MENTAL HEALTH AND STRESS PREDICTION USING LARGE LANGUAGE MODELS

A Thesis Submitted
In Partial Fulfillment of the Requirements for the
Degree of

MASTER OF TECHNOLOGY

In

Artificial Intelligence
by

ABHISHEK PANDEY
(Roll No. 2K22/AFI/01)

Under the Supervision of

DR. SANJAY KUMAR

(Department of Computer Science & Engineering)

To the

Department of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India

May, 2024

ACKNOWLEDGEMENTS

I am highly indebted to **Dr. Sanjay Kumar** for his guidance and constant supervision as well as for providing necessary information regarding the project & alsofor his support in completing this research work. I would like to express my gratitudeto the **Head of the Department (Computer Science and Engineering, Delhi Technological University)** for their kind cooperation and encouragement whichhelped me in the completion of this research work. I would like to express my specialgratitude and thanks to all the Computer Science and Engineering staff for giving me such attention and time. My thanks and appreciation also go to my colleagues and the people who have willingly helped me with their abilities.

Abhishek Pandey
2K22/AFI/01
Department of CSE

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I, Abhishek Pandey, Roll No. 2K22/AFI/01 student of M.Tech (Artificial Intelligence), hereby certify that the work which is being presented in the thesis entitled "**Mental Health and Stress Prediction using Large Language Models**" in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Artificial Intelligence in the Department of ComputerScience and Engineering, Delhi Technological University is an authentic record of my work carried out during the period from August 2022 to June 2024 under the supervision of Dr. Sanjay Kumar, Asst. Prof, Department of Computer Science and Engineering.The content presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Place: Delhi **Candidate's Signature**

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE

Certified that **Abhishek Pandey** (Roll No. 2K22/AFI/01) has carried out the research work presented in the thesis titled "**Mental Health and Stress Prediction using Large Language Models**", for the award of Degree of Master of Technology from the Department of Computer Science and Engineering, Delhi Technological University, Delhi under my supervision. The thesis embodies the result oforiginal work and studies carried out by the student herself and the contents of thethesis do not form the basis for the award of any other degree for the candidate or submission from any other University /Institution.

Dr. Sanjay Kumar
(Supervisor) Department of CSE
Date: Delhi Technological University

Mental Health and Stress Prediction using Large Language Models

ABSTRACT

In recent times, the evolution of Large Language Models (LLMs) has brought about transformative breakthroughs in many real-life applications including mental health. The continuous advancement of artificial intelligence and natural language processing techniques has led to noteworthy achievements, with LLMs showcasing considerable potential in the detection and prediction of various mental health issues. We present insights into the role of Large Language Models in mental health detection, especially focusing on anxiety, depression, and stress detection. We first present a taxonomy for the categorization of current research based on several methods used, including prompt engineering, fine-tuning, and instruction fine-tuning. The core of this research focuses on the methodologies employed in recent studies where LLMs have been utilized for detecting mental health and analyzing the performance of various models on tasks like anxiety detection, depression detection, and stress detection. This study includes an analysis of different models, datasets, and algorithmic approaches, along with the integration of LLMs into healthcare systems, focusing on examining the strengths and limitations of different techniques highlighting the challenges, opportunities, and future gaps in mental health using large language models. A range of performance metrics including accuracy, precision, recall, and F1-score have been employed to assess the overall efficacy of the models.

LIST OF RESEARCH PAPERS

1. Abhishek Pandey and Sanjay Kumar, "**Large Language Models in Mental Healthcare Applications: A survey**" [Accepted and Presented] in 'International Conference on Computing and Machine Learning (CML 2024)' March 2024.

2. Abhishek Pandey and Sanjay Kumar, "**Mental Health and Stress Prediction Using Large Language Models**" [Accepted] at '8th 2024 IEEE Symposium on Wireless Technology & Applications (ISWTA 2024)' July 2024.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

AI	Artificial Intelligence
DL	Deep learning
GA	Generative AI
ML	Machine learning
ALBERT	A Light BERT
LLM	Large Language Model
GPT	Generative Pre-trained Transformers
LR	Logistic Regression
SVM	Support Vector Machine
KNN	K- Nearest Neighbor
GPU	Graphics Processing Unit
MLM	Masked Language Model
NLP	Natural Language Processing
RoBERTa	Robustly Optimized BERT Approach
LSTM	Long Short Term Memory
BERT	Bidirectional Encoder Representations from Transformers
TF-IDF	Term Frequency – Inverse Document Frequency

CHAPTER 1

INTRODUCTION

1.1 Overview

Mental health problems are common worldwide and affect about one person in every eight. Conditions such as depression, stress-related disorders, and anxiety account for most cases of these, but they can lead to a very poor quality of life or even death by suicide [1]. Anxiety disorders affect approximately 300 million people globally while depression is reported globally in over 280 million individuals [2]. Some individuals do not seek any form of green light mainly because they perceive them. This is often not the case with anxiety disorders, where sufferers exhibit exaggerated apprehension about everyday occurrences [3]. It is important that these manifestations should not be confused with extreme dissatisfaction only resulting in unattended or ignored mental health problems hence reduced access. One of the new trends is to use machine learning (ML) and natural language processing (NLP) for the automatic analysis of mental health behaviors in social media texts [3] [4]. The major challenge faced by researchers working on ML in about anxiety from social media has been distinguishing between normalcy and mental disorder issues when it comes to written texts. Research appeared to have focused more on general sentiments or distinct emotions than anything else in this field [5]. However, it is becoming evident that models focusing solely on emotions are insufficient for detecting mental health behaviors. Since the last decade, research in computational social science and natural language processing (NLP) has focused on using online textual data, such as social media content, to identify mental health problems. However, the majority of these studies have concentrated on developing machine learning (ML) models that are domain-specific, i.e., a single model for a single task, like depression detection, stress detection, or suicide risk assessment. Prior approaches to NLP for mental health mostly used text categorization tasks to model mental health analysis on social media, where pre-trained language models (PLMs) accomplished state-of-the-art (SOTA)

performance. For certain downstream tasks, even conventional pre-trained language models like BERT require fine-tuning [6]. Studies on the multitask setting have also looked at things like sadness and anxiety prediction. The foundation of LLMs lies in the early advancements in AI and natural language processing (NLP). Initial models focused on basic language understanding and generation but were limited by the lack of depth in linguistic interpretation. However, with the improvement in machine learning and artificial intelligence, these models have become more sophisticated. In recent years, various deep learning-based techniques have been pivotal in solving many real-life and business problems such as plant disease detection, medical imaging, social network analysis, computer vision, and many others [5,6,7,8,9,10,11]. The introduction of models like GPT (Generative Pre-trained Transformer) marked a significant leap. GPT and its successors, including GPT-2, GPT-3, and the latest GPT-4, developed by OpenAI [3] [4] [6] represent a paradigm shift in language understanding and generation. These models are trained on huge amounts of textual content, allowing them to generate human-like text responses. The evolution of ChatGPT, specifically, showcases how a language model can interact conversationally, simulating a human-like chat experience. LLMs can converse with people and provide support to those in need by utilizing their extensive knowledge base and language processing skills. Recent studies have shown that there is increasing interest in using LLMs for mental health applications, such as identifying depression symptoms and understanding how immigration affects mental health [12, 13].

1.2 Motivation

The motivation for this research on mental health and stress prediction using large language models (LLMs) addresses the global mental health crisis more effectively and efficiently. Millions of people worldwide suffer immensely from mental health illnesses like depression, anxiety, and stress, which have a major negative impact on their quality of life and financial resources. Conventional approaches to identifying and tracking these diseases frequently depend on subjective self-reports and resource-intensive clinical evaluations, which can be laborious, unreliable, and out of reach for a large number of

people, especially those living in rural or underdeveloped areas. As a result of people regularly sharing their ideas, feelings, and day-to-day experiences on social media, a wealth of user-generated content has emerged, offering a rich supply of real-time data reflecting people's mental states. Recent developments in LLMs, such as BERT, GPT-3, and RoBERTa, provide strong instruments for handling and evaluating unstructured text data, making it possible to extract important patterns and insights that were previously hard to find. These models are perfect for anticipating different mental health disorders and identifying early indicators of mental discomfort since they have sophisticated context and semantic understanding. Using LLMs presents a chance to create automated and scalable continuous mental health monitoring systems that might enhance results, enable early intervention, and lessen the stigma attached to obtaining mental health care. These systems can also improve the level of personalisation of care by customising treatments to meet the requirements and communication styles of each individual. To enable the proper use of AI in mental health, however, ethical considerations such as data privacy, permission, and bias reduction must govern the deployment of LLMs in this delicate domain. The goal of this project is to use LLMs to change the way mental health care is delivered by making it more accurate, accessible, and tailored to the requirements of different populations.

