International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Depression Detection Using Machine Learning and Deep Learning Techniques

*Neeharikaa Vara Sree Yenugutalaa*

Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh, India
neeharikasep12@gmail.com

**ABSTRACT:**

Depression is a mental health condition characterized by prolonged feelings of sadness, hopelessness, and disinterest in activities. Detection of depression involves identifying and diagnosing the condition. Detecting depression early is crucial for effective treatment and improving overall well-being, as it can significantly impact an individual's quality of life. Deep learning is a branch of AI and machine learning that involves training artificial neural networks to perform tasks and make predictions by simulating the structure and function of the human brain. It plays a vital role in detecting depression by analyzing various data sources, such as text, speech, and images. Machine learning is a type of AI that enables computers to learn and make decisions without being explicitly programmed. This is done by analyzing patterns and relationships in data, allowing computers to improve over time. Machine learning can detect depression by analyzing patterns in various forms of data such as language, tone, or behavior markers. With the help of large datasets, algorithms can predict depression in individuals. For Detecting Depression we use various Machine Learning and Deep Learning Techniques like Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), and Naive Bayes (NB).

**Keywords:** Depression Detection, Machine Learning, Random Forest, Naive Bayes, Support Vector Machine, Convolutional Neural Network, Recurrent Neural Network

## 1. INTRODUCTION

Depression detection refers to the process of identifying and assessing signs and symptoms related to depression in individuals. This important task involves various methods for finding signs and symptoms of this condition in individuals, including traditional techniques like clinical assessments and self-report questionnaires, as well as more modern approaches using machine learning and deep learning algorithms. Advanced algorithms use data from various sources such as social media, text messages, and physiological indicators to identify patterns associated with depression. By analyzing language, sentiment, and behavior, these algorithms provide useful insights into an individual's mental health status. This technological advancement has the potential to provide early intervention and support for those who are struggling with depression, ultimately improving the overall well-being of individuals. Machine learning plays a crucial role in depression detection by examining diverse data sources like text, speech, physiology, and behavior. It deciphers linguistic cues and sentiments in written content, revealing depression indicators. In speech, it spots changes in tone and rhythm, reflecting emotional distress. Physiological signals like heart rate variability and facial expressions divulge further patterns. Surveys and questionnaires, digitally analyzed, provide precise mental health assessments. Mobile apps and wearables enhance monitoring, capturing daily activities and interactions. By combining machine learning with traditional methods, timely identification and support can be provided for those facing depression, leading to more effective interventions and better outcomes in mental health care. Deep learning is instrumental in detecting depression by extracting intricate patterns from complex data. Through neural networks, it can process diverse sources such as text, speech, and images to discern subtle emotional cues indicative of depression. For example, in text analysis, recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) excel at understanding context and capturing nuances in written expression. In image analysis, convolutional neural networks (CNNs) decode facial expressions, body language, and other visual cues associated with depression. Deep learning's ability to learn hierarchies of features enables it to uncover nuanced indicators, significantly enhancing the accuracy and effectiveness of depression detection methods. When it comes to machine learning (ML) and deep learning (DL), algorithms fall into two main categories: supervised learning and unsupervised learning. This research primarily focuses on utilizing supervised learning algorithms to improve the detection of depression in individuals by allowing the algorithms to learn complex patterns and connections within the data.

## 2. Literature Survey

Musleh, D. A. employed varied N-gram ranges and TF-IDF methods for feature extraction, implementing diverse NLP techniques during data pre-processing. It is important to note that our findings are context-specific to Arabic-speaking Twitter users and may not apply to other cultural or linguistic backgrounds. The unique nature of Twitter data allowed us to capture real-time emotional expressions, presenting a valuable opportunity for the timely detection of potential mental health concerns. Notably, our Random Forest classifier exhibited a noteworthy accuracy of 82.39%, underscoring the promising effectiveness of the model in identifying signs of depression within Arabic tweets [1].

Hasib, K. M., explored the use of SVM, RF, RNN, CNN, and other techniques for automated depression detection using data from social networks. Potentially inaccurate data from social networks may hinder the accuracy of automated depression diagnosis. These techniques offer superior insights for automated depression detection compared to traditional methods. Utilizing social media data provides a potentially rich source of information for understanding users' mental states and activities [2].

