MENTAL HEALTH CHATBOT WITH AI

By

Muhammad Usman Iftikhar	**01-132202-032**
Ahmad Faraz	**01-132212-004**
Sadeed Ullah	**01-132212-038**

Supervised by

Engr. Muhammad Kashif Naseer

Department of Computer Engineering

Bahria University Islamabad

2025

MENTAL HEALTH CHATBOT WITH AI

By

Muhammad Usman Iftikhar	**01-132202-032**
Ahmad Faraz	**01-132212-004**
Sadeed Ullah	**01-132212-038**

Supervised by

Engr. Muhammad Kashif Naseer

Submitted to the Department of Computer Engineering in the partial fulfillment of the requirements for the degree of Bachelor of Computer Engineering.

Department of Computer Engineering

Bahria University Islamabad

2025

CERTIFICATE

MENTAL HEALTH CHATBOT

SDG No	Description of SDG	SDG No	Description of SDG
SDG 1	No Poverty	SDG 9	Industry, Innovation, and Infrastructure
SDG 2	Zero Hunger	SDG 10	Reduced Inequalities
SDG 3	Good Health and Well Being ✓	SDG 11	Sustainable Cities and Communities
SDG 4	Quality Education	SDG 12	Responsible Consumption and Production
SDG 5	Gender Equality	SDG 13	Climate Change
SDG 6	Clean Water and Sanitation	SDG 14	Life Below Water
SDG 7	Affordable and Clean Energy	SDG 15	Life on Land
SDG 8	Decent Work and Economic Growth	SDG 16	Peace, Justice and Strong Institutions
		SDG 17	Partnerships for the Goals

Range of Complex Problem Solving			
No	**Attribute**	**Complex Problem**	
1	Range of conflicting requirements	Involve wide-ranging or conflicting technical, engineering and other issues.	
2	Depth of analysis required	Have no obvious solution and require abstract thinking, originality in analysis to formulate suitable models.	✓
3	Depth of knowledge required	Requires research-based knowledge much of which is at, or informed by, the forefront of the professional discipline and which allows a fundamentals-based, first principles analytical approach.	✓
4	Familiarity of issues	Involve infrequently encountered issues.	
5	Extent of applicable codes	Are outside problems encompassed by standards and codes of practice for professional engineering.	
6	Extent of stakeholder involvement and level of conflicting requirements	Involve diverse groups of stakeholders with widely varying needs.	✓
7	Consequences	Have significant consequences in a range of contexts.	
8	Interdependence	Are high level problems including many component parts or sub-problems	
Range of Complex Problem Activities			
No	**Attribute**	**Complex Activities**	
1	Range of resources	Involve the use of diverse resources (and for this purpose, resources include people, money, equipment, materials, information and technologies).	✓
2	Level of interaction	Require resolution of significant problems arising from interactions between wide ranging and conflicting technical, engineering or other issues.	
3	Innovation	Involve creative use of engineering principles and research-based knowledge in novel ways.	✓
4	Consequences to society and the environment	Have significant consequences in a range of contexts, characterized by difficulty of prediction and mitigation.	
5	Familiarity	Can extend beyond previous experiences by applying principles-based approaches.	

UNDERTAKING

I certify that research work titled **"Mental Health Chatbot With AI"** is my own work. The work has not been presented elsewhere for assessment. Where material has been used from other sources it has been properly acknowledged / referred.

Muhammad Usman Iftikhar
01-132202-032

Ahmed Faraz
01-132212-004

Sadeed Ullah
01-132212-038

DEDICATION

In the name of Allah, the Most Gracious, the Most Merciful. All praise is due to Allah, Lord of the worlds, whose guidance and blessings made this work possible. This is dedicated to our parents, whose unconditional love, support, and encouragement aided, and motivated us throughout this endeavor. We are forever grateful for their patience and understanding, without which we have failed to accomplish this feat. We would also like to extend our gratitude to our siblings and friends who continuously uplifted our spirits when we needed it most. We shall always cherish their contribution to our achievement.

