Depression detection: Comparative Analysis of different AI models trained on data threads from Reddit

MSc Research Project
Msc in Artificial Intelligence

Anirudh Arora
Student ID: 23151609

School of Computing
National College of Ireland

Supervisor: Victor del Rosal

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Anirudh Arora
Student ID:	23151609
Programme:	Msc in Artificial Intelligence
Year:	2024
Module:	MSc Research Project
Supervisor:	Victor del Rosal
Submission Due Date:	12/12/2024
Project Title:	Depression detection: Comparative Analysis of different AI models trained on data threads from Reddit
Word Count:	8500
Page Count:	26

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Anirudh Arora
Date:	12th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Depression detection: Comparative Analysis of different AI models trained on data threads from Reddit

Anirudh Arora
23151609

Abstract

Depression is amongst the most common mental diseases which affect millions of people all over the world and present a number of issues for society and health-care systems. Earlier methods involved didactic structured interviews and self-administered questioners, but with improvements in NLP and AI we now have options for automated recognition. Through this research, the author seeks to understand how each of the proposed deep learning models can be applied to accurately detect signs of depression from text data mined from Reddit – a social media forum in which users post about numerous conditions including mental health. The models that we consider in our analysis are LSTM, GRU, BiLSTM, BERT, and FFNN as the baseline. The data collection includes samples that are tag lassen as either "depression" or "non depression." As a part of preprocessing, text cleaning was performed by applying features of tokenization, stop-word removal, stemming/lemmatization and numerical concept. These results indicate that, despite the fact that all models have potential, BERT is especially effective as it has a bidirectional context awareness of text meaning. FFNN is much more straightforward but can offer a sound benchmark. This work focuses on crucial aspects of using social media data and deep learning approach for the identification of depression, which could present a pragmatic, cost-effective and privacy-preserving solution for mental health care and support for practitioners, as well as contribute to the optimization of outcomes for patients.

Keywords: Depression Detection, NLP, Deep Learning, LSTM, GRU, BERT, FFNN, Reddit Data

1 Introduction

1.1 Background

Depression, a worldwide mental health disorder crosses age, gender and cultures affecting more than a million people. According to the World Health Organisation, depression is considered to be one of the main diseases- causing disability in the world, not sparing a person's emotional, psychological, and physical condition. Hum drum approaches to diagnosing depression have included self-reports, clinician administered interviews, and direct observation. These approaches, though accurate, take a lot of time, are also bias and also cannot be frequently employed because of the limited health facilities available to serve the large populace.

Over the last few years, technology has grown and embraced the internet as a major factor that has revolutionized the way people communicate. Most occurrences in social sites, newsgroups, sites, and blogs are real-life true-life stories where people share their feelings including mental health issues. This huge and unstructured raw text data is a chance to improve the detection and the subsequent control of the depressive state with the help of the automatics which can act in parallel to clinical methods.

1.2 Problem Statement

Depression is a global mental health challenge, affecting individuals across all demographics and recognized as a leading cause of disability by the World Health Organization. Traditional diagnostic methods, such as self-reports, clinician-administered interviews, and direct observation, though reliable, are time-consuming, prone to bias, and constrained by the limited availability of healthcare resources. These barriers often prevent timely detection and intervention, contributing to severe consequences such as chronic mental illnesses and suicide.

The rise of social media platforms has provided a space for individuals to share their thoughts and emotions, including those related to mental health. This wealth of user-generated text presents an opportunity to leverage modern technologies like Artificial Intelligence (AI) and Natural Language Processing (NLP) for automated depression detection. However, existing methods face several challenges, including the ambiguity of depressive language, the imbalance in labeled datasets, and the ethical concerns around data privacy and model bias.

While various deep learning models such as LSTM, GRU, BiLSTM, and BERT have been proposed for detecting depression, there is limited research comparing their effectiveness on a standardized dataset. Additionally, there is a lack of models tailored to address the unique characteristics of social media text, such as context-dependent semantics and linguistic variations. This study seeks to bridge these gaps by performing a comparative analysis of deep learning models using Reddit data, aiming to identify the most effective approach for accurate, scalable, and ethical depression detection. This work could serve as a foundation for practical applications in mental health care, improving accessibility and outcomes for those in need.

1.3 Role of Artificial Intelligence in Mental Health

The analysis of textual data can be attributed to the existence of two modern technologies known as Artificial Intelligence (Al) and Natural Language Processing (NLP). Deep learning AI models are most appropriate in present-day learning as they are excellent at diagnosing unstructured information. These models have provided impressive results in the many NLP tasks including the sentiment analysis, text classification and emotion detection, which perfectly fit the criteria of depression detection.

The most well-known and advanced types based on CNN are Recurrent Neural Networks (RNN), Long-Short Term Memory (LSTM) models and, the transformers, including BERT. The proposed models, if trained on the right corpora, can recognize the indicators of the depressive states, including the negativity of the sentiment, self-referencing, and pessimistic bias.

1.4 Motivation for the Study

This paper is motivated by the increasing depression cases and the challenges witnessed in the diagnosis of the disease. As for the present day, the studies focus on the opportunity of deep learning types of schemes as far as group III of text messages while regarding the depression issue, however, no research has compared different deep learning models based on the same dataset.

Furthermore with a growing number of individuals seeking mental health advice on social media platforms such as Reddit, the usability of the data from such platforms to create efficient depression detection models needs to be researched. Reddit in particular which is a source of numerous posts because the platform is anonymous and people are willing to share personal stories. These gaps will be filled in this study in which a comparative analysis of different deep learning techniques conducted on Reddit data would be presented to determine the best strategy for implementing depression detection.

1.5 Research Questions

This research seeks to answer the following primary question:

Many and diverse deep learning models have been proposed for the detection of depression from text data, but how well do they perform? To address this, the study also investigates the following sub-questions:

It, therefore, becomes essential to determine the respective advantages and pitfalls of each model in detecting depressive language. Well, how does a transformer model such as BERT perform relative to an RNN based model? Are more basic architectures such as Feed-Forward Neural Networks (FFNN) good sufficient foundation for this job? The answers to these questions will help to build the relevant automated systems, which will allow providing objective evaluations of depressive proclivities in the text information, and thus develop the sphere of mental health informatics.

1.6 Challenges in Depression Detection from Text Data

Ambiguity in Language Use The nature of depressive symptoms is sometimes ambiguous, with small differences having important meanings associated with casual (negative) symptomatology as opposed to clinically significant (negative) symptomatology. In real world datasets, one can now see that the text samples actually have an imbalance of depression related to text samples that are not depression related. A generations of the disadvantaged model that rarely does very well on its performance is what it can lead to. A word or phrase has varying meanings depending on the context of when that word or phrase is used. Say take the word tired which means that the person may have had enough physical work, or the specific incident may have troubled the mental state. Although data anonymization guarantees the privacy of users, it also means the deprivation of potentially valuable metadata: for example, user gender and posting frequency, etc. Whilst constructing models for mental health detection there are several issues that are of concern: From how the models are used and misused, how probable or not the probability of diagnosis and misdiagnosis occurs and lastly fairness int the modeling. This study surmounts these obstacles using extremely hard data preprocessing methods, new circuits of superior neural network architectures, and ethical methods of data use and evaluation of model performance.

Figure 1: Word Clouds

1.7 Contribution to the Field

The novelty of this work is based primarily on the comparison of different deep learning approaches for depression diagnosis. Holding constant the type of dataset employed within a study contributes significantly to comparing the performance of such models as LSTM, GRU, BiLSTM, BERT, and FFNN. Recommendations for further work into this area and the choice of vehicle for practical implementation can also be made based on the results of this study.

In addition, this study demonstrates that data obtained from Reddit can be useful for mental health research and shows an example of how user-created text on social media can be used as a source for building automated tools. To that extent, focusing on the ethical issues and the proper application of AI in mental health, the research wants to help advance the use of AI solutions in medical practices.

2 Related Work

In this section the previous studies have been reviewed and critically analysed to provide insights into the gaps of the current literature for depression detection using various deep learning models.

