A MAJOR PROJECT – II REPORT
ON
MULTILINGUAL DEPRESSION DETECTION IN ONLINE SOCIAL MEDIA ACROSS EIGHT INDIAN LANGUAGES

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF

MASTER OF TECHNOLOGY
IN
COMPUTER SCIENCE & ENGINEERING

Submitted By
RAJDERKAR VIRAJ JAYANT
2K22/CSE/17

Under The Supervision Of
DR. ARUNA BHAT
(Associate Professor)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042
May 2024

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE’S DECLARATION

I, **Rajderkar Viraj Jayant (2K22/CSE/17)**, hereby certify that the work which is being presented in the major project report II entitled “**Multilingual Depression Detection in Online Social Media Across Eight Indian Languages**” in partial fulfilment of the requirements for the award of the Degree of Master of Technology, submitted in the **Department of Computer Science and Engineering**, Delhi Technological University is an authentic record of my own work carried out during the period from August 2022 to April 2024 under the supervision of **Dr. Aruna Bhat.**

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor (s) **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project titled "**Multilingual Depression Detection in Online Social Media Across Eight Indian Languages**", submitted by **Rajderkar Viraj Jayant**, Roll No. **2K22/CSE/17**, Department of **Computer Science & Engineering**, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of degree of **Master of Technology** (M.Tech) in Computer Science and Engineering is a genuine record of the project work carried out by the student unde my supervision. To the best of my knowledge this work has not been submitted in par or full for any Degree to this University or elsewhere.

Place: Delhi
Date:

Dr. Aruna Bhat
Associate Professor
Delhi Technological University

ACKNOWLEDGEMENT

I am extremely grateful to my project guide, **Dr. Aruna Bhat**, Associate Professor, Department of Computer Science and Engineering, Delhi Technological University, Delhi for providing invaluable guidance and being a constant source of inspiration throughout my research. I will always be indebted to her for the extensive support and encouragement she provided.

I am highly indebted to the panel faculties during all the progress evaluations for their guidance, constant supervision and for motivating me to complete my work. They helped me throughout by giving new ideas, providing necessary information and pushing me forward to complete the work.

Rajderkar Viraj Jayant
(2K22/CSE/17)

ABSTRACT

Detecting depression via social media platforms has emerged as an important field of studies, leveraging machine learning techniques to analyze user-generated posts across diverse linguistic and cultural contexts. This provides a unique research opportunity into depression detection on social media, specialising in multilingual analysis across various Indian languages. The methodology encompasses data collection, preprocessing, model improvement, and experimental design. Social media tweets, sourced from Twitter, have been gathered and categorized into depressive and non-depressive subsets based on keyword analysis. Ethical issues had been paramount in the course of the data collection and labelling part, ensuring privacy and compliance with moral recommendations. Data preprocessing strategies, consisting of textual content cleaning, normalization, and tokenization, were employed to put together the dataset for analysis. Word embedding with pre-trained FastText vectors improved the semantic representation of the entered text, contributing to model overall performance. The improvement of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, as an ensemble method, enabled powerful sentiment analysis throughout the eight Indian languages.

Experimental results confirmed the efficacy of these models in detecting depressive styles in social media posts. High accuracy, precision, recall, and F1-scores were observed throughout languages such as Gujarati, Kannada, Hindi, Bengali, Telugu, Marathi, Malayalam, and Tamil. The models exhibited strong performance, indicating their adaptability to numerous linguistic landscapes and their potential to become aware of depressive content with reliability. Comparative evaluation found out that each LSTM and GRU fashions executed well, with mild variations in accuracy and recall. Ensemble strategies combining LSTM and GRU predictions confirmed improved precision and recall in a few instances, highlighting the capability for boosting overall performance through ensemble approaches. However, variations in precision and recall throughout languages underscored the significance of thinking about language-precise nuances in model improvement.

Despite the promising outcomes, several limitations have been noticed, such as dataset representativeness and language nuances. The effectiveness of the models can be improved by using a balanced dataset. Furthermore, the diverse linguistic landscape of Indian languages brought complexities which could require further exploration in future research endeavors. In conclusion, this study contributes precious insights into depression detection on social media in multilingual contexts. The developed models show strong performance across diverse Indian languages, imparting potential in depression detection and intervention. Recommendations for further studies consist of expanding datasets, developing language-specific models, and exploring superior architectures to deal with the mentioned challenges and enhance the effectiveness of depression detection models on social media systems.

Keywords: Depression Detection, Social Media Analysis, Multilingual Sentiment Analysis, Indian Languages, LSTM and GRU Models, Ensemble Learning, Translation Approach, Text Preprocessing, Mental Health Detection, Low-Resource Languages, Social Media Data Mining, Cross-Lingual Sentiment Analysis, Word Embeddings, Natural Language Processing

TABLE OF CONTENTS

LIST OF TABLES .... ix.

LIST OF FIGURES .... ix.

LIST OF ABBREVIATIONS .... x.

1 INTRODUCTION 1

1.1 Background and Context 1

1.2 Problem Statement 3

1.3 Research Objectives 4

1.4 Research Questions 6

1.5 Thesis Structure 9

2 LITERATURE REVIEW 12

2.1 Existing Research on Depression Detection 12

2.1.1 Thai Twitter and Depression Detection 13

2.1.2 Challenges and Complexities in Detection 15

2.1.3 Chinese and Cross-Language Studies 16

2.1.4 Key Themes and Techniques in Depression Detection 16

2.1.5 Challenges in Multilingual Contexts 17

2.1.6 Key Findings from Previous Studies 21

3 METHODOLOGY 26

3.1 Data Sources and Collection 26

3.1.1 Twitter as a Data Source 26

3.1.2 Dataset Acquisition and Content 26

3.1.3 Data Collection Process 27

3.2 Data Preprocessing 29

3.2.1 Text Cleaning and Normalization 29

3.2.2 Tokenization and Word Analysis 30

3.2.3 Word Cloud Visualization....30
3.3 Model Development....31
3.3.1 Why LSTM and GRU?....36
3.3.2 Training and Optimization Strategies....38
3.3.3 Ensemble Approaches....39
3.4 Word Embedding with Pre-trained FastText Vectors....40
3.5 Training and Model Architecture....42
3.5.1 Model Architecture....43
3.6 Experimental Design....48
3.6.1 Dataset Configuration....48
3.6.2 Handling Imbalanced Data....49
4 EXPERIMENTAL SETUP AND RESULTS....51
4.1 Experiments Conducted....51
4.2 Results and Analysis....54
4.3 Comparisons Between Models....55
4.4 Insights and Future Directions....59
4.5 Discussion of Limitations....60
5 CONCLUSION AND FUTURE SCOPE....64
5.1 FUTURE SCOPE....65
6 REFERENCES....67
LIST OF PUBLICATIONS....74

LIST OF TABLES

Table 2.1: Summary of papers published in year 2022 13

Table 2.2: Summary of papers published in year 2021 15

Table 231: Summary of papers published in year 2020 17

Table 2.4: Summary of papers published in year 2019 18

Table 2.5: Summary of papers published in year 2017 and 2018 19

Table 4.1: Our Results 53

Table 4.2: Comparative analysis (Summary of papers) 55

Table 4.3: Comparative Analysis(Precision, Recall, F1 Score) 58

LIST OF FIGURES

Figure 2.1: Year wise number of publications 14

Figure 3.1: Wordcloud Hindi 32

Figure 3.2: Wordcloud Gujarati 32

Figure 3.3: Wordcloud Bengali 33

Figure 3.4: Wordcloud Marathi 33

Figure 3.5: Wordcloud Telugu 34

Figure 3.6: Wordcloud Tamil 34

Figure 3.7: Wordcloud Malayalam 35

Figure 3.8: Wordcloud Kannada 35

Figure 3.1: LSTM Architecture 36

Figure 3.2: GRU Architecture 37

Figure 3.3: Ensemble Architecture 39

Figure 4.1: Comparison of various papers: Accuracy 57

Figure 4.2: Comparison of various papers: Precision 57

Figure 4.3: Comparison of various papers: Recall 58

Figure 4.4: Comparison of various papers: F1-Score 58

LIST OF ABBREVIATIONS

1. **NLP** - Natural Language Processing
2. **AI** - Artificial Intelligence
3. **ML** - Machine Learning
4. **LSTM** - Long Short-Term Memory
5. **GRU** - Gated Recurrent Unit
6. **RNN** - Recurrent Neural Network
7. **CNN** - Convolutional Neural Network
8. **BERT** - Bidirectional Encoder Representations from Transformers
9. **FastText** - Facebook AI's word embedding model
10. **SVM** - Support Vector Machine
11. **ROC** - Receiver Operating Characteristic
12. **AUC** - Area Under the Curve
13. **TPR** - True Positive Rate
14. **FPR** - False Positive Rate
15. **F1 Score** - Harmonic mean of precision and recall
16. **PCA** - Principal Component Analysis
17. **TF-IDF** - Term Frequency-Inverse Document Frequency
18. **POS** - Part of Speech
19. **API** - Application Programming Interface
20. **CSV** - Comma-Separated Values
21. **JSON** - JavaScript Object Notation
22. **HTML** - HyperText Markup Language
23. **URL** - Uniform Resource Locator
24. **HTTP** - HyperText Transfer Protocol
25. **MSE** - Mean Squared Error
26. **MAE** - Mean Absolute Error
27. **RMSE** - Root Mean Squared Error
28. **GAN** - Generative Adversarial Network
29. **SST** - Sentiment Strength Detection
30. **NLP** - Natural Language Processing
31. **Naive Bayes** - A probabilistic classifier based on Bayes' theorem

1 INTRODUCTION

1.1 Background and Context

Depression is one of the most widespread health situations of modern-day lifestyles, affecting hundreds of thousands worldwide, irrespective of age, gender, or monetary fame. The World Health Organization identifies depression as a pinnacle contributor to incapacity, with almost 280 million humans suffering from depressive disorders. While “depression” is most commonly used for describing diminished stages of emotional fitness and happiness, the quantity of its outcomes extends a long way beyond.

People who develop depressive issues may additionally suffer in their mood over a long period, pursuits, and capacity for amusement. Other feasible signs include adjustments in urge for food, weight, and sleep patterns, loss of vanity, declining capacity to pay attention, and thoughts of self-harm or suicide.

Beyond the non-public effect, depression has widespread societal prices, essential to decreased productiveness, extended healthcare prices, and strained personal relationships. Amid this growing mental health crisis, social media has emerged as a powerful platform for people to express their thoughts, studies, and feelings. Platforms inclusive of Twitter, Facebook, Instagram, and Reddit offer a totally unique public discussion board in which individuals talk freely, regularly in real-time. This openness has contributed to a more seen discourse on health, with customers sharing private testimonies, trying to find manuals, or discussing broader issues associated with emotional well-being.

However, this new form of communication moreover gives situations and possibilities for mental health research. Social media's huge reach and various client base create a wealth of information that may provide valuable insights into human conduct, temper, or even symptoms of health problems like depression. Yet, the sheer extent of facts, the variety of language and expression, and the casual nature of social media communication can complicate the technique of figuring out big patterns.

To address those challenges, Natural Language Processing (NLP) has grown to be a vital device in analyzing text-based total statistics from social media. NLP permits researchers the way to observe big volumes of text, extracting treasured information related to language patterns, sentiment, and emotional cues. By leveraging NLP, researchers can understand signs of depression in social media posts, providing a way for early detection that could cause well timed intervention and help.

The use of NLP for intellectual fitness evaluation on social media is gaining traction among researchers and healthcare experts. By studying linguistic functions which consist of phrase and sentence study, and the overall sentiment of posts, it is possible to come across early signs and symptoms of depressive signs. This functionality has the ability to revolutionize intellectual health surveillance, thinking of a greater proactive approach to figuring out and addressing depression.

Despite those improvements, large demanding situations stay. The casual language used on social media, the kind of expressions all through cultures and languages, and concerns about privacy and ethical considerations are a number of the hurdles that researchers should navigate. Additionally, the detection models have to study to understand a huge type of depressive signs and symptoms and symptoms, considering that depression manifests in another way at some point of individuals and cultures.

In summary, social media offers a promising but complicated panorama for early depression detection. The combination of NLP and social media statistics offers a very specific approach to know-how and addressing the depression disaster, with the potential to keep lives and decrease the weight of depression on human beings and society. However, researchers need to stay conscious of the demanding situations and ethical implications at the equal time as striving to expand strong, dependable models for depression detection in a diverse and ever-evolving on line surroundings.

1.2 Problem Statement

The growing occurrence of depression has spurred the development of modern tactics to its detection, especially inside online environments like social media. Natural Language Processing (NLP) has emerged as a important tool for reading textual content-primarily based content material, allowing researchers to discover signs and symptoms of depression through language patterns, sentiment evaluation, and unique textual content-pushed strategies. While giant improvement has been made in this challenge, important gaps persist, specially close to detecting depression across a several range of linguistic and cultural contexts.

Most of the existing studies in depression detection on social media has focused on English-language datasets. This bias in the direction of English shows the wider dominance of English-speakers users on systems like Twitter, Facebook, and Reddit, that are generally used in such research. However, this slim attention overlooks the of non-English-speakers populations, fundamental to a huge hollow in studies that limits the generalizability of cutting-edge-today models.

India, with its rich linguistic diversity and complicated cultural panorama, exemplifies the challenges in extending depression detection studies beyond English. The Indian charter acknowledges 22 respectable languages This variety offers a unique set of demanding situations for NLP-based total depression detection. Existing models trained greater on English-language records won't be effective while implemented to Indian languages, that have awesome grammatical systems, idiomatic expressions, and cultural references.

The lack of comprehensive datasets for Indian languages complicates the improvement of robust depression detection models. Data series in low-useful resource languages is inherently difficult because of restricted availability, linguistic variations, and cultural nuances. This shortage of statistics influences the capability to create accurate models and will increase problems of the reliability of models that have learned totally on English records. The want for full-size information curation, language-specific preprocessing, and culturally sensitive strategies provides to the complexity of the task.

Moreover, cultural variations play a sizeable part in how depression is perceived and expressed. In many Indian contexts, intellectual health problems, such as depression, convey a stigma that would have an impact on how humans speak their feelings on social media. This cultural variability influences the language used to explain depressive symptoms, making it hard to develop everyday models for depression detection. Models that work properly for English-speaking populations won't seize the perfect expressions and cues found in Indian languages.

The implications of those gaps are profound. Without reliable models for multilingual depression detection, huge segments of the worldwide population continue to be underserved. Early detection of depression can cause well timed intervention and superior outcomes, but this calls for models that are powerful for specific languages and cultures. The project isn't honestly to translate present strategies but to expand new methodologies that account for linguistic variety and cultural context.

Given this history, the focal point of this thesis is to cope with those gaps by way of growing a entire method to depression detection in low-resource Indian languages. This entails the improvement of multilingual datasets, language-particular models, and culturally sensitive strategies. By doing so, this research interests make contributions to a inclusive and effective technique for depression detection on social media, with the potential to decorate intellectual fitness outcomes for various populations.

1.3 Research Objectives

Depression detection in social media is a developing situation with huge implications for public health, specially in a culturally and linguistically numerous context like India. Despite improvements in Natural Language Processing (NLP) and machine learning knowledge of, most current depression detection research is on English-language information. This thesis seeks to cope with the space in studies and property for Indian languages, aiming to increase the accuracy and reliability of depression detection models in a multilingual setting.

The primary goal of this thesis is to create a robust approach for detecting depression in eight Indian languages. This intention reflects the need for inclusivity in mental health research, spotting the linguistic and cultural variety in India. Here's an in depth breakdown of the ideal goals:

Develop a Comprehensive Approach for Depression Detection

The first intention is to develop a entire approach to depression detection during eight Indian languages. This method will leverage advanced NLP strategies, which can be crucial for reading massive volumes of text from social media systems. Given the several grammatical systems and unique vocabularies of Indian languages, the method ought to be adaptable to every language's unique traits. The studies will discover strategies to efficiently approach and analyze text, making sure accurate detection of depressive symptoms throughout unique linguistic contexts.

Implement an Ensemble Model Incorporating LSTM and GRU Architectures

The 2nd goal is to enforce an ensemble model that mixes Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. These deep getting to know models are broadly diagnosed for their effectiveness in handling sequential data, making them best for studying textual content from social media. The ensemble method is designed to enhance overall overall performance via integrating the strengths of each LSTM and GRU. By combining these architectures, the model can seize complicated patterns and relationships in the text, leading to greater correct depression detection.

Use Pre-Trained FastText Embeddings

The third objective is to use pre-trained FastText embeddings to enhance model overall performance and address language-precise worrying situations. FastText embeddings are acknowledged for their capability to seize phrase representations at the subword degree, it really is in particular useful in multilingual contexts. Given the linguistic range in Indian languages, the usage of pre-train d embeddings allows the model to better understand phrase models and similarities. This functionality is critical for analyzing text with numerous vocabulary and casual language usually placed on social media.

Provide Insights into the Complexities of Analyzing Mental Health on Social Media

The fourth goal is to offer insights into the complexities of studying intellectual health on social media in a multilingual context. This consists of addressing the particular demanding situations posed by the usage of cultural variations, casual language, and the stigma associated with mental health in some areas. The thesis will discover how these elements affect the detection of depressive symptoms and communicate techniques for overcoming those stressful situations. By gaining a deeper data of those complexities, the research objectives to make contributions to the wider area of multilingual depression detection.