ACKNOWLEDGEMENT

First and foremost, we express our deepest gratitude to **Engr. Muhammad Kashif Naseer** for their invaluable guidance, encouragement, and expertise throughout this project. Your mentorship has been instrumental in shaping the development of this mental health chatbot. We extend our heartfelt thanks to *Bahria University* for providing us with the resources and platform to bring this idea to life. Special thanks to **team members** for their unwavering support and dedication. Our efforts, insights, and enthusiasm have been crucial in overcoming challenges and achieving our goals. We are profoundly grateful to the **mental health professionals and researchers** who offered their time, feedback, and expertise, ensuring the chatbot is both practical and empathetic in addressing users' needs. Thank you all for being part of this journey and for helping us transform this vision into reality.

ABSTRACT

An increasing number of individuals face mental health problems, including anxiety, stress, and depression because academic, social, and personal demands keep rising. The conventional mental health services encounter multiple obstacles such as restricted accessibility as well as cost barriers and societal stigma that limit people from obtaining timely assistance. This project developed an AI-enabled mental health chatbot mobile app that provides scalable, accessible, and interactive methods for mental health evaluation and support for users. This Mobile app presents the Depression Anxiety Stress Scales (DASS-21) questionnaire in a conversational format, allowing users to conduct daily self-assessments. User interaction with this chatbot runs through its empathetic transformer-based system, which determines user sentiment using sentiment analysis to produce weekly mental health indexes. The application stores generated insights in Firebase, enabling users to track their progress over time and identify patterns in their mental health. Additionally, the app offers customized coping strategies and resources based on individuals' responses, ensuring that users receive targeted support that suits their specific needs. This personalized approach not only fosters a deeper understanding of one's mental well-being but also empowers users to take proactive steps towards improvement. Users update their self-analyses, conducted in the mental health application on a regular basis, they get a better idea of their emotional patterns. These self-estimates records also will be available to appointed therapists providing more rational decisions and individual care under the control of a human. Users benefit from the app because it implements intelligent communication to detect mental health conditions while building self-awareness and increasing involvement with mental health services.

Keywords: DASS-21, BERT, BART, Emotion classifier, Text Generation, Mental Health, Chatbot, Mobile Application

TABLE OF CONTENTS

LIST OF FIGURES

Number *Page*

LIST OF TABLES

Chapter 1
Introduction

1.1 Background

1.1.1 The Growing Burden of Mental Health Disorders

Mental health issues such as stress, together with anxiety and depression, now serve as primary sources of global disability [1], which affects an immense population across various demographic groups. The World Health Organization (WHO) reports that mental illnesses affect approximately one-fifth of all adults within a single year, proving that mental health problems exist throughout the population [2]. Mental health issues move further than individual sufferers because they affect families and communities as well as economies [3]. Mental illnesses place substantial weight on individuals through personal turmoil and simultaneously affect productivity, health system expenses, and vitality reduction.

Many mental health problems exist while various obstacles continue to prevent people from obtaining prompt and beneficial treatment approaches. The prejudice directed at people with mental health problems continues to be the most significant barrier because society views them as deficient or psychologically flawed [4]. Public prejudice creates emotional distress and loneliness that motivates many suffering people to avoid seeking medical intervention. Insufficient mental health knowledge among people exacerbates the problem, as they struggle to identify symptoms of mental illness and comprehend appropriate treatment options [5]. The reduced level of understanding about mental health lets people face their mental conditions independently without help.

Professional mental health service access functions as a major impediment that hinders treatment. Most mental health experts, along with psychologists, psychiatrists, and counsellors, remain scarce across numerous regions, specifically in low- and middle-income countries [6]. The insufficient funding directed at mental health services together with the separation of mental healthcare from primary healthcare results in an exacerbated mental health services shortage [7]. The resulting delay in care duration or complete inability to access services from specialists causes people's conditions to deteriorate while their disability levels rise.

1.1.2 Emergence of AI in Mental Health Care

Current Artificial Intelligence (AI) and Natural Language Processing (NLP), together with big data advancements, have transformed mental health intervention methods by providing accessible and effective innovative solutions [8]. Advanced algorithms supported by modern technologies established the foundation for creating text-based chatbots that perform natural, human-like communication. These chatbots offer major improvements to mental health service delivery because they provide emotional support and responsive counselling for users who need assistance from professionals in a society with limited traditional outreach .