1.3 Objectives

- Creating an LLM-based framework for analysis of mental health on social media content, which is capable of achieving higher accuracy in the detection and classification of mental health states and risk factors.
- Integrating interpretability mechanisms into the LLM framework enables an explanation of the rationale behind predictions and provides valuable insights for targeted interventions and support.
- To ensure effectiveness and reliability in real-world applications, fine-tuning the LLM model on the Reddit dataset of social media posts which is labeled for specific mental health indicators.

- Evaluating the performance of the LLM framework in terms of accuracy, robustness, and interpretability, ensuring its effectiveness and reliability for real-world applications. Exploring the ethical implications of utilizing LLMs for mental health analysis, addressing concerns around privacy, bias, and potential misuse.
- Conducting an exhaustive evaluation of the LLM framework's performance, focusing on its accuracy, durability, and interpretability, to verify its suitability and dependability for practical applications.

1.4 Challenges

- Datasets: The limited number of datasets and types of Large Language Models (LLMs) present a significant challenge in mental healthcare. These limitations affect the volume and variety amount of data that can be used for training these models, which may impact the effectiveness of the model and the extent to which they can be employed in case of mental health scenarios [3, 13, 14].
- Resource availability: The availability of resources is one of the biggest constraints in mental health care using LLMs. High computation power and powerful tools are required to create and use the models in the unique setting of mental health [14, 23].
- Ethical Use and Dependence: Human interaction can be replaced by engaging LLMs which can raise ethical concerns when utilized in mental health care. In comparison to this, it raises concerns about the level of care and compassion that LLMs can offer [3, 4, 5, 6, 7].
- Accuracy and Reliability: Ensuring the accuracy of the information provided by LLMs is critical, especially while dealing with delicate issues like mental health, due to the sensitive and potentially impactful nature of the advice or information given [21, 23]. To ensure that these models provide safe, effective, and contextually appropriate guidance in mental health care, they must maintain a higher imperative of accuracy and reliability in the models [18, 19, 20, 22].
- Bias and Representation: LLMs may acquire biases present in their training data and may raise concerns about fairness and representation in mental health care [22].

Since LLMs learn from vast datasets often drawn from the internet they may unintentionally provide biased responses. In the context of mental health, this could result in unfair treatment recommendations, misdiagnosis, or lack of inclusivity for diverse populations [24, 25].

1.5 Thesis Organisation

This thesis is organized into several chapters to provide a structured and comprehensive discussion of the research:

- Chapter 1 covers the introduction of the study, gives outlines of the motivation of the research work, and sets forth the objectives of the study.
- Chapter 2 covers the background of the study and offers a comprehensive review of the existing literature on different methods for mental health detection. It also highlights the technical background and potential utilities of these methods and examines the datasets employed in prior research.
- Chapter 3 discusses the case for the creation of a dataset and details the assumptions, like hyperparameters and optimizers, mentioning all the steps performed during this study.
- Chapter 4 highlights the experimental setup and result analysis of the study performed, and it is dedicated to the performance of all the techniques used previously.
- Chapter 5 wraps up the study, discussing potential future research and conclusion of the overall study.

CHAPTER 2

RELATED WORK

2.1 Background

2.1.1 Mental Health

Large language models are advanced machine learning models trained on massive datasets to understand and generate human-like text. They use architectures based on transformers, which allow them to process and generate text by considering the context and semantics of words within a sequence. Some of the LLMs are:

- **BERT (Bidirectional Encoder Representations from Transformers)**: A model that processes text bi-directionally, understanding the context of a word from both its preceding and following text [52].
- **GPT-3 (Generative Pre-trained Transformer 3)**: A model that is trained on a variety of internet text datasets that is known for generating logical and contextually appropriate language [51].
- **RoBERTa (Robustly Optimized BERT Approach)**: An optimized version of BERT, enhancing its performance by training on more data and fine-tuning hyperparameters [50].

2.2 Technical Background

2.1.2 Different techniques for mental health detection

Several techniques involving collections of data and analytical methods are employed during technology-enabled diagnosis of mental health cases, each technique for identification of mental health status or particular problems. Below are the methods employed in mental health detection:

a) Machine Learning Techniques

- **Classification Algorithms**: Various algorithms like Logistic Regression, Random Forests, and Support Vector Machines (SVM) are trained on different labeled datasets to classify text or other data such as 'depressed' or 'not depressed' [25, 26].
- **Neural Networks**: To process text and identify patterns associated with mental health conditions various deep learning models, like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), can be utilized for identification and classification [28, 29, 30].

b) Natural Language Processing (NLP) Techniques

- **TF-IDF (Term Frequency-Inverse Document Frequency)**: It is a statistical measure that is used to evaluate the importance of a word in a document to a corpus. It helps in identifying significant terms that may indicate mental health issues [31, 32, 32].

$$TF\text{–}IDF(t, d, D) = TF(t, d) \cdot IDF(t, D), \text{ where:}$$

$$TF(t) = \frac{\text{Number of times term t appears in a document}}{Total\ no\ of\ terms\ in\ a\ document\ d}$$

$$IDF(t, D) = log_e\left(\frac{\text{Total number of documents in the corpus}}{\text{Number of documents with term t in them}}\right)$$

- **Word Embeddings**: Techniques like Word2Vec and GloVe represent words in a continuous vector space where semantically similar words are close to each other. These embeddings capture the context of words in a text [30].
- **Transformer Models**: Models such as BERT (Bidirectional Encoder Representations from Transformers), GPT-3 (Generative Pre-trained Transformer 3), and RoBERTa are capable of understanding context and semantics at a deeper level, making them highly effective for detecting nuanced signs of mental health issues in the text [34, 35].

c) Large Language Models (LLMs)

Large Language Models (LLMs) are a form of Artificial Intelligence created with the aim of understanding, generating, and interacting with humans at high and sophisticated levels. These models are core to various recent natural language processing (NLP) and have broad applications such as text generation, question answering, and language translation. LLMs are designed on the architecture of neural networks, known as the Transformer model which was introduced in the paper "Attention is All You Need" by Vaswani et al., By providing a novel approach to modeling sequences, Transformers transformed natural language processing (NLP) by mainly depending on attention and self-attention mechanisms.

2.3 Literature Survey

Xu et al. [3] examined the use of various LLMs, including Alpaca, GPT-3.5, GPT 4, and FLAN-T5 for online textual data-based mental health prediction tasks. It explores zero-shot prompting, few-shot prompting, and instruction fine-tuning techniques. The study offered insightful information about the possibilities and constraints of LLMs in applications related to mental health. In terms of balanced accuracy, their optimally tuned models, Mental-FLAN-T5 and Mental-Alpaca beat the best prompt design of GPT-3.5 by 10.9% and GPT-4 by 4.8%. Yang et al. [4] presented a comprehensive approach to enhance the interpretability of mental health analysis using social-media data and Large Language Models (LLMs). Vajre et al. [11] proposed a new language model for mental health PsychBERT that has been pre-trained on a substantial corpus of scientific literature on mental health and social media data. Danner et al. [13] presented a novel artificial intelligence (AI) application for depression detection, using advanced transformer networks to analyze clinical interviews. The paper focused on BERT-based models, GPT-3.5, and ChatGPT-4, showing state-of-the-art results in detecting depression from linguistic patterns and contextual information. The study emphasized the potential of AI in revolutionizing mental health care through early depression detection and intervention. Liu et.al [14] proposed an AI model called ChatCounselor, a large language model (LLM)