Overall, the research dreams mentioned in this thesis purpose to decorate the sector of depression detection via addressing the specific goals of low-useful resource Indian languages. By growing a comprehensive method, implementing an ensemble model, the use of pre-professional FastText embeddings, and presenting insights into the complexities of social media analysis, the study seeks to make a significant contribution to the early detection of depression in a multilingual and culturally numerous setting. The effects of this studies have the capacity to decorate intellectual fitness monitoring, facilitate early intervention, and in the long run enhance the well-being of human beings in various linguistic groups.

1.4 Research Questions

In exploring the complex and nuanced vicinity of depression detection throughout diverse languages and cultures, unique studies questions function guiding factors for research. These questions now not simplest shape the route of the take a look at however additionally assist identify regions wherein additional studies is needed. The following studies questions underpin this thesis, specializing in detecting depressive signs in social media content material throughout a couple of Indian languages the use of advanced system mastering and natural language processing (NLP) techniques.

1. **What Are the Key Linguistic Features That Indicate Depressive Symptoms in Social Media Posts Across Multiple Indian Languages?**

The first studies question seeks to emerge as privy to the linguistic markers which can be indicative of depressive signs and symptoms in social media posts. This includes studying the vocabulary, grammatical structures, and typical tone of the textual content. Key features should range from specific words and phrases typically associated with depressive moods to more subtle language styles along with reduced phrase range, stepped forward use of terrible adjectives, or common references to unhappiness and hopelessness. Given the linguistic variety throughout Indian languages, it's far vital to discover those competencies in a culturally sensitive manner, considering how depression is expressed in specific languages and areas.

To deal with this query, the studies will behavior an in depth linguistic evaluation of social media posts from clients with recognized depressive signs. This assessment will focus on identifying habitual linguistic styles that may characteristic reliable signs of depression. By expertise the ones key features, researchers can improve the accuracy of depression detection models and growth language-unique guidelines for future research.

2. **How Effective Are the LSTM and GRU Models in Detecting Depression in Multilingual Datasets?**

The 2d research question examines the effectiveness of famous deep studying models, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), in detecting depression in multilingual datasets. These models are diagnosed for their capability to process sequential information and seize complicated relationships and interior textual content. The study targets to determine how properly those models perform when implemented to datasets containing social media posts in multiple Indian languages.

To address this question, the studies will implement LSTM and GRU models, and train them on a numerous dataset of social media posts. The look at will study the models' accuracy, precision, take into account, and different relevant metrics to decide their effectiveness in detecting depressive symptoms. Additionally, the studies will find out how the ones models address the particular demanding situations of multilingual datasets, inclusive of language-unique idioms, code-switching, and variations in text shape.

3. **Can an Ensemble Approach Improve Depression Detection Accuracy Compared to Individual Models?**

The 3rd study question explores whether or not an ensemble method, which mixes the outputs of more than one models, can beautify depression detection accuracy in assessment to the use of character models on my own. Ensemble strategies are frequently applied in machine learning knowledge of to decorate usual overall performance with the aid of leveraging the strengths of diverse models. The look at will test out whether or no longer combining LSTM and GRU models into an ensemble yields better results in detecting depression for the duration of multiple Indian languages.

To cope with this question, the research will create an ensemble model that integrates the outputs of LSTM and GRU models. The study will study the overall performance of the ensemble approach with that of person models, assessing upgrades in accuracy, robustness, and reliability. This assessment will help decide if ensemble strategies provide a remarkable benefit in multilingual depression detection.

4. **What Challenges Arise When Working with Low-Resource Indian Languages, and How Can They Be Addressed?**

The fourth study query examines the particular disturbing situations associated with jogging in low-useful resource Indian languages for depression detection. These stressful conditions can also encompass limited availability of categorized datasets, variations in linguistic shape, cultural nuances, and the need for culturally touchy strategies to data analysis. The research objectives to understand the ones demanding situations and suggest answers for addressing them.

To solution this query, the study will delve into the problems encountered at the same time as gathering and studying facts in low-aid Indian languages. This might also contain exploring techniques for constructing comprehensive datasets, designing language-unique preprocessing strategies, and thinking about moral implications associated with privacy and cultural sensitivity. By expertise the ones stressful conditions, the studies can offer precious insights for future studies and make a contribution to a extra inclusive approach to depression detection in multilingual contexts.

1.5 Thesis Structure

A nicely-described thesis form is essential for directing readers thru the studies adventure, from the initial challenge announcement to the final conclusions. The shape have to logically gift the critical factor components of the study, making sure that each phase builds upon the previous one to create a cohesive narrative. Here's an accelerated model of the proposed form, detailing the content of each bankruptcy and explaining how they make contributions to the overall purpose of the thesis:

Chapter 1: Introduction
The first chapter sets the degree for the complete thesis by manner of providing ancient past data on the importance of early depression detection inside the context of intellectual health. It discusses the significance of social media as a supply of information for this motive, thinking about its feature in allowing humans to specific their thoughts and feelings. The introduction additionally offers the problem declaration, emphasizing the distance in research concerning low-useful resource Indian languages. This bankruptcy outlines the research targets and questions, giving a easy feel of what the thesis objectives to perform. Additionally, it consists of a pinnacle degree view of the thesis structure to prepare readers for the imminent sections.

Chapter 2: Literature Review

This chapter delves into modern studies on depression detection, providing a comprehensive assessment of earlier research inside the place. It discusses the several methodologies used to find out depressive signs and symptoms on social media, including Natural Language Processing (NLP) and machine studying techniques. The literature evaluate also explores the challenges related to multilingual analysis, highlighting key findings from relevant studies. By reading this body of research, the bankruptcy establishes the context for the thesis and identifies areas in which similarly exploration is wanted, specifically in low-useful resource Indian languages.

Chapter 3: Methodology

The approach bankruptcy describes the technique taken to gather, preprocess, and study the information. It starts offevolved thru outlining the facts assets, collectively with the social media platforms used to accumulate text-based absolutely information. This segment explains the statistics collection techniques, addressing any ethical considerations which consist of client privateness and consent. The bankruptcy then statistics the preprocessing strategies finished to the statistics, consisting of tokenization, stopword elimination, and normalization, to put together it for evaluation. Additionally, the approach phase covers the development of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, explaining why those architectures have been selected. It moreover introduces the ensemble technique, which mixes the outputs of various models to improve detection accuracy.

Chapter 4: Experimental Setup and Results

The fourth bankruptcy outlines the experiments conducted to test the effectiveness of the depression detection models. It provides sure data at the experimental setup, which include the configuration of the training and sorting out datasets, hyperparameters, and assessment metrics. This section presents the effects of the experiments, using tables, graphs, and charts to demonstrate key findings. The bankruptcy additionally includes a radical assessment of the outcomes, comparing the performance of various models and discussing any trends or styles found at some point of the experiments.

Chapter 5: Conclusion

The final chapter summarizes the critical issue findings from the thesis, reinforcing the maximum huge contributions to the sphere. It offers pointers for researchers, practitioners, and policymakers interested by depression detection on social media, emphasizing the importance of multilingual and culturally sensitive techniques. The end moreover outlines ability destiny research instructions, indicating areas in which similarly investigation must yield precious insights.

By following this structure, the thesis gives a complete exploration of multilingual depression detection on social media. Each chapter builds on the preceding one, developing a logical development from the challenge announcement to the final conclusions. This based method contributes to a deeper know-how of the demanding conditions and possibilities in the situation, paving the manner for future studies and development in this essential area of take a look at.

2 LITERATURE REVIEW

The exploration of psychological health via social media has garnered large interest in contemporary years, leading to a diverse sort of studies endeavors throughout languages and systems. This examines key studies that have applied diverse methodologies and datasets to understand the complexities of depression detection, highlighting each the conditions and opportunities on this subject.

2.1 Existing Research on Depression Detection

Social media systems along with Twitter, Reddit, Facebook, and Instagram have come to be large assets of information for depression detection research. These structures provide a very particular window into human beings's mind, feelings, and social interactions, allowing researchers to research styles and pick out out signs and symptoms of depressive signs and symptoms. Below are some of the remarkable research on this vicinity.

Paper	Language	Social Media	Tools	Remarks
[1]	Thai	Twitter	NLTK, Google Translation API	Data is labelled using inferences from PHQ-9 questionnaire and demographic information. Used NLTK for NLP.
[2]	English	Twitter + Reddit	Tweepy, PRAW, BERT, Word2Vec, GloVe	Data labelling based on hashtags and social groups of the post. Used Tweepy and PRAW. Authors show that BERT outperforms other Word2Vec and GloVe.
[16]	English	Twitter	NLTK, KNN, SVM, LR, LSTM	Does a technical study of oversampling techniques by using balanced and unbalanced dataset for all 5 models.

[17]	English	Twitter	TextCNN, word2vec	A novel approach involving a pre-trained TextCNN which uses word2Vec and extracts the emotion is proposed.
[18]	English	Mastodon	spaCy Library	Uses an uncommon social platform. Studies dialogue structural elements. And the relationship between emotion and depression.
[19]	Chinese	Weibo	Jieba, NB, KNN, and LR	The researchers used a novel combination of Naive Bayes, KNN, and logistic regression techniques to create their model and claim greater accuracy compared to previous algorithms.
[20]	English	Twitter	Bi-GRU, MLP, HAN	A new hybrid approach which models data and extracts explanations from it. Uses MLP and HAN, enhances classification performance and supports its prediction.
[21]	English	Twitter		A deep RNN-LSTM approach was used to predict textual data with an accuracy of 99% and a reduced false positive rate. The model was compared with other ML approaches.
[22]	Chinese	Weibo	XLNet Embedding	Authors publish a large depression detection dataset for Weibo (WU3D). They also provide a method to fuse multimodal information to detect individual-level depression.
[23]	English	Twitter		A deep RNN-LSTM approach was used to predict textual data with an accuracy of 99% and a reduced false positive rate. The model was compared with other ML approaches.
[24]	English	Twitter	N-gram, TF-IDF, LDA	A numerical score was generated for each user based on their posts. This and other Linguistic features (N-Gram+TF-IDF) and LDA achieves an accuracy of 89% using SVM.

Table 2.1: Summary of papers published in year 2022

2.1.1 Thai Twitter and Depression Detection

One test that sticks out inside the discipline of depression detection is the research on Thai Twitter, performed by means of the usage of [1]. Take a look at utilising the

PHQ-9 questionnaire, a widely used device for measuring depression severity, at the side of demographic records and Twitter IDs, to categorize individuals into tremendous ranges of depression. The researchers used numerous system learning models, which includes Decision Tree and Deep Learning, to categorise customers based on their social media activity. This method highlighted the complexity of depression detection and showed the effectiveness of the usage of system learning strategies for this motive.

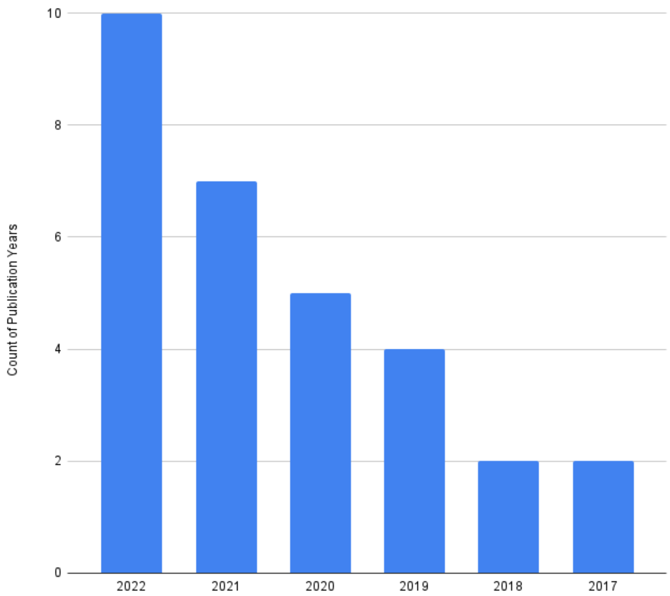

Figure 2.1: Year wise number of publications

English-Language Studies on Twitter and Reddit:

Another big take a look at through [2] centered on datasets from Reddit and Twitter in English. This research accomplished tremendous accuracy expenses the usage of advanced models like Long Short-Term Memory (LSTM), Bi-directional LSTM (Bi-LSTM), and Bidirectional Encoder Representations from Transformers (BERT). These models are acknowledged for his or her capacity to technique sequential statistics and seize tricky relationships in textual content, making them high-quality for reading social media content material material. The success of those models underscored the significance of word embedding strategies in accurately taking photographs textual records, which in turn is important for powerful depression detection.

2.1.2 Challenges and Complexities in Detection

The study on Thai Twitter found out several challenges in depression detection, which incorporates diverse tiers of depressive signs and symptoms and signs and the need for context-based absolute expertise. Similarly, research on English-language datasets confirmed the significance of multimodal evaluation, as noted in research like [3], which blanketed numerous data assets to beautify detection accuracy. This technique is specially useful in taking the numerous strategies of clients specific depressive signs and symptoms on social media.

Paper	Language	Social Media	Tools	Remarks
[25]	English	Twitter	VADER, NRC	A union of 3 methods was done: Polarity calculation between -1 to 1, emotion detection of 8 basic classes, self declaration of depression by users in their posts.
[26]	English	Twitter + Reddit	Bag of polarities	A novel approach with the assumption that the words' discriminative value depends on who uses (traits of the user) them and on the polarity of the messages in which they occur.
[27]	Bengali	Bengali blogs, Facebo ok	Keras tokenizer, RNN, LSTM	The authors used RNN with LSTM and claim to achieve an accuracy of 98% using a Bengali language-based analysis system for analysing words, which is difficult to write in English.
[28]	Japanese	Blog Data	SVD, count vectorisor	The study was conducted for the Japanese language and shows that restricting the parts of speech to nouns, verbs, adjectives, adjectival verbs, and adverbs improves results. Their model's F1 score was 95.9, and its accuracy was 95.6.
[29]	Arabic	Twitter	SVM, KNN, NB, DT, RF	Studies Arabic. The results show that KNN outperforms other models followed by SVM.
[30]	English	Twitter		The models in this study performed well even when tested on datasets unrelated to the training datasets. Also, dynamic sampling could increase

				the accuracy of the less-populous class.
[31]	English	Twitter		Uses auto-summarisation of relevant user posts and thereby focuses only on the most relevant information during model train-training. CNN-GRU

Table 2.2: Summary of papers published in year 2021

2.1.3 Chinese and Cross-Language Studies

Expanding past Western structures, research like [4] tested depression detection on Weibo, a popular Chinese social media platform. This studies emphasised the want for culturally specific techniques, recognizing that language and cultural context play a crucial function in how depression is expressed. The bypass-language comparisons carried out through [6], the use of Twitter and Reddit, further highlighted the complexities of multilingual assessment and the significance of language-specific models.

2.1.4 Key Themes and Techniques in Depression Detection

Several commonplace concern matters and strategies emerge from the present studies on depression detection in social media. These include using machine learning and deep learning. getting to know models, which encompass LSTM, Bi-LSTM, and BERT, to investigate text-based total information. The research additionally spotlight the importance of characteristic extraction strategies, which incorporates word embeddings, in know-how character-generated content material.

The literature overview additionally factors to the want for multimodal analysis, as this technique permits researchers to combine specific sorts of information to improve detection accuracy. Furthermore, the research emphasises the challenges of running with multilingual datasets and the importance of addressing cultural nuances to grow powerful depression detection models.

In summary, the existing studies on depression detection in social media has provided valuable insights into the methodologies and annoying situations involved on this location. While full-size development has been made, specifically with machine and deep learning to know strategies, similarly research is needed to cope with the complexities of multilingual and culturally numerous contexts. The next section will explore the particular demanding situations of operating with more than one languages and the implications for depression detection models.

2.1.5 Challenges in Multilingual Contexts

Depression detection in multilingual contexts poses substantial situations for researchers, specifically whilst handling low-useful useful resource languages. The unique linguistic and cultural additives of every language, mixed with the intricacies of herbal language processing (NLP), create a complex panorama for reading social media records. This phase explores the crucial component of situations in multilingual depression detection, specializing in linguistic range, cultural nuances, and records scarcity.

Linguistic Diversity

Linguistic range is a main task in multilingual depression detection. Languages range significantly in their grammar, syntax, vocabulary, and semantics. For instance, languages like Chinese have precise character-primarily based absolute scripts, even as others like Hindi use a aggregate of scripts and showcase complicated inflection patterns. These variations could make it difficult to growth a one-length-suits-all model for depression detection.

Furthermore, the way depression is expressed varies across languages. In a few cultures, direct references to mental fitness challenges are commonplace, while in others, human beings use greater subtle or metaphorical language to carry their feelings. This linguistic variety requires researchers to increase models which may be touchy to the ones models, making sure that they could as it must come across depressive symptoms and signs and symptoms regardless of the language used.