Wysa, which operates using AI and NLP technology [9], helps users engage in authentic mental health discussions through its text-based communication interface. Engineers have engineered these programs to interpret user messages through natural responses, demonstrating empathy for users . Camod can evaluate users' messages and deliver personalized support through adaptive responses that address specific mental health problems [10]. This customized approach leads to a better connection between the user and the experience because it creates a therapist-like understanding, which many people seek.

These AI-driven interventions deliver an accessible environment with no judgement to users as one of their most valuable features. Traditional mental health services face challenges as many users shy away from seeking help due to concerns about stigma and prejudice. Users gain access to anonymous interactions through chatbots because they can freely express themselves without risking labelling or misinterpretation [11]. People who are first timers with mental health conversations and worry about support requests find tremendous value in this service.

The ability of chatbots to be accessible at all times acts as a revolutionary advancement within mental health service provision. These platforms deliver round-the-clock support so users can obtain instant help anytime, allowing them to overcome traditional mental health service accessibility issues [12]. The convenience of crisis support remains crucial because people need emergency mental healthcare, yet medical professionals might not instantly reach them. Offering real-time help through chatbots functions as a trustworthy backup for disturbed users who need direction in coping strategies while promoting feelings of reassurance .

AI-driven chatbots deliver support services beyond their data analysis capabilities as well as data acquisition functions. The platforms collect user feedback together with interaction data to detect essential mental health patterns, which helps expand general knowledge about population mental health struggles [13]. The collected data allows health- care professionals to create new mental health solutions which better serve their client population needs. The analysis of chatbot-generated data helps generate knowledge that assists mental health research in achieving better treatment approaches.

The promising mental health intervention potential of AI and NLP technologies re- quires understanding their concurrent ethical concerns as well as technical boundaries [14]. Chatbots lack the capability to address serious mental health crises, along with complex psychological matters that require professional assistance. Users need to grasp the role of these tools as supplementary resources next to professional care instead of consid- ering them care replacements. Users need their privacy and security guarantees to feel protected while they navigate the digital platforms.

1.1.3 Significance of Early Intervention

Most mental health disorders appear in young people during their adolescent and early adulthood years; research suggests that nearly 75% of all mental health disorders form before age 25 [15]. This statistic highlights the significant importance of prompt mental health intervention, as it can significantly alter the trajectory of an individual's mental health development. After proper detection of mental health issues, schools should start treatment promptly since this approach yields better results and protects against serious long-term mental health problems while enhancing life satisfaction.

Young adults need urgent mental health care because suicide ranks as a major contributor to deaths within this specific age range. The Centres for Disease Control and Prevention's (CDC) documented substantial growth in suicide death rates among 10- to 24-year-olds throughout the past 10 years [16]. This alarming situation provides compelling evidence that young people need accessible, proactive mental health care programs and intervention methods. Traditional mental health services, along with negative social perceptions about mental health, bar many people from getting help and are less attractive to youths.

AI-powered chatbots demonstrate enormous potential as a solution to bring improved mental health care access for young adults during the current deficiency period [17]. Digital platforms offer instant mental health support, which users can use to speak about their problems anytime and from wherever they wish. Public chatbots offer young people an unmatched level of secrecy, which attracts them because they prefer not to show their concerns to professionals nor experience peer criticism. Young people feel comfortable sharing their feelings through chatbots because these platforms create a secure platform that eliminates the common obstacles found in conventional therapy.

AI-powered chatbots use natural language processing to create customised interventions with user input and understand specific requirements for each person. The delivery of personalised interactions through these resources leads to improved user satisfaction, which in turn increases their likelihood of accessing the support services they need. Chatbots provide users with clinical and evidence-based stress management skills, as well as personalised mindfulness practices and automatic self-assessment tests that promote individual growth in mental health management [18]. Through practical resources and skill provision, chatbots transform youth into mentally resilient individuals who maintain better well-being.

The ability to scale AI-powered chatbots enables healthcare organisations to widely distribute mental health support, which ensures that stakeholders in underserved areas receive the necessary services [19]. Countries particularly in rural regions and areas with limited mental health professional availability often result in delayed appointment services for patients seeking care. Young adults can get immediate mental health support from chatbots, which make essential resources available across locations. Emergency responses become crucial, especially during crisis situations, because patients need immediate support that can prove vital to their survival [20].