Li, X. investigated abnormal organization in the functional connectivity network of mild depression. The study focuses on mild depression and may not generalize to other forms of depression or mental health conditions. The approach leverages EEG data, providing a non-invasive and potentially objective method for diagnosing mild depression. Utilizing graph theory and deep learning techniques (CNNs) enhances the analysis of functional connectivity, allowing for a more comprehensive understanding of brain network abnormalities [3].

To consolidate insights from prior research, Liu conducted a synthesis of studies employing Support Vector Machines (SVMs), Bayes, latent Dirichlet allocation (LDA), decision trees, and neural networks techniques for detecting depressive symptoms in social media text data. It is essential to acknowledge a potential limitation highlighted across these studies, namely sampling bias, as they heavily rely on data sourced from social media, which may not be fully representative of the broader population. Despite this concern, machine learning (ML) methods showcased a promising ability to identify depressive symptoms at an early stage, enabling timely intervention and support. The integration of ML approaches has the potential to complement traditional methods in mental health assessment, offering an additional resource for public mental health practitioners to enhance their understanding and address mental health challenges effectively [4].

In the work by Ashraf, a model was developed specifically for analyzing Arabic users' tweets to detect depression. It is crucial to note that the effectiveness of the proposed machine learning model is contingent on the quality and diversity of the available data. Machine learning, as implemented in this model, plays a pivotal role in elevating the accuracy and precision of diagnostics, offering the potential for more reliable detection of mental distress such as depression. The model's emphasis on early detection holds significant promise, as it can facilitate timely support and intervention for individuals experiencing depression, underscoring the importance of proactive mental health strategies in the digital context [5].

Amanat's work focuses on constructing an early depression diagnostic system through the application of machine learning techniques, specifically implementing Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN) models for accurate depression prediction from text. Notably, while the reliance on textual data may restrict the model's capacity to capture non-verbal cues or alternative communication forms, the system attains an impressive accuracy rate of 99.0% in predicting depression from text. This noteworthy performance surpasses frequency-based deep learning models, underscoring the efficacy of Amanat's approach in enhancing the accuracy of early depression detection [6].

Sajja's research focuses on developing a machine-learning model for predicting feelings of anxiety and depression. The success of the model is acknowledged to be contingent upon the quality and diversity of the speech data employed for training. By leveraging machine learning approaches, Sajja aims to create data-driven models that facilitate the study and prediction of anxiety and depression. The implementation of automation in decision-making based on test data is highlighted as a potential avenue for achieving more efficient and reliable predictions, emphasizing the role of technology in advancing mental health assessment and support [7].

Varsha identifies and detects signs of depression in individuals through the analysis of their comments, posts, or texts on social media platforms. The accuracy of the model developed by Varsha is recognized to be influenced by the quality and diversity of the social media data collected for training. The utilization of social media data for depression detection is highlighted as a potentially vast and accessible source of information. Varsha employs data mining and machine learning algorithms to enhance the efficiency and accuracy of detecting emotions, with a particular emphasis on identifying signs of depression. This approach underscores the importance of technology in leveraging vast online datasets for mental health analysis and detection [8].

Lora's research focuses on the accurate classification of positive and negative emotions expressed in Twitter tweets. The effectiveness of the models developed in this study is acknowledged to be influenced by the diversity and quality of the Twitter dataset used for training. Leveraging Twitter data is highlighted as advantageous for real-time analysis of emotions and sentiments expressed by users. Notably, Lora's study employs a diverse set of models, encompassing both traditional machine learning and advanced deep learning techniques. This comprehensive approach ensures a thorough evaluation of classification techniques, contributing to a more nuanced understanding of emotion detection in social media content [9].

Aleem's research is centered around the precise classification of positive and negative emotions expressed in Twitter tweets. Recognizing that the effectiveness of machine learning algorithms for depression detection can vary based on the quality and diversity of available data, Aleem addresses this consideration in the context of emotion classification on social media. The study showcases the potential of machine learning in efficiently analyzing extensive healthcare data for mental health applications, particularly in the realm of depression detection. Aleem's work provides a structured overview of various machine learning algorithms used in this domain, serving as a valuable resource for both researchers and practitioners engaged in mental health studies [10].