solution designed to offer support for mental health. It is distinct for its authority in actual conversations between professional psychologists and clients, allowing them to have counseling abilities along with particular skills. Their performance was analyzed by the counseling bench using a bunch of real-world counseling questions, showing that it outperformed existing open-source models and approached the performance level of ChatGPT. This capability improvement has only become possible because the model was trained on top-quality highly domain-specific data. Belyaeva et al. [15] revealed the HeLM (Health Large Language Model for Multimodal Understanding) model which allowed large language models to measure disease risk using multiple clinical modalities. Specifically, HeLM was significantly expanded on LLMs due to its capabilities of processing and understanding complex clinical data, such as tabular data and high dimensional time-series data, including spirometry. Qi et al. [19] introduced a new benchmark to frequently evaluate the supervised learning and large language models efficiency in identifying mental health problems from Chinese social media texts. New datasets were introduced for cognitive misinterpretation and suicide risk classification and compared conventional supervised learning models with LLMs like GPT-3.5 and GPT-4 using different training strategies. Mao et al. [17] presented a paper that focuses on utilizing a combination of Time Distributed Convolutional Neural Network (T CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM) for the analysis of semantic features in speech for depression severity prediction. It primarily focused on developing and testing models that can accurately classify the severity of depression using data derived from speech and text, showcasing the potential of integrating multiple data modalities in mental health diagnostics. Aragon et al. [18] focused on identifying mental disorders by analyzing emotional patterns in social media posts. They introduced novel representations based on fine-grained emotions (sub-emotions) to detect depression and anorexia. The approach was based on examining fine-grained emotional expressions (sub-emotions) and their variability over time in social media documents, offering a new perspective on mental health analysis [21]. The study demonstrates the effectiveness of large language models in identifying and classifying mental health across various social

media posts. Table 2.1, highlights the various key findings of our literature review that are underlined, and an equivalent review of the various models suggested for mental health detection.

Table 2.1 Literature Survey of recent large language models based mental health detection.

Author(s) & Year	Models Used	Performance parameters	Datasets	Advantages	Disadvantages
Xuhai Xu et al. [1], 2023	Alpaca, Alpaca LoRA, FLAN-T5 and GPT-3.5	Mental-Alpaca and Mental-FLAN-T5 show better performance	Dreaddit, SDCNL, CSSRS - Suicide	Instruction fine-tuning improves LLMs' performance across multiple mental health tasks.	Absence of a model fairness evaluation.
Kailai Yang et al., 2023	CNN, BiLSTMAtt, BERT/RoBERTa, Mental BERT	ChatGPT presents better performance compared to traditional methods	MELD, DailyDialog, RECCON, CAMS Dreaddit, T-SID, SAD	Enhances interpretability in mental health analysis using ChatGPT, providing explanations for its predictions.	With little prompt modifications, ChatGPT performance also changes.
Kailai Yang et al. [2], 2023	LLaMA2	MentaLLaMA shows strong capabilities in mental health classification and the generation of high-quality explanations.	IMHI Dataset	Advances interpretability in mental health analysis by generating detailed explanations alongside classifications.	The effectiveness of MentaLLaMA depends on the availability of high-quality, training data.
Vedant Vajre et al. [3], 2021	BERT	F1 score (PsychBERT): 0.98, F1 score (LSTM, Logistic Regression): 0.95, Accuracy: 0.990 and	Datasets are constructed by collecting conversations.	It demonstrates better performance in detecting mental health behaviors from social media text compared to state-of-the-art models.	Heavily depends on the quality of the training dataset.

		Validation Loss: 0.045			
Michael Danner et al. [4], 2023	BERT, GPT3.5	DAIC-WOZandExtended-DAICdataset: BERT Based Prec- 0.83 Recall- 0.82 F1-score- 0.82	Wizard-of-Oz (DAIC-WOZ) and Extended DAIC	Improved accuracy in depression detection using GPT-based models, outperforming traditional methods.	Concerns regarding data privacy and the ethical handling of sensitive information.
June M. Liu et al. [5], 2023	GPT3.5, GPT-4, Vicuna-v1.3-7B	ChatCounselor demonstrates enhanced performance in psychological counseling assessments.	Psych8k, a dataset comprising real-life counseling conversations.	Model shows improved performance in the counseling domain particularly in generating interactive and meaningful responses.	Data privacy and ethical handling of information is a challenge.
Anastasya Belyaeva et al. [6], 2023	(Flan-PaLMChilla 62b, XGBoost, Logistic Regression	Major Depression: HeLM-0.63 XGBoost- 0.54 LogRegression-0.62	UK Biobank	The framework shows improved performance when combining different data modalities.	There's a risk of overfitting to specific data types or traits.
Hongzhi Qi et al. [10], 2023	BERT, LSAN, ChatGLM2-6B,GLM-130B, GPT-3.5, GPT-4	BERT Prec- 88.42 Recall- 77.78 F1-score- 82.76 GPT-3.5 Prec- 84.26 Recall- 73.339 F1-score- 78.45	Datasets are created by crawling comments from the “Zoufan” blogs on the Weibo social platform.	Comparison between supervised learning and state-of-the-art LLMs, providing valuable insights into their respective strengths.	Potential limitations in LLMs for fine-grained, multi-label classification tasks,
Kaining Mao et al. [8], 2023	Bi-LSTM and T-CNN	Acc – 0.9685 Prec- 0.9709	DAIC-WOZ	Combining Bi-LSTM & T-CNN captures temporal & spatial features in speech, enhancing	It requires more computational resources and makes the model

				prediction performance.	harder to interpret.
Mario Ezra Aragon et al.[9], 2023	BOSE (BagofSub-Emotions) AND D-BOSE	F1– 0.64 Prec- 0.67 Recall- 0.70	eRisk 2018	Sub-emotion analysis, providing a better understanding of emotional expressions in social media posts.	Sub-emotion analysis adds complexity and makes it challenging.

2.3.1 Criteria for research paper selection:

- **Time Period:** We concentrated on the latest advancements in the mental health domain within the context of LLMs, specifically focusing on significant progress made between 2020 to 2023. Additionally, we incorporated references from earlier literature to provide a comprehensive and insightful overview of the current landscape.
- **Keywords:** A targeted keyword search strategy was employed, incorporating terms such as 'Large Language Models', 'Generative AI', 'Mental Health', 'ChatGPT', 'Prompt engineering', and others. This strategy ensured a focused and comprehensive review of the most recent developments in the intersection of LLMs and mental health.
- **Inclusion/Exclusion Criteria:** Our research exclusively included conference and journal papers written in English. We prioritized studies that explored the intersection of LLMs and mental health, emphasizing impactful and relevant developments in the field.

Through this study, we aim to explore:

- Various datasets, models, and different training techniques;
- Applications of mental health, conditions, and validation measures;
- Ethical concerns, privacy, safety, and other challenges.
- Gaps in the current toolkit and its applicability in clinical settings.

2.4 Datasets

The selection and composition of datasets play an important role in model designing and prediction of mental health disorders using Large Language Models (LLMs). Datasets in this context typically consist of textual data sourced from a variety of platforms, including social media posts [19, 21]. The diversity in the datasets is critical, they must encompass a wide range of demographic backgrounds, linguistic styles [19, 22, 24], and mental health disorders to enhance the model's accuracy and reduce biases. The quality, size, and representativeness of these datasets directly influence the LLM's ability to accurately detect and predict mental health issues, underscoring their importance in the field of mental healthcare informatics. Table 2.2, presents the information about different datasets along with their mental condition and the source from where it has been extracted.

Table 2.2 Datasets used in mental health from different sources.

Dataset Name	Mental Condition	Platform/ Data Source	Distinct Labels	Dataset size
Derradit [2]	Stress detection	Reddit	Yes, No	Train: 2837 Val: 300 Test: 414
DR	Depression detection	Reddit	Yes, No	Train: 1003 Val: 430 Test: 405
CLP	Depression detection	Reddit	Yes, No	Train: 456 Val: 196 Test: 299
SWMH	Mental disorder detection	Reddit	Suicide, Anxiety, Bipolar disorder, Depression	Train: 34822 Val: 8705 Test: 10,882
CAMS	Depression/Suicide detection	Reddit	Alienation, None	Train: 2207 Val: 320 Test: 625

DepSeverity	Depression Severity detection	Reddit	Minimal, mild, moderate, and severe	Train: 2842 Val: -S Test: 711
Loneliness	Loneliness detection	Reddit	Yes, No	Train: 2463 Val: 527 Test: 531

2.5 Potential Utilities in Mental Health Care

Large language models (LLMs) can interpret and simulate human behavior using natural language texts in large volumes so that they can be utilized in advancing various mental health care functions. They can, among other things, help us interpret and predict behaviors, and pick out psychological triggers for stress while serving as emotional crutches when the person is affected by distressful issues. In addition to this; when well regulated from all possible angles, especially on matters related to ethics and privacy regulations; they can serve as adjuncts supporting clinical roles in various ways. For example; they can help with diagnostic procedures at their preliminary stages as well as management of psychopathologies or even by facilitating adherence strategies during psychotherapy [36, 37].