The large potential of AI-powered chatbots in mental health support requires understanding their restricted capabilities. The use of chatbots should always complement professional mental health assistance because they lack the ability to address severe mental health crises or diverse psychological complexities. The tools comprise auxiliary services which enhance standard healthcare systems instead of functioning as direct replacements. Universities must continue their studies to guarantee that these digital assistants receive the necessary skills to address different mental health issues properly while maintaining ethical standards.

1.1.4 Recent Year Emergence

Public health institutions currently view mental health as one of their primary con- cerns. The recent surge in anxiety, together with stress and depression, requires open access to immediate mental support systems. The development of conversational agents through artificial intelligence advancements allows people to receive mental health assistance

[12]. A mental health chatbot mobile app utilizing AI stands as the core subject of this thesis because it seeks to deliver persistent monitoring, guidance, and support to end users.

1.2 Problem Statement

Current mental health intervention faces obstacles which include restricted availability and high expense and social discrimination against patients. Medical and mental health-care providers seek new solutions which provide instant monitoring services be- yond traditional clinical space restrictions. A proposed mobile application solves these problems by letting therapists observe and monitor patient progress through a secured system using AI-powered tools for assessment and both therapist and patient communication with individual recommendations.

Millions of people around the world grapple with mental health issues such as stress, anxiety, and depression, which can significantly impair their quality of life and overall well-being. Despite the widespread nature of these conditions, many individuals remain reluctant to seek treatment due to a variety of barriers. Stigma surrounding mental health continues to be one of the most formidable obstacles, as societal perceptions often label those who seek help as weak or unstable. This stigma can lead to feelings of shame and isolation, discouraging individuals from reaching out for the support they desperately need.

In addition to stigma, financial constraints pose a significant barrier to accessing mental health services. Many individuals find that the cost of therapy and other mental health interventions is prohibitively high, particularly in regions where mental health care is not covered by insurance or where there are limited public health resources. This financial burden can prevent individuals from seeking help, leading to a cycle of untreated mental health issues that can worsen over time. Furthermore, the lack of mental health professionals in many areas exacerbates the problem, as individuals may face long wait times for appointments or may be unable to find qualified providers in their vicinity.

The existing mental health intervention faces three main barriers: limited access, high service costs, and the social rejection of patients. Healthcare providers from medical and mental fields actively search for new solutions to exceed traditional clinic space limitations for instant monitoring services. A proposed mobile application addresses these issues through an AI-enabled platform that ensures patient security for monitoring and assessment with customisable tools that enhance therapeutic communication between professionals and clients.

People from different parts of the globe struggle with mental disorders comprising depression and stress and anxiety that diminish their daily functioning as well as their mental health quality. Regular treatment for widespread conditions remains out of reach for many people, although the disorders affect numerous populations. Society considers mental health seekers to have flaws or instability because they carry unwarranted judgements about psychiatric problems. This social misconception stands as a major barrier for people who need help. The resultant stigma causes people to experience social isolation together with feelings of shame, so they avoid seeking essential support.

The high cost of mental health care and social views create substantial obstacles for people who need such services. Therapy alongside mental health treatments often costs too much money for many individuals who do not have insurance coverage of mental health benefits or who live in regions with insufficient public health services. Being unable to afford mental health treatments prevents people from seeking help, thus causing untreated mental health problems to grow worse with time. The scarcity of mental health professionals throughout various locations increases the challenge since patients encounter extended waiting periods to book appointments or encounter deficits in available suitable providers serving their regions.

The COVID-19 pandemic generated two major impacts on mental health: it elevated mental health difficulties because of prolonged stress while simultaneously disrupting standard mental health support services. Numerous therapy patients lost their support network when lockdowns along with social distancing guidelines became mandatory. The care gap disruption has generated a fatal requirement for novel solutions, which must fill the unmet support needs of vulnerable people.

People now understand the necessity for accessible mental health solutions that provide nonjudgmental care while meeting patients in their current locations. Digital mental health interventions represent emerging solutions for traditional therapy, which include teletherapy and mobile applications and AI-powered chatbots. These tools create opportunities to immediately provide help combined with necessary resources to people who cannot find the courage to attend face-to-face sessions. Technology enables mental health organisations to create service access for people who can now receive support in their homes.