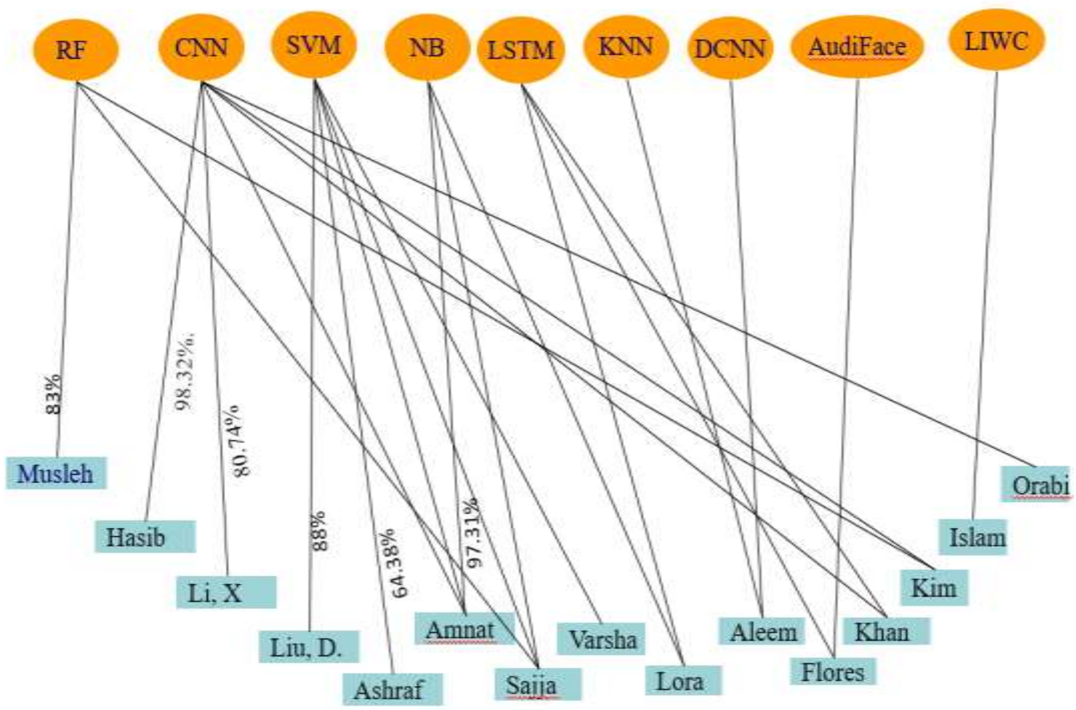
Flores's research focuses on the development of AudiFace, a multimodal deep learning model designed for efficient and accurate depression screening. The effectiveness of AudiFace is recognized to be influenced by the quality and diversity of the input data, along with the characteristics of the participants involved in the study. Notably, AudiFace demonstrates substantial improvements in depression screening capabilities, achieving the highest F1 scores across the majority of datasets. A distinctive feature of AudiFace is its successful incorporation of multiple modalities, encompassing temporal facial features, audio, and transcripts, which collectively enhance the screening process. This multimodal approach underscores the potential of integrating diverse sources of information for more robust and comprehensive depression screening [11].

Khan's work aims to furnish an extensive review of Facial Emotion Recognition (FER), encompassing techniques such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Support Vector Machines (SVM), Random Forest (RF), and other methodologies. It is acknowledged that the review's scope may be influenced by the selection of literature and datasets considered. Despite this, Khan's review provides a comprehensive overview of FER, incorporating both traditional machine learning (ML) and modern deep learning (DL) methods. Notably, the analysis extends to publicly available FER datasets, offering benchmark results and aiding researchers in the selection of appropriate evaluation metrics. This approach enhances the utility of the review as a valuable resource for those involved in the field of Facial Emotion Recognition [12].

Kim's research focuses on the development of a framework for the automatic detection of depression using acoustic features in the Korean language. It is noted that the study's findings may be specific to the Korean language and its cultural nuances, possibly limiting their generalizability to other languages or cultures. Nevertheless, Kim's work introduces a noteworthy framework for automatic depression detection through speech analysis, highlighting its potential for widespread accessibility, particularly through smartphones. The study demonstrates that deep-learned acoustic characteristics surpass the performance of conventional approaches and pre-trained models in the accurate detection of depression, showcasing the efficacy of advanced techniques in this context [13].