2.6 Limitations in existing work:

While large language models (LLMs) offer promising capabilities for mental health prediction and support, there are several significant limitations and challenges in their current application. These limitations stem from technical constraints, ethical considerations, and practical deployment challenges. Here's an overview of the main limitations in the existing work in mental health prediction using LLMs:

a. **Data Privacy and Security:** Mental health data is extremely sensitive. Ensuring the privacy and security of this data when used in training and deploying LLMs is crucial. There are significant risks associated with data breaches, which could

have severe consequences for individuals whose mental health data is exposed [38, 39, 40].

b. **Bias and Fairness:** LLMs can perpetuate and even amplify biases present in their training data. If the training data is not representative of the diverse populations the model serves, it can lead to biased predictions [36]. This is particularly problematic in mental health applications, where misdiagnoses or inappropriate treatment recommendations can have serious repercussions [37]. Many LLMs are trained primarily on data from specific linguistic and cultural contexts, which can limit their effectiveness and fairness when applied to diverse global populations.

c. **Accuracy and Reliability:** Despite their sophistication, LLMs may still struggle with understanding the deep and nuanced contexts necessary for accurate mental health assessments [41, 42]. Misinterpretation of language nuances, sarcasm, or indirect speech can lead to incorrect assessments. While LLMs can identify general patterns, personalizing these patterns to individual cases without significant amounts of personal data (which raises privacy issues) remains a challenge [43].

d. **Ethical and Regulatory Challenges:** The deployment of LLMs in mental health raises ethical questions, including the potential for increasing dependency on technology for emotional support and the risks of automated systems making critical health decisions [3, 4, 5].

e. **Dependency and Dehumanization:** There's a risk that reliance on LLMs for mental health prediction and support could lead to a reduction in human interaction in therapy and support processes, which are crucial for effective mental health care. While LLMs can mimic certain types of human interactions, they cannot replicate human empathy and the deeper emotional connections often necessary in therapeutic relationships.

f. **Interpretability and Transparency:** LLMs, especially those based on deep learning, are often considered "black boxes" because it can be difficult to understand how they arrive at certain predictions or decisions [43]. This lack of

transparency can be a major issue in mental health settings where understanding the reasoning behind a diagnosis or treatment recommendation is crucial [44].

g. **Technical and Resource Limitations:** Training and deploying LLMs require significant computational resources, which can be a barrier for many organizations and countries [46]. Moreover, maintaining the accuracy and relevance of LLMs requires continuous updates and retraining, which can be resource-intensive.

The limitations of LLMs in mental health prediction highlight the need for ongoing research, development, and careful consideration of ethical, privacy, and regulatory issues [47]. Balancing the benefits of advanced AI technologies with their potential risks is crucial in this sensitive area. Addressing these limitations will require a multidisciplinary approach involving expertise in AI, mental health, ethics, and law.

2.7 Problem Statement

Mental health disorders are a major global health issue, impacting millions of individuals and imposing significant socio-economic costs on society. Traditional methods of detecting and assessing mental health and stress rely primarily on self-reported measures or clinical evaluations, which can be subjective, costly, and not scalable. With the increasing prevalence of mental health issues and the demand for continuous monitoring and early intervention, there is a pressing need for more accessible, objective, and efficient tools.

CHAPTER 3

PROPOSED METHODOLOGY

This section discussed the proposed methodology for mental health and stress prediction using large language models. We have utilized the datasets collected from different platforms and used state-of-the-art traditional machine learning techniques and NLP techniques like TF-IDF, N-grams, BERT, RoBERTa, ALBERT, and some of the LLMs models like T5, FLAN-T5 for the mental health and stress prediction. In the coming sub-sections, we have discussed all the processes of data acquisition, pre-processing, network model training, and validation. Fig 3.1 shows the overall process of the proposed methodology from data collection, data preprocessing, data labeling, and classification.

To provide a structured overview of the current research landscape, we propose a taxonomy scheme that categorizes the primary applications of LLMs into overarching groups. These categories include Prompt Engineering, Chatbots, Fine-Tuning, Pre-training, Emotional & Linguistic Context and Long-term memory. This taxonomy allows for a comprehensive examination of each category and its potential in greater detail.

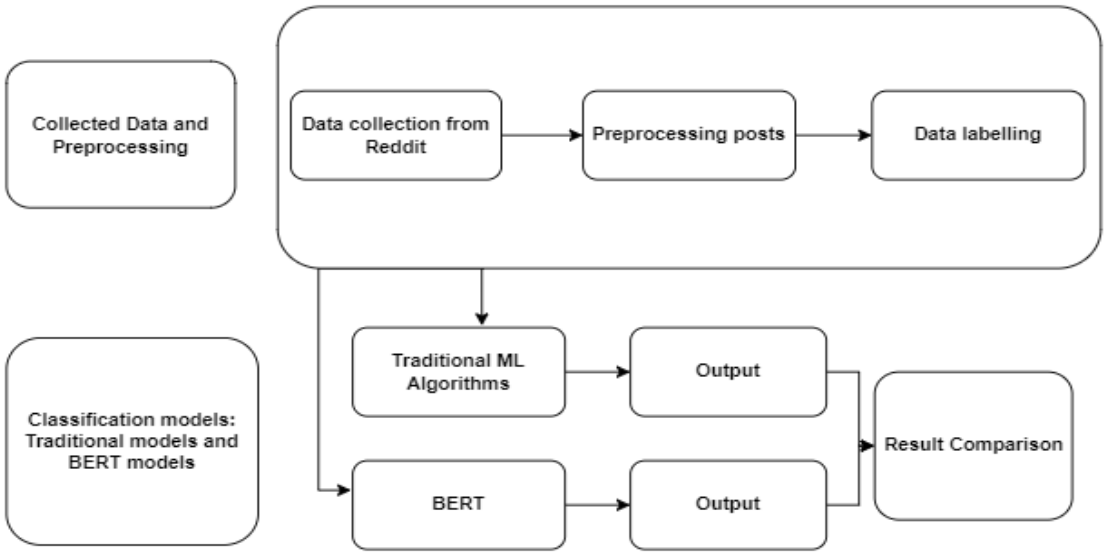

Fig. 3.1 Workflow of the proposed model for mental health and stress prediction: data collection and preprocessing, model training and classification, result evaluation

To address the challenges and meet the objectives outlined in the problem statement for predicting mental health states and stress levels using large language models (LLMs), a comprehensive methodology needs to be designed. This methodology would encompass data collection, model development, ethical considerations, and integration into practical applications. Here's a detailed breakdown:

3.1 Data Collection

Datasets are typically collected from various social media platforms like Reddit and Twitter, various posts from these platforms are crawled and used for analysis of the mental health status of a person. Various datasets like Derradit, DR, CAMS, CLP, SWMH, DeepSeverity, and loneliness are extracted from Reddit and are utilized for the prediction and classification of mental health status. Each datasets have features like posts and questions which were used for further process of prediction. Also ensuring that data spans from different demographics, and cultural backgrounds to minimize bias. On social media platforms like Reddit where a person can freely express their feelings, it was helpful in fairly making our classification using different techniques [48].

3.2 Data Preprocessing

Once data collection is completed, the next step is to preprocess it. Working directly on unprocessed data may lead to poor performance of the model and may affect the model's accuracy. Data cleaning and preprocessing are performed which was used in training. It includes handling missing values, normalization, tokenization, and removing noise [49]. Additionally, some of the sensitive information was redacted to protect the privacy of the users.