Digital solutions bring about two benefits: they normalise mental health discussions while simultaneously minimising social stigma around mental health issues. Online platforms and resources that gradually attract more users will transform mental health perceptions within society toward an accepting and open environment. The openness encourages others to seek mental health assistance while sharing their situations, thus creating supportive conditions for people with mental health issues.

Digital mental health solutions deliver large-scale availability of resources so underserved communities can obtain support they might otherwise not receive. Teletherapy alongside online support groups offers care access to persons living in rural and underserved regions since they do not have physical distance limitations. Quick access to intervention during emergencies becomes crucial because it may save lives.

1.3 Aims and Objectives

A set of precise goals and objectives has been devised in order to govern the course of this study and guarantee that its purpose is understood. While the objectives outline the precise actions needed to accomplish those goals, the aims emphasize the study's overall goals. The main goals and objectives that serve as the basis for this work are delineated in the following points:

- To design and develop an AI-powered chatbot that assists users in managing anxiety, stress, and depression.
- To integrate the DASS-21 questionnaire in a conversational format.
- To employ sentiment analysis for weekly score generation.
- To store and manage user data securely for therapist review.
- Develop a text-to-text conversational chatbot using AI and machine learning techniques.
- Diagnose mental health conditions through the DASS-21 test and categorize clients into mild, moderate, or severe categories.
- Provide mood tracking using a customize dataset with five columns: (text, id, sad, happy, angry.)

1.4 Scope of the Project

This project focuses on the development of a mobile application equipped with advanced Artificial Intelligence (AI) capabilities to monitor, assess, and support users' mental health in real time. It targets widespread issues such as anxiety, depression, and stress conditions that are often underreported due to stigma or lack of access to mental health professionals. The app serves as a confidential and judgment-free platform where users can interact daily with an empathetic chatbot, which evaluates their emotional well-being using the DASS-21 questionnaire in a conversational format.

The core functionality of the application involves AI-driven sentiment analysis and mood tracking. By analyzing user input using a fine-tuned BERT model, the chatbot detects emotional patterns and assigns severity scores that help categorize users' mental states. These scores are aggregated weekly and securely stored using Firebase, providing users with progress insights over time and enabling mental health professionals to review logs when needed. In this way, the app not only assists users in managing their emotional health but also supports therapists in offering more personalized and timely interventions.

Beyond basic chat support, the application offers personalized mental health tips, coping strategies, and mood-enhancing suggestions based on each user's emotional history and DASS-21 results. This feature empowers users to take proactive steps in their mental health journey, helping them develop emotional resilience, self-awareness, and improved daily functioning. The chatbot also fosters engagement through its friendly UI, daily reminders, and AI-powered responsiveness, making it a sustainable solution for long-term mental health care.

The project is scalable and future-ready, designed with flexibility to integrate real-time therapist communication, multilingual support, and even wearables (for physiologicalindicators like heart rate or sleep patterns). As mental health continues to emerge as a critical global issue, this solution has the potential to bridge service gaps in underserved regions, reduce stigma, and deliver accessible care through intelligent, ethical, and user-centered digital intervention.

Chapter 2
Literature Review

2.1 Related Work

The rising interest in AI mental health applications has driven advancements to improve accessibility and reduce stigma [21] and develop effective support systems. The review explores existing literature and technological advancements which pertain to AI-based mental health assistance .

2.1.1 AI Chatbots in Mental Health

The research team at Fitzpatrick et al. [22] created Woebot as a standalone AI bot which provides Cognitive Behavioral Therapy (CBT) services to users with anxiety and de- pression symptoms. The research established that users achieved substantial mental well-being advancements through Woebot interventions during short sessions. The ef- fectiveness and possible implementation of chatbot-based interventions receive support from available evidence. Wysa represents an empathy-driven chatbot which uses evidence-based therapeutic methods like CBT and Dialectical behavior therapy (DBT) and motivational interviewing to interact with users [6]. The researchers studied both empathetic interactions alongside extended periods of connection which produced the best therapeutic results.