Paper	Languag e	Social Media	Tools	Remarks
[32]	Chinese	Weibo	TF-IDF, LPA, word2vec	A novel technique for LFA is proposed which uses a combination of TF-IDF, word2Vec, label propagation algorithm.
[33]	Bengali	Facebook + Twitter	TF-IDF, CountVectorizer	They tested 6 common ML algorithms and found that each produced accuracy

				levels of around 90% for test data and 91% for train data.
[34]	English	Reddit	Word2VecEmbed, GloVe, FastextEmbed	They primarily focused on shortening the classification time. Discovered that Word2VecEmbed+Meta features worked effectively.
[35]	English	Reddit	Bipolar Feature extraction	Developed a novel modified Bayesian classifier, which improved the performance when it was used with a bipolar feature vector, resulting in an F1 score of 82.75%.
[36]	English	Twitter	TF-IDF	The top 100 words used by depressive Twitter users have been analysed.

Table 231: Summary of papers published in year 2020

Cultural Nuances

Cultural nuances play a critical role in depression detection, as they've an impact on how people speak intellectual fitness and express emotions. In many cultures, mental fitness remains a stigmatized topic, important for human beings to keep away from direct references to depression of their social media posts. This cultural sensitivity can bring about varied expressions of depressive symptoms and signs and symptoms, making it hard to create correct detection models.

For instance, studies like [4], which tested depression detection on Weibo in Chinese, highlighted the importance of information and cultural context. In some cases, cultural references and idiomatic expressions can bring vast meaning that won't be proper away obvious to the ones out of doors the tradition. Researchers ought to bear in mind those cultural nuances even as growing models to ensure they successfully seize the underlying sentiment of the text.

Data Scarcity

Low-useful resource languages pose additional demanding situations because of information shortage. Unlike English, which has vast datasets and sources for NLP, many distinct languages have restricted records to be had for training and analysis. This scarcity can restriction the improvement of robust models for depression detection, as the dearth of records makes it tough to construct correct and reliable algorithms.

Research via the usage of [6], which explored move-language comparisons on Twitter and Reddit, emphasized the complexities of multilingual evaluation and the want for language-precise tactics. In low-aid languages, information collection and labeling may be time-eating and beneficial resource-substantial, requiring researchers to employ innovative methods to conquer those demanding situations. Additionally, data scarcity can impact model generalizability, as models professional on restricted information may not perform well across distinct contexts.

Paper	Language	Social Media	Tools	Remarks
[37]	Spanish, English	Twitter	Word2vec, Google Translate, skip gram	Studies cross lingual depression detection. Compares manual (contextual factors added) and automated translation.
[38]	English	Reddit	TF-IDF, Bag of words	A voting classifier and a straightforward feedforward neural network were developed and compared. The voting classifier performs better.
[39]	English	Twitter, Reddit	1,2,3 gram CLM	The Twitter-based CLM model outperforms the Reddit-based model by a wide margin. Provides possible explanation and comparison of the 2 platforms.
[40]	English	Twitter	SS3	This study uses the SS3 model which outperformed state-of-the-art approaches in both incremental chunk-by-chunk and incremental post-by-post classification.

Table 2.4: Summary of papers published in year 2019

Translation and Multilingual Models

To deal with the demanding situations of multilingual contexts, researchers frequently turn to translation-primarily based strategies. Translating social media statistics right into a commonplace language, such as English, can simplify evaluation and permit for using present NLP models. However, translation introduces its very very own set of challenges, which encompass loss of context, misinterpretation of idiomatic expressions, and cultural inaccuracies.

The want for multilingual models is plain, as they offer a more whole technique to depression detection. These models should be able to handling linguistic variety even as accounting for cultural nuances and records shortage. Techniques like pre-train d phrase embeddings, switch mastering, and pass-lingual training are regularly hired to enhance model typical performance in multilingual contexts.

Paper	Language	Social Media	Tools	Remarks
[41]	Chinese	Weibo	TF-IDF, NB, KNN, DT, SVM	Created a unique emotional lexicon for social media users. Multi kernel SVM, they claim outperforms other ML methods such as NB, KNN, Decision Tree and SVM.
[42]	Chinese	Wechat	Deep learning, Word segmentation	Proves that deep learning can be used for perinatal depression screening and is more objective than the present manual screening procedures.
[43]	Hindi, English	Twitter	TF-IDF, Multinomial Naive Bayes	The proposed model is able to classify tweets in English, Hindi, and Hinglish languages, helping to reduce depression rates globally.
[36]	English	Twitter	NLTK toolbox, WordNet	Describes the differences in the online behaviours of depressed and non-depressed users.

Table 2.5: Summary of papers published in year 2017 and 2018

In precise, the demanding situations in multilingual depression detection are multifaceted, encompassing linguistic variety, cultural nuances, records scarcity, and the complexities of translation. Addressing those worrying conditions calls for progressive techniques and a deep statistics of the cultural and linguistic context. By acknowledging those complexities, researchers can enlarge greater effective models for depression detection in multilingual contexts, ultimately contributing to a broader and additional inclusive understanding of mental health on social media.

2.1.6 Key Findings from Previous Studies

A thorough evaluation of prior research in depression detection on social media famous great insights into the strategies and models used, the languages examined, and the various achievement costs finished. These findings underscore the want for tailor-made strategies even as handling linguistic and cultural contexts. This phase explores the key takeaways from previous studies, with a focal point on model range, language-specific demanding situations, and the feature of cultural context.

Model Diversity and Performance

Studies in depression detection have hired a big selection of machine learning and deep gaining knowledge of models, demonstrating various ranges of fulfillment at some stage in one-of-a-type datasets and languages. The comparative evaluation of those models offers a treasured perspective at the effectiveness of different strategies. For instance, studies on Bengali depression detection models has proven quite a few accuracy stages depending on the method used. A take a look at using BiGRU (Bidirectional Gated Recurrent Unit) finished an accuracy charge of 86%, demonstrating the efficacy of recurrent neural networks for sequential statistics assessment. In evaluation, some other study using GRU (Gated Recurrent Unit) on Bangla social media records achieved a decrease accuracy of 75.7%, indicating that model desire and configuration substantially impact performance.

Similarly, studies on Hindi depression reputation have applied specific models with diverse achievement fees. One study the usage of CNN-BiLSTM (Convolutional Neural Network with Bidirectional Long Short-Term Memory) for emotion recognition in code-blended social media finished an accuracy of 83.21%. In contrast, a take a look at the usage of BERT (Bidirectional Encoder Representations from Transformers) with a six-transformer approach to investigate blended English-Hindi records provided a distinct angle, highlighting the flexibility of transformer-primarily based models in dealing with complicated linguistic systems.

Language-Specific Challenges

The findings from previous studies highlight the specific demanding situations related to wonderful languages. For instance, studies in Dravidian languages like

Malayalam and Tamil revealed lower accuracy expenses, emphasizing the want for language-particular strategies. We have a study on Deep Learning-Mixed Malayalam-English textual content analytics achieved an accuracy of 76.33% the usage of GRU, indicating that mixed-language contexts present additional complexities.

The model in model performances for the duration of languages demonstrates the importance of thinking about linguistic nuances and cultural context in depression detection. A test on sentiment assessment in more than one South Indian languages, which includes Malayalam, Tamil, and Kannada, confirmed that each language required a tailor-made method to acquire most green effects. For instance, while LSTM (Long Short-Term Memory) completed an accuracy of 84.28% in Malayalam, it only performed 50 .93% in Tamil, underscoring the impact of linguistic variety on model effectiveness.

Cultural Context and Expression

Another key locating from previous research is the affect of cultural context on depression detection. Different cultures have various attitudes within the direction of mental fitness, affecting how people explicit depressive signs and signs. This model necessitates culturally sensitive strategies even as studying social media statistics for signs of depression.

Studies that centered on systems like Weibo in Chinese highlighted the importance of understanding cultural nuances. These studies found out that cultural references and idiomatic expressions play a substantial function in how depression is communicated, that can impact the accuracy of detection models. Similarly, research on move-language comparisons concerning Twitter and Reddit hooked up the complexities of multilingual assessment, reinforcing the need for culturally aware methodologies.

Implications for Depression Detection

The findings from those research have critical implications for depression detection on social media. The variety in model performances during languages underscores the need for language-precise and culturally touchy strategies. Tailored strategies, along with language-specific function extraction and culturally informed statistics processing, are vital for enhancing the accuracy and reliability of depression detection models.

Moreover, the diverse success costs of various models suggest that a one-size-fits-all approach won't be powerful in multilingual contexts. Researchers should keep in mind an appropriate characteristics of each language and lifestyle to broaden robust models that may efficiently discover depressive signs. The key findings from previous studies offer a foundation for future studies, highlighting the importance of linguistic variety, cultural context, and tailored techniques in depression detection on social media.

Implications for Multilingual Depression Detection

The implications of previous research on multilingual depression detection are profound, in particular in the context of social media evaluation. As studies have discovered, the detection of depressive signs and symptoms in the course of exclusive languages and cultural contexts affords specific worrying situations, requiring researchers to undertake several strategies. These implications can guide future studies and workout, emphasizing the want for language-precise and culturally touchy approaches.

Importance of Language-Specific Models

One of the important implications for multilingual depression detection is the importance of growing language-particular models. Studies have verified that models designed for English-language datasets frequently carry out poorly whilst applied to low-useful aid languages due to linguistic models. This model in overall performance underscores the need of making models which are tailored to the ideal linguistic characteristics of each language. Language-particular models permit for a more nuanced understanding of text-based total definite facts, taking pics of subtleties and context frequently not noted by way of typical models.

Addressing Cultural Nuances

Another crucial implication is the need to deal with cultural nuances in depression detection. Cultural context plays a enormous function in how individuals unique feelings and talk intellectual fitness. In many cultures, mental health remains a touchy topic, leading to oblique or coded expressions of depressive signs and signs.

Researchers have to bear in mind those cultural nuances even as developing detection models, making sure that they may successfully interpret the language and sentiment of numerous cultural companies.

For example, studies on Chinese social media structures like Weibo highlighted the significance of information cultural references and idiomatic expressions. Similarly, research on code-blended languages which incorporates Hinglish (a mix of Hindi and English) located out the complexities of detecting depressive symptoms in multilingual contexts. By incorporating cultural recognition into their tactics, researchers can improve the accuracy of depression detection models and reduce the danger of misinterpretation.

Need for Robust and Inclusive Models

The various overall performance of models across languages indicates the want for robust and inclusive depression detection models. To benefit this, researchers want to pass beyond simple translation-based absolutely procedures and recognition on developing models that would adapt to distinct linguistic and cultural contexts. This can also involve growing new algorithms, the usage of transfer studying, or using multi-project getting to know to improve model generalizability.

Inclusive models are critical for addressing the kind of languages and cultures on social media. By growing models that may artwork at some stage in multiple languages, researchers can extend the achieve of depression detection efforts and provide assist to a broader type of people. This inclusivity additionally contributes to decreasing the stigma related to intellectual health in certain cultures, as it fosters a greater open and information method to discussing depression.

The Role of Social Media in Mental Health Research

The implications for multilingual depression detection additionally increase to the broader position of social media in intellectual health studies. Social media systems provide a totally unique opportunity to collect actual-time records on human behavior and emotions. However, this wealth of data also comes with moral challenges, along with privacy and consent. Researchers need to make certain that their techniques follow moral recommendations on the identical time as leveraging social media's ability to decorate mental health effects.

Social media evaluation permits researchers to pick out out trends and patterns in mental fitness, supplying valuable insights for early detection and intervention. By expertise the consequences of multilingual depression detection, researchers can contribute to a more complete and culturally touchy approach to mental fitness studies. This has the potential to tell public fitness regulations and aid systems, ultimately main to better mental fitness results for people throughout languages and cultures.

In summary, the results of preceding studies for multilingual depression detection are substantial and accomplishing. The several strategies employed in advance studies show the want for language-specific and culturally sensitive techniques in social media-based totally mental health studies. By considering linguistic variety and cultural nuances, researchers can broaden more powerful and inclusive depression detection models, contributing to a deeper records of health on social media.

3 METHODOLOGY

The method section of a thesis on multilingual depression detection calls for a complete rationalization of the information preprocessing strategies, model development, and the experimental layout. This phase pursuits to give an explanation for the steps taken to make certain the take a look at validity and reliability, which include ethical concerns and coping with multilingual information.

3.1 Data Sources and Collection

The basis of a entire study on depression detection in social media is based closely on the variety of the information assets. This segment info the statistics sources and series strategies employed on this look at, with a focus on using Twitter as a number one supply of customer-generated content material material. The preference of Twitter is defined together with the steps taken to gather, machine, and make sure the integrity of the data.

3.1.1 Twitter as a Data Source

Twitter is one of the most broadly used social media systems, with tens of thousands and thousands of clients worldwide posting tweets each day. This giant amount of consumer-generated content material fabric makes Twitter a really perfect supply for studies regarding sentiment evaluation, behavioral dispositions, and exceptional subjects related to mental fitness. The platform's open nature and real-time content material fabric make it specifically useful for collecting facts on a giant range of subjects, which encompass depression.

Twitter's numerous client base includes humans from various linguistic and cultural backgrounds, allowing researchers to observe social media behavior in tremendous contexts. This variety became a crucial element in selecting Twitter as the primary records source for this study on multilingual depression detection.

3.1.2 Dataset Acquisition and Content

The dataset used for this was taken from Kaggle, a famous platform for sharing datasets and wearing out information technology competitions. The Kaggle dataset used for this research comprised a group of tweets that have been translated into multiple languages the usage of Google Translate. This translation step becomes important to create a multilingual dataset appropriate for the examiner's goals.

The dataset consisted of subsets: depressive tweets and non-depressive tweets. Depressive tweets contained content material indicative of depressive symptoms, even as non-depressive tweets represented ordinary Twitter discourse. The categorization of those subsets became based on keywords, linguistic evaluation, and guide annotation to make certain accuracy. This balanced dataset furnished a consultant training surroundings for sentiment analysis models and helped to decrease biases during model improvement.

3.1.3 Data Collection Process

The records collection started out with the extraction of tweets from Kaggle, that specialize in content material that would propose depressive signs. This covered tweets containing keywords generally related to depression, which encompass "sad," "depressed," "hopeless," and exceptional associated phrases. The dataset changed into then augmented with greater tweets to create a balanced illustration of depressive and non-depressive content material.

Ethical troubles finished a massive position in the information collection machine. Since Twitter is a public platform, the gathered records changed into already publicly accessible; however, it changed into vital to protect the privacy and confidentiality of users. Personal information, consisting of usernames and profile snap shots, modified into removed from the dataset to ensure anonymity. Additionally, the take a look at adhered to ethical tips and incredible practices to reduce any potential harm to individuals whose tweets were included in the dataset.

Diversity and Representativeness

One of the challenges in collecting statistics from social media is making sure diversity and representativeness. To deal with this, the dataset was carefully curated to encompass tweets from several linguistic and cultural backgrounds. The use of Google Translate to create a multilingual dataset allowed the observer to study depression detection throughout precise languages, presenting a broader mindset on sentiment and behavior.

The dataset also protected a combination of tweets from first-rate geographic places, time durations, and contexts. This variety turned into critical for developing robust models that would generalize across top notch settings. By collecting a consultant pattern of tweets, the take a look at aimed to capture the several processes in which humans express depressive symptoms and signs on social media.

In summary, the records resources and collection strategies used on this study were designed to provide a comprehensive and various dataset for multilingual depression detection. The attention on Twitter, mixed with the interpretation-based totally approach, allowed the take a look at to explore patron-generated content material fabric from diverse linguistic and cultural backgrounds. The moral issues and efforts to make certain range contributed to the integrity and reliability of the dataset, forming a stable foundation for the subsequent stages of the research.

Collecting statistics for depression detection on social media requires a cautious and methodical technique, especially whilst handling sensitive subjects and personal records. The technique need to ensure that the facts is correct, representative, and amassed ethically. This phase describes the data series method used inside the study, highlighting the strategies for categorizing tweets, the measures taken to make certain accuracy, and the stairs followed to cope with moral concerns.

Categorization of Tweets

The records collection method commenced with the extraction of tweets from Twitter, specializing in content that might be indicative of depressive symptoms. To gain this, a keyword-primarily based totally method changed into used to perceive tweets that contained phrases usually related to depression. These key terms blanketed phrases and phrases like "depressed," "unhappy," "hopeless," "traumatic," and others that could advocate emotional distress.

The extracted tweets have been then classified into subsets: depressive tweets and non-depressive tweets. Depressive tweets contained content that recommended depressive symptoms, even as non-depressive tweets represented general Twitter discourse without any clear symptoms of depression. This categorization turned into vital for creating a balanced dataset that is probably used to train and check the depression detection models.

The facts collection approach became designed to apprehend clients' privateness and avoid any accidental consequences. Although the statistics become publicly available, the study took steps to make certain that it changed into used responsibly and in a manner everyday with ethical studies practices. By following those ethical pointers, the study aimed to keep the consider of the general public and the integrity of the studies.

In summary, the data collection process concerned a systematic approach to extracting and categorizing tweets for depression detection, with a sturdy emphasis on moral worries and accuracy. This thorough system laid the idea for the subsequent levels of the study, bearing in mind a sturdy evaluation of depression detection in multilingual contexts.

3.2 Data Preprocessing

Data preprocessing is a vital step in making ready social media datasets for analysis, mainly in studies focused on depression detection. It includes a sequence of responsibilities to easy, normalize, and rework uncooked facts proper into a format suitable for ML and Natural language processing (NLP). Proper facts preprocessing permits lessen noise, improve facts consistency, and ensure that the models used for analysis collect extremely good input. This section describes the important thing steps in information preprocessing for this study, together with textual content cleaning, normalization, tokenization, and further information visualization strategies.