Islam's research involves conducting depression analysis on Facebook data to comprehend users' moods and attitudes during online communication. The study acknowledges a potential limitation in its reliance on a specific data source (Facebook), suggesting that its findings may not generalize to other social networks or online platforms. Nevertheless, Islam's work presents a novel approach to analyzing social network data, offering insights into users' feelings and sentiments, with a particular focus on depression. The study showcases the effectiveness of machine learning techniques in detecting depression, suggesting scalable solutions for addressing mental health issues on social media platforms. This research contributes to the evolving landscape of leveraging digital data for understanding and addressing mental health concerns in the online realm [14].

Orabi's research tackles the intricate task of detecting mental illnesses through social media platforms, which often serve as reflections of users' personal lives. The chosen deep neural network architectures' effectiveness is acknowledged to be influenced by the specific characteristics of Twitter data, potentially limiting generalization to other social media platforms. Orabi's work addresses the challenge of mental illness detection on social media by leveraging the abundant personal information available. The research evaluates and identifies the most effective deep neural network architecture for detecting signs of mental illnesses, with a specific focus on depression, within the constraints of limited text data [15].




**Literature Survey Graphical Representation**

## 3. Methodology

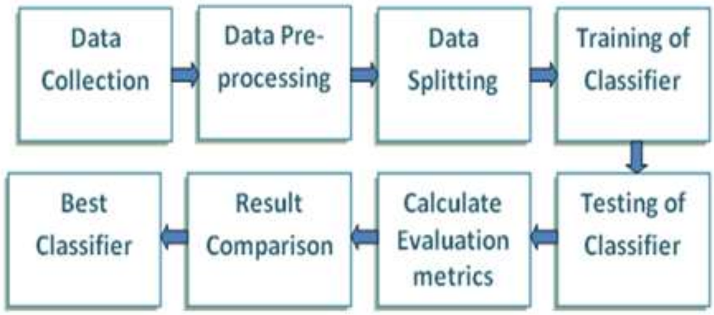



***Fig -1: Generalised Steps for Depression Detection***

The fig-1 shows the generalized steps for Depression Detection. Here are general steps you can follow when attempting to detect depression.

**Data Collection:**

Gather a diverse dataset containing text data labeled with depression indicators. Include a mix of positive (indicating depression) and negative (non-depression) samples.

**Data Pre-processing:**

Perform text preprocessing steps such as lowercasing, removal of stop words, and punctuation. Address imbalances in the dataset if needed (e.g., through oversampling or undersampling).

**Data Splitting:**

Split the dataset into training and testing sets (e.g., 80% training, 20% testing).Ensure a balanced representation of classes in both sets.

**Feature Extraction:**

Use techniques like TF-IDF (Term Frequency-Inverse Document Frequency) to convert text data into numerical features.

**Model Training:**

Choose machine learning classifiers such as Naive Bayes, SVM, or Random Forest. Train the selected classifiers on the training set using the extracted features.

**Model Testing**:

Apply the trained models to the testing set to predict depression labels. Generate predictions based on the learned patterns from the training data.

**Model Evaluation:**

It is important to evaluate the model's performance by using metrics such as accuracy, precision, recall, and the F1 score. Additionally, analyzing the confusion matrix can help to gain insight into false positives and false negatives.

**Result:**

Analyze the model's ability to correctly classify depression and non-depression instances. Please evaluate the balance between sensitivity and specificity in a given scenario.

**Best Classifier:**

Compare the performance of different classifiers.

Select the classifier with the highest overall performance based on evaluation metrics.