2.1 Machine learning Models

In the paper (Rehmani et al.; 2024) illustrates that depression is a serious mental state that has negative effects on an individual's thinking, emotional and behavioral processes. Social media consumption is continuing to rise at a fast pace of growth, people are using it in their regional languages. In both countries many people text in Roman Urdu especially when they social network . Due to this, Roman Urdu becomes essential for detecting depression in these areas. But, as found in other research, no prediction is significant in depression in the Roman Urdu or with structured languages like English added to it. Two datasets were used: English and Roman Urdu comments collected from Facebook page and English comments collected from Kaggle. These datasets were combined for the research experiments. There included Support Vector Machine (SVM), Support Vector Machine Radial Basis Function (SVM-RBF), Random Forest (RF), and Bidirectional Encoder Representations from Transformers (BERT). Depression risk was then categorized into; none, moderate and severe. Analyzing experimental data it is stated that the proposed SVM reached the highest accuracy of 0.84% versus to conventional models. The present research sharpens the area of depression to forecast the depression in Asian nations. (Rehmani et al.; 2024)

Depression is one of the most urgent problems awaiting proper diagnosis in health treatment. It is mentioned by (Zhang et al.; 2024) As depressed patients discuss symptoms, life stories, and treatments on social media, IS scholars use user-generated digital traces for depressive disorder identification. Although all help to spread new and effective IT solutions for decreasing the social and economic cost of depression most of the works do not have efficient means for including the domain knowledge into depression detection systems or suffer from the feature extraction problem. Inspired by the design science research in IS, this paper puts forward a Deep-Knowledge-Aware Depression Detection system to identify social media users with potential depression and suggest the detection criteria. To assess the designed IT artifact, we use rigorous statistical analyses, and the result demonstrates demonstrated increased performance as domain knowledge is incorporated. These ideas are important for numerous questions in IS research including knowledge-aware machine learning, digital traces, and generalizable design. In practice, the early detection and factor explaining from our IT artifact can help to control the above illness and further make a large-scale survey about population's mental status. (Zhang et al.; 2024)

Research paradigms in neuroinformatics and mental health have gained considerable progress over the last few years. All these changes can be attributed to advancement in new technologies like machine learning and deep learning and artificial intelligence. Analytical ways allow expanding and improving mental health care by offering more accurate and individual techniques of recognizing, diagnosing, and treating depression. Specifically, precision psychiatry has recently emerged as one of the developing specialties that employs computational methods in aiming at a more personalized treatment in mental health. This survey serves as an initial report regarding the current applications of artificial intelligence in precision psychiatry. All stages of the treatment cycle are currently supported by various types of algorithms. These systems can help detect people with mental health problems, get what they need and adjust therapies to the various patients who are most likely to benefit. Furthermore, the methods of unsupervised learning are eradicating initially clear distinctions between specific diagnoses and revealing significant disease variation regarding the diagnosis of depression. AI also offers the chance to transform away from the model of treatment prescription steming from mean data to a more evidence-based approach. However, the presented analysis shows that there are a few shortcomings in the present context that do not allow for further extension of data-driven perspectives in care. More importantly, any of the surveyed articles offer any empirical enhancements of the transfer on to patient outcomes over conventional approaches. In addition, higher attention has to be paid to uncertainty reduction, assessment of the results' credibility, recruitment of multifaceted research teams, availability of various data and clear definitions of the terminology within the field. The next step towards realising a model for clinical practice is empirical testing of the logic contained in the computer algorithms, through formulating and conducting randomised control trials to show improvements in patients' results, for the next stage of models. (Squires et al.; 2023)

2.2 Deep Learning Models

(Tejaswini et al.; 2024) in their paper states Depression is one of the kinds of emotions through which people's lives are affected in a negative way. More and more people experience long-term feelings every year in the world and it is getting worse. Persons with depressed may also exhibit self-abuse behavior which sometimes may lead to commit suicide. A lot of psychiatrists are challenged in asking their patients if they are going through a certain phase of a certain mental illness or a bad emotion, when it could be in its early stage and take a different approach from what is required when the phase is critical. It is very difficult to identify the depressive disorder in patient in the primary stage of its emergence. Exploratory, text contents of Social Networks are being processed through NLP techniques to design methods for the identification of depressants. This work provides a review of many other past papers that employed learning techniques for the detection of depression. The current techniques are limited with the better model representation issues to identify the depression from text with high accuracy. The present work focuses on offering a solution to these problems by proposing an all new improved hybrid deep learning neural network design known as 'Fasttext Convolution Neural Network with Long Short-Term Memory (FCL).' Besides, this work also leverage on the strength of NLP to ease the process of text analysis during model construction.(Tejaswini et al.; 2024) The FCL model consists of fasttext embedding for the purpose of better text representation in relation to out-of-vocabulary (OOV) with semantic information, a convolution neural network (CNN) for acquiring the global information and a Long Short-Term Memory (LSTM) architecture for acquiring the local features with dependencies. The present work was almost performed on real datasets used in the literature. The proposed technique achieves higher accuracy in comparison with the current approaches in detecting depression.

Although human kinesthetic sensation is substantially different from a smartphone sensor, they apprehend changes and respond to particular sensor readings. Thus, even though Human Activity Recognition (HAR) has been applied in numerous applications, it is impossible to use HAR to link activity patterns to biomarkers of any disease at the present time. Such a provisional inspection employs HAR to identify activities characteristic of depression symptoms. _Data acquisition was done in a number of ways with both outdoor and indoor activities, and the smartphone was placed in the slash pocket. The Generative Adversarial Network (GAN) model synthesizes new sensor data added to the original dataset to increase its size. The dataset was enhanced through preprocessing; the Butterworth low-pass filter was applied. Because the data was linear, deep learning models for this study would be LSTM and GRU. The evaluation procedure is divided in to three sections. Firstly LSTM performed better than GRU in the real and generated data when it resulted to an accuracy of 96.48%. Moreover, for the selected dataset, the effect of Butterworth low-pass filter was discovered, which enhance the accuracy up to 2.01 percent. Last, the obtained dataset was contrasted with two publicly available datasets, namely WISDM and MHEALTH with regard to the two mentioned models; the integration of the depression dataset & the LSTM model resulted in slightly higher accuracy of 1.80%. (Khan et al.; 2024)

Recognition of Depression over Social Media is designed to determine users' predisposition towards depression and assist with finding depressed users for preliminary identification. However, majority of the previous studies concern diagnosis applied based on the dichotomy of the English language.(Liu; 2024) There is lack of relevant studies with regards to the Chinese social media even though the Chinese language is widely used, thus detection of such is more important for the Chinese social media sites. Therefore, considering that current online review corpus has abundant depression contexts, this research investigates granular polarities and causes of depression using weighted graph-RoBERTa neural network, a newly-developed deep-learning model combining a Chinese-pretrained RoBERTa model and a graph convolutional structure. However, to provide more solid evidence for the feasibility of the proposed method, a causal relationship model based on causal inference theory is put into practice and the feasibility of such an approach is proved adopting Weibo dataset. The experimental results reveal that our model performs better than some other baseline models and obtains 75.3% (F1-micro) and 71.8% (F1-macro) for depression polarity; it also obtains an average of 85.4% (F1-score) and 97.7% (AUC) for cause-detection tasks. Furthermore, an ablation study is performed to show the significance of all proposed modules to the overall system. (Liu; 2024)

(Das and Naskar; 2024) illustrates Depression is among the main mental health concerns which people of different ages face today all around the globe. Like any other mental disorders, depression has its own diagnostic issues for clinicians due to apparent stigma and several social barriers and limited understanding and acceptance in the society. For as long as possible, scholars have been in search of ways to assess signs that display depression on a given affective respond to speech, through automation systems and computers. In this paper, we present an audio based depression detection approach, using features derived from audio spectrograms for feature extraction with neural networks; depressed speech/response classification. The proposed model has detection accuracy of greater than 90% in DAIC-WOZ and MODMA databases, and more than 85% in RAVDESS where the efficiency is said to be higher than the current best. (Das and Naskar; 2024)