3.2.1 Textual Preprocessing

Text cleaning is one of the important steps in preprocessing, as raw text data contains noise in the form of HTML tags, special characters, numbers, punctuations, emojis, and

non-standard encoding. Text cleaning was done to ensure that the text was in a clear and consistent structure.

3.2.2 Tokenization:

Tokenization is the process of dividing the text into individual elements or tokens. Tokens in the context of natural language processing (NLP) are usually characters or words. Since transformer models function at the token level the input text was tokenized.

3.2.3 Padding and Truncation:

BERT requires fixed-length input sequences. If the input text is larger than the maximum sequence of length supported by BERT, then it needs to be truncated. Conversely, if the input text is shorter, then it needs to be padded with special tokens to match the maximum sequence of length.

3.2.4 Stemming and Lemmatization:

Stemming is the process of reducing words to their base or root form by removing suffixes from the words, whereas lemmatization is also similar to stemming which involves morphological analysis to remove inflectional endings only and return the base form a word. Fig. 3.2 shows the top 20 words extracted from TF-IDF techniques for both the stressed and non-stressed classes [49].

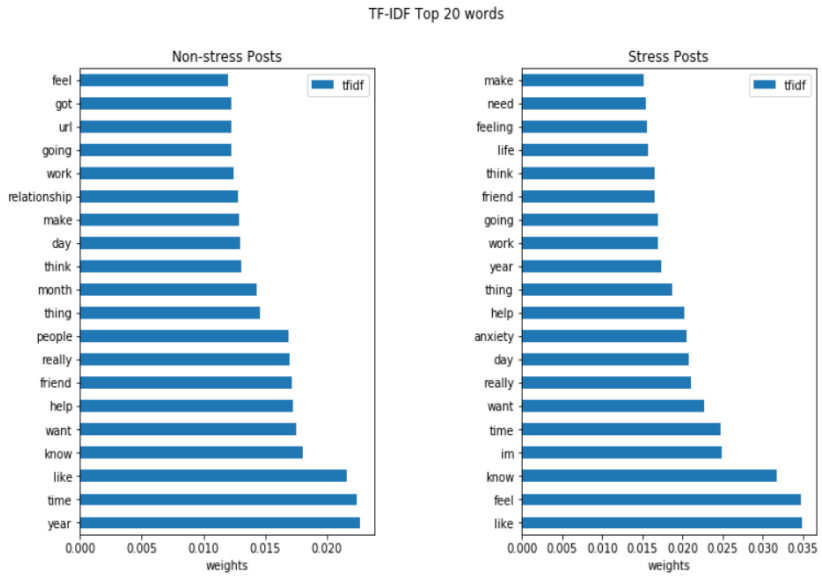

Fig. 3.2 TF-IDF Top 20 words

3.3 Model Selection and Development

3.3.1 ML methods for mental health

Machine learning techniques have a wide range of applications in mental health research, including the analysis of social media content to detect early indicators of mental distress and the development of predictive models for various mental health disorders. For example, research has demonstrated that ML algorithms can successfully classify textual data, such as social media posts or transcripts from therapy sessions, to identify markers of mental health conditions such as depression, anxiety, and PTSD [23, 24, 25, 26].

3.3.2 Transformers methods for mental health:

Natural language processing (NLP) techniques, especially sophisticated models like TF-IDF, BERT, and RoBERTa, are essential in this field [31, 32]. These models facilitate the extraction of valuable insights from unstructured text data by comprehending the context and semantics of the language used.

3.3.3 Large language models used

For mental health and stress prediction using text analysis, several large language models (LLMs) stand out due to their advanced natural language understanding capabilities. These models can be particularly effective in deciphering the nuances of human language, which is crucial for accurately interpreting the subtle cues that might indicate mental health issues or stress levels [3, 4, 5]. Fig.3.2 presents the generalized workflow for mental health classification using large language models. Here are some of the prominent LLMs that can be used for this purpose:

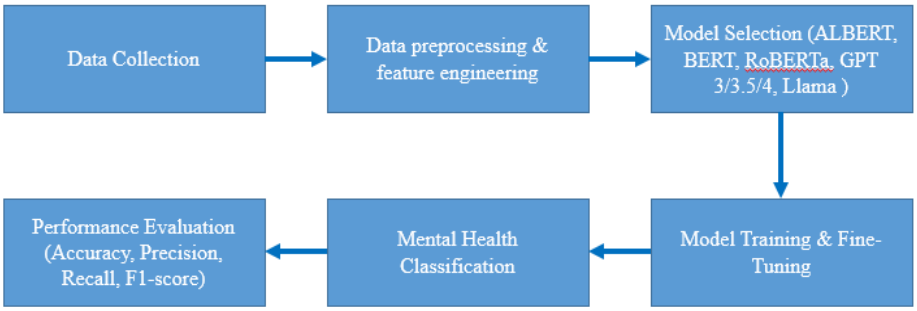

Fig. 3.3 Generalized workflow for mental health classification using large language models

a. BERT (Bidirectional Encoder Representations from Transformers)

BERT is a transformer-based machine learning technique for natural language processing (NLP) pre-training. It is designed to help computers understand the meaning of ambiguous language in the text by using surrounding text to establish context. The BERT model can be fine-tuned with additional output layers to create state-of-the-art models for a wide range of tasks, such as question answering, sentiment analysis, and language inference, making it suitable for analyzing mental health-related text. It processes words with all other words in a sentence simultaneously, boasting approximately 110 million parameters [52].

Fig. 3.4 (a) Word cloud: non-stress weights

Fig. 3.4 (b) Word cloud: stress weights

Fig. 3.4 (a) and 3.4 (b) shows the word cloud of stressed weights and non-stressed weights.

b. GPT-3 (Generative Pre-trained Transformer 3)

GPT-3 is one of the most advanced language models due to its large number of parameters (175 billion) and is developed by OpenAI. GPT-3 excels at understanding and generating human-like text, making it highly effective for tasks that require generating conversational responses or interpreting complex narratives. Its ability to generate coherent, context-aware text in a conversational manner makes it suitable for interactive applications, such as virtual mental health advisors or support bots.

c. RoBERTa (Robustly Optimized BERT Approach)

RoBERTa iterates on BERT's key hyper-parameters and training procedures but trains with much larger mini-batches and learning rates, and removes BERT's next-sentence pre-training objective. RoBERTa was shown to achieve state-of-the-art results on many NLP benchmark tasks, and its enhanced training methodology could be particularly useful in extracting more nuanced sentiment or thematic information from text related to mental health [50].

d. ALBERT (A Lite BERT)

ALBERT is a version of BERT that has been optimized to reduce model size and complexity. It achieves this through techniques such as factorized embedding parameterization and cross-layer parameter sharing. ALBERT maintains similar performance levels to BERT but is more efficient, making it suitable for environments where computational resources are limited [51].

e. T5 (Text-To-Text Transfer Transformer)

T5 converts all NLP problems into a unified text-to-text format where the task is to convert one type of text into another, making it highly versatile and adaptable.

This approach simplifies the process of applying the model to various tasks, including those related to mental health, such as sentiment analysis, emotion recognition, and therapeutic chatbots [5].

f. FLAN-T5 (Flexible Language Attention from Networks)

FLAN-T5, is becoming a powerhouse in the world of large language models (LLMs). It is built on the Transformer architecture and it has taken a step further than its predecessor, the T5 model. FLAN-T5 is trained on a huge amount of textual data and code, giving the exceptional abilities to understand and process human language [3, 4]. The main strength of this model is, it focuses on attention mechanisms, allowing it to pinpoint the most relevant parts of text for a specific set of tasks. It is particularly useful for tasks like translation and question answering, where understanding textual context is very crucial. It is designed to be fine-tuned for a broad range of natural language processing (NLP) applications.

3.4 Model Training and Validation

3.4.1 Training

Dividing the dataset into training data, and then tokenizing the text using tokenizer that matches the model architecture. As we know BERT, ALBERT, and RoBERTa are used typically in transfer learning contexts, the process involves fine-tuning a pre-trained model on a specific dataset.

3.4.2 Validation

Validating the model's performance using a separate set of validation data. Metrics such as accuracy, precision, recall, F1-score, and BLEU/ROUGE would be useful to evaluate the model comprehensively. Depending upon the performance of the validation set, the learning rate was adjusted.