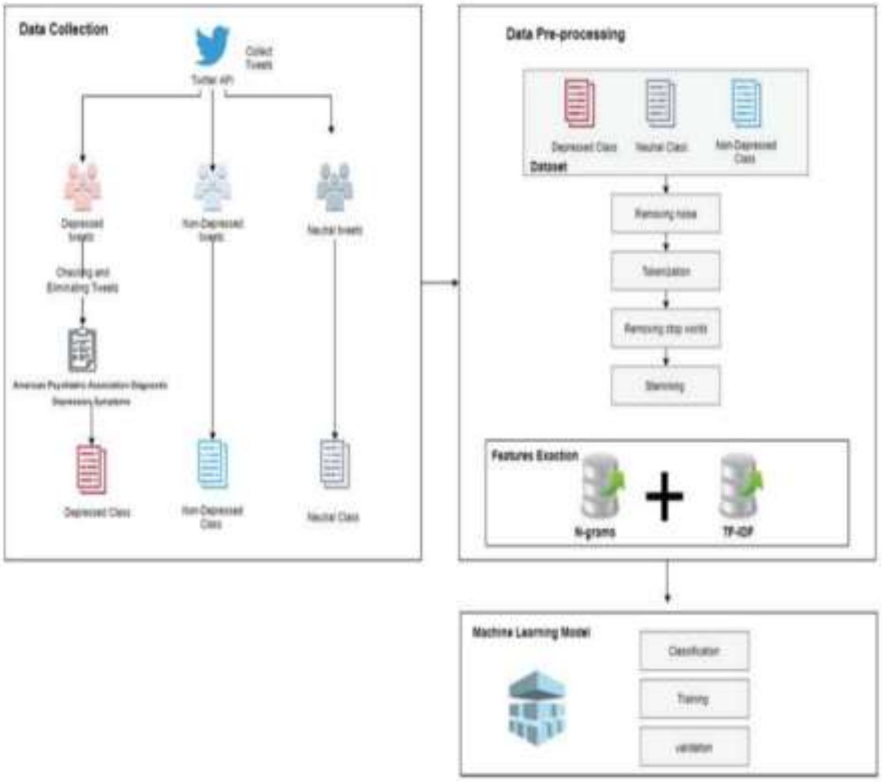



*Fig. – 2: Structure of Machine Learning Models*

**Dataset Collection:**

The CES-D survey was administered to Arabic Twitter users, and the scale was translated into Arabic for the study.The CES-D scale is a short, self-reporting scale that contains 20 questions. It is designed to measure depressive symptomatology. In addition to the CES-D survey, the dataset was collected by using a dictionary of depressive phrases. " The study found that 45% of Arabic Tweets expressed sadness or depression, while 9% described difficulties sleeping, 9% of the Tweets indicated that they had gotten too much sleep, and 8% of the Tweets revealed suicidal thinking; a sensitive topic and dangerous sign of depression amongst users.

**Data Preprocessing:**

In this stage, the dataset was cleaned and prepared using natural language processing techniques, including normalization, stop word filtering, tokenization, and stemming.

**Features Extraction:**

Different N-gram ranges and TF-IDF methods were used to extract the required features. Six supervised machine-learning models were used to train the dataset, including SVM, RF, LR, KNN, AdaBoost, and NB. As a result, they found the feature combination that gave optimal accuracy.

**Term Frequency-Inverse Document Frequency (TF-IDF):**

It is commonly used for text classification. Term frequency is the number of times a word appears within a document. Inverse document frequency is a weight term scheme that gives tokens that appear more frequently in documents a lower impact, or weight, and gives tokens that occur less frequently a higher weight.

**N-gram:**

N-grams extract characters or words from a text and are used in stemming spelling checking, and text compression. N-grams are a commonly used approach to identify similarities between sequences of N items, such as words or characters. The N value is an integer and is set to be a unigram n = 1 (one word or character), and a bigram n = 2 (two words).

**Generating the Models**

Six supervisor machine-learning models were implemented using Python. Sklearn, and Grid- Search CV libraries to determine the classification models. The classifiers were as follows:

**Naive Bayes (NB)**

NB is a probabilistic classifier that uses the Bayes theorem, where all features (attributes) are assumed to be independent of each other.

**Support Vector Machine (SVM)**

SVM is a classifier that uses risk minimization theory to find the optimal separating hyperplane within the feature space. The Simple Support Vector Machine (SVM) algorithm is commonly used for both linear regression and classification problems.