3.2.1 Text Cleaning and Normalization

Text cleaning is step one in data preprocessing, designed to dispose of inappropriate and noisy elements from the dataset. This technique concerned:

Removing Special Characters and Symbols: Tweets frequently include precise characters, symbols, and different non-textual factors that don't contribute to the evaluation. This step involved removing characters like punctuation marks, emojis, and other non-alphanumeric symbols to make sure the text become smooth and prepared for similarly processing.

Removing URLs and Hashtags: Social media content material often includes URLs, hashtags, and other external references. These factors can add noise to the dataset and might not be applicable to the assessment of depressive signs and symptoms. The preprocessing step eliminated all URLs and hashtags, focusing completely on the text of the tweets.

Text Normalization: Text normalization is essential for maintaining consistency at some point of the dataset. This involved converting all textual content to lowercase to make sure uniformity, no matter how customers typed their tweets. By standardizing the text, the look at decreased ability variability and ensured a regular instance of the records.

3.2.2 Tokenization and Word Analysis

Tokenization is the system of breaking down text into person tokens, usually phrases or phrases. This step is critical for growing a vocabulary that can be used to train and study machine analyzing models. The following tokenization and word evaluation techniques had been employed:

Tokenization with Keras Tokenizer: The Keras Tokenizer have become used to break down the wiped clean textual content into person tokens. This step involved parsing every tweet and splitting it into wonderful phrases or terms, developing a mounted illustration of the information. Tokenization is a foundational step in NLP, permitting in addition assessment and modeling.

Word Frequency Analysis: After tokenization, an extensive analysis of word frequencies have become finished to recognize the composition of the dataset. This analysis involved counting the occurrences of each token and identifying the maximum common phrases in both depressive and non-depressive tweets. By analyzing word frequencies, the study gained insights into the linguistic trends of the dataset, helping to discover styles associated with depressive symptoms.

3.2.3 Word Cloud Visualization

Word cloud visualization is a qualitative method used to discover textual statistics and find out not unusual phrases and phrases. This step became covered in the preprocessing approach to provide a visual illustration of the most regularly used words in depressive tweets. The WordCloud software software become used to create word clouds, which visually displayed the maximum not unusual terms inside the dataset.

The phrase cloud visualization served several capabilities:

Identifying Significant Phrases: The word cloud highlighted the phrases and phrases maximum often related to depressive tweets, providing insights into common issue topics and language patterns. This seen illustration became valuable for guidance further assessment and model improvement.

Enhancing Understanding of the Dataset: The word cloud supplied a qualitative knowledge of the dataset, complementing the quantitative evaluation conducted through tokenization and word frequency assessment. It helped researchers pick out key phrases that could advocate depressive signs and symptoms and guided the development of more powerful feature extraction strategies.

3.3 Model Development

The development of strong and dependable models is a key issue of depression detection in social media. Given the complicated nature of text statistics and the want to capture nuanced styles that mean depressive signs and symptoms and symptoms, the choice of model architectures and training strategies is important. This phase describes the improvement of the models used in this study, specializing in the use of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures, and explains why these deep gaining knowledge of models had been selected for depression detection.

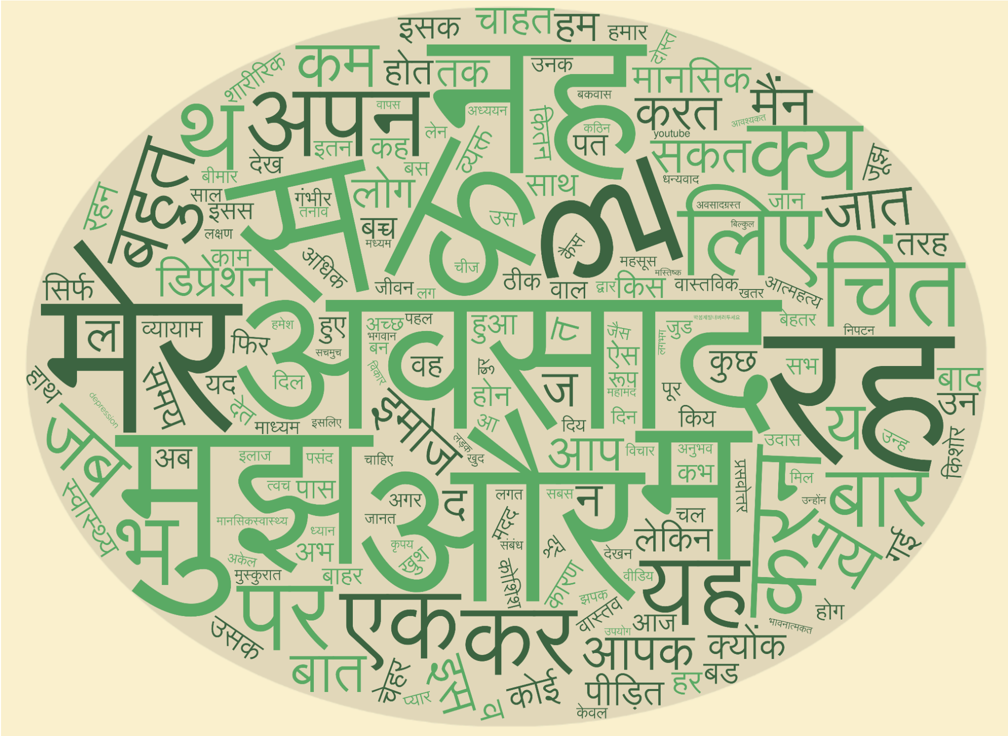

Figure 3.1: Wordcloud Hindi

Figure 3.2: Wordcloud Gujarati

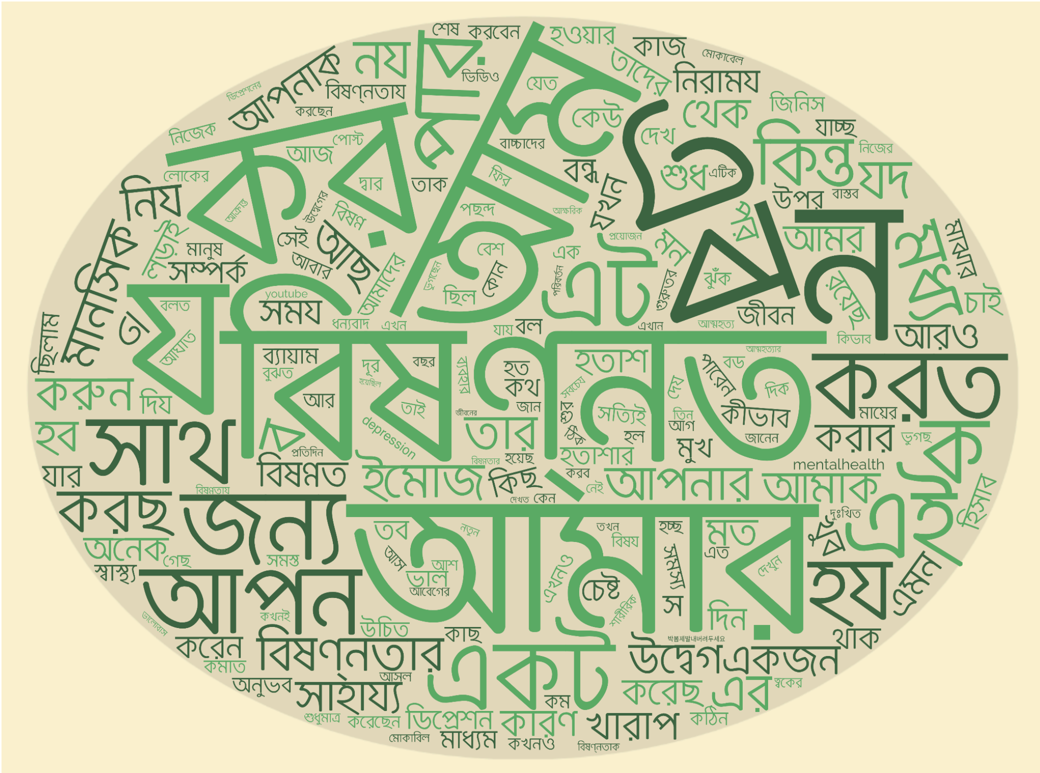

Figure 3.3: Wordcloud Bengali

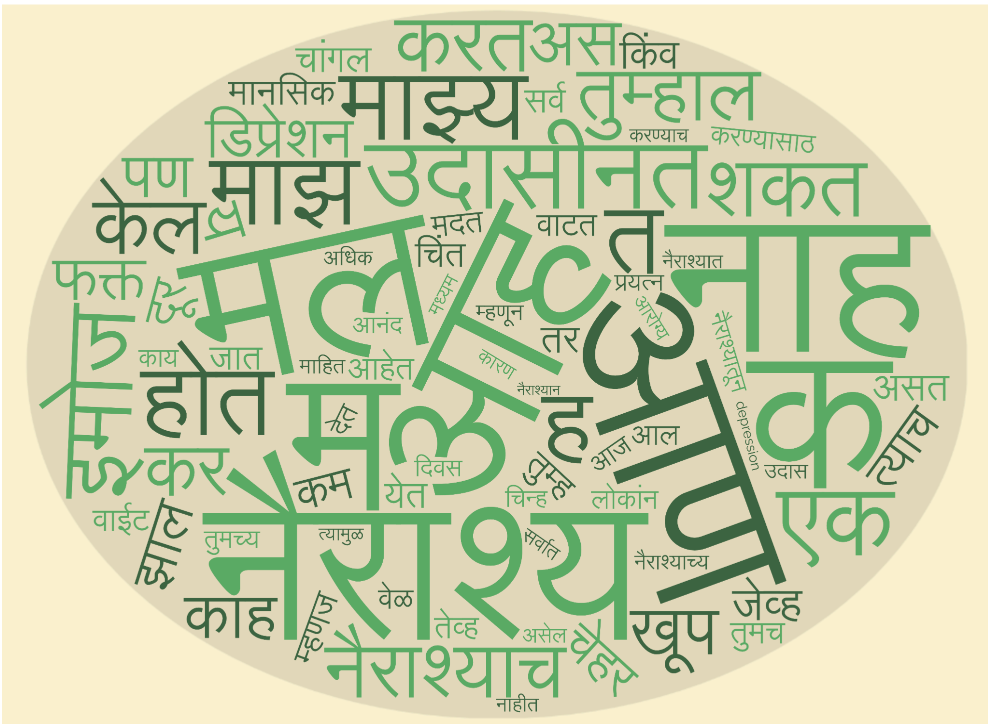

Figure 3.4: Wordcloud Marathi

Figure 3.5: Wordcloud Telugu

Figure 3.6: Wordcloud Tamil

Figure 3.7: Wordcloud Malayalam

Figure 3.8: Wordcloud Kannada

3.3.1 Why LSTM and GRU?

LSTM and GRU are types of Recurrent Neural Networks (RNNs) designed to process sequential records. Unlike conventional neural networks, RNNs have internal loops that permit them to keep statistics in the course of time steps, making them perfect for studying textual content data, wherein context and order matter a wide variety. The sequential nature of social media posts, which incorporates tweets, makes LSTM and GRU nicely-proper for detecting patterns which can suggest depressive symptoms.

LSTM: Long Short-Term Memory (LSTM) is a kind of RNN that consists of memory cells capable of storing information for prolonged periods. This characteristic allows LSTM to capture prolonged-range dependencies in sequential information, allowing it to apprehend

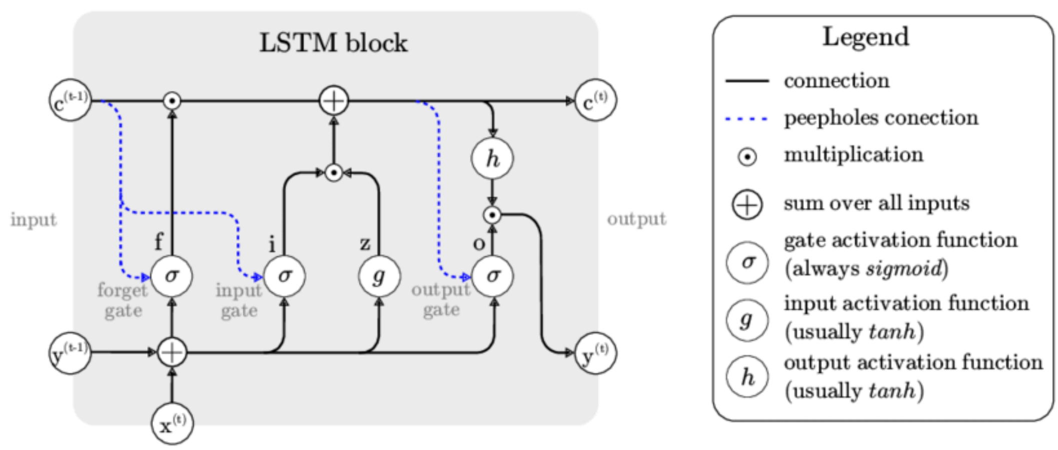

Figure 3.1: LSTM Architecture [21]

styles that unfold through the years. LSTM is in particular beneficial for analyzing social media information, where the context and flow of textual content play a extensive function in knowledge sentiment.

GRU: Gated Recurrent Unit (GRU) is a variation of RNN that simplifies the shape by way of combining the neglect and enter gates into a single replace gate. GRU is computationally efficient and may keep contextual records without requiring as a first-rate deal memory as LSTM. This performance makes GRU a robust choice for depression detection, mainly whilst managing huge datasets or whilst computational belongings are restrained.

Model Architecture and Components:

The improvement of the LSTM and GRU models means constructing specific architectures tailored to the challenge of depression detection. The key additives of these architectures blanketed:

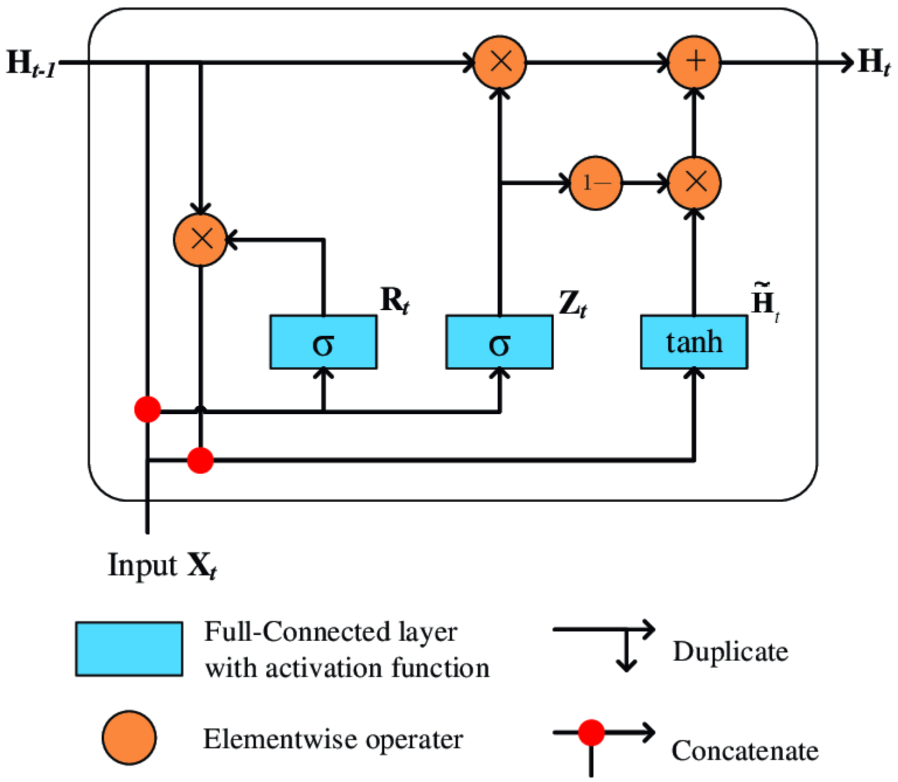

Figure 3.2: GRU Architecture [22]

Embedding Layers: To convert words into numerical representations, word embedding layers were used. These layers transformed textual content information into dense vectors, allowing the models to recognize the semantic relationships among phrases. pre-trained FastText embeddings have been hired to decorate the instance of enter textual content, as FastText captures subword data and is appropriate for multilingual contexts.

Recurrent Layers (LSTM/GRU): The center of the model shape consisted of 1 or extra recurrent layers, the use of both LSTM or GRU units. These layers processed the sequential textual content statistics and recognized patterns that might indicate depressive symptoms. The recurrent layers have been designed to cope with variable-length sequences, permitting the models to research tweets of various lengths.

Dense Output Layer: The final layer of the models have become a dense output layer, liable for making predictions based totally on the abilities extracted by means of manner of the recurrent layers. This output layer used an activation function, which include sigmoid or softmax, to categorise tweets as depressive or non-depressive. The dense layer's output was used to decide the probability of a tweet indicating depressive signs and signs.