3.4.3 Model Iteration

Based on the validation and test results, further refined the model iteratively. Sometimes, collecting more data or revisiting data annotation quality can also enhance performance.

3.4.4 Fine-Tuning for Mental Health Prediction

It is one of the crucial steps in the context of using pre-trained models like BERT, ALBERT, RoBERTa, and other transformer architecture. In this adapting a model that has been pre-trained on a specific task. It involves choosing the right pre-trained model, and adjusting the model architecture to fit the specific tasks, it also involves setting several key parameters like learning rate, no of epochs, and batch size.

3.4.5 Feature Engineering

Based on the data, creating new features from the text that might be relevant to the problems being solved, such as text length, word count, or presence of specific words, or developing and testing various features that could improve the model's predictive accuracy, such as scores, linguistic features (e.g., usage of first-person pronouns, language complexity), and temporal patterns in text entries.

3.4.6 Hyperparameters

Hyperparameter selection is one of the important steps during fine-tuning. Various key hyperparameters like learning rate, epochs, batch size, no of parameters, optimizers, and length embeddings are to be chose carefully to improve the accuracy of model. Epochs means the no of training cycles and batch size is chosen depending on the hardware, and it is adjustable. Smaller batch size often requires a lower learning rate [43]. Table 3.1 presents the hyperparameters of BERT used in this study. In BERT technique, 110M parameters were used with batch size of 32 and AdamW was used as optimizer in this study.

Table 3.1 Hyperparameters for BERT

Models	***BERT Base-uncased***
Parameters	110M
Batch size	32, 64
Epochs	50
Optimizers	AdamW
Learning rate	2×10^{-5}
Length Embeddings	27

3.5 Model Evaluation

This proposed methodology combines rigorous data handling, advanced AI techniques, and strong ethical practices to develop an LLM-based system for mental health and stress prediction. By ensuring the model is robust, unbiased, and integrated thoughtfully into clinical settings, it aims to enhance the accessibility and quality of mental health care.

Training Large Language Models

Training LLMs involves two main phases:

1. **Pre-training:** The model is trained on a large corpus of text data. This stage is unsupervised or self-supervised, meaning the model learns to predict parts of the text from other parts. For example, masked language modeling (used in BERT) involves hiding certain words from the input and training the model to predict them based on the context provided by other words [41].
2. **Fine-tuning:** Once the model has learned a general understanding of language from the pre-training phase, it can be fine-tuned on specific tasks. This involves additional training,

usually on a smaller, task-specific dataset. The fine-tuning adjusts the model's parameters to optimize its performance on tasks such as sentiment analysis, question answering, or document summarization [43].

CHAPTER 4

EXPERIMENTAL SETUP & RESULT ANALYSIS

4.1 Experimental Setup

In the specified sub-section, the experimental setup for the model is meticulously detailed, focusing on the software and hardware components essential for its operation.

4.1.1 Software Requirements

i) **Platform:** The model operates within the Google Colab environment, a cloud-based platform that provides a seamless interface for executing Python code.

ii) **APIs and Drivers:** Essential APIs and drivers are required to ensure compatibility and functionality within the Google Colab ecosystem.

iii) **Software:** Google Colab is the chosen software platform due to its accessibility and integration with various Google services.

iv) **Language:** Python 3.9 is the programming language selected for its widespread use and support within the data science community.

4.1.2 Hardware Requirements

i) **Processing Power**: The model utilizes Tensor Processing Units (TPUs) provided by Google Colab, which offer high-speed computation capabilities essential for processing large datasets and complex algorithms. Memory: Memory allocation is managed on demand by Google Colab, allowing for flexible scaling based on the model's requirements.

ii) **Secondary Storage:** A hard disk or Solid-State Drive (SSD) that meets the requirements of Windows OS 10 is recommended for local storage needs.

4.1.3 Libraries/Packages

i) **Pytorch:** An open-source machine learning library favored for its ease of use and efficiency in creating and training neural networks.

ii) **TensorFlow**: A Framework for building and training machine learning models.

iii) **Numpy:** A fundamental package for scientific computing in Python, providing support for large, multi-dimensional arrays and matrices.

iv) **Pandas:** A library offering data structures and data analysis tools, ideal for manipulating numerical tables and time series.

v) **Matplotlib:** A plotting library for Python and its numerical mathematics extension, NumPy, useful for creating static, interactive, and animated visualizations.

vi) **Scikit-learn (sklearn)**: Machine learning library for Python, useful for evaluation metrics and some preprocessing tasks.

vii) **Transformers (Hugging Face)**: Library for state-of-the-art NLP models like BERT, GPT-3, RoBERTa, etc.

This comprehensive setup ensures that the model is well-equipped to handle the computational demands of detecting human breathing patterns through sound analysis. The combination of Google Colab's cloud-based resources with the power of Python and its associated libraries creates a robust framework for conducting this innovative research.

4.2 Dataset Description

The experiment used a dataset extracted from Reddit, where users shared their opinions and feelings on Reddit from posts, transformer models were used to identify the mental health status of users. The collection has been carefully selected to include a wide range of posts from different places, topics, and from different users ensuring a wide range of linguistic and cultural subtleties. The dataset captures the diverse range of mental health issues by integrating posts from various locations, and users. The dataset was divided into training, validation, and testing sets after data cleaning and preprocessing to remove over-fitting from the model.

4.3 Performance Evaluation Matrix

4.3.1 Evaluation metrics

- **Accuracy**: The ratio of correctly predicted instances to the total instances. It provides a general sense of model performance but may not be sufficient for imbalanced datasets.
- **Precision**: The ratio of true positive predictions to the total predicted positives. It indicates the model's ability to avoid false positives.
- **Recall (Sensitivity)**: The ratio of true positive predictions to the actual positives. It measures the model's ability to identify all relevant instances.
- **F1 Score**: The harmonic mean of precision and recall, providing a single metric that balances both aspects.
- **ROC-AUC (Receiver Operating Characteristic - Area Under Curve)**: Measures the model's ability to distinguish between classes across different threshold settings. Table 4.1 shows the key performance parameters used in the entire study.

Table 4.1 Key performance indicators used

Performance Metrics	**Formula used**
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1-score	$\frac{2 * Precision * Recall}{Precision + Recall}$

4.3.2 Task-Specific Metrics

- **Confusion Matrix**: Provides a detailed breakdown of true positives, false positives, true negatives, and false negatives for each class, helping to identify specific areas of model weakness.
- **Macro/Micro-Averaged Metrics**: For multi-class classification, macro-averaged metrics treat all classes equally, while micro-averaged metrics aggregate the contributions of all classes to compute the average metric.

4.4 Result Analysis

Initially, the dataset was undertaken preprocessing steps, including tokenization, lowercasing, stemming, lemmatization, and data cleaning, which included handling null values, categorical values, removing outliers, hashtags, links, mentions, emoji, and special characters. Then, Pre-trained models were imported from the hugging-face library during the model-development phase. These models were then trained using the provided dataset, incorporating a fine-tuning approach whereby the AdamW optimizer and Softmax classifier. The models' performance was assessed using the testing dataset using performance metrics like accuracy, precision, recall, F1-score, classification report, BLEU, and ROGUE score. Table 4.2 presents the F1 score of various techniques of classification of mental health for all the models including traditional ML algorithms (such as Logistics Regression, Naïve Bayes, SVM, and Random Forest), NLP techniques (such as TF-IDF and N-gram) and transformer-based techniques (like BERT, ALBERT, and RoBERTa) implemented in chapter 3.

In brief, RoBERTa model outperforms BERT and ALBERT in terms of F1 scores among the transformer models evaluated. These results demonstrate the effectiveness LLMs, in particular RoBERTa. In Table 4.2, result comparison of all the techniques has been written. F1-score was used for the comparison of the performance of all the techniques. As per the table, transformer model RoBERTa outperforms all the the other transformer models as well as Machine learning and Natural Language Processing Techniques.