**Random Forest (RF)**

RF is a combination of tree predictors that predicts the class label by randomly generating a forest. The forest consists of multiple decision trees, each holding the value of an independent random vector. The trees are then equally distributed among all trees, and the final classification is based on the majority vote.

**Adaboost**

AdaBoost is an approach applied to textual or numeric data types. Split the data repeatedly and continue to re-assign different weights to the training data. This ensures that misclassified data from the first split will be reassigned correctly during the next data split. The process will keep going until the optimal data split is determined.

**K-Nearest Neighbors (KNN)**

K-Nearest Neighbors (KNN) is a classification algorithm that uses the distance, such as Euclidean distance, between data points to classify new and unknown data points based on the closest existing data points.. The distance function calculates the distance between two points, and the "K" value shows the size of the neighborhood. Unknown data points are classified based on simple voting.

**Logistic Regression (LR)**

LR is a binary logistic model. LR is used based on one or more features to estimate the probability of binary response. The author Nadeem describes LR as a discrete choice model, as it is not technically qualified as a classification method. The relationship between the binary variables and the features is clarified through the use of the below. However, for multi-class text classifications, a multinomial extension must be used.

**Evaluation Criteria:**

A comparison of all classifiers is presented and a discussion of the effect of TF-IDF on text classification. The comparison of all classifiers is based on the results of the accuracy, recall, precision, and F1-score model evaluation results. These results were obtained from the optimal parameters of the grid.

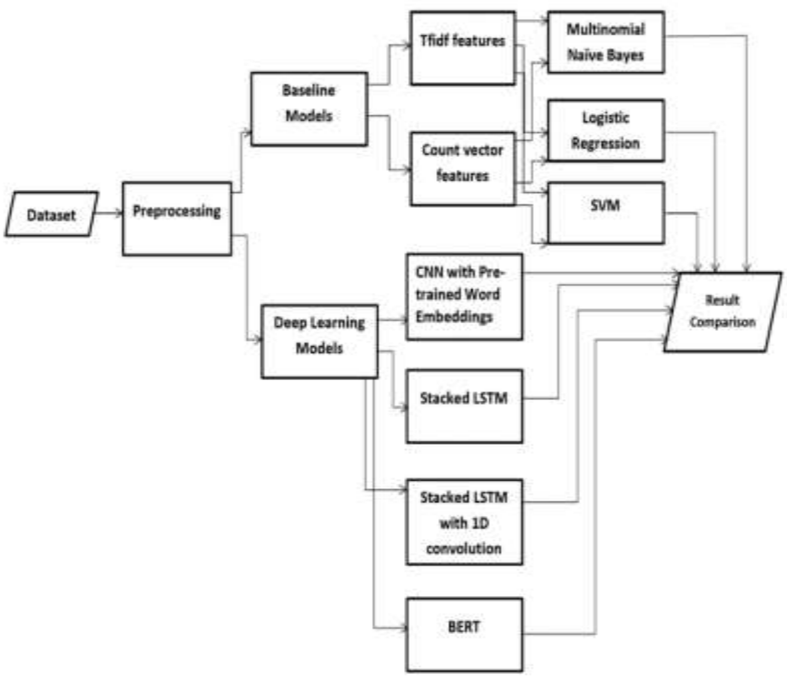



*Fig. – 3: Structure of Base Line and Deep Learning Models*

**Dataset Collection:** For the dataset, the sentiment140 dataset has been taken. The dataset contains 1048576 rows with six columns. However, only two columns have been used for this work. They are "label" (the polarity of the tweet) and "tweet" (the text of the tweet). Here zero means negative emotion tweets and 4 means positive emotion tweets.

**Dataset Preprocessing**

Data preprocessing is an essential phase after successfully collecting the dataset. In the real world, there can be some unnecessary, missing, and noisy data for their huge size. By preprocessing, it is possible to remove this unnecessary data to reduce the size of huge data. Besides, preprocessing can reduce computation complexity. So, it needs to be preprocessed to remove noisy data and reduce complexity.