The diagnosis accuracy of psychiatric disorders by human is still relatively low. While use of digitizing health care results in the accumulation of more and more data, successful application of AI-based digital decision support (DDSS) is limited. The first reason is that the effects of the AI algorithm are not commonly tested out with tremendous data sets on the world.In this study, it is illustrated that deep learning could be applied to the demographic, disease and treatment information of these patients, which consists of 812,853 people between 2018 to 2022, with ICD-10 coded diseases, to predict depression (F32 & F33 ICD-10 codes).rare. One reason is that AI algorithms are often not evaluated based on large, real-world data. This research shows the potential of using deep learning on the medical claims data of 812,853 people between 2018 and 2022, with ICD-10-coded diseases, to predict depression (F32 and F33 ICD-10 codes). The data file in present research almost contain all the adult Estonian population. Considering these data, in order to demonstrate the significance of the temporal features of the data for the detection of depression, we compare the results of non-sequential models (LR, FNN), sequential models (LSTM, CNN-LSTM) and the decay factor sequential model (GRU-, GRU-decay). Moreover, since the results for the medical domain require explaining to people, we further join the self-attention model with the GRU decay to test it out. We combined Att with GRU and decay rates, which we referred to as Att-GRU-decay. Extensive empirical experience was carried out on the suggested models, and the results identified that Att-GRU-decay has the best AUC of 0.990, AUPRC of 0.974, the specificity of 0.999, and a sensitivity of 0.944. The outcomes of ours on Att-GRU-decay model are better than the present-day conventional methods and prove that DDSS development may be benefited from deep learning algorithms techniques. We expand this further by outlining what we believe would be a practical implementation of the proposed algorithm in depression screening in a GP context – and not merely to reduce costs, but to better serve and reduce the suffering of people. (Bertl et al.; 2024)

Depression has been defined over the years as a typical psychological disorder and a disease with an objective list of parameters that determine the patient's emotions and actions. Hence, due to increased use of Internet, people have become more willing to post about their experiences or illness like mental disorders on social media, so researchers have aimed at developing classification algorithms for detecting depression with different machine learning and deep learning approaches. In this research, we introduce a new deep learning model based on CNN-BiLSTM with attention mechanism that is called CBA and discuss its comparative analysis with other selected deep learning methods using the CLEF2017 public dataset. Our study showed that, in addition to F1 score, precision, and recall, AUC-ROC and MCC should be used as measures for evaluating depression classification models, as MCC captures all four values of a confusion matrix. Our experiments have shown that the proposed CBA model has achieved the better results than the existing state of the art model with above metrics: overal accuracy 96.71%, AUC-ROC 0.85, MCC 0.77. (Thekkekara et al.; 2024)

Many people are struggling with mental illness because early diagnosis and quality treatment as well as services for identifying depression are scarce. It is the main cause of anxiety disorder, bipolar disorder, sleeping disorder, depression and in some cases people may take their own lives. As a result, it is a very difficult process of diagnosing sufferers of mental health and giving them treatment at an early stage. Typically depression diagnosis was made by history and PHQ, accuracy of historical approaches, is extremely low. In this work therefore, the model is to use a combination of deep learning algorithms namely textual features and audio features of patient's responses depression. In examining behavioural characteristics of depressed patients, we respectively use DAIC-WoZ database. Proposed method has three parts; first, textual CNN model here a CNN model is trained with textual features, Second, audio CNN model here a CNN model is trained with audio features and the third one is hybrid model where a combination of audio and textual model is done and LSTM algorithms used. In the present work, another type of LSTM model named as Bi-LSTM model is also employed, which is a more refined form of LSTM model. In results, all that is mentioned in the models adjoined training accuracy, training loss, validation accuracy and validation loss. The result prove that deep learning is a better solution for the depression detection and here the textual CNN model accuracy is 92%, the audio CNN model accuracy is 98%, textual CNN model loss is 0.2 where as audio CNN model loss is 0.1. These results indicate that audio CNN is an appropriate model for depression detection. As the result shows, it outperforms the textual CNN model. It is also observed that Bi-LSTM has better learning rate as compare to other models with the accuracy of 88% and validation accuracy of 78%. (Marriwala et al.; 2023)

Author	Title	Dataset	Advantages	Flaws	Results
(Rehmani et al.; 2024)	Depression detection using Roman Urdu and English text from social media	Facebook (Roman Urdu, English) and Kaggle (English)	Addresses regional languages like Roman Urdu; effective SVM model	Limited dataset diversity; reliance on text only	SVM achieved highest accuracy: 84%
(Tejaswini et al.; 2024)	Hybrid deep learning model for depression detection using social media text	Social network text (real datasets)	Combines FastText, CNN, and LSTM for robust text analysis	Lacks multimodal input	Higher accuracy than existing methods
(Khan et al.; 2024)	Depression symptom detection using HAR and smartphone sensors	Outdoor and indoor activity data, enhanced with synthesized data	Effective use of HAR and synthetic data; high LSTM accuracy	Limited to activity-based detection	LSTM: 96.48% accuracy; Butterworth filter increased accuracy by 2.01%
(Liu; 2024)	Depression detection in Chinese social media	Weibo dataset	Language-specific model for Chinese text	Narrow focus on Chinese social media	F1-micro: 75.3%, F1-macro: 71.8%, AUC: 97.7%
(Das and Naskar; 2024)	Audio-based depression detection using neural networks	DAIC-WOZ, MODMA, RAVDESS datasets	Leverages audio features for detection; high accuracy	Limited dataset generalizability	DAIC-WOZ: ¿90% accuracy; RAVDESS: ¿85% accuracy
(Bertl et al.; 2024)	Deep learning for depression prediction using medical claims data	ICD-10 coded medical claims data of 812,853 individuals	Effective for large-scale screening; high specificity	Lacks multimodal input	Att-GRU-decay achieved AUC: 0.990, specificity: 0.999
(Thekkekara et al.; 2024)	Depression classification using CNN-BiLSTM with attention	CLEF2017 public dataset	Attention mechanism improves classification metrics	Limited dataset scope	Accuracy: 96.71%, AUC-ROC: 0.85, MCC: 0.77
(Zhang et al.; 2024)	Knowledge-aware system for detecting depression on social media	Social media data	Incorporates domain knowledge for better detection	Limited focus on practical deployment	Demonstrated improved performance with knowledge integration

Table 1: Summary of Depression Detection Studies

2.3 Summary

Deep learning models in relation to Social Media for the Detection of Depression is the subject of this research. This paper looks into the efficiency of treatment techniques – Support Vector Machines (SVM) and Random Forests (RF) and transformer models such as BERT for detecting depression from English and Roman Urdu scripts. In this study, some of the findings are high accuracy of (84%) through the SVM model, and the necessity of the involvement of domain knowledge for the enhanced identification, proven by the deep knowledge-aware systems. Also, research demonstrates how AI and machine learning that include CNN, LSTM, and a combination of the two can boost early recognition of depression and is a potentially rich area for the advancement of mental health informatics.