2.1.2 Sentiment Analysis for Mental Health

Sharma and Mehra analyzed implementations of healthcare sentiment analysis with an emphasis on its value of psychological medical observation [9]. The sentiment analysis system identifies specific linguistic indicators within user input which positively or negatively correlate with emotional expressions. Such a design enables systems to track changes in a user's psychological condition chronologically. The evaluation of mental health condition severity can be achieved through NLP techniques according to different researchers' findings [4]. Sentiment analysis links to chatbots through an integration that allows systems to sustain interactive exchanges which adapt according to user sentiment.

2.1.3 Use of Psychological Scales in Chatbots

Standardized mental health assessment depends on the Depression Anxiety Stress Scales (DASS-21) tool which proves effective at validating mental health status. A standardized DASS-21 assessment built into a chatbot system provides uniform symptom monitoring. Modern mobile health applications integrate DASS-21 and similar psychological screening tools, yet most systems need conversational protocols which this proposed research seeks to develop [8].

2.1.4 Gaps in Existing Solutions

Despite the success of existing AI chatbots, challenges remain:

- Lack of personalization and adaptability.
- Limited integration of psychological assessment tools.
- Data privacy and therapist oversight mechanisms.

In the Table 2.1, Digital technologies and artificial intelligence already promote mental health, the topic of a number of research. The study by Fitzpatrick et al, which is remarkable in that it examined the effectiveness of an AI-based chatbot that was aimed at offering mental health support, is also worth noting. Their research revealed that mental health of the users improved noticeably after exposure to the chatbot with time. The participants whose symptoms of anxiety and sadness were moderate or mild reported the chatbot, the responses of which simulated the dynamics of the therapy relationship, based on cognitive-behavioral approaches, the most helpful. The study shows the promise of conversational AI as an affordable, scalable alternative/supplement to traditional treatment.

Another interesting research was conducted by Inkster and colleagues who focused on mobile apps to treat depression. Their study learned that such applications could be rather helpful when they decrease depressive symptoms in particular circumstances. Most importantly, the involvement of the users was discovered as such an important element, as the users who actively used the app and listened to its suggestions achieved the greatest gains. The design of mental health apps, according to this research, must make usability, customisation, and motivating techniques their top priority so that the user interest and commitment to it can be maintained in long-term usage.

Sharma and Mehra used such an analysis to study the healthcare sphere, specifically, mental health monitoring using the method of sentiment analysis, thereby contributing to the profession. Based on their study, the sentiment value of people obtained by using their textual data in terms of journal entries, chat record or Facebook posts showed a positive correlation with their mental health. Good mood has been correlated with more healthy affects, whereas the more negative affect measures the more psychiatric distress tended to come up. It implies that feet analysis may become an effective non-invasive method of detecting emotional disorders at the early stages and provide early treatment.

Moreover, recent advances concerning natural language processing and emotion detection have expanded the variety of AI applications in mental health care. Research has shown that the use of emotion classification algorithms with chatbots enhances the compassionate response of a chatbot, thus increasing the level of confidence of users and the perceived help. Moreover, responses can now be tailored according to user history, mood habits and even clinical screening results (such DASS-21 scores) due to machine learning algorithms such as supervised categorization and reinforcement learning.

All these studies contribute to the growing body of evidence showing that the AI-powered technology to assist with mental health is efficient, scalable, and inexpensive as well as allude to its widespread applications, in particular chatbots, sentiment analysis tools, and mobile applications. To ensure that sensitive information is handled in a moral manner, and that the required response is taken, when needed, researchers recommend careful usage of such tools, preferably made with the help of medical professionals.

Table 2.1: Studies on AI in Mental Health

S. No.	Author	Method	Key Findings
1	**Fitzpatrick et al**	AI Chatbot for Mental Health	Demonstrated improvement in users' mental well-being [23]
2	**Inkster et al**	Mobile Apps for Depression	Apps are effective in managing symptoms if user engagement is high [24]
3	**Sharma and Mehra**	Sentiment Analysis in Healthcare	Positive correlation between sentiment scores and mental health status [25]

The review indicates that AI-based solutions are becoming increasingly effective and widely accepted for mental health support.