3.3.2 Training and Optimization Strategies

The training technique for the LSTM and GRU models concerned the usage of the preprocessed dataset to train the models to understand depressive styles. The following training and optimization techniques had been implemented:

Early Stopping: To prevent overfitting and make sure efficient training, early stopping mechanisms had been used. This approach consists of tracking a validation metric (consisting of loss or accuracy) in the route of training and stopping the training manner whilst the metric stops improving for a awesome huge style of epochs. Early preventing lets in avoid overfitting via stopping the models from becoming too specialised to the training statistics.

Dropout and Regularization: To decorate model generalization, dropout and regularization techniques have been applied. Dropout entails randomly setting a fragment of the neurons to zero throughout training, lowering overfitting with the resource of preventing the model from becoming overly depending on particular neurons. Regularization, which incorporates L2 regularization, provides a penalty to the loss characteristic to deter huge weights, promoting more strong and generalizable models.

Optimization Algorithms: The models were professional the usage of optimization algorithms like Adam or RMSprop, recognised for their adaptive mastering costs and efficiency in training deep neural networks. These algorithms allowed for smooth convergence and powerful training, making sure that the models carried out excessive overall performance without immoderate training time.

3.3.3 Ensemble Approaches

In addition to man or woman LSTM and GRU models, ensemble techniques were explored to decorate depression detection accuracy. An ensemble technique involves combining the predictions of more than one models to create a higher and accurate final prediction. This method can decorate performance via manner of leveraging the strengths of diverse models at the same time as lowering the impact of individual model weaknesses.

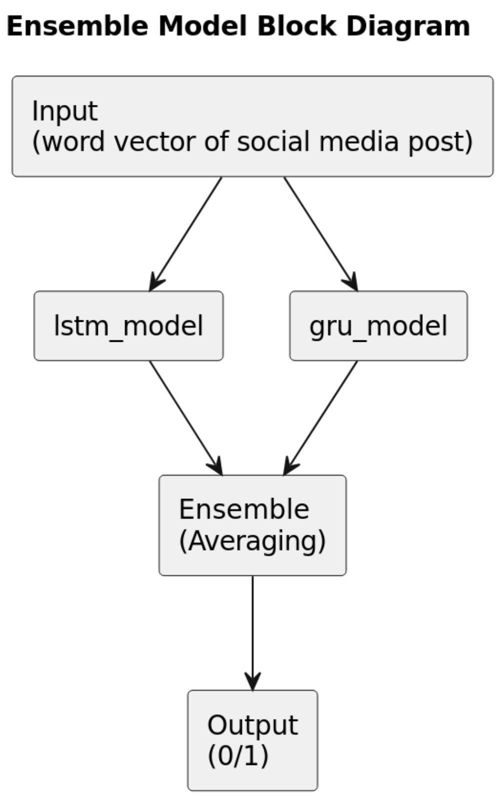

Figure 3.3: Ensemble Architecture

In summary, the development of models for depression detection concerned the careful preference and construction of LSTM and GRU architectures, with a focus on their ability to technique sequential statistics and capture complicated patterns in textual content. The training and optimization strategies implemented within the course of model development ensured that the models were strong, dependable, and able to as it need to be detecting depressive symptoms in social media information. The use of ensemble approaches in addition contributed to improved performance, demonstrating the effectiveness of mixing multiple models for greater accuracy and generalization.

3.4 Word Embedding with Pre-trained FastText Vectors

Word embedding is a pivotal step within the system of growing machine studying models for natural language processing (NLP), in particular at the same time as managing textual content data in several languages. It involves transforming phrases into dense vectors that seize semantic statistics, allowing models to apprehend relationships among phrases and their contextual meanings. This segment delves into the function of phrase embedding in this study and explains why pre-trained FastText vectors were selected to represent terms, with a focal point on their benefits in multilingual contexts.

What Is Word Embedding?

Word embedding is a manner used to transform textual content facts into numerical representations. Unlike traditional techniques that depend upon one-heat encoding, phrase embedding produces dense vectors that include semantic statistics about phrases. These dense vectors permit device getting to know models to seize relationships among phrases, apprehend context, and apprehend patterns in textual statistics.

The concept of word embedding is primarily based on the distributional hypothesis, which suggests that phrases occurring in comparable contexts typically tend to have comparable meanings. By changing phrases into dense vectors, word embedding creates a example that displays those semantic similarities, allowing models to make experience of the text in a greater sophisticated manner.

Why FastText?

FastText is a famous word embedding technique superior by Facebook's AI Research (FAIR) group. Unlike other embedding techniques like Word2Vec or GloVe, FastText represents terms as collections of subword devices, allowing it to recognize words at a granular stage. This method has several key blessings, especially in multilingual contexts:

Subword-Level Representation: FastText breaks down words into subword machines, shooting semantic records even for terms which are unusual or unseen in the training records. This function is specially beneficial in languages with complex morphology or in which terms are derived from smaller linguistic additives.

Robustness to Out-of-Vocabulary Words: Because FastText represents phrases as subwords, it could cope with out-of-vocabulary (OOV) terms more successfully. This is crucial in social media assessment, wherein users regularly use slang, abbreviations, or misspellings.

Suitability for Multilingual Data: FastText is well-perfect for multilingual contexts, as it can capture linguistic subtleties during one in all a kind languages. This is specifically vital in research like this, wherein the dataset consists of tweets in a couple of languages, every with particular vocabulary and structure.

FastText in Model Development

In this study, pre-trained FastText vectors have been used to represent words throughout model improvement. The choice of FastText become pushed through its capability to seize semantic connections and linguistic subtleties, offering a miles better illustration of the entered text. The pre-trained FastText embeddings had been obtained from models professional on massive multilingual corpora, ensuring a whole vocabulary that covered several languages and dialects.

The use of pre-professional embeddings offered numerous advantages in model development:

Improved Semantic Representation: FastText embeddings captured semantic relationships among phrases, permitting the models to recognize contextual records and perceive patterns related to depressive signs and symptoms. This advanced example contributed to improved model performance and accuracy.

Reduction in Training Time: By using pre-trained embeddings, the models started out with a stable baseline illustration of phrases, lowering the need for big training. This efficiency allowed the look at recognition of other factors of model improvement, which incorporates structure format and optimization techniques.

Flexibility and Adaptability: FastText's subword-diploma illustration provided flexibility, permitting the models to comply to models in word forms and cope with new or unusual terms. This flexibility become vital in a social media context, in which language use is dynamic and constantly evolving.

Implications for Multilingual Depression Detection

The use of pre-professional FastText vectors had extensive implications for multilingual depression detection. By supplying a sturdy illustration of phrases in a couple of languages, FastText helped bridge the distance amongst high-quality linguistic contexts. This capability end up crucial for developing models that could generalize across languages and because it should be discover depressive symptoms and signs in a multilingual dataset.

FastText's subword-degree instance also allowed the models to seize nuances and diffused variations in language use, contributing to a more complete analysis of social media content. This nuanced know-how of textual content facts became critical for growing effective depression detection models, especially in a culturally and linguistically numerous context.

In particular, phrase embedding with pre-knowledgeable FastText vectors done a vital role in model development for this study. The use of FastText supplied a robust semantic illustration, superior model standard performance, and facilitated the analysis of multilingual textual content facts. These blessings have been instrumental in developing models for depression detection that could function efficiently in a several and multilingual environment, ultimately contributing to a greater accurate and inclusive method to social media-based totally definitely mental fitness studies.

3.5 Training and Model Architecture

Training deep reading models for depression detection on social media calls for careful attention of the architecture, records go with the flow, and optimization strategies. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models are specifically suitable to this challenge due to their potential to approach sequential facts and seize complex styles. This segment info the shape of the LSTM and GRU models used within the observe, together with the training method, which encompass key strategies and strategies hired to ensure powerful learning.

3.5.1 Model Architecture

The structure of LSTM and GRU models typically consists of numerous layers, each serving a selected feature. In this examine, the models have been designed with an embedding layer, recurrent layers (using LSTM or GRU), and a dense output layer.

Embedding Layer: The embedding layer is the number one factor of the model shape. It converts terms into dense vectors, providing a numerical illustration for the subsequent layers. In this study, pre-trained FastText embeddings were used, which give a sturdy illustration of words, shooting semantic connections and linguistic subtleties. This layer plays a important function in permitting the model to apprehend the contextual relationships between phrases.

Recurrent Layers: The recurrent layers shape the middle of the LSTM and GRU models. These layers system the sequential textual content statistics, permitting the models to capture dependencies and contextual facts. The LSTM shape consists of a forget about about gate, enter gate, and output gate, providing an advanced mechanism for retaining records over the years. GRU, a variation of LSTM, combines some of those gates, making it computationally efficient at the same time as retaining similar talents.

Dense Output Layer: The dense output layer is the final element of the structure, liable for producing the model's predictions. It generally uses an activation feature, which include sigmoid or softmax, to generate opportunities that imply whether or not a tweet is depressive or non-depressive. This layer performs a key position in sentiment assessment, because it determines the model's output based at the capabilities extracted with the resource of the recurrent layers.

Training Process

The training process entails feeding the preprocessed records into the models to allow them to study styles and traits associated with depressive signs and symptoms. The following key elements were a part of the training way:

Training Data: The training information consisted of the preprocessed dataset, consisting of depressive and non-depressive tweets. The dataset became divided into training and validation devices, allowing the models to analyze from one set even as being evaluated on the alternative. This division facilitates display overfitting and guarantees the models can generalize beyond the training information.

Loss Function and Optimization: The loss function measures the difference a number of the model's predictions and the real consequences, guiding the model's gaining knowledge of approach. In this study, a binary skip-entropy loss feature have become used, as it's far commonly employed for binary magnificence duties like depression detection. Optimization algorithms, along with Adam or RMSprop, had been used to modify the model's weights and biases, minimizing the loss function over successive iterations (epochs).

Early Stopping: To prevent overfitting, early preventing mechanisms were carried out. This method entails tracking a validation metric, along side loss or accuracy, at some stage in training and preventing the training manner even as the metric stops enhancing for a precise quantity of epochs. Early preventing allows keep away from overfitting with the aid of preventing the models from becoming too specialized in the training facts.

Regularization and Dropout: Regularization strategies, together with dropout and L2 regularization, were used to decorate the models' generalization functionality. Dropout includes randomly placing a fragment of the neurons to zero throughout training, reducing the hazard of overfitting. L2 regularization adds a penalty for big weights, encouraging the models to broaden more strong and generalizable answers.

Handling Imbalanced Data

In depression detection, imbalanced data is a common trouble, with depressive tweets regularly being less common than non-depressive ones. This imbalance can cause biased models that carry out properly on the majority class however poorly at the minority elegance. To address this mission, strategies collectively with augmentation and oversampling have been used to create an extra balanced dataset. These strategies ensured that the models had enough exposure to depressive tweets, letting them discover ways to hit upon depressive styles because it should be.

Evaluation Metrics

The training method included a robust evaluation device to check the models' standard performance. Common metrics together with accuracy, precision, recollect, and F1-rating were used to assess how properly the models ought to differentiate amongst depressive and non-depressive tweets. These metrics provided a comprehensive view of the models' abilities and allowed for awesome-tuning at a few degree in the training system.

In summary, the training and model shape for this take a look at have been cautiously designed to permit powerful depression detection on social media. The LSTM and GRU architectures, with their embedding layers, recurrent layers, and dense output layer, were nicely-appropriate to the project of processing sequential text data. The training machine, with its use of early preventing, regularization, and dropout, ensured that the models have been sturdy and generalizable. The dealing with imbalanced information and the use of appropriate assessment metrics contributed to the models' accuracy and reliability, supplying a robust basis for powerful sentiment evaluation and depression detection in a multilingual context.

Early Stopping and Optimization Strategies

Deep learning knowledge of models, like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), are effective equipment for reading sequential facts, however they'll be also at risk of overfitting, in particular when working with complicated datasets like the ones derived from social media. Overfitting takes place whilst a model learns the training statistics too well, taking pictures noise along side the signal, which reduces its capability to generalize to new, unseen records. To address this, early stopping and several optimization strategies are implemented during model training. This section discusses these strategies and their role in enhancing the generalization potential of LSTM and GRU models.

Early Stopping

Early stopping is a manner used to keep away from overfitting through the usage of tracking a validation metric at some stage in model training and halting the training manner at the same time as the metric not improves. This method facilitates ensure that the model does no longer over-research the training statistics, maintaining its ability to perform properly on unseen data.

How Early Stopping Works

Early preventing consists of splitting the dataset into training and validation machines. While the training set is used to replace the model's weights and biases, the validation set is used to reveal the model's overall performance with out updating the model itself. A commonplace metric, including validation loss or validation accuracy, is observed for the length of each epoch of training. If the validation metric stops enhancing for a certain huge type of epochs, the training approach is stopped, indicating that similarly training may additionally moreover motive overfitting.

Benefits of Early Stopping

Prevents Overfitting: By stopping training whilst the validation metric now not improves, early stopping reduces the threat of overfitting. This guarantees that the model stays generalizable and performs properly on new information.

Improves Training Efficiency: Early preventing can shorten the training machine by stopping useless epochs after the model has reached its maximum proper overall performance. This can store time and computational resources.

Enhances Generalization: Early preventing allows preserve a stability amongst model accuracy and generalization, allowing the model to carry out nicely across distinct datasets.

Dropout

Dropout is an optimization approach that reduces overfitting by means of randomly setting a fraction of the neurons to zero all through training. This approach prevents the model from becoming overly depending on precise neurons, selling a much better and generalizable model.

How Dropout Works

Dropout is usually finished to the dense layers in a neural community, wherein a percent of the neurons are randomly deactivated inside the course of training. This random deactivation forces the model to rely upon distinct subsets of neurons for every training generation, reducing the chance of overfitting.

Benefits of Dropout

Reduces Overfitting: By introducing randomness in neuron activation, dropout reduces the model's dependency on unique neurons, minimizing the hazard of overfitting.

Promotes Robustness: Dropout encourages the model to use a broader set of capabilities, primary to a far higher and strong model.

Improves Generalization: With dropout, the model is much less probably to memorize the training statistics, allowing it to generalize higher to new statistics.

Regularization

Regularization is any other optimization technique that gives a penalty to the loss function to deter big weights, making strong and generalizable models. L2 regularization, additionally referred to as Ridge regularization, is a commonplace technique applied in deep learning knowledge of.

How Regularization Works

L2 regularization includes a penalty to the loss characteristic primarily proportional to the square of the model's weights. This penalty discourages massive weights, decreasing the danger of overfitting. By keeping the weights in test, regularization promotes smoother and extra strong models.

Benefits of Regularization

Prevents Large Weights: Regularization discourages big weights, that can result in overfitting, making sure that the model stays robust.

Enhances Generalization: By penalizing big weights, regularization encourages the model to generalize better to unseen facts.

Improves Training Stability: Regularization can make contributions to more robust training, reducing the threat of erratic model behavior.

Combining Strategies for Improved Performance
In exercising, early stopping, dropout, and regularization are regularly combined to create a strong training approach. Early stopping guarantees the model does now not overfit by using tracking validation metrics. Dropout introduces randomness to save you overfitting at the neuron degree, whilst regularization discourages big weights, promoting model stability. Together, the ones strategies create a balanced technique to training deep learning knowledge of models, improving generalization and ensuring maximum effective general performance.

In summary, early prevention and various optimization techniques play a critical position in growing LSTM and GRU models for depression detection. These strategies help save you from overfitting, enhance training performance, and beautify the models' ability to generalize to new data. By imposing those techniques, the study ensured that the ensuing models have been sturdy, solid, and capable because it ought to be detecting depressive patterns in social media records.

3.6 Experimental Design

The experimental layout in the study focused on depression detection on social media entails a cautiously installed method to make certain that the effects are each reliable and legitimate. This section offers an improved assessment of the key additives of experimental layout, collectively with dataset configuration, dealing with imbalanced information, and the selection of assessment metrics. These elements make contributions to the study's robustness and allow for an entire evaluation of the models' overall performance.

3.6.1 Dataset Configuration

The dataset used for training and testing models in this study turned into derived from Twitter, containing tweets which have been classified into depressive and non-depressive subsets. The dataset changed into configured as follows:
Training and Testing Subsets: The dataset became broken up into training and testing subsets. The training subset, typically round 80% of the dataset, became used to train the models, permitting them to research the styles and trends that differentiate

depressive from non-depressive tweets. The final 20% formed the testing subset, used to evaluate the models' performance and investigate their generalization to new records.

3.6.2 Handling Imbalanced Data

Imbalanced dataset is a not unusual venture in depression detection, as depressive tweets are frequently a great deal much less frequent than non-depressive tweets. To address this trouble, numerous techniques had been employed to create a balanced dataset:
Data Augmentation: Data augmentation entails generating extra samples to increase the illustration of the minority. In this observation, data augmentation techniques which consist of text manipulation, paraphrasing, or including noise had been used to create greater depressive tweets. This approach helped increase the training records's range and reduced the chance of overfitting to a small subset of depressive tweets.

Oversampling: Oversampling consists of duplicating samples from the minority to stabilise the dataset. This approach applied to the training set to make sure that depressive tweets have been correctly represented at some point of model training. Oversampling helps create a extra balanced training environment, allowing the models to analyze from a enough type of depressive tweets.

Synthetic Data Generation: Synthetic information generation, the usage of techniques like Synthetic Minority Over-sampling Technique (SMOTE), creates synthetic samples for the minority elegance. This approach generates new information factors primarily based on present ones, further balancing the dataset and improving the models' ability to generalize.