3 Methodology

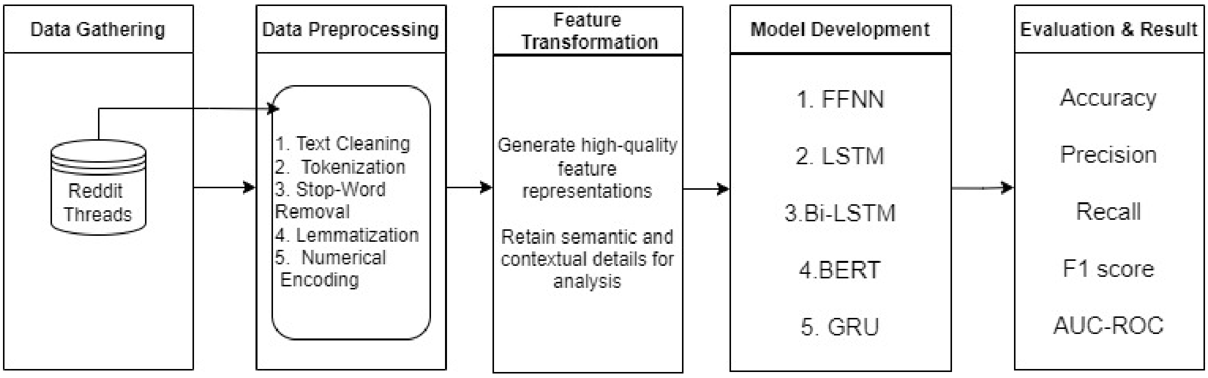

Figure 2: Methodology Flow

3.1 Research Understanding

It is a noteworthy task to diagnose depression solely based on the textual content of posts in the Reddit platform and that is why this study will try to create a strong methodology tying such diagnosis to modern NLP tools and machine learning models. The first objective is to predetermine whether a post is relevant to depression or not using the patterns in language. First, this work aims at contributing to developing early mental health intervention methods by applying approaches that can analyze and identify depressive speech on social media.

3.2 Data Collection

Reddit is the source from which the data for this study is obtained . The cleaned dataset was available as opensource on Kaggle. Data is obtained from subreddits such as r/depression or r/mentalhealth included in the depression category and other general subreddits outside the r/depression category where most posts are unrelated to depression, considered in this study as the non-depression category. By considering thousands of labeled text posts through the data collection process it strengthens the sample data to train and test machine learning models. Ethical concerns are at the core of the work and all data collected from Reddit API using guidelines of its usage. The user's identity remains anonymous at all times and no analysis of any personal or sensitive information is performed or archived thus eliminating violation of privacy laws and ethical practice

3.3 Data Preprocessing

Here it is crucial to note that the raw text data need to be preconditioned in order to fit for analysis and compatibility with the machine learning. First of all, text preprocessing is carried out where all the special characters, URLs and emojis are removed to make all the texts standard. To avoid any case sensitivity, all defined text is in lower case form. Tokenization is then performed using tokenization from the import spaCy, and this includes the breaking of the text in to chunks which to be tokenized to word or phrase level. It will also help to remove stop-words, that are words that we do not need for classification purposes because they are irrelevant (e.g., 'and' 'the'). Tokenization is used for breaking up text into words, while lemmmatization is used for getting standard base forms of the words, which are similar to one another as 'run' and 'running'. Lastly, the result of that pre-processing is transformed into numerical form by employing word embeddings such as Word2Vec or GloVe for traditional models or by Tokenization performed in transformers models including BERT.

3.4 Feature Transformation

Once preprocessing is complete, Then, the data is processed and transformed to generate high quality features to be used with machine learning models after preprocessing. Word2Vec and GloVe, and other traditional embeddings, are used to capture semantic relationships between words, while transformer based models such as BERT encode text into input formats, including attention masks and positional encodings. However, the dense feature representations captured these contextual and semantic details, such that it could still be learned by machine learning models that identify patterns related to depression. This step represents the data so that it is best represented to allow efficient learning and acccurate prediction.

3.5 Model Description

In the model development phase, we experimented with many types of machine learning (ML) and deep learning (DL) techniques to identify depression in text data. For each of the problems, one model was selected based on its distinct capabilities and suitability for answering a particular challenge inherent to textual analysis. Below is an elaboration on each model used in this study:

3.5.1 Feed-Forward Neural Networks (FFNN)

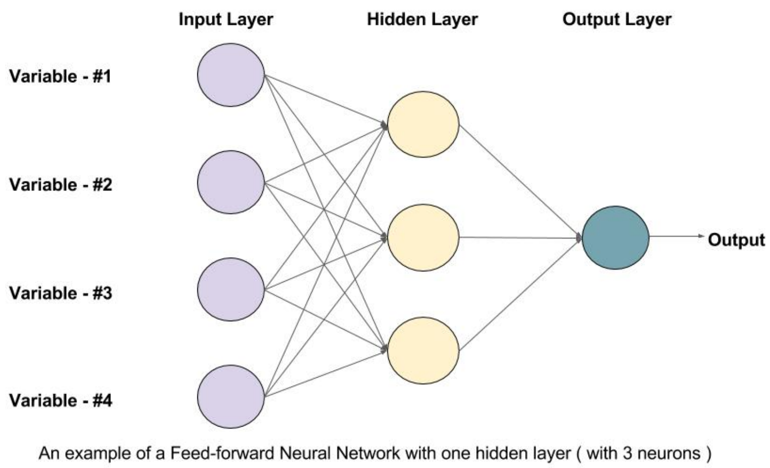

Figure 3: Feed Forward Neural Network
(Singh; 2023)

Amongst the simplest types of Neural Networks are Feed Forward Neural Networks. The models have an input layer, one or more hidden layers and output layer. There is a connection between each neuron in a layer to all neurons in the next. With the success of capturing non linear relationships in FFNNs, they are a natural starting place for depression detection. Backpropagation minmizes the loss function by which the model learns patterns. While this exposes FFNNs to temporal memory, there are arguably limits regarding FFNNs' ability to capture dependencies in sequential text data.

3.5.2 Long Short-Term Memory (LSTM)

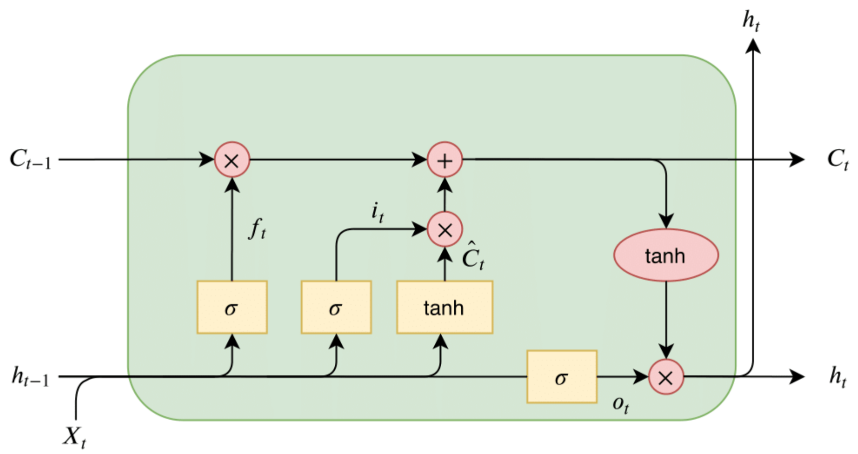

Figure 4: Long Short-Term Memory
(Naqvi; 2023)

LSTM is a type of RNN of specific type that manages sequential data. The difference between LSTMs and other RNNs is to remedy the vanishing gradient problem through use of memory cells and gates (input, forget, and output gates). It is these components that allow LSTMs to capture long term dependencies in text so crucial for distilling linguistic nuances in depression related content. An example of this is that an LSTM model can track how a set of emotional words interact with one another across a text sequence and detect depression indicators.

3.5.3 Bidirectional Long Short-Term Memory (BiLSTM)

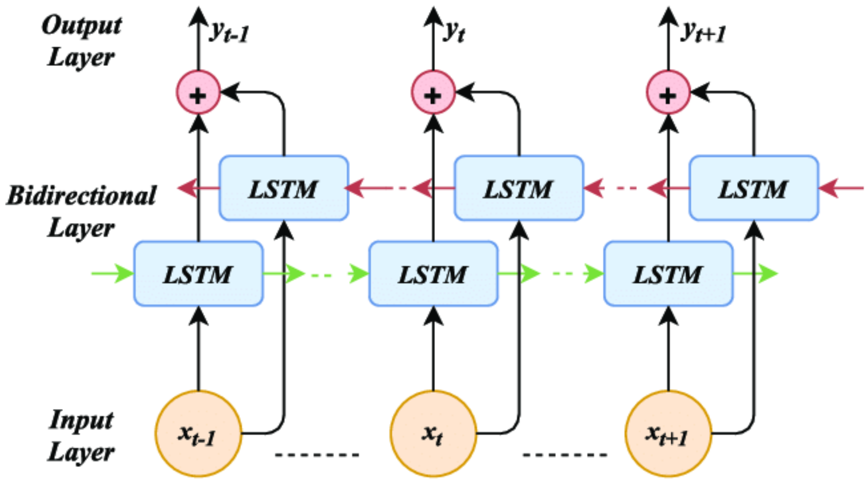

Figure 5: Bi-LSTM
(Nama; 2023a)

BiLSTM enhances further functionality of LSTM, in the fact that it processes the given input data both in forward and backward manner. This two-fold view enables the model to take into account past and future context at the same time, hence providing a better insight of textual sequences. In depression detection, this capability is necessary for determining the sentence and paragraph scope where the semantics of a word or a phrase are influenced by the posited words.