Table 4.2 Comparison of F1 Scores for the task of mental health classification using various techniques

Category	***Classifier***	***F1-score***
Traditional ML Algorithms	Logistic Regression	0.94
	Naïve Bayes	0.93
	Linear SVC	0.93
	Random Forest	0.86
NLP techniques	TF-IDF	0.85
	N-gram	0.84
Transformer based models	BERT	0.95
	RoBERTa	0.97
	ALBERT	0.94

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

The application of large language models (LLMs) in the field of mental health and stress prediction using social media textual data offers significant potential to revolutionize mental health care. By leveraging the advanced natural language processing capabilities of models like BERT, GPT-3, and RoBERTa, researchers, and practitioners can analyze vast amounts of unstructured text data to identify early signs of mental distress and various mental health conditions such as depression, anxiety, and PTSD.

This thesis has demonstrated that LLMs can be effectively fine-tuned to understand the linguistic nuances and context within social media posts, allowing for more accurate predictions of mental health statuses. These models can provide scalable, real-time monitoring and support, making mental health resources more accessible, especially in remote or underserved areas. The integration of such models into existing mental health care frameworks can enhance early detection, improve patient outcomes, and facilitate timely interventions.

However, the deployment of LLMs in mental health applications is not without challenges. Ethical considerations, including data privacy, consent, and bias, must be addressed to ensure the responsible use of these technologies. Furthermore, the models' generalizability across different demographics and cultural contexts remains an area for ongoing research and development.

5.2 Future Scope

1. **Enhanced Model Robustness and Accuracy**

   The future of research should be focused on enhancing the robustness as well as the accuracy of the large language models (LLMs) in predicting mental health problems. It is necessary to refine the existing models to better handle the evolving language patterns and linguistic diversity of the English language. Various advanced techniques like transfer learning and continual learning can be employed to keep the model up-to-date with new linguistic trends.

2. **Multimodal Data Integration**

   By integrating multiple media, such as text, images, videos, and music, it is possible to gain an insight into someone's psychological state of mind. It will be insightful to know how LLMs can be combined with other AI models that can analyze these data types to increase the accuracy and depth in predicting mental health cases.

3. **Real-Time Monitoring and Intervention**

   Building real-time monitoring and intervention systems that can significantly increase the utilization of large language models in mental health care. Therefore, in the future a system can be designed that can automatically trigger alerts, connecting users with emergency services.

4. **Personalization and User Adaptation**

   Personalized mental health support systems that adapt to individual users' language and behavioral patterns can improve user engagement and the effectiveness of interventions. Future studies should investigate methods for personalizing LLM-based predictions and recommendations, ensuring that they are tailored to the unique needs and circumstances of each user.

5. **Ethical and Fair AI Development**

   Addressing ethical issues is paramount for the future deployment of LLMs in mental health. Research should continue to develop and implement frameworks that ensure data privacy, obtain informed consent, and mitigate biases in model

predictions. Transparency and accountability in AI systems must be prioritized to build trust and ensure equitable access to mental health resources.

6. **Cross-Cultural and Linguistic Generalization**

   Expanding the applicability of LLMs across different cultures and languages is essential for global mental health support. Future work should focus on training models on diverse datasets that encompass various cultural and linguistic backgrounds, ensuring that the predictions are accurate and relevant for all populations.

7. **Collaboration with Healthcare Professionals**

   In future research, various interdisciplinary teams like psychologists, psychiatrists, and AI experts, collaborate and develop tools that are clinically validated and aligned with best practices in mental health care.

8. **Policy and Regulation Development**

   With the increasing use of large language models in mental health care, it becomes necessary to develop policies and regulations that will guide its use. Future research would involve creating rules and standards for AI tools that have an ethical concern based on human rights principles guaranteeing good quality of care given by professionals within this field of endeavor.

In brief, the future of mental health and stress prediction using large language models is promising, with the potential to transform mental health care delivery. By addressing the current challenges and exploring new research directions, we can develop more effective, ethical, and inclusive AI-powered mental health solutions that benefit individuals and communities worldwide.

REFERENCES

1. Mental health by the numbers, NAMI. Available at: https://nami.org/mhstats (Accessed: 14 December 2024).
2. Mental Illness National Institute of Mental Health, https://www.nimh.nih.gov/health/statistics/mental-illness (Accessed: 14 Dec 2024).
3. Xu, X., Yao, B., Dong, Y., Yang, H., Hendler, J. A., Dey, A. K., & Wang, D.: Mental-LLM: Leveraging large language models for mental health prediction via online text data. arXiv (2023) https://doi.org/10.48550/arxiv.2307.14385
4. Yang, K., Zhang, T., Kuang, Z., Xie, Q., & Ananiadou, S. (2023). MentaLLaMA: Interpretable Mental Health Analysis on Social Media with Large Language Models. arXiv (2023) https://doi.org/10.48550/arxiv.2309.13567
5. A. Mallik, S. Kumar, “Word2Vec and LSTM based deep learning technique for context-free fake news detection”. Multimedia Tools and Applications, 83, 919–940, (2024)
6. A. Kumar, sD.K. Jain, A. Mallik, S. Kumar, “Modified node2vec and attention based fusion framework for next POI recommendation”. Information Fusion, 101, 101998 (2024)
7. S. Kumar, A. Mallik, " COVID-19 Detection from Chest X-rays Using Trained Output Based Transfer Learning Approach". Neural processing letters, 55(3), 2405-2428 (2023)
8. S. Kumar, A Mallik, S.S Sengar, “Community detection in complex networks using stacked autoencoders and crow search algorithm”. The Journal of Supercomputing, 79(3), 3329-3356 (2023)
9. A. Bhowmik, S. Kumar, N. Bhat, “Evolution of automatic visual description techniques-a methodological survey”. Multimedia Tools and Applications, 80(18), 28015-28059 (2021)
10. D. Kurchaniya, S. Kumar, "Two stream deep neural network based framework to detect abnormal human activities", Journal of Electronic Imaging 32 (4), 043021-043021 (2023)
11. E. Mahajan, H. Mahajan, S. Kumar, “EnsMulHateCyb: Multilingual hate speech and cyberbully detection in online social media”, Expert Systems with Applications 236, 121228, (2024)
12. Vajre, V., Naylor, M., Kamath, U. S., & Shehu, A. PsychBERT: A Mental Health Language Model for Social Media Mental Health Behavioral Analysis. 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). (2021) https://doi.org/10.1109/bibm52615.2021.9669469

13. Danner, M., Hadžić, B., Gerhardt, S., Ludwig, S., Uslu, I., Shao, P., Weber, T., Shiban, Y., & Rätsch, M. Advancing Mental Health Diagnostics: GPT-Based Method for Depression Detection (2023). https://doi.org/10.23919/sice59929.2023.10354236
14. Liu, J. M., Li, D., He, C., Ren, T., Liao, Z., & Wu, J. ChatCounselor: a large language model for mental health support. arXiv (2023). https://doi.org/10.48550/arxiv.2309.15461
15. Belyaeva, A., Cosentino, J., Hormozdiari, F., Eswaran, K., Shetty, S., Corrado, G. S., Carroll, A., McLean, C. Y., & Furlotte, N. A. Multimodal LLMs for health grounded in Individual-Specific data. In Lecture Notes in Computer Science (pp. 86–102). (2023). https://doi.org/10.1007/978-3-031-47679-2_7
16. Qiu, J., Li, L., Sun, J., Peng, J., Shi, P., Zhang, R., Dong, Y., Lam, K., Lo, F. P., Xiao, B., Yuan, W., Wang, N. L., Xu, D., & Lo, B. Large AI models in health informatics: applications, challenges, and the future. IEEE Journal of Biomedical and Health Informatics, 27(12), 6074–6087. (2023). https://doi.org/10.1109/jbhi.2023.3316750
17. Mao, K., Zhang, W., Wang, D. B., Li, A., Jiao, R., Zhu, Y., Wu, B., Zheng, T., Lei, Q., Lyu, W., Ye, M., & Chen, J. Prediction of depression severity based on the prosodic and semantic features with bidirectional LSTM and time distributed CNN. IEEE Transactions on Affective Computing, 14(3), 2251–2265. (2023). https://doi.org/10.1109/taffc.2022.3154332
18. Aragón, M. E., López-Monroy, A. P., Gonzalez-Gurrola, L. G., & Montes, M. Detecting mental disorders in social media through emotional patterns - the case of anorexia and depression. IEEE Transactions on Affective Computing, 14(1), 211–222. (2023). https://doi.org/10.1109/taffc.2021.3075638
19. Qi, H., Zhao, Q., Song, C., Zhai, W., Luo, D., Liu, S., Yang, Y., Wang, F., Zou, H., Yang, B. X., Li, J., & Fu, G. Supervised learning and large language model benchmarks on mental health datasets: Cognitive distortions and suicidal risks in Chinese social media. arXiv. (2023). https://doi.org/10.48550/arxiv.2309.03564
20. Monica Agrawal, Stefan Hegselmann, Hunter Lang, Yoon Kim, and David Sontag. 2022. Large language models are few-shot clinical information extractors. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 1998–2022.
21. Arfan Ahmed, Sarah Aziz, Carla T Toro, Mahmood Alzubaidi, Sara Irshaidat, Hashem Abu Serhan, Alaa A Abd-Alrazaq, and Mowafa Househ. 2022. Machine learning models to

detect anxiety and depression through social media: A scoping review. Computer Methods and Programs in Biomedicine Update (2022), 100066.