Evaluation Metrics
To study the models' performance, a comprehensive set of assessment metrics were brought into use. These metrics supplied an intensive evaluation of ways the models should locate depressive tweets and differentiate them from non-depressive ones. The following metrics had been hired:

Accuracy: Accuracy measures the percentage of accurate predictions made with the model, indicating the overall effectiveness of the model. While useful, accuracy may be deceptive when handling imbalanced information, as it can no longer replicate the overall performance at the minority magnificence.

Precision: Precision measures the share of actual positives amongst all positive predictions, indicating how some of the tweets identified as depressive had been simply depressive. Precision is vital whilst the price of fake positives is immoderate.
Recall: Recall measures the share of right positives among all real positives, indicating how many depressive tweets were efficiently diagnosed via the model. Recall is vital while the fee of false negatives is high, as within the case of depression detection.

F1-Score: The F1-rating combines precision and recall metric, offering a stability between the two. It is beneficial for assessing model performance in instances wherein there's an imbalance between commands or whilst precision and don't forget are equally crucial.

In summary, the experimental layout for this study provides a complete technique to dataset configuration, handling imbalanced records, and the choice of evaluation metrics. Techniques like oversampling have been used to deal with imbalances, at the same time as assessment metrics like accuracy, precision, recall, and F1-rating supplied an in depth evaluation of the models' performance.

4 EXPERIMENTAL SETUP AND RESULTS

Experimental setup plays an important function in obtaining reliable consequences in an study that uses machine-learning knowledge to stumble on depression on social media. This section provides a whole review of the experiments accomplished to check the technique, collectively with the configurations, hyperparameters, and data-validation techniques used. The consequences of these experiments are then furnished and analyzed to assess the general overall performance of the advanced models.

4.1 Experiments Conducted

The experiments done for this have been designed to assess the effectiveness of Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and ensemble models in detecting depressive styles in social media posts. The experimental setup grew to be carefully deliberate to make sure that the effects had been dependable, legitimate, and generalizable. This section gives a comprehensive description of the experiments, collectively with the shape of the training and testing datasets, model configurations, move-validation techniques, and optimization strategies.

Training and Testing Datasets

The first step inside the experimental setup came to be to put together the datasets for training and sorting out. The dataset used for this take a look at consisted of tweets from Twitter, categorized into depressive and non-depressive subsets. The branch of the dataset into training and checking out subsets allowed for a managed surroundings wherein to train and study the models.
Training Subset: The training subset comprised approximately 80% of the entire dataset. This subset became used to train the models, letting them look at the styles and features related to depressive tweets. To ensure a balanced instance of depressive and non-depressive tweets, stratified sampling became employed, in which the statistics emerge as divided based on the elegance labels and sampled proportionally.

Testing Subset: The last 20% of the dataset customary the checking out subset. This subset changed into used to assess the models' common performance after training. By testing on a separate dataset, the experiments ought to verify the models' functionality to generalize to new statistics, supplying a greater accurate diploma in their effectiveness.

Model Configurations

The LSTM and GRU models used in this look at had been built with specific configurations designed for depression detection. These configurations were based totally on first-rate practices for deep learning models in natural language processing (NLP), with a focus on taking pix sequential patterns in text.

Embedding Layer: The embedding layer become the preliminary layer within the model, answerable for changing words into dense vectors. Pre-trained FastText vectors had been used for word embedding, offering a strong semantic illustration of the entered text. FastText's subword-primarily based complete illustration helped capture the nuances and contextual relationships amongst phrases, improving the models' capability to find out depressive styles.
Recurrent Layers: The recurrent layers common the center of the LSTM and GRU models. These layers processed the sequential text facts, allowing the models to seize dependencies and context. The variety of recurrent layers and the extensive sort of machines constant with layer were optimized to ensure a balance amongst model complexity and general overall performance.

Dense Output Layer: The dense output layer was the final challenge of the model, accountable for generating the kind output. This layer used an activation function, which includes sigmoid or softmax, to generate probabilities that indicated whether or not or not a tweet became depressive or non-depressive.
Hyperparameters in conjunction with analyzing cost, batch length, and the huge number of layers had been tuned to optimize the models' performance. This tuning

technique involved experimentation with all kinds of configurations to find the best setup for each model.

Early Stopping: Early stopping became used to screen validation loss at some stage in training. If the validation loss stopped improving for a positive variety of epochs, the training way come to be halted. This method helped save you overfitting by way of preventing training earlier than the models have become overly specialized within the training data.

Dropout: Dropout is an optimization method that includes randomly putting a fragment of the neurons to 0 inside the path of training. This approach reduces overfitting with the aid of manner of preventing the models from turning into overly depending on particular neurons. Dropout was carried out to the recurrent and dense layers to enhance model robustness.

Regularization: Regularization, collectively with L2 regularization, is used to feature a penalty to the loss function based at the square of the model's weights. This penalty discouraged large weights, making extra stable and generalizable models.
These optimization techniques performed a crucial position in making sure that the models could generalize to new statistics and perform well at the checking out subset. They additionally contributed to reducing training time and enhancing the general reliability of the experimental outcomes.

In particular the experimental setup and outcomes provided a complete assessment of the method and the overall performance of the developed models. The experiments had been designed to test the effectiveness of LSTM, GRU, and ensemble models in detecting depressive patterns in social media posts. The setup included a balanced division of training and finding out datasets, cautiously constructed model configurations, move-validation techniques to ensure robustness, and optimization strategies to save you from overfitting. These additives contributed to a nicely-rounded experimental design that yielded dependable and legitimate outcomes for the study's goals.

4.2 Results and Analysis

The consequences of the experiments for detecting depressive patterns in social media posts provide widespread insights into the effectiveness of the models. This segment delves into the precise evaluation of these effects, exploring normal overall performance metrics, comparisons among models, and the overall implications for multilingual depression detection.

Performance Metrics

To verify the general performance of the models, numerous metrics have been used to measure their accuracy, precision, remember, and F1-rating. These metrics offer a comprehensive assessment of the way the models can discover depressive tweets and distinguish them from non-depressive ones.

Language	Accuracy	Precision	Recall	F1 Score
Gujarati	0.91669	0.91058	0.89615	0.9028
Kannada	0.92232	0.92664	0.8936	0.90753
Hindi	0.92513	0.92843	0.89936	0.91181
Bengali	0.91922	0.91842	0.89437	0.90492
Telugu	0.91866	0.93089	0.88704	0.90416
Marathi	0.91247	0.91511	0.8831	0.89649
Malayalam	0.91388	0.907	0.89309	0.89952
Tamil	0.91416	0.91904	0.88777	0.90065

Table 4.1: Our Results

Accuracy: This metric shows the percentage of correct predictions made by the model, presenting a preferred view of its typical overall performance. High accuracy suggests that the model is effective in differentiating depressive from non-depressive tweets.

Precision: Precision measures the share of actual positives among all detected positive predictions. High precision suggests that the model has a low charge of false positives, it really is important in averting misclassification of non-depressive tweets as depressive.

Recall: Recall measures the proportion of real positives amongst all real positives. An immoderate recall shows that the model effectively identifies depressive tweets, reducing the threat of false negatives.

F1-Score: The F1-score is the harmonic mean of precision and recall, presenting a balanced view of the model's universal performance. It is mainly beneficial in cases where there might be a change-off between precision and recall.

Analysis of Results

The outcomes from the experiments reveal that the models carried out good overall performance in the course of special Indian languages. The following table 4.1 summarizes the important standard overall performance metrics for every language.

These consequences indicate that the models completed continually properly for the Indian languages. The high accuracy, precision, recall, and F1-score suggest that the model structure and training techniques have been powerful in capturing the characteristics of depressive tweets. This degree of general overall performance demonstrates the robustness of the models and their capability for realistic packages in depression detection on social media.

4.3 Comparisons Between Models

To apprehend the strengths and weaknesses of various models, comparisons had been made among Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and ensemble models. These comparisons furnished insights into the relative overall performance of every technique:

LSTM vs. GRU: The LSTM models typically showed better accuracy and recall, at the same time as the GRU models exhibited faster training times and computational performance. This distinction is predicted, as LSTM has an inclination to hold more contextual facts due to its complicated structure, whilst GRU is designed for overall performance.

Ensemble Models: The ensemble technique, which mixes predictions from multiple models, verified advanced overall performance in some cases. By leveraging the strengths of diverse models, ensemble techniques can reduce the effect of model weaknesses and enhance widespread accuracy and robustness.

Paper	Language	Model	Accuracy
[12]	Bengali	BiGRU	0.86
[13]	Bengali	GRU	0.757
[14]	Bengali	BiGRU	90.3
[15]	Gujarati	Linear SVC	-
[16]	Hindi	Linear SVC	-
[16]	Hindi	CNN-BiLSTM	0.8321
[17]	Hindi	BERT	0.7143
[18]	Malayalam	GRU	0.7633
[19]	Tamil	BERT	-
[20]	Malayalam	BERT	-
[20]	Malayalam	LSTM	0.8428
[20]	Tamil	LSTM	0.5293
[20]	Kannada	LSTM	0.5745

Table 4.2: Comparative analysis (Summary of papers)

These comparisons advocate that while LSTM and GRU models are each effective, the selection among them can also rely upon the specific necessities of having a look at, together with computational property and training time. Ensemble models can offer extra advantages in terms of accuracy and robustness, mainly in complex duties like depression detection.

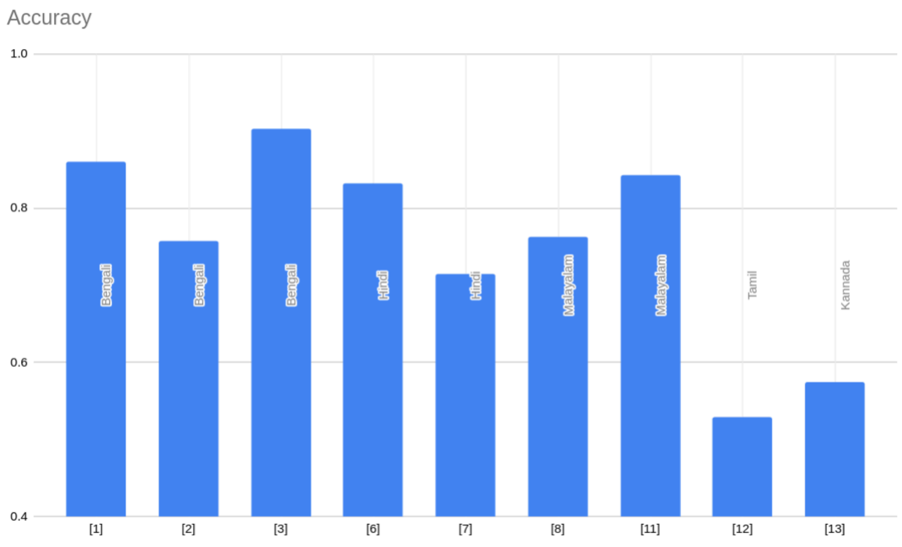

Figure 4.1: Comparison of various papers: Accuracy

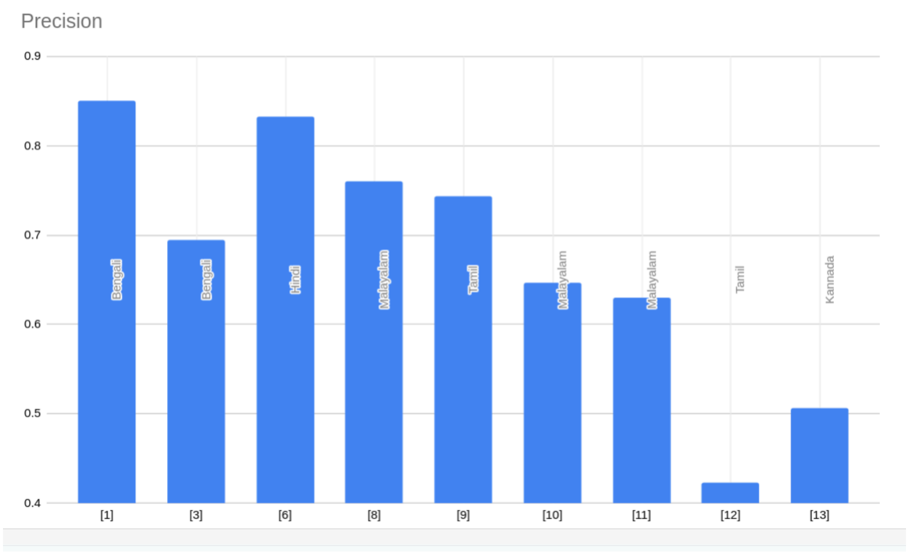

Figure 4.2: Comparison of various papers: Precision

Language-Specific Variations

The consequences from the experiments highlighted that whilst the models usually executed properly throughout distinct Indian languages, there had been variations in precision and recall. These variations suggest that a few models may also battle with unique linguistic nuances, indicating the need for tailor-made strategies in some instances.

Variations in Precision and Recall: The variations in precision and recall throughout languages can be attributed to differences in language structure, vocabulary, and cultural context. For example, a few languages can also use more idiomatic expressions, even as others may also have precise grammatical patterns that influence the models' ability to come across depressive patterns.

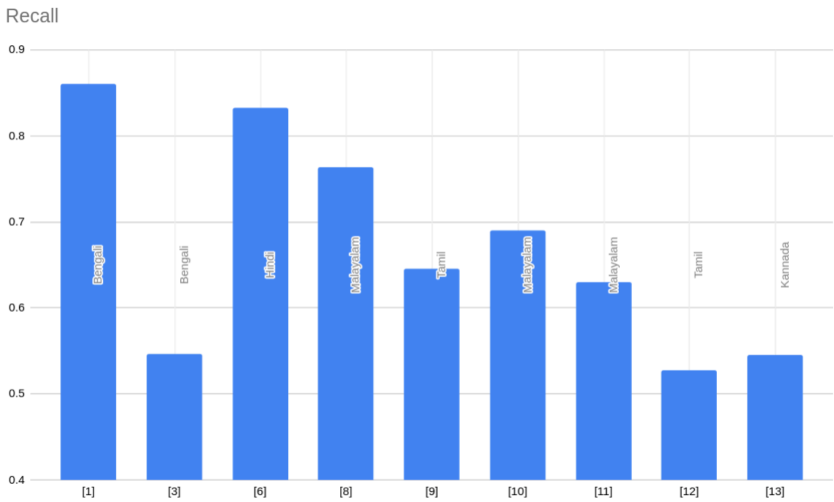

Figure 4.3: Comparison of various papers: Recall

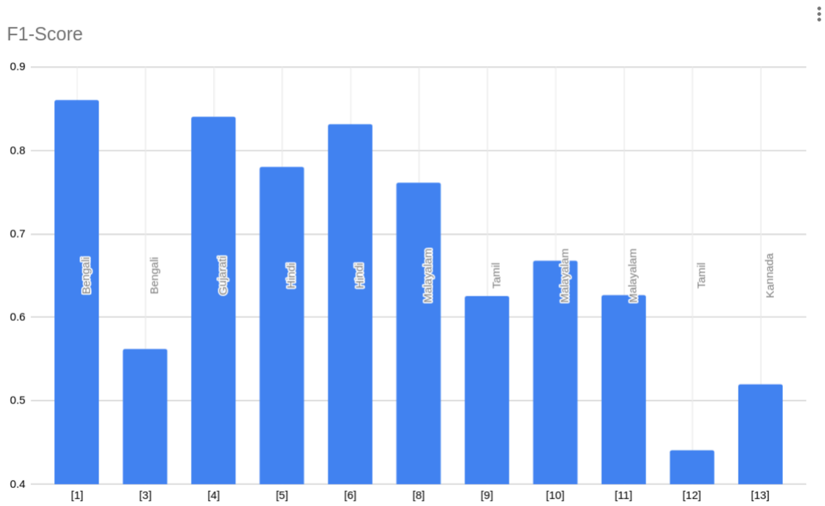

Figure 4.4: Comparison of various papers: F1-Score

Language-Specific Considerations: To address these models, language-particular issues may be required. This ought to involve extra preprocessing steps, custom phrase embeddings, or specialised models designed to address particular linguistic characteristics. Tailored processes can assist enhance the performance of depression detection models in multilingual contexts.

Paper	Precision	Recall	F1 Score
[12]	0.85	0.86	0.86
[13]	-	-	-
[14]	0.694	0.547	0.562
[15]	-	0.84	-
[16]	-	0.78	-
[16]	0.8323	0.8321	0.8319
[17]	-	-	-
[18]	0.7607	0.7633	0.7617
[19]	0.744	0.645	0.625
[20]	0.647	0.69	0.668
[20]	0.6303	0.6304	0.627
[20]	0.4232	0.5272	0.441
[20]	0.5062	0.5455	0.5193

Table 4.3: Comparative Analysis(Precision, Recall, F1 Score)

4.4 Insights and Future Directions

The comparisons between LSTM, GRU, and ensemble techniques, along with the determined language-unique variations, provide valuable insights into the design and improvement of depression detection models for social media. The following insights can manual future studies and improvement in this discipline:

Flexibility in Model Choice: The flexibility to select among LSTM, GRU, and ensemble methods permits researchers to choose the satisfactory model based on the precise necessities in their look at. This adaptability is crucial for developing strong models that can perform nicely throughout exclusive contexts.