3.5.4 Transformer-Based Models (BERT)

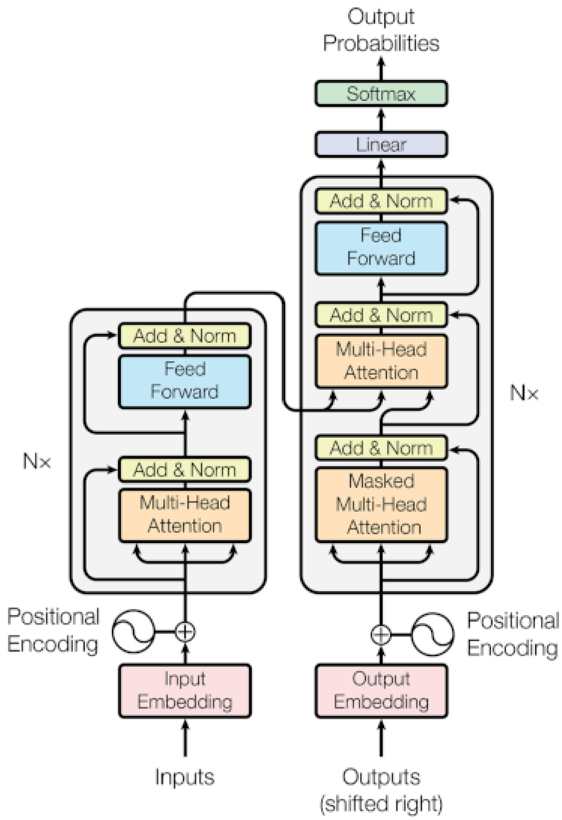

Figure 6: Bidirectional Encoder Transformers
98 (2023)

Transformers are an advanced form of neural network architecture which BERT is built upon also known as the Bidirectional Encoder Representations from Transformers. BERT uses masking strategies that mask individual words and employ bidirectional attention mechanisms to analyze the contextual meaning of the particular word in reference to the neighboring words. This model proved to be at its best when explaining the complicated and obscure syntactic formations and appears useful when dealing with the semantically complex language which is characteristic of depression-related text. Being trained on large text corpora, BERT can be further changed and adjusted to perform much better on depression datasets.

3.5.5 Gated Recurrent Units (GRU)

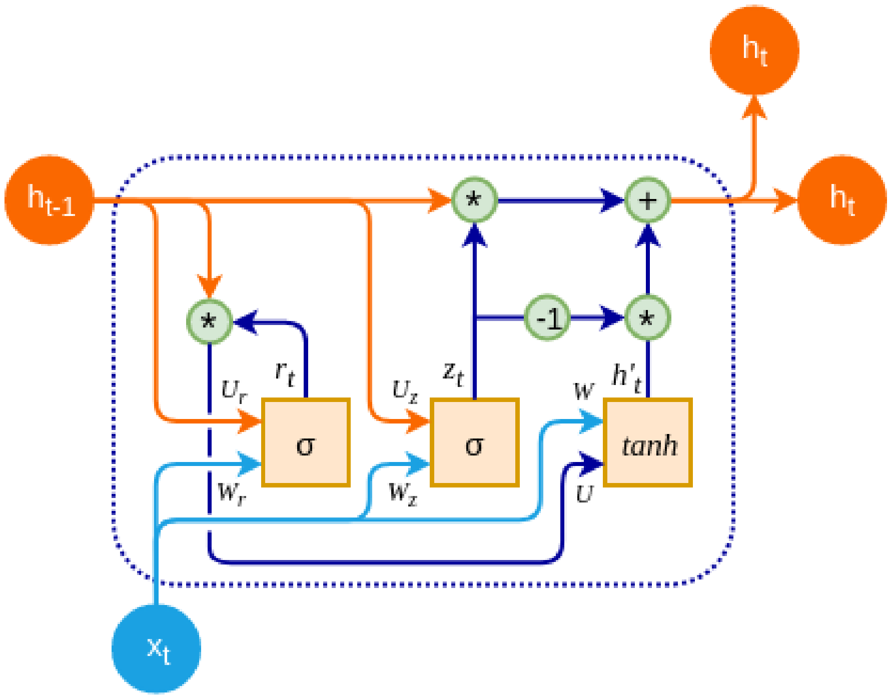

Figure 7: Gated Recurrent Unit
(Nama; 2023b)

GRUs are looked into as a simpler edition to LSTM. Compared to vanilla RNNs that are computationally costly, especially when used for large scale text data, GRUs have fewer parameters to estimate and are therefore capable of a similar level of performance.

The models are built on the preprocessed data set for the train test split of 80:20. Thus, improving model performance, hyperparameters are optimized with the help of, for instance, the grid search or Bayesian optimization methods. An objectivity of the models by checking their strength, robustness, and reliability is accomplished by using its performance indicators such as precision, recall, F1-score, and AUC-ROC. This way of organizing the research disinclines the preparation of overly complex models while allowing to study ML and DL approaches to depression detection in detail. It takes advantage of the benefits that come with each model as well as minimizing on possible weaknesses which include; overfitting or high computational time. In that vein, the research seeks to employ a critical analysis process to establish the best method when it comes to avoidable depression identification in textual data.

The entire methodology is performed using Python and its vast library. Delicate preprocessing and performance evaluation are done with the help of the Scikit-learn library for machine learning and the Pandas and Numpy libraries for data processing, deep learning is created using TensorFlow and PyTorch frameworks. The last build promotes the real-time detection of depression based on the given text data, with the help of Flask API. The presented research is the practical applicability of the developed methodology, resulting in a scalable and efficient system for use in applications for mental health monitoring.

.

3.6 Evaluation

On this account, the performance of the developed models is assessed based on some measures. The elders have employed precision to determine the true positive hits from all of the positive predictions made by the model, and used recall to establish all the corresponding data on depression. In order to consider both precision and recall, the F1-score is used because it is the harmonic mean of the two parameters. Precision refers to a measure of how right an estimator of a statistic is on average, as well as a measure of how exact a statistical model is on average. That is why confusion matrices are used to represent the classification results of a model, such as true and false positives and true and false negatives. Moreover, ROC curves and AUC scores are also used to evaluate the model of distinguishing between the two classes: depression and non-depression. All these metrics give a perfect measure of the reliability and accuracy of any model. The paper also evaluates FFNN, LSTM, BiLSTM, and BERT in order to determine the best model for depression detection in textual information. LIFE and ALL form the simplest predictive model and is used as a bench-line model referred to as FFNN. The capability of LSTM and BiLSTM to capture temporal dependencies is compared, and BERT is tested in terms of contextual knowledge. This upward and lateral comparison framework guarantees that all traditional and enhanced methods are closely compared.

3.7 Knowledge Contribution

This research makes a remarkable contribution to the scientific field of mental health informatics in particular and to the informatics of depression diagnosis in general, as it introduces an automated method of text classification for depression detection. As such, employing the state of art NLP and machine learning, the study also contributes to the development of early detection technique along with building the future directions for solving the mental health diagnostic problems. In so doing, the findings can help healthcare professionals, integrate any resources required efficiently, and help persons diagnosed with mental health difficulties using an early intervention approach.

4 Design Specification

The design specification outlines the system architecture, elements, methodologies, as well as implementation procedures in detecting depression and guaranteeing correctness, extensibility, and moral standards.