Data preprocessing which has been done involved the following steps:

- Converting uppercase letters to lowercase
- Removing punctuations
- Removing hashtags, URLs, retweet mentions and user mentions
- Replacing multiple spaces with one space

**Baseline Models:**

Three baseline models were developed for this study using both Tf-idf and count vectors. Tweets were transformed into feature matrices using the Tfidf Vectorizer with specific parameters. For count vectors, Count Vectorizer was employed. The models include Multinomial Naive Bayes, Linear SVM, and Logistic Regression, utilizing respective sci-kit-learn libraries. The data was split with 90% for training and 10% for testing.

**Deep Learning Models:**

**Convolutional Neural Network (CNN) with Pre-trained Word Embeddings:** This text describes the implementation of a Convolutional Neural Network (CNN) for binary text classification using pre-trained word embeddings from the Google News Word2Vec model. The CNN utilizes five different filter sizes, followed by GlobalMaxPooling1D layers and a series of Dense and Dropout layers with relu activation. The training process shows increasing accuracy on both the training and validation sets, indicating good performance. However, the validation loss is slightly higher than the training loss, suggesting potential for further improvement. The model was trained with a batch size of 34, a dropout probability of 0.1, and Adam optimization over 3 epochs. The dataset was split into 70% for training, 10% for validation, and 20% for testing.

**Stacked LSTM Model:** This text introduces a Stacked LSTM model, an extension of the original LSTM, with multiple hidden layers containing multiple memory cells. The training process shows fluctuating accuracy on both training and validation sets, indicating potential overfitting. The model was trained with a dropout probability of 0.2, Adam optimization over 3 epochs, and Tanh activation. The dataset was split into 70% for training, 10% for validation, and 20% for testing. Validation accuracy dips at epoch 2, suggesting some loss in predictive ability for new data. Additionally, validation loss is higher than training loss, indicating possible overfitting.

**Stacked LSTM with 1D Convolution:** An additional 1D convolutional layer was added to the previous model to speed up training. This model used a dropout probability of 0.3, a kernel size of 5, a pool size of 4, and madam optimization, and was trained for 3 epochs with Tanh activation. The training process shows increasing accuracy on both training and validation sets, indicating good performance. Training loss is slightly higher than validation loss, suggesting potential for further improvement. This model outperforms the previous one.

**BERT-Based Model:** BERT (Bidirectional Encoder Representations from Transformers) is a model pre-trained on English Wikipedia and Brown Corpus by Google. It enhances computer understanding of language for tasks like complex search queries. Implemented with TensorFlow 2.0, we used a batch size of 32, a drop-out rate of 0.2, and trained for 5 epochs. The model achieved an accuracy of 83.2%.

**Evaluation Parameters:**

Results of Baseline Models and Deep Learning Models Parameters such as accuracy, precision, recall, specificity, and the F1-score were used.

## 4. Results

| Sno | Method | Dataset | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|---|---|
| 1 | RF | CES-D survey | 83% | 82.40% | 82.67% | 82.52% |
| 2 | SVM | DAIC-WoZ | 80% | 77% | 87% | - |
| 3 | CNN | BDI-II | 80.1% | - | - | - |
| 4 | LSTM | WOZ-DAIC | 93% | - | - | 87% |
| 5 | SVM | DAIC | 64% | 80.25% | - | 69.52% |
| 6 | RNN | data set of tweets from the Kaggle website | 99% | 99% | 98% | 99% |
| 7 | SVM | 600 records in the input data set | 94% | - | - | - |
| 8 | SVM | collected almost 10,000 data from Facebook | 75% | 77% | 80% | 78% |
| 9 | CNN | sentiment140 | 84% | 83% | 84% | 83% |
| 10 | SVM | PHQ-9 | 95% | 95% | - | 97% |
| 11 | Audi Face | DAIC-WOZ | - | - | - | 93% |
| 12 | CNN | Cohen Knede | - | - | - | - |
| 13 | CNN | 153 patients, 165 healthy data | 78.14% | 76.86% | 77.90% | 77.27% |
| 14 | LIWC | Facebook users' comments | 80% | - | - | - |
| 15 | CNN | 1,145 Twitter users | 87.95% | 87.43% | 87.02% | 86.96% |

The above table incorporates the results of 15 different reference papers. It includes the methods used and performance metrics of the respective papers