Importance of Ensemble Techniques: Ensemble techniques can drastically beautify robustness and improve precision and keep in mind. This method may be in particular useful whilst handling complex information or whilst person models showcase one-of-a-kind strengths and weaknesses.

Need for Language-Specific Approaches: The variations discovered across languages underscore the importance of language-precise approaches. Tailored preprocessing, custom phrase embeddings, and specialised models may be vital to enhance overall performance in sure contexts.

Addressing Language Diversity: As social media structures appeal to customers from numerous linguistic backgrounds, addressing language variety is important for the success of depression detection models. Researchers should bear in mind the precise traits of each language and cultural context to broaden models which can be inclusive and generalizable.

In precise, the comparisons and insights derived from the experiments provide a complete view of the overall performance of LSTM, GRU, and ensemble models in detecting depressive tweets on social media. The outcomes spotlight the advantages of ensemble techniques, the need for language-precise strategies, and the importance of flexibility in model. These insights can provide direction for future research and make a contribution to the development of more strong and effective depression detection models in multilingual contexts.

4.5 Discussion of Limitations

In any study that entails ML and NLP, obstacles are inherent due to the complexities of the records, the scope of the experiments, and the variety of outcomes. This segment discusses the important issues encountered during the look at depression detection in social media, specializing in troubles related to dataset representativeness, language nuances, and handling imbalanced records. By addressing these limitations, researchers can better apprehend the constraints of the modern methodology and perceive regions for destiny improvement.

Dataset Representativeness

One of the primary boundaries on this study is related to dataset representativeness. Although the dataset became designed to be balanced and representative, it may not completely capture the linguistic and cultural variety within every language. This predicament has several implications:

Limited Diversity in Data Sources: The dataset used on this study was sourced from a selected social media platform, particularly Twitter. While Twitter affords a wealthy supply of user-generated content material, it may not replicate the overall spectrum of linguistic and cultural variations across exceptional areas and groups. This trouble may want to impact the generalizability of the models to other structures or contexts.

Geographic and Cultural Variability: Social media users come from various geographic locations and cultural backgrounds, each with specific language use and communique styles. The dataset's representativeness may be restrained if it does not correctly account for this variability. This obstacle may want to result in models that perform nicely in sure contexts however war in others.

To deal with this quandary, future research ought to consciousness on increasing the dataset to encompass a broader range of social media systems, geographic regions, and cultural contexts. Additionally, data augmentation techniques will be used to growth variety inside the dataset, providing a extra comprehensive representation of the linguistic and cultural panorama.

Language Nuances

Another limitation recognized in this study is the presence of language nuances that might affect model performance. The variations in precision and don't forget throughout specific Indian languages propose that some models would possibly conflict with particular linguistic traits. This drawback has the following implications:

Complexity of Language Structures: Different languages have precise grammatical systems, syntax, and vocabulary. These models can pose challenges for machine learning models, which might be skilled on datasets that do not completely seize those complexities. For example, a few languages may use idiomatic expressions or code-switching, that may confuse models trained on extra statistics.

Need for Language-Specific Approaches: The presence of language nuances indicates that a one-length-suits-all technique might not be enough for depression detection in multilingual contexts. Tailored tactics that keep in mind the particular linguistic traits of every language may be important to improve model accuracy and reliability.

To deal with this obstacle, researchers may focus on growing language-unique models or tactics that account for linguistic nuances. This may contain additional preprocessing steps, specialized word embeddings, or models designed to deal with precise language capabilities. Further research into the particular characteristics of each language could also assist identify regions wherein models need additional training or refinement.

Handling Imbalanced Data

Handling imbalanced statistics is a common mission in research focused on depression detection, as depressive tweets are often less frequent than non-depressive ones. Despite strategies like facts augmentation and oversampling, addressing imbalanced data remains difficult. This difficulty has the following implications:

Impact on Model Performance: Imbalanced facts can cause biased models that perform nicely on the bulk elegance however poorly at the minority elegance. This imbalance can result in models with high accuracy however low remember, indicating that they struggle to become aware of depressive tweets.

Effectiveness of Techniques: The effectiveness of strategies like records augmentation and oversampling may additionally range depending at the dataset's composition and the diploma of imbalance. In some instances, those techniques might not be sufficient to create a balanced dataset, leading to models that don't generalize well.

To deal with this limitation, researchers should discover additional strategies for coping with imbalanced records, inclusive of synthetic facts technology or re-sampling techniques. Additionally, evaluation metrics like F1-score, precision and recall, can provide a greater accurate degree of model performance whilst handling imbalanced records.

5 CONCLUSION AND FUTURE SCOPE

The discussion of limitations highlights several key areas where the research encountered challenging situations and in which in additional research is required. Dataset representativeness, language nuances, and coping with imbalanced statistics are major constraints that could affect the generalizability and accuracy of depression detection models. By acknowledging these barriers and addressing them using tailored methods and additional studies, researchers can enhance the robustness and reliability of future research on this subject.

Despite these limitations, the experimental setup tested the effectiveness of LSTM and GRU models in detecting depressive tweets throughout eight Indian languages. The evaluation provided precious insights into the overall performance of various architectures and the effect of language-precise nuances on model accuracy. Although challenges remain, the proposed methodology and experimental layout suggest that this method can make contributions to a deeper expertise of depression detection in multilingual contexts and function as a basis for further studies and improvement in the body of research related to mental illness and depression detection on social media.

The outcomes from the experiments provided precious insights into the effectiveness of Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and ensemble models in detecting depressive tweets throughout numerous Indian languages. This segment interprets the important outcomes, emphasizing the implications for model layout, language adaptability, and future improvement inside the subject of depression detection on social media. The experiments verified that the models executed high accuracy, precision, recall, and F1-rating, indicating their functionality to reliably distinguish depressive tweets from non-depressive ones. These values show that the model architecture and training strategies were powerful in capturing the capabilities and patterns associated with depressive content.

The performance of the models throughout eight Indian languages shows that they have been adaptable to various linguistic contexts. This adaptability is specially crucial in culturally numerous landscapes like India, where linguistic variations can notably effect social media content. Linguistic Flexibility: The results indicated that the models were able to work across a couple of languages, demonstrating their linguistic flexibility. This functionality is important for depression detection in multilingual settings, in which language-precise nuances and cultural differences play a full-size role. Cultural Sensitivity: The adaptability of the models to distinctive linguistic contexts indicates that they might seize cultural nuances in social media posts. This sensitivity is vital for correctly detecting depressive patterns, as cultural models can have an impact on how human beings express their feelings on-line.

5.1 FUTURE SCOPE

Despite the promising accuracy and reliability, the variations in precision and recall throughout distinct languages indicate that some models would possibly battle with precise linguistic nuances. This commentary underscores the need for tailored processes that consider the specific characteristics of each language. Language-Specific Challenges: The variations in overall performance across languages advise that a few models can also require additional tuning to deal with specific linguistic features. This want for personalisation suggests that a one-length-fits-all method may not be sufficient for multilingual depression detection. Nuances and Cultural Context: The statement that some models would possibly conflict with language-precise nuances highlights the significance of thinking about cultural context in model design. Tailored strategies that account for these differences ought to enhance the accuracy and reliability of depression detection in multilingual settings.

The interpretation of results leads to several implications for further research and development within the field of depression detection on social media:
Expanding Language Coverage: The potential of the models to work across diverse Indian languages suggests that further studies may want to explore extra languages and cultural contexts. They may want to further validate the adaptability of the models and enhance their generalization talents. Improving Language-Specific Models: The results in overall performance across languages recommend that future studies ought to focus on developing language-particular models. This approach would involve tailoring the models to cope with unique linguistic and cultural characteristics, enhancing their accuracy and reliability.

Exploring Advanced Architectures: While LSTM and GRU have been powerful, exploring greater advanced architectures, including transformer-based models, ought to cause further enhancements in performance. These advanced models ought to offer more desirable skills for processing complicated sequential facts and adapting to diverse linguistic contexts.

In summary, the analysis of outcomes demonstrates that the LSTM, GRU, and ensemble models have been powerful in detecting depressive tweets across diverse Indian languages. The promising accuracy, precision, recall, and F1-score imply that the models had been dependable and adaptable to one-of-a-kind linguistic contexts. However, the results in overall performance throughout languages spotlight the want for tailored methods and future studies can improve language-precise models. These insights contribute to a deeper know-how of multilingual depression detection and suggest pathways for future studies and improvement in this discipline.

6 REFERENCES

[1] J. Angskun, S. Tipprasert, and T. Angskun, "Big data analytics on social networks for real-time depression detection," *Journal of Big Data*, vol. 9, no. 1, p. 69, 2022.

[2] E. E. Michalak, G. Murray, A. H. Young, and R. W. Lam, "Burden of bipolar depression: impact of disorder and medications on quality of life," *CNS Drugs*, vol. 22, pp. 389–406, 2008.

[3] C. Malone and A. Wachholtz, "The relationship of anxiety and depression to subjective well-being in a mainland Chinese sample," *Journal of Religion and Health*, vol. 57, pp. 266–278, 2018.

[4] A. T. Beekman, B. W. Penninx, D. J. Deeg, E. de Beurs, S. W. Geerlings, and W. van Tilburg, "The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective," *Acta Psychiatrica Scandinavica*, vol. 105, no. 1, pp. 20–27, 2002.

[5] K. Zeberga, M. Attique, B. Shah, F. Ali, Y. Z. Jembre, and T. S. Chung, "A novel text mining approach for mental health prediction using Bi-LSTM and BERT model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[6] S. Zad, M. Heidari, H. James Jr, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *2021 IEEE World AI IoT Congress (AIIoT)*, 2021, pp. 0255–0261.

[7] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

[8] R. A. Calvo, D. N. Milne, M. S. Hussain, and H. Christensen, "Natural language processing in mental health applications using non-clinical texts," *Natural Language Engineering*, vol. 23, no. 5, pp. 649–685, 2017.

[9] S. D'Alfonso, "AI in mental health," *Current Opinion in Psychology*, vol. 36, pp. 112–117, 2020.

[10] M. Conway and D. O'Connor, "Social media, big data, and mental health: current advances and ethical implications," *Current Opinion in Psychology*, vol. 9, pp. 77–82, 2016.

[11] T. Zhang, A. M. Schoene, S. Ji, and S. Ananiadou, "Natural language processing applied to mental illness detection: a narrative review," *NPJ Digital Medicine*, vol. 5, no. 1, p. 46, 2022.

[12] A. Le Glaz, Y. Haralambous, D. H. Kim-Dufor, P. Lenca, R. Billot, T. C. Ryan, et al., "Machine learning and natural language processing in mental health:

systematic review," *Journal of Medical Internet Research*, vol. 23, no. 5, p. e15708, 2021.

[13] A. Le Glaz, Y. Haralambous, D. H. Kim-Dufor, P. Lenca, R. Billot, T. C. Ryan, et al., "Machine learning and natural language processing in mental health: systematic review," *Journal of Medical Internet Research*, vol. 23, no. 5, p. e15708, 2021.

[14] I. A. Nanomi Arachchige, P. Sandanapitchai, and R. Weerasinghe, "Investigating machine learning & natural language processing techniques applied for predicting depression disorder from online support forums: A systematic literature review," *Information*, vol. 12, no. 11, p. 444, 2021.

[15] K. Saqib, A. F. Khan, and Z. A. Butt, "Machine learning methods for predicting postpartum depression: Scoping review," *JMIR Mental Health*, vol. 8, no. 11, p. e29838, 2021.

[16] S. Gupta, L. Goel, A. Singh, A. Prasad, and M. A. Ullah, "Psychological analysis for depression detection from social networking sites," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[17] B. Cui, J. Wang, H. Lin, Y. Zhang, L. Yang, and B. Xu, "Emotion-Based Reinforcement Attention Network for Depression Detection on Social Media: Algorithm Development and Validation," *JMIR Medical Informatics*, vol. 10, no. 8, p. e37818, 2022.

[18] C. Li, C. Braud, and M. Amblard, "Multi-Task Learning for Depression Detection in Dialogs," *arXiv preprint arXiv:220810250*, 2022.

[19] J. Liu and M. Shi, "A hybrid feature selection and ensemble approach to identify depressed users in online social media," *Frontiers in Psychology*, vol. 12, p. 6571, 2022.

[20] H. Zogan, I. Razzak, X. Wang, S. Jameel, and G. Xu, "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media," *World Wide Web*, vol. 25, no. 1, pp. 281–304, 2022.

[21] A. H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, "Deep learning for depression detection of twitter users," in *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, 2018, pp. 88–97.

[22] Y. Wang, Z. Wang, C. Li, Y. Zhang, and H. Wang, "Online social network individual depression detection using a multitask heterogeneous modality fusion approach," *Information Sciences*, vol. 609, pp. 727–749, 2022.

[23] A. Amanat, M. Rizwan, A. R. Javed, M. Abdelhaq, R. Alsaqour, S. Pandya, et al., "Deep learning for depression detection from textual data," *Electronics*, vol. 11, no. 5, p. 676, 2022.

[24] P. Kumar, P. Samanta, S. Dutta, M. Chatterjee, and D. Sarkar, "Feature Based Depression Detection from Twitter Data Using Machine Learning Techniques," *Journal of Scientific Research*, vol. 66, no. 2, 2022.

[25] R. Safa, P. Bayat, and L. Moghtader, "Automatic detection of depression symptoms in twitter using multimodal analysis," *The Journal of Supercomputing*, vol. 78, no. 4, pp. 4709–4744, 2022.

[26] J. de Jesús Titla-Tlatelpa, R. M. Ortega-Mendoza, M. Montes-y Gómez, and L. Villaseñor-Pineda, "A profile-based sentiment-aware approach for depression detection in social media," *EPJ Data Science*, vol. 10, no. 1, p. 54, 2021.

[27] M. Rafidul Hasan Khan, U. S. Afroz, A. K. M. Masum, S. Abujar, and S. A. Hossain, "A deep learning approach to detect depression from Bengali text," in *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020, Volume 2*, Springer, 2021, pp. 777–785.

[28] Y. M. Y. Niimi, "Machine learning approach for depression detection in Japanese," in *Proceedings of the 35th Pacific Asia Conference on Language, Information and Computation*, 2021, pp. 350–357.

[29] A. A. Alaskar, "Depression Detection from Arabic Tweets Using Machine Learning Techniques," *J. Comput. Sci. Software Dev.*, vol. 2, pp. 1–10, 2022.

[30] R. Chiong, G. S. Budhi, S. Dhakal, and F. Chiong, "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts," *Computers in Biology and Medicine*, vol. 135, p. 104499, 2021.

[31] H. Zogan, I. Razzak, S. Jameel, and G. Xu, "DepressionNet: A novel summarization boosted deep framework for depression detection on social media," *arXiv preprint arXiv:210510878*, 2021

[32] G. Li, B. Li, L. Huang, S. Hou, et al., "Automatic construction of a depression-domain lexicon based on microblogs: text mining study," *JMIR Medical Informatics*, vol. 8, no. 6, p. e17650, 2020.

[33] D. B. Victor, J. Kawsher, M. S. Labib, and S. Latif, "Machine learning techniques for depression analysis on social media-case study on Bengali community," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 1118–1126.

[34] F. M. Shah, F. Ahmed, S. K. S. Joy, S. Ahmed, S. Sadek, R. Shil, et al., "Early depression detection from social network using deep learning techniques," in *2020 IEEE Region 10 Symposium (TENSYMP)*, 2020, pp. 823–826.

[35] S. H. Hosseini-Saravani, S. Besharati, H. Calvo, and A. Gelbukh, "Depression detection in social media using a psychoanalytical technique for feature extraction and a cognitive based classifier," in *Advances in Computational Intelligence: 19th Mexican International Conference on Artificial Intelligence, MICAI 2020, Mexico City, Mexico, October 12–17, 2020, Proceedings, Part II*, Springer, 2020, pp. 282–292.

[36] R. U. Mustafa, N. Ashraf, F. S. Ahmed, J. Ferzund, B. Shahzad, and A. Gelbukh, "A multiclass depression detection in social media based on sentiment analysis," in *17th International Conference on Information Technology–New Generations (ITNG 2020)*, Springer, 2020, pp. 659–662.

[37] L. Coello-Guilarte, R. M. Ortega-Mendoza, L. Villaseñor-Pineda, and M. Montes-y Gómez, "Crosslingual depression detection in twitter using bilingual word alignments," in *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 10th International Conference of the CLEF Association, CLEF 2019, Lugano, Switzerland, September 9–12, 2019, Proceedings 10*, Springer, 2019, pp. 49–61.

[38] D. Vršnak, M. Paulinovic, and V. Kuzina, "Early Depression Detection Using Temporal and Sentiment Features," *Text Analysis and Retrieval 2019 Course Project Reports*, 2019, p. 66.

[39] S. Marnauzs and J. Kalita, "A Domain Independent Social Media Depression Detection Model," 2019.

[40] S. G. Burdisso, M. Errecalde, and M. Montes-y Gómez, "A text classification framework for simple and effective early depression detection over social media streams," *Expert Systems with Applications*, vol. 133, pp. 182–197, 2019.

[41] Z. Peng, Q. Hu, and J. Dang, "Multi-kernel SVM based depression recognition using social media data," *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 43–57, 2019.