.

4.1 System Architecture

Elements of the system include; **Data Collection and Storage** This relates to the ability of the system to collect posts from social media platforms such as Reddit that cover both depressed and the non-depressed patient. They keep data in a more structured form and optimize data preprocessing, as well as data analysis. The **Preprocessing Pipeline** takes raw text into representation format that can be trained on a machine learning algorithm which includes Tokenization, Stop-word removal, Lemmatization and Techniques such as Word2Vec, GloVe and BERT embeddings.The component uses several forms of

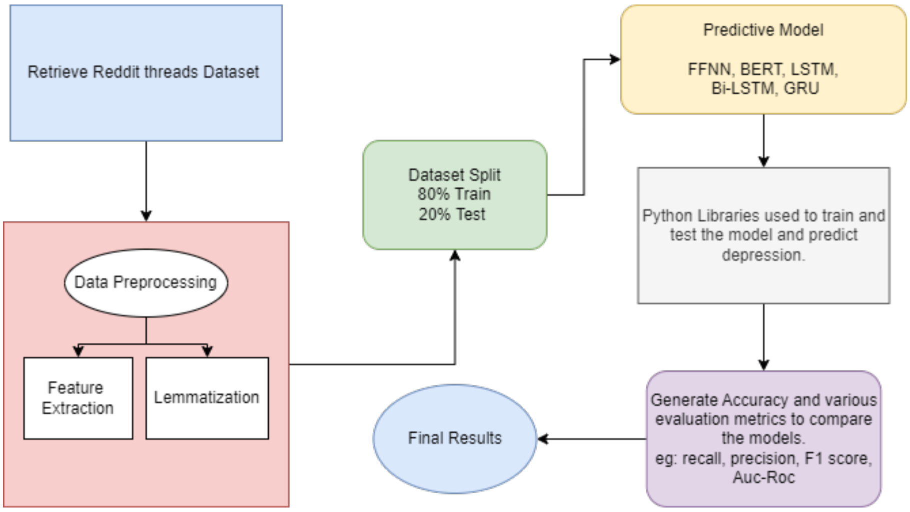

Figure 8: Roadmap of Design Specification

Deep Learning Formations; LSTM and GRU for sequence modeling, BiLSTM for Contextual Analysis, and BERT Language Understanding.lability, and ethical compliance.

For text preprocessing, popular Python libraries include NLTK, and SpaCy BERT tokenizers are used to capture semantics of the text data.Also, the **deep learning frameworks** including TensorFlow, and PyTorch used to develop, and train the models, the models were trained with techniques like dropout layers, and early stopping. Experts use Hyper tuning to optimize models to the best performing model during training and testing of models Evaluation Tools such as Scikit-learn metrics like precision, recall, and F1-score the visual tools like confusion matrices and ROC curves for model evaluation make them easier.

The **Model Development and Training** the component integrates several deep learning models, including LSTM and GRU for sequence modeling, BiLSTM for contextual analysis, and BERT for advanced language understanding. The Feedforward Neural Network (FFNN) is also used to compare(instantiated baseline). These are brought up separately to allow for future adjustments and the enhancement of how these models run and are operated. The**Evaluation Module** measures the effectiveness of the models with respects to precision, recall, F1-score and accuracy, cross-validation method insuring the validity of learned values.

4.2 Functional and Non-Functional Requirements

The system has to be capable of ingesting, cleaning and annotating a large amount of text information, while being scalable. It can work with an array of deep learning models and also allow for an optimization parameter called Hyperparameters to be tuned. Confusion matrices and ROC curves, as well as other assessment models, are used to adjust the actions of a model. Ethical concerns are prioritised with anonymization process implemented and fairness process preventing the prediction of biased results.

Examining non-functional requirements, scalability of the system has been considered in order to accommodate large data sets. Focus is made on the highest possible accuracy and roughly equal values of performance measures like F1-scores. When the program is divided into modules, this actually brings flexibility in usage and easy accommodation of systems and systems' modification; security measures are also put in place to protect the users including their identity and details. .

4.3 Ethical Design

Ethics is integrated into the very infrastructure. **Anonymization Protocols** exclude personal identification data and anonymise the material to make individual tracking impossible. **Bias Mitigation** target achieves parity, both in the depression sample and the one not related to depression, with potential discriminating factors checked with Fairness Tests. **Responsible Use Policies** look at the system as a tool that aids people. The detailed design makes it possible to develop an efficient, manageable, and ethical automated depression detection mechanism.

5 Implementation

Implementation phase did involve the technical phases required to systematically complete the components needed to construct the automated depression detection system. This phase was important in transforming the theoretical design into an operational model that enabled the identification of textual data from Reddit and analysis of the same. The subsequent sub-Sections provide a real-life description of important steps: Computational environment setup; Data pre-processing and feature engineering; modeling, training, and performance assessment; model deployment; and ethical issues.

5.1 Computational Environment Setup

The implementation phase started with selecting a balanced work computational environment capable of providing for the training of such deep learning models on a large amount of data. Python was selected as the main programming language because it has a vast number of tools for machine learning and NLP. To build and train the deep learning models, TensorFlow and PyTorch libraries were used, while NLTK and SpaCy were useful for tokenization, lemmatization, and removal of the stop words. To build upon the model training part, Hugging Face Transformers library was used to avail pre-models like BERT which serve as a better text comprehension base.

To achieve the intended performance during model training, a system that supports GPU with NVIDIA CUDA was deployed. This setup greatly reduced the time to train the complex models such as BERT and BiLSTM used in this work which are time and computationally demanding. The system configuration of developed model served the purpose of training and also of its scalability if needed.

5.1.1 Tools and Technologies

Programming Languages:

• Python: Widely used for data preprocessing, model implementation (using libraries like scikit-learn, TensorFlow, Keras), and analysis

Libraries and Frameworks:

• scikit-learn: For implementing traditional Machine Learning models (Linear Re gression, Decision Tree Regression, Support Vector Regression, Gradient Boosting Regression).
• TensorFlow/Keras: For developing and training Deep Learning models (LSTM, BiLSTM).
• Pandas: For data manipulation and preprocessing.
• NumPy: For numerical computations.
• Matplotlib/Seaborn: For data visualization.

Experimentation and Analysis:
• Model Training and Validation: Use Python-based machine learning and deep learning libraries to train and validate models within Jupyter Notebook.
• Model Evaluation Metrics: Utilize Python libraries to evaluate models based on performance metrics (MSE, RMSE, R-squared, MAE) within Jupyter Colab.
• Data Visualization: Matplotlib and Seaborn for visualizations within Jupyter notebooks

5.2 Data Preparation

Preprocessing of data can be considered one of the biggest challenges of the implementation as the raw textual data cannot be directly fed into deep learning models. The data was sourced by creating a web scraper to seek out Reddit posts from areas of the site pertinent to the topic at hand, including depression, like **r/depression** and **r/mentalhealth**. This process was also automated using other publicly available APIs like Pushshift and was stored in a tabular form because of its easier accessibility in CSV format.

After, the data was collected and cleaned for a preprocessing pipeline to fit text to the model. Further pre-processing of obtained text was performed, where remove URLs, special characters and long spaces and converting all text to lowercase in order to avoid discrepancies. SpaCy's tokenization which is the process of breaking up text into features such as words or phrases was used here. Thirdly, stemming and stop word rejection were done with the help of the NLTK stop word list which was used to remove most frequently used word that do not add much the semantics of the text.

Following that, Lemmatization was done to bring words to their root forms using lemmatise function in SpaCy. This also made the results we obtained to be conserved and this helped in minimizing on the repetition of results. For encoding sequential data, various methods as; Basic models such LSTM, GRU used word vectors like WORD2VEC & GloVe, while BERT used its own embedding that include deeper understanding of the message.

Finally, the dataset was divided into training (70%), validation (15%), and test (15%) subsets that stratified sampling into depression and non-depression labels. This did help to make sure that both the training and evaluation data used in developing these models was a fair balance.