2.2 Overview of Mental Health Chatbots

AI products that function as chatbots serve as effective solutions to provide instant, judgement-free mental health assistance to fill a major care accessibility problem. Wysa and Woebot [5] use Natural Language Processing (NLP) along with machine learning technologies to develop dynamic, user-friendly programs that support users dealing with stress and anxiety and depression symptoms. The chatbots maintain human-likeconversation abilities through which users can freely share their mental state under a protective confidential framework. The applications offer flexible scalability to serve wider groups of users who avoid mental health help because of stigma or accessibility problems.

The study examines existing commercially available mental health applications which use chatbots by investigating their written descriptions about the different techniques for delivering mental health support. Several investigations have examined the effectiveness of these applications, though researchers have overlooked extensive evaluations about user perceptions and experiences with Wysa—the mental health conversational agent[6]. Numerous studies about digital mental health tools differ in depth and breadth, thus creating a substantial knowledge gap regarding user interactions with them.

Due to this research gap, we ran a thematic analysis, according to Braun Clarke, to obtain user experiences among those with mental health disorders regarding Wysa. A review analysis revealed four main sentiment groups, which included positive, negative, and mixed feelings with neutral ones. The classification methods showcase both the wide array of user encounters and the complex patterns that users follow when engaging with AI mental health assistance.

Research output demonstrates that users demonstrate practical insights about mental health interventions, which enhances current theoretical perspectives. Ascertaining user feelings and how satisfaction differs among them remains vital for designing enhanced mobile applications. Users express positive sentiments about fast and easy access to support, but negative sentiments come from rocketing chatbot restrictions and perceived missing human emotional connection. Users who express both positive and negative feelings about the app reveal useful insights into improvement opportunities since the application delivers useful features, yet its design will benefit from targeted enhancements.

The research aims to address existing knowledge weaknesses through a detailed examination of Wysa, which offers personalised mental health assistance via mobile chats[6].

This paper focuses on the user experience during AI-based mental health system adoption to demonstrate why user involvement determines the success of such tools. Research findings from this analysis will guide future mental health application development and create tools that fulfil the needs of individuals who require support. In healthcare, patient-centred care stands as a core principle which establishes the necessity of learning individual needs to create improved health results for patients. In the field of mental health, the principle becomes vital because personal experiences alongside preferences and values determine how effectively patients respond to treatment. The review of literature reveals that medical practitioners need to adapt mental health interventions according to personalised patient requirements for both clinical success and relevance to their recipients.

Natural Language Processing (NLP) technology advancements allow applications to make more precise diagnoses of clinical needs linked to anxiety and depression[9]. The applications examine user content through text or speech to detect symptoms, which leads them to suggest proper interventions. The AI-driven tools maintain medical care alignment to strengthen mental health assistance networks because they provide timely support to patients in need. We must conduct a critical assessment to ascertain the actual effectiveness of these mental health strategies. Developers must incorporate user feedback and assessment processes into AI service development to make continuous improvements that align with patient requirements and preferences.

Patient preferences should become a critical consideration during AI-based mental health solution creation and delivery processes. Technological systems must adapt to meet user expectations and patient values because this practice drives patient satisfaction along with active engagement. AI support effectively backs mental health care when employing patient value assessment because this practice makes interventions meet both treatment standards and the personal needs of those being served. Developers who focus on user preferences produce mental health applications which better address their needs while promoting active engagement between users and mental health care.

Developers must carefully handle clinical and User Experience (UX) biases when developing AI-based mental health solutions. Developers must monitor equity and user acceptability, ensuring that these digital tools are accessible to all affected user groups. It is crucial because the mental health support seekers include people from various life contexts with different personal histories. AI-driven mental health interventions gain effectiveness through developer efforts and proactive measures that combat biases and support inclusivity.

The latest improvements in NLP technologies lead computer applications to achieve better accuracy in identifying patients' clinical needs regarding anxiety and depression diagnosis. Applications check user content via text or speech for symptoms, after which they generate proper treatment recommendations. The tools enable continued medical care alignment to strengthen health assistance net- works through their ability to instantly assist patients in need. We need to conduct a comprehensive review to accurately assess the actual value that mental health strategies deliver. WEFS must integrate user assessment and feedback methods for AI service en- hancement to allow developers to create successive developments that satisfy patients' needs.