22. Denecke, K., Vaaheesan, S., & Arulnathan, A. A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability test. IEEE Transactions on Emerging Topics in Computing, 9(3), 1170–1182. (2021). https://doi.org/10.1109/tetc.2020.2974478
23. T, L., Jain, D., Rapole, S. R., Curtis, B., Eichstaedt, J. C., Ungar, L., & Guntuku, S. C. Detecting Symptoms of Depression on Reddit. (2023). https://doi.org/10.1145/3578503.3583621
24. Bird, J. J., & Lotfi, A. Generative Transformer Chatbots for Mental Health Support: A Study on Depression and Anxiety. ACM (2023). https://doi.org/10.1145/3594806.3596520
25. Jo, E., Epstein, D. A., Jung, H., & Kim, Y. Understanding the Benefits and Challenges of Deploying Conversational AI Leveraging Large Language Models for Public Health Intervention. ACM. (2023). https://doi.org/10.1145/3544548.3581503
26. Ma, Z., Ma, Y., & Su, Z. Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. arXiv, 2023, 1105–1114. (2023). https://arxiv.org/abs/2307.15810
27. Yao, X., Mikhelson, M., Watkins, S. C., Choi, E., Thomaz, E., & De Barbaro, K. Development and evaluation of three chatbots for postpartum mood and anxiety disorders. arXiv. (2023). https://doi.org/10.48550/arxiv.2308.07407
28. Haque, A., Reddi, V. S. K., & Giallanza, T. Deep Learning for Suicide and Depression Identification with Unsupervised Label Correction. In Lecture Notes in Computer Science (pp. 436–447) (2021). https://doi.org/10.1007/978-3-030-86383-8_35
29. Wang, J., Shi, E., Yu, S., Wu, Z., Ma, C., Dai, H., Yang, Q., Kang, Y., Wu, J., Hu, H., Yue, C., Zhang, H., Liu, Y., Li, X., Ge, B., Zhu, D., Yuan, Y., Shen, D., Liu, T., & Zhang, S. Prompt Engineering for Healthcare: Methodologies and applications. arXiv. (2023). https://doi.org/10.48550/arxiv.2304.14670
30. Yang, K..Towards Interpretable Mental Health Analysis with ChatGPT. ArXiv. (2023).
31. M. Danner, B. Hadzic, S. Gerhardt, S. Ludwig, I. Uslu, P. Shao, T. Weber, Y. Shiban, and M. Ratsch, “Advancing Mental Health Diagnostics: GPT-Based Method for Depression Detection”. In 2023 62nd Annual Conference of the Society of Instrument and Control Engineers (SICE) , September 2023, (pp. 1290-1296). IEEE.

32. V. Vajre, M. Naylor, U. Kamath, and A. Shehu. “PsychBERT: a mental health language model for social media mental health behavioral analysis”. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) December 2021, (pp. 1077-1082). IEEE.
33. N. Proferes, N. Jones, S. Gilbert, C. Fiesle and M. Zimmer “Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics”. Social Media+ Society, 7(2), 2021, p.20563051211019004.
34. Dhelim, Sahraoui, et al. "Detecting Mental Distresses Using Social Behavior Analysis in the Context of COVID-19: A Survey." ACM Computing Surveys 2023.
35. K. Yang, T. Zhang, H. Alhuzali and S. Ananiadou, “Cluster-level contrastive learning for emotion recognition in conversations”. IEEE Transactions on Affective Computing, 2023.
36. J. Devlin, M. Chang, W., K. Lee and K. Toutanova “Bert: Pre-training of deep bidirectional transformers for language understanding”. arXiv preprint arXiv:1810.04805, 2018.
37. S. Zanwar, D. Wiechmann, Y. Qiao, and E. Kerz, “Exploring Hybrid and Ensemble Models for Multiclass Prediction of Mental Health Status on Social Media”. arXiv preprint arXiv:2212.09839, 2022.
38. I.J. Dristy, A.M. Saad, and A.A. Rasel, “Mental Health Status Prediction Using ML Classifiers with NLP-Based Approaches”. In 2022 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET), December 2022, (pp. 1-6). IEEE.
39. V.M. Deshmukh, B. Rajalakshmi, S. Dash, P. Kulkarni and Gupta, S.K. “Analysis and characterization of mental health conditions based on user content on social media”. In 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI),, January 2022, (pp. 1-5). IEEE.
40. M.E. Villa-Pérez, L.A. Trejo, M.B. Moin and E. Stroulia “Extracting Mental Health Indicators From English and Spanish Social Media: A Machine Learning Approach” . IEEE Access, 11, 2023, pp.128135-128152.
41. M. Nouman, H. Sara, S.Y. Khoo, M.P. Mahmud and A.Z. Kouzani, “Mental Health Prediction through Text Chat Conversations”. In 2023 International Joint Conference on Neural Networks (IJCNN), June 2023, (pp. 1-6). IEEE.
42. B.L. Cook, A.M. Progovac, P. Chen, B. Mullin, S. Hou and E. Baca-Garcia “Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms

in a text-based mental health intervention in Madrid". Computational and mathematical methods in medicine, 2016.

43. J.M. Liu, D. Li, H. Cao, T. Ren, Z. Liao and J. Wu "Chatcounselor: A large language models for mental health support". arXiv preprint arXiv:2309.15461, 2023.
44. K. Yang, S. Ji, T. Zhang, Q. Xie, Z. Kuang and S. Ananiadou, "Towards interpretable mental health analysis with ChatGPT". arXiv, 2023.
45. D. Van Le, J. Montgomery, K.C. Kirkby and J. Scanlan. "Risk prediction using natural language processing of electronic mental health records in an inpatient forensic psychiatry setting". Journal of biomedical informatics, 86, pp.49-58, 2018.
46. R. Stewart and S. Velupillai "Applied natural language processing in mental health big data". Neuropsychopharmacology, 46(1), 2021, p.252.
47. J. Pennington, R. Socher and C.D. Manning "Glove: Global vectors for word representation". In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), October 2014, (pp. 1532-1543).
48. T. Mikolov, K. Chen, G. Corrado and J. Dean "Efficient estimation of word representations in vector space". arXiv preprint arXiv:1301.3781, 2013.
49. V. Efimov "Large language models: ROBERTA — a robustly optimized BERT approach". Medium . https://towardsdatascience.com/roberta-1ef07226c8d8 (2023).
50. RoBERTa.(n.d.). https://huggingface.co/docs/transformers/model_doc/roberta, (2024)
51. ALBERT.(n.d.). https://huggingface.co/docs/transformers/model_doc/albert, (2024)
52. BERT.(n.d.). https://huggingface.co/docs/transformers/model_doc/bert, (2024)

LIST OF PUBLICATIONS/ACCEPTANCE & THEIR PROOFS

1. Abhishek Pandey and Sanjay Kumar, "Large Language Models in Mental Healthcare Applications: A survey" [Accepted and Presented] in 'International Conference on Computing and Machine Learning (CML 2024)' March 2024.
2. Abhishek Pandey and Sanjay Kumar, "Mental Health and Stress Prediction Using Large Language Models" [Accepted] at '8th 2024 IEEE Symposium on Wireless Technology & Applications (ISWTA 2024)' July 2024.