj Digit. Med.* **5**, 46 (2022).
<https://doi.org/10.1038/s41746-022-00589-7>
8. Islam MR, Kabir MA, Ahmed A, et al. Depression detection from social network data
using machine learning techniques. *Health Information Science and Systems*. 2018
Dec;6(1):8. DOI: 10.1007/s13755-018-0046-0. PMID: 30186594; PMCID:
PMC6111060.
9. <https://link.springer.com/article/10.1007/s13755-018-0046-0>

10. Vankayala Tejaswini, Korra Sathya Babu, and Bibhudatta Sahoo. 2024. Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model. ACM Trans. Asian Low-Resour. Lang. Inf. Process. 23, 1, Article 4 (January 2024), 20 pages. <https://doi.org/10.1145/3569580>
11. Hatoon S. ALSAGRI, Mourad YKHLEF, Machine Learning-Based Approach for Depression Detection in Twitter Using Content and Activity Features, IEICE Transactions on Information and Systems, 2020, Volume E103.D, Issue 8, Pages 1825-1832, Released on J-STAGE August 01, 2020, Online ISSN 1745-1361, Print ISSN 0916-8532, <https://doi.org/10.1587/transinf.2020EDP7023>, https://www.jstage.jst.go.jp/article/transinf/E103.D/8/E103.D_2020EDP7023/article/-char/en