## 5. Conclusion

In conclusion, leveraging machine learning algorithms, including deep learning models such as CNN and LSTM, alongside traditional methods like SVM, LR, and RF, presents a promising avenue for the early detection of depression through the analysis of cough sounds. The development of mobile applications integrating these frameworks can make diagnostic tools more accessible, particularly for individuals with limited financial resources or residing in areas with insufficient medical infrastructure.While the mentioned models boast accuracy rates exceeding 90%, it is essential to emphasize the need for thorough validation, regulatory approval, and ethical considerations before deploying these technologies in clinical settings. Depression is a complex mental health condition, and any diagnostic tool must be sensitive, specific, and reliable. Additionally, the involvement of healthcare professionals in interpreting and validating results remains crucial.The potential of technology to democratize mental health diagnosis is significant, offering the possibility of early intervention and support for individuals experiencing depression. However, responsible development, validation, and ethical implementation are paramount to ensure the effectiveness and safety of these tools in contributing to the overall well-being of individuals affected by depression.Looking toward the future, collaborative efforts between researchers, clinicians, and technologists can contribute to refining and expanding the capabilities of depression detection tools. Continuous refinement of machine learning models through larger and more diverse datasets, as well as ongoing validation studies, will be essential. Moreover, establishing standardized protocols for ethical deployment and ensuring compliance with regulatory frameworks will be crucial for the responsible integration of these technologies into mental health care.

**References**

[1] Musleh, D. A., Alkhales, T. A., Almakki, R. A., Alnajim, S. E., Almarshad, S. K., Alhasaniah, R. S., ... & Almuqhim, A. A. (2022). Twitter Arabic Sentiment Analysis to Detect Depression Using Machine Learning. *Computers, Materials & Continua*, *71*(2).

[2] Hasib, K. M., Islam, M. R., Sakib, S., Akbar, M. A., Razzak, I., & Alam, M. S. (2023). Depression Detection From Social Networks Data Based on Machine Learning and Deep Learning Techniques: An Interrogative Survey. *IEEE Transactions on Computational Social Systems*.

[3] Li, X., La, R., Wang, Y., Hu, B., & Zhang, X. (2020). A deep learning approach for mild depression recognition based on functional connectivity using electroencephalography. *Frontiers in neuroscience*, *14*, 192.

[4]Liu, D., Feng, X. L., Ahmed, F., Shahid, M., & Guo, J. (2022). Detecting and measuring depression on social media using a machine learning approach: systematic review. *JMIR Mental Health*, *9*(3), e27244.

[5] Ashraf, A., Gunawan, T. S., Riza, B. S., Haryanto, E. V., & Janin, Z. (2020). On the review of image and video-based depression detection using machine learning. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, *19*(3), 1677-1684.

[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*, *11*(5), 676.

[7] Sajja, G. S. (2021). Machine learning-based detection of depression and anxiety. *Int. J. Computer Appl*, *183*, 20-23.

[8] 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.

[9] Lora, S. K., Sakib, N., Antora, S. A., & Jahan, N. (2020). A comparative study to detect emotions from tweets analyzing machine learning and deep learning techniques. *International Journal of Applied Information Systems*, *12*(30), 6-12.

[10] Aleem, S., Huda, N. U., Amin, R., Khalid, S., Alshamrani, S. S., & Alshehri, A. (2022). Machine learning algorithms for depression: diagnosis, insights, and research directions. *Electronics*, *11*(7), 1111.

[11] Flores, R., Tlachac, M. L., Toto, E., & Rundensteiner, E. (2022, December). AudiFace: Multimodal Deep Learning for Depression Screening. In *Machine Learning for Healthcare Conference* (pp. 609-630). PMLR.

[12] Khan, A. R. (2022). Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges. *Information*, *13*(6), 268.

[13] Kim, A. Y., Jang, E. H., Lee, S. H., Choi, K. Y., Park, J. G., & Shin, H. C. (2023). Automatic Depression Detection Using Smartphone-Based Text-Dependent Speech Signals: Deep Convolutional Neural Network Approach. *Journal of Medical Internet Research, 25,* e34474.

[14] 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*, 1-12.

[15] Orabi, A. H., Buddhitha, P., Orabi, M. H., & Inkpen, D. (2018, June). Deep learning for depression detection of Twitter users. In *Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic* (pp. 88-97).