The decisions surrounding AI-based mental health solutions should incorporate patient choices as an essential component both in design phases and distribution services. User expectations and patient values need to be integrated into technological systems since this approach leads directly to increased satisfaction from patients while improving their active involvement. AI technologies prove beneficial to mental health care through patient value assessments, whereby healthcare interventions simultaneously fulfil professional guidelines as well as the individual requirements of patients. The development effort of applications through a user-orientated design results in better mental health software, which both improves user satisfaction and enhances their connection with healthcare.

AI-based mental health solutions need designers to develop them with close attention to the handling of clinical and UX biases because of their clinical importance. The mental health providers who seek support belong to different life circumstances and bring individual personal stories to the process. AI-driven mental health interventions work because developers use inclusive work strategies and systemic anti-bias programs to make them more effective.

2.3 Role of AI in Mental Health Care

New mental health support grows stronger through AI, as it recognizes problems early, matches patients' needs precisely, and gives users 24/7 access to help systems. AI systems that use NLP and sentiment analysis technology can read user input to spot emotional states and give personal mental health help [9]. The new digital technology aids mental health care by fixing its old system weaknesses and creating more extensive treatment options. The following list identifies how AI changes mental healthcare today.

2.3.1 Enhanced Accessibility

AI-powered chatbots work all day to offer mental health support that reaches people in distant areas without proper healthcare services. The promotion of accessibility works best for people who need mental health help because they do not have suitable care near them. These AI tools give prompt support that fills gaps in care for people who need help.

2.3.2 Early Detection and Intervention

These systems can inspect text submissions to quickly spot the initial phases of mental health concerns, such as anxiety, depression, and stress. Through text interpretation, the chatbots identify mood shifts faster to help prevent condition growth. If a person receives proper mental health treatment right when a problem appears, they can reach better results and lower their chances of major problems happening.

2.3.3 Personalized Support

AI chatbots process input data through machine learning and NLP to tailor psychological support that matches each user's specific requirements. Our system uses cognitive-behavioral, mindfulness, and mood-monitoring tools that we cus- tomize after watching users talk to them. The system delivers better service by matching support to each user's situation, which leads to improved mental health results [9].

2.3.4 Scalability

Computers with artificial intelligence have the potential to handle thousands of medical communication sessions at once, which traditional therapy lacks by design due to therapy practitioner availability thresholds. Their ability to process numerous users simultaneously enables them to serve large mental health challenges in a financially efficient manner. AI supports a wider demographic benefits mental health systems by addressing care needs for more people.

2.3.5 Anonymity and Reduced Stigma

People avoid professional help because mental health stigma and anonymous interactions work together to limit their willingness to seek it. Mental health support systems benefit from AI chatbots because users receive free judgement while sharing emotions anonymously, thus creating more mental health resource engagement. Anonymous communication empowers individuals to seek help without fear of naming or criticizing, thereby promoting mental health awareness throughout society [3].

2.3.6 Integration with Existing Therapies

Artificial intelligence technologies align with existing therapy by running scheduled maintenance for standard practice activities, including both progress recording and symptom evaluation. Therapists have the chance to utilize chatbot generated insights, which leads to better treatment plan outcomes for mental health professionals. By merging technology with expert human services, mental health professionals achieve better clinical results through an integrated medical strategy.

2.3.7 Data-Driven Insights

AI examines user data collected under privacy protocols to recognize shared issues affecting multiple populations. Analyzing user data leads to valuable insights, which function as a basis for specific mental healthcare solutions and public policies that let healthcare providers serve community requirements better. The application of data-based approaches helps direct resources to their most beneficial use and advances mental health care methods.

2.3.8 Immediate Crisis Support

AI-controlled chatbots detect crisis moments through their programmed analytics and both offer immediate help or automatically direct cases to emergency responders when necessary. Users receive necessary help during distressing situations through this built-in feature which enables immediate emergency response during critical times.

2.3.9 Cost Efficiency

Many individuals cannot afford traditional therapy because of its high-cost re- strictions. AI chatbots create an affordable medical care option which broadens accessibility to mental health services across society. AI-enabled technology reduces the expense of mental healthcare, which paves the way for universal care availability so more people can obtain necessary support.

2.4 Challenges and Ethical Considerations

Several obstacles persist