[42] Y. Chen, B. Zhou, W. Zhang, W. Gong, and G. Sun, "Sentiment analysis based on deep learning and its application in screening for perinatal depression," in *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, 2018, pp. 451–456.

[43] M. J. C. M. B. M. J. Arif and V. D. K., "Linguistic Analysis of Hindi-English Mixed Tweets for Depression Detection," *Journal of Mathematics*, vol. 2022, p. 1–7, 2022.

[44] A. Liberati, D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gøtzsche, J. P. Ioannidis, et al., "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *Annals of Internal Medicine*, vol. 151, no. 4, pp. W–65, 2009.

[45] G. Shen, J. Jia, L. Nie, F. Feng, C. Zhang, T. Hu, et al., "Depression detection via harvesting social media: A multimodal dictionary learning solution," in *IJCAI*, 2017, pp. 3838–3844.

[46] K. Kroenke, R. L. Spitzer, and J. B. Williams, "The PHQ-9: validity of a brief depression severity measure," *Journal of General Internal Medicine*, vol. 16, no. 9, pp. 606–613, 2001.

[47] G. Shen, J. Jia, L. Nie, F. Feng, C. Zhang, T. Hu, et al., "Depression detection via harvesting social media: A multimodal dictionary learning solution," in *IJCAI*, 2017, pp. 3838–3844.

[48] M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, "A review of feature extraction in sentiment analysis," *Journal of Basic and Applied Scientific Research*, vol. 4, no. 3, pp. 181–186, 2014.

[49] P. McNamee and J. Mayfield, "Character n-gram tokenization for European language text retrieval," *Information Retrieval*, vol. 7, pp. 73–97, 2004.

[50] J. Ramos, et al., "Using tf-idf to determine word relevance in document queries," in *Proceedings of the First Instructional Conference on Machine Learning*, vol. 242, Citeseer, 2003, pp. 29–48.

[51] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods," *Journal of Language and Social Psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[52] D. Larsen-Freeman and J. DeCarrico, "Grammar," in *An Introduction to Applied Linguistics*, Routledge, 2019, pp. 19–34.

[53] N. A. Ilyosovna, "The importance of English language," *International Journal on Orange Technologies*, vol. 2, no. 1, pp. 22–24, 2020.

[54] J. Chen and J. Y. Nie, "Automatic construction of parallel English-Chinese corpus for cross-language information retrieval," in *Sixth Applied Natural Language Processing Conference*, 2000, pp. 21–28.

[55] J. R. Saffran, E. L. Newport, and R. N. Aslin, "Word segmentation: The role of distributional cues," *Journal of Memory and Language*, vol. 35, no. 4, pp. 606–621, 1996.

[56] J. Angskun, S. Tipprasert, T. Angskun, "Big data analytics on social networks for real-time depression detection," *Journal of Big Data*, vol. 9, no. 1, p. 69, 2022, Springer.

[57] K. Zeberga, M. Attique, B. Shah, F. Ali, Y. Jembre, T.-S. Chung, "A novel text mining approach for mental health prediction using Bi-LSTM and BERT model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, Hindawi.

[58] R. Safa, P. Bayat, L. Moghtader, "Automatic detection of depression symptoms in twitter using multimodal analysis," *The Journal of Supercomputing*, vol. 78, no. 4, p. 4709–4744, 2022, Springer.

[59] G. Li, B. Li, L. Huang, S. Hou, et al., "Automatic construction of a depression-domain lexicon based on microblogs: text mining study," *JMIR Medical Informatics*, vol. 8, no. 6, p. e17650, 2020, JMIR Publications Inc., Toronto, Canada.

[60] S. Gupta, L. Goel, A. Singh, A. Prasad, M.A. Ullah, "Psychological analysis for depression detection from social networking sites," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, Hindawi.

[61] J. de Jesús Titla-Tlatelpa, R.M. Ortega-Mendoza, M. Montes-y-Gómez, L. Villaseñor-Pineda, "A profile-based sentiment-aware approach for depression detection in social media," *EPJ Data Science*, vol. 10, no. 1, p. 54, 2021, Springer Berlin Heidelberg.

[62] H. Zogan, I. Razzak, X. Wang, S. Jameel, G. Xu, "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media," *World Wide Web*, vol. 25, no. 1, p. 281–304, 2022, Springer.

[63] M.R.H. Khan, U.S. Afroz, A.K.M. Masum, S. Abujar, S.A. Hossain, "A deep learning approach to detect depression from Bengali text," *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020, Volume 2*, p. 777–785, 2021, Springer.

[64] D.B. Victor, J. Kawsher, M.S. Labib, S. Latif, "Machine learning techniques for depression analysis on social media-case study on bengali community," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, p. 1118–1126, 2020, IEEE.

[65] Y.M. Niimi, Y. Miyaji Yuka, "Machine learning approach for depression detection in japanese," *Proceedings of the 35th Pacific Asia Conference on Language, Information and Computation*, p. 350–357, 2021.

[66] A.A. Alaskar, "Depression Detection from Arabic Tweets Using Machine Learning Techniques," *J. Comput. Sci. Software Dev.*, vol. 2, p. 1–10, 2022.

[67] D.B. Victor, J. Kawsher, M.S. Labib, S. Latif, "Machine learning techniques for depression analysis on social media-case study on Bengali community," *2020 4th*

International Conference on Electronics, Communication and Aerospace Technology (ICECA), p. 1118–1126, 2020, IEEE.

[68] M.K. Kabir, M. Islam, A.N.B. Kabir, A. Haque, M.K. Rhaman, "Detection of Depression Severity Using Bengali Social Media Posts on Mental Health: Study Using Natural Language Processing Techniques," *JMIR Formative Research*, vol. 6, no. 9, p. e36118, 2022, JMIR Publications Inc., Toronto, Canada.

[69] A.H. Uddin, D. Bapery, A.S.M. Arif, "Depression analysis of bangla social media data using gated recurrent neural network," *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, p. 1–6, 2019, IEEE.

[70] L. Gohil, D. Patel, "Multilabel classification for emotion analysis of multilingual tweets," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, p. 4453–4457, 2019.

[71] T.T. Sasidhar, B. Premjith, K.P. Soman, "Emotion detection in hinglish (hindi+ english) code-mixed social media text," *Procedia Computer Science*, vol. 171, p. 1346–1352, 2020, Elsevier.

[72] A. Wadhawan, A. Aggarwal, "Towards emotion recognition in Hindi-English code-mixed data: A transformer based approach," *arXiv preprint arXiv:2102.09943*, 2021.

[73] S. Thara, P. Poornachandran, "Social media text analytics of Malayalam–English code-mixed using deep learning," *Journal of Big Data*, vol. 9, no. 1, p. 45, 2022, Springer.

[74] A. Hande, S.U. Hegde, R. Priyadharshini, R. Ponnusamy, P.K. Kumaresan, S. Thavareesan, B.R. Chakravarthi, "Benchmarking multi-task learning for sentiment analysis and offensive language identification in under-resourced dravidian languages," *arXiv preprint arXiv:2108.03867*, 2021.

[75] P.H.V. Pavan Kumar, B. Premjith, J.P. Sanjanasri, K.P. Soman, "Deep Learning Based Sentiment Analysis for Malayalam, Tamil and Kannada Languages," 2021.

[76] G. Van Houdt, C. Mosquera, G. Nápoles, "A Review on the Long Short-Term Memory Model," *Artificial Intelligence Review*, vol. 53, p. , Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[77] Y. Fang, S. Yang, B. Zhao, C. Huang, "Cyberbullying Detection in Social Networks Using Bi-GRU with Self-Attention Mechanism," *Information*, vol. 12, p. 171, Apr. 2021, doi: 10.3390/info12040171.

LIST OF PUBLICATIONS

[1] V. Rajderkar and A. Bhat, “Multilingual Depression Detection in Online Social Media Across Eight Indian Languages”, communicated and accepted in 2024 3rd International Conference for Innovation in Technology (INOCON)
(IEEE, Scopus Indexed)
Date of Acceptance: 25/01/2024
Date of Registration: 29/01/2024
Date of Presentation: 06/03/2024
Place: Sai Vidya Institute of Technology, Bengaluru, Karnataka

[2] V. Rajderkar and A. Bhat, “Narrative Review On Depression Detection In Online Social Media” in 2nd IEEE 2nd International Conference on Computer, Communication and Control (IC4-2024)
(IEEE, Scopus Indexed)
Date of Acceptance: 10/01/2024
Date of Registration: 11/01/2024
Date of Presentation: 09/02/2024
Place: Medicaps University, Indore, Madhya Pradesh

[3] V. Rajderkar and A. Bhat, “A Systematic Literature Review on Sentiment Analysis in Online Social Media Data for Depression Detection” communicated to Multimedia Tools and Applications, Springer.
(A Scopus Indexed Journal)

PROOFS: Publication 1: INOCON 2024: Acceptance email:

IEEE 3rd INOCON 2024 - ACCEPTANCE NOTIFICATION
1 message

Microsoft CMT <email@msr-cmt.org> Thu, 25 Jan, 2024 at 4:57 pm
Reply to: D k Gupta <deepak_gupta@gibds.org>
To: Viraj Rajderkar <virajraj022@gmail.com>

Dear Viraj Rajderkar

Paper ID / Submission ID : 1274

Title : Multilingual Depression Detection in Online Social Media Across Eight Indian Languages

We are pleased to inform you that your paper has been accepted for the Oral Presentation as a full paper for the- "IEEE 3rd International Conference on Innovation in Technology (INOCON), will be held in Sai Vidya Institute of Technology, Bengaluru, Karnataka, India.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication and will be indexed by Ei Compendex and Scopus Indexing.

Complete the Registration Process (The last date of payment is 30 JAN 2024)

Registration Link for Payment :

For Indian Authors : https://rzp.io/l/faDpY4K

For Foreign Authors : https://in.explara.com/e/ieee-inoconf
(Click on Stripe payment to make payment and then pay)

After the last date of registration IEEE PDF Express and E copyright information will be given for registered authors.

Note :

1. Any changes with the Author name, Affiliation and content of paper will not be allowed after acceptance. (if not added then can add and update in CMT)
2.This is Hybrid Conference, both online and physical presentation mode is available,

The reviews are below.

======= Review 1 =======

*** Relevance and timeliness: Rate the importance and timeliness of the topic addressed in the paper within its area of research.

Good (4)

*** Technical content and scientific rigour: Rate the technical content of the paper (e.g.: completeness of the

analysis or simulation study, thoroughness of the treatise, accuracy of the models, etc.), its soundness and scientific rigour.

Valid work but limited contribution. (4)

*** Novelty and originality: Rate the novelty and originality of the ideas or results presented in the paper.

Some interesting ideas and results on a subject well investigated. (3)

*** Quality of presentation: Rate the paper organization, the clearness of text and figures, the

Good (4)

*** Technical content and scientific rigour: Rate the technical content of the paper (e.g.: completeness of the

analysis or simulation study, thoroughness of the treatise, accuracy of the models, etc.), its soundness and scientific rigour.

Marginal work and simple contribution. Some flaws. (2)

*** Novelty and originality: Rate the novelty and originality of the ideas or results presented in the paper.

Minor variations on a well investigated subject. (2)

*** Quality of presentation: Rate the paper organization, the clearness of text and figures, the completeness and accuracy of references.

Okay. Well Written

*** Strong aspects: Comments to the author: what are the strong aspects of the paper

In this research, a new scenario of the technology has been introduced; research may be interesting for readers in this area.

*** Weak aspects: Comments to the author: what are the weak aspects of the paper?

1. The presentation should be substantially improved.

2. The presentation of the proposed algorithm is too conceptual, and details of how to operate the proposed algorithm in practice should be clearly elaborated.

3. All formulas must well written instead of copy paste of image.

Note: No major revisions in paper as per reviews and Similarity index is fine.

Thanks & Regards,
Publication Chair
3rd INOCON 2024
inoconf@gmail.com
+91-7483158331(WhatsApp)

Registration Payment Receipt:

VIDUSHI RESEARCH INNOVATION & PUBLICATION

Invoicing and payments powered by

Payment Receipt Transaction Reference: pay_NUPpfcT81G0mab

This is a payment receipt for your transaction on 3rd INOCON 2024

AMOUNT PAID **₹ 9,000.00**

ISSUED TO
virajraj022@gmail.com
+919967717418

PAID ON
29 Jan 2024

Registred Author Name
Viraj Rajderkar

DESCRIPTION	UNIT PRICE	QTY	AMOUNT
Non IEEE Member Full Registration	₹ 9,000.00	1	₹ 9,000.00
	Total		**₹ 9,000.00**
	Amount Paid		₹ 9,000.00

Presentation Certificate:

Sri Sai Vidya Vikas Shikshana Samithi ®

SAI VIDYA INSTITUTE OF TECHNOLOGY

Approved by AICTE, New Delhi, Affiliated to VTU, Recognized by Govt. of Karnataka
Accredited by NBA, New Delhi (CSE, ISE, ECE, MECH & CIVIL), NAAC 'A' Grade
RAJANUKUNTE, BENGALURU 560 064, KARNATAKA

IEEE   IEEE BANGALORE SECTION

3rd International Conference for Innovation in Technology (INOCON 2024)

Certificate

This is to certify that Dr./Prof./Mr./Ms. **Viraj Rajderkar** has presented paper entitled **Multilingual Depression Detection in Online Social Media Across Eight Indian Languages** in 2024 3rd International Conference for Innovation in Technology (INOCON), jointly organised by Departments of CSE, ISE, ECE, CSE (AI & ML), CSE (DS) at Sai Vidya Institute of Technology, Bengaluru, India held during 4th – 6th March, 2024.

Dr. Shantakumar B Patil
Secretary and Outreach Chair

Dr. M S Ganesha Prasad
General Chair

Publication 2: IEEE IC4 Acceptance email:

IC4 Notification of Paper Status - Accept

1 message

Microsoft CMT <email@msr-cmt.org> Wed, 10 Jan, 2024 at 3:40 pm
Reply to: Dr. Kailash Chandra Bandhu <kailash.bandhu@medicaps.ac.in>
To: Viraj Rajderkar <virajraj022@gmail.com>

Dear Viraj Rajderkar

Submission Id: 195
Title: Narrative Review On Depression Detection in Online Social Media

Thank you for submitting your paper titled "Narrative Review On Depression Detection in Online Social Media" to our SCOPUS indexed 2nd IEEE 2nd International Conference on Computer, Communication and Control. After careful review, we are pleased to inform you that the status of your paper is changed to Accept and selected for presentation at 2nd International Conference on Computer, Communication and Control.

We appreciate the time and effort you put into preparing your paper, and we are impressed by the quality of your work. We also want to remind you that our IEEE conference takes plagiarism very seriously, and we require that all accepted papers be original and free from any plagiarism (below 15%).

Therefore, we kindly ask that you ensure that your paper is free from any plagiarism before the conference presentation and publication. Please make sure to cite all sources properly and ensure that any direct quotes are clearly indicated. Further, you are required to submit the camera ready version (Strictly in IEEE template after incorporating the suggestions made by reviewer) and furnish the registration process by paying the required fee latest by 10th Jan' 2024 (Revised deadline). (The reviewer comments and plagiarism report may be accessed by log on to your CMT account).

Please register the paper after paying applicable fee, through the below link:
https://conference.medicaps.ac.in/

Upload the required details on the below link for registration confirmation:
https://docs.google.com/forms/d/e/1FAIpQLSdOIz522CXC2l7CiCLfcUCoJ1f6PvMoXndNJSdNY57Ius-hmg/viewform

We look forward to your presentation and welcoming you to our conference.

You can log into your CMT site at: https://cmt3.research.microsoft.com/IC42023

Note: Due to multiple authors request, the registration deadline has revised. Those who have already done the registration please ignore it.

Regards,
Convenor
2nd IEEE IC4-2024
ic4@medicaps.ac.in
9669385567

Registration Payment Receipt:

P-195

MEDI-CAPS UNIVERSITY

A.B. Road, Pigdamber, Rau, Indore (M.P.) - 452001, Indore- 452001 (M.P.) Ph No : 0731-4259500,

Other Receipt

Rec. No. : 10058

Student Name : Viraj Rajderkar

Father's Name :

Date : 11-Jan-2024

Class :

Session : 2023-2024

SN	Particulars	Amount
1	IEEE (IC-4) CONFERENCE	8,000.00

Total Payable 8,000.00

by Online/Cheque/DD/PO 8,000.00

Amount In Words : eight thousand rupees only

Bank Name : 335403681847 Cheque No. : NEFT Date : 20/12/2023

20-12-2023 UPI/335403681847/VIRAJ JAYANT RAJDERKAR

RECEIVER'S SIGN

Presentation Certificate:

MEDI-CAPS UNIVERSITY

2ND IEEE INTERNATIONAL CONFERENCE

ON COMPUTER, COMMUNICATION AND CONTROL (IC4-2024)

Certificate of Presentation

This is to certify that

Viraj Rajderkar

has presented the paper entitled

Narrative Review on Depression Detection in Online Social Media

in the 2nd **IEEE International Conference on Computer, Communication and Control (IC4-2024)** held during **Feb. 08-10, 2024.**

Dr. Ratnesh Litoriya
Convener

Dr. Pramod S. Nair
Conference Chair

Dr. Dilip K. Patnaik
Vice-Chancellor

Sponsored by IEEE turnitin