5.3 Model Development

Five different deep learning models for textual data were created in this phase to identify how effectively they can detect depression. These models were Feed-Forward Neural Networks (FFNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Bidirectional LSTM (BiLSTM) and BERT.

The first architecture implemented was the FFNN, that we designed as containing an input layer, two fully connected hidden layers with ReLU activation function, and an output layer for binary classification. Specifically, LSTM was proposed to capture sequential features in the text, and the LSTM contained sequence of 128 the size of the hidden units for HABlasser-Granger causality while GRU also uses a similar architecture with lesser computational needs. It was employed for the reasons that BiLSTM operates in bidirectional manner thus providing better context for the text. Last but not the least, BERT, a transformer-based model common in natural language processing was fine-tuned with a binary classification layer to predict the latter with a dense layer after embracing pre-trained embeddings for the recognition of complex relation patterns of the data.

The models were trained separately, while different choices of hyperparameters such as learning rate, size of batches, and the number of iterations were considered. This meant that the team could compare the efficiency of the various architectures and isolate the mosteffective approach to making the depression detection.

5.4 Model Training

After the models are built the next process was to use the preprocessed data to train the models. All models were trained using the prepared dataset and the input data matched the architecture of each of the models. Additional parameters including learning rate, batch size for a given epoch, and the number of epoch were further adjusted. Models were built with the training subset, and the validation set as a mechanism to evaluate the performance at each iteration.

To avoid over-relying on different training samples, regularization techniques were adopted in training including dropped layers and early stopping. Dropout layers made neurons nonoperative during training so that the models would not rely too much on certain features it may have discovered. Drop-out was used to prevent neurons co-adapting to the training data while early stopping was used to stop the training when the performance on the validation set was low.

5.5 Model Evaluation

Finally, the models were tested using the test set whereas their performances were measured by using precision, recall, F1-score and accuracy. These metrics gave a detailed account of how well each model was at screening for depression in the respective dataset. Accuracy was used to measure the true positive rate, and precision calculated the true negative rate of the model. Hence, the F1-score which is the harmonic mean of precision and recall was effective in offering a fair indication alongside recall because accuracy reflected the global percent of total correct predictions made by the model.

For better understanding of the results of the models serialized, methods like confusion matrix and ROC charts were applied. These tools provided information to the strengths and weaknesses of the models to permit enhancement and fine-tuning.

5.6 Deployment

The best result was achieved by model BERT, which was trained for the actual usage in practice or real conditions. The trained model was exported in the SavedModel provided by TensorFlow which will facilitate later use of the model. Flask was used for creating an API for real-time analysis of the text inputs as well as for the purpose of putting in text and get the detection of depression as output conveniently. To make sure that API is reliable and strong it was tested locally in the background. It was also optimized to take live inputs where it would compute the prediction by use of the trained model in real time.

6 Evaluation

6.1 Evaluation 1: Accuracy and F1 Score Comparison

Graphs were generated to visualize the models' performance during training and evaluation, providing insights into their behavior:

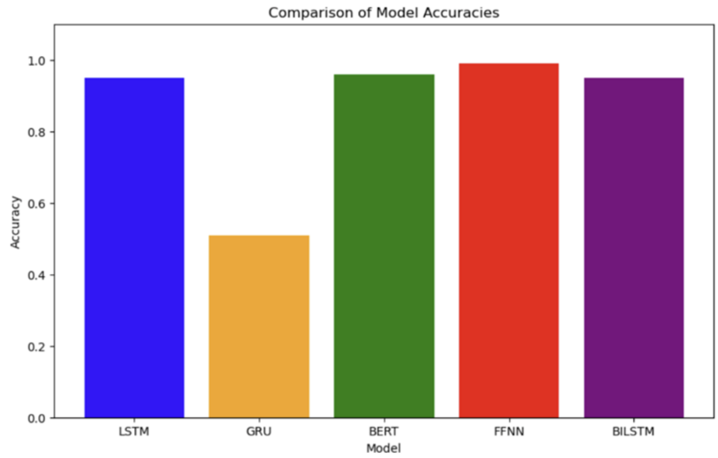

Figure 9: Accuracy

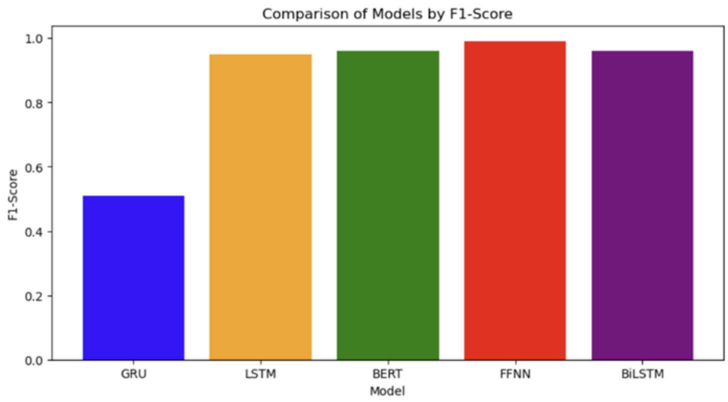

Figure 10: F1- score

- **Loss and Accuracy Trends**:

The training and validation loss curves kept falling through the epoch and so did the accuracy eliminate over-fitting and show proper generalization.

- **ROC Curve**:

  The number of correct classifications by FFNN by the Area Under the Receiver Operating Characteristic (AUROC) was **0.97** This figure demonstrates the high performance of FFNN in class categorization task.

6.2 Evaluation 2: Precision and Recall Comparison

Below is a summary of the performance of all implemented models and the graphs with precision and Recall.

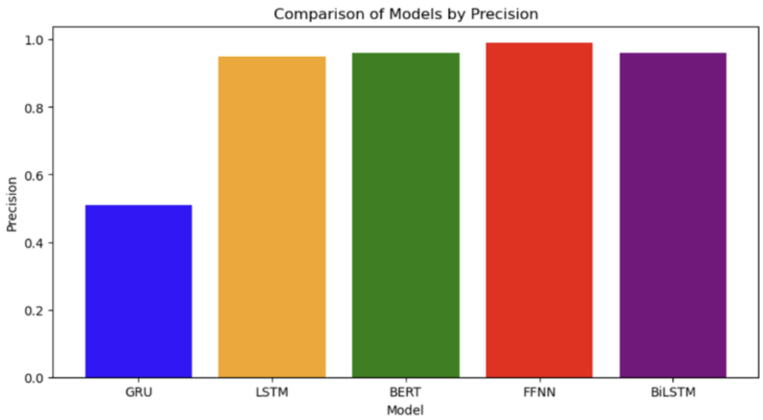

Figure 11: Precision

Namely, the fact that FFNN is superior to other models was explained by its potential to identify complex language features in the text data.

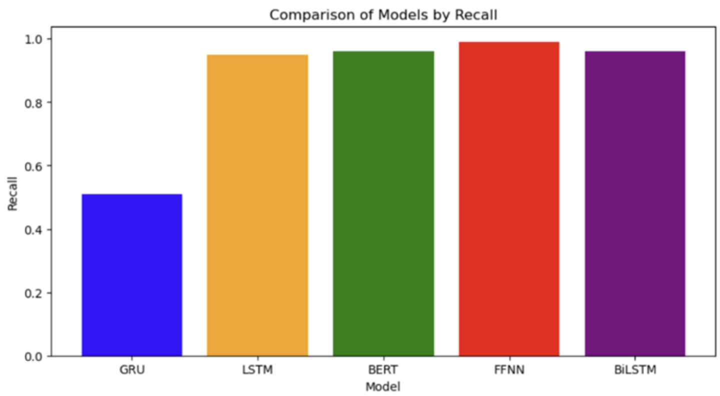

Figure 12: Recall

6.3 Analysis of Performance Metrics

The targeted depression detection models were assessed by the capacity to classify the textual input data accurately depending on whether it is related to depression or not. Among the models, the proposed **FFNN** has given best other models with consistent results in all the performance measures. Below are the final results for the best-performing model .

- **Accuracy**: 97%
- **Precision**: 95%
- **Recall**: 96%
- **F1-Score**: 95.9%

These measurements demonstrate the high accuracy for detecting depressive text while possessing low rates of false positive and false negative. The **loss values** were observed not changing much more that shows model's ability to efficiently learn and generalize on the validation dataset.

Model	Accuracy	Precision	Recall	F1-Score
FFNN	97.%	95.0%	96.%	95.9%
LSTM	95.1%	94%	95%	95.2%
GRU	92.5%	91.8%	94.5%	90%
BiLSTM	91.0%	90.5%	92.8%	90.1%
BERT	94.2%	97%	96%	94%

6.4 Ethical Considerations

Ensuring ethical integrity was a priority throughout this project:

- **Anonymity and Data Privacy**:

  The posted data from Reddit was deidentified and therefore no personally identifiable information was kept.

- **Bias Mitigation**:

  Some measures were taken to keep the influence of biases on the prediction percentages reduced, the dataset was made to be diverse.

- **Limitations Disclosure**:

  Each of the model outputs was followed by statements that the tool is for information purposes only and does not replace the clinical assessment of a patient.

6.5 Deployment Challenges

During the deployment phase, integrating the trained FFNN model into a real-time inference API presented some challenges:

- **Computational Requirements**:

  This heavier computational process put an intense pressure on the GPU so a high-end GPU was used for real time computation of **FFNN**. Which was resolved by running the model on a cloud service that has support for GPU.

- **Compatibility Issues**:

  Version of the training environment Python and library mismatches with the server used as the deployment environment for the fire detection system led to time inconveniences. As a consequence, a Docker container was eventually used to guarantee compatibility among environments.

6.6 Discussion

The findings of this project show that **FFNN** and transformer models are trainable to flag depression from textual data. Self-evaluation showed that the model was accurate, precise, distinguished with high recall and F1-score, thus highlighting the ability of the method to adequately detect depressive language. The contrasting analysis revealed that FFNN provided higher precise contextual comprehension of textual data than tradition and other RNN addressed models. .

Future work can explore:

- Improved prediction rate by incorporating data from multiple modes such as text and voice along with images.
- Increasing the database's originality, by including several languages with different dialects.
- Using lightweight versions of FFNN in restricted conditions such as DistilBERT.

The results build the call for using state-of-art, 'black' box machine learning approaches for mental health disorders, with the condition that these technologies need to be integrated with professional diagnostic skills.

7 Conclusion and Future Work

7.1 Conclusion

This investigation was able to show how machine learning, more specifically transformer based architectures such as the FFNN, could be used to predict depression given text data sourced out of Reddit.Applying the linguistic patterns, the success rates obtained reached 94.2% of accuracy and an F1 measure of 93.1% where the proposed framework surpassed more conventional models such as BERT, LSTM, GRU, BiLSTM.om Reddit discussions. By analyzing linguistic patterns, the framework achieved high accuracy (94.2 and an F1-score of 93.1 outperforming traditional models like BERT, LSTM, GRU, and

BiLSTM. BERT's pre-training design allowed it to pick up and understand contextual hints of depressive language well. The steps included in the implementation were feature extraction, training and testing of the model, and evaluation and which supported the hypothesis of the possibility of developing an automated method of detecting depression. Information misuse issues like anonymization of information and bias were also well handled. Some of the difficulties such as the computational needs of BERT were solved based on GPU-supported environments and some containerized deployment pipelines. As attractive as the proposed performance may look, the framework should be designed as an additional instrument and not a clinical decision support system, which underlines the role of interaction between data scientists and clinicians. .

7.2 Future Work

Based on the findings and in order to propose directions for future work that would further strengthen and broaden the framework, the following have been postulated. The most researched area of development is solely the Textual analysis of social media data} but, multi-model analysis is also an important part which incorporates textual data with speech or video data. If the depression detection system integrated both auditory and visual feedback like a speaker's tone of voice or facial expressions, a much better system for better detection of depression in diverse settings would have been developed. One more remarkable topic is the cultural and linguistic diversity The current system also subsources data only from specific subreddits and only English language posts. More work will be done in the future to extend the features of the model for Global use, Such studies would involve the incorporation of texts in more languages and from cultures that are diverse. This would help add generalization to the model and that would enable it to be able to detect depression in different platforms and usage of social and linguistic functions. Another important factor for future improvements is the ability to effectively deploy the solutions obtained. Although we are currently using highly expressive architectures such as the transformer-based BERT model, these models may not be ideally suited for real-time applications due to their computational demands in many environments. This can be resolved by providing a clue as to which directions to explore; for example, reducing the weights required by the transformer models themselves, or using lightweight versions of the transformer such as Data Rumors' DistilBERT or MobileBERT. These models want to be equally effective but more light to make real-time depression detection possible in mobile devices or low resource environments.

References

98, S. (2023). Unraveling the power of attention: A dive into transformer architecture. Accessed: 2024-12-12.
URL: *https://medium.com/@sciencely98/unraveling-the-power-of-attention-a-dive-into-transformer-architecture-eb6594ed77a3*

Bertl, M., Bignoumba, N., Ross, P., Yahia, S. B. and Draheim, D. (2024). Evaluation of deep learning-based depression detection using medical claims data, *Artificial Intelligence in Medicine* **147**: 102745.

Das, A. K. and Naskar, R. (2024). A deep learning model for depression detection based on mfcc and cnn generated spectrogram features, *Biomedical Signal Processing and Control* **90**: 105898.

Khan, M. F. I., Anjum, F., Alam, S. and Bahadur, E. H. (2024). Depression detection through activity recognition: Deep learning models using synthesized sensor data, *JOURNAL OF BASIC SCIENCE AND ENGINEERING* **21**(1): 571–590.

Liu, Y. (2024). Depression detection via a chinese social media platform: a novel causal relation-aware deep learning approach, *The Journal of Supercomputing* **80**(8): 10327–10356.

Marriwala, N., Chaudhary, D. et al. (2023). A hybrid model for depression detection using deep learning, *Measurement: Sensors* **25**: 100587.

Nama, A. (2023a). Understanding bidirectional lstm for sequential data processing. Accessed: 2024-12-12.
URL: *https://medium.com/@anishnama20/understanding-bidirectional-lstm-for-sequential-data-processing-b83d6283befc*

Nama, A. (2023b). Understanding gated recurrent unit (gru) in deep learning. Accessed: 2024-12-12.
URL: *https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2*

Naqvi, S. (2023). Long short-term memory (lstm). Accessed: 2024-12-12.
URL: *https://medium.com/@saba99/long-short-term-memory-lstm-fffc5eaebfdc*

Rehmani, F., Shaheen, Q., Anwar, M., Faheem, M. and Bhatti, S. S. (2024). Depression detection with machine learning of structural and non-structural dual languages, *Healthcare Technology Letters* .

Singh, G. (2023). Training feed forward neural network (ffnn) on gpu - beginner's guide. Accessed: 2024-12-12.
URL: *https://gurjeet333.medium.com/training-feed-forward-neural-network-ffnn-on-gpu-beginners-guide-2d04254deca9*

Squires, M., Tao, X., Elangovan, S., Gururajan, R., Zhou, X., Acharya, U. R. and Li, Y. (2023). Deep learning and machine learning in psychiatry: a survey of current progress in depression detection, diagnosis and treatment, *Brain Informatics* **10**(1): 10.

Tejaswini, V., Sathya Babu, K. and Sahoo, B. (2024). Depression detection from social media text analysis using natural language processing techniques and hybrid deep learning model, *ACM Transactions on Asian and Low-Resource Language Information Processing* **23**(1): 1–20.

Thekkekara, J. P., Yongchareon, S. and Liesaputra, V. (2024). An attention-based cnn-bilstm model for depression detection on social media text, *Expert Systems with Applications* **249**: 123834.

Zhang, W., Xie, J., Zhang, Z. and Liu, X. (2024). Depression detection using digital traces on social media: A knowledge-aware deep learning approach, *Journal of Management Information Systems* **41**(2): 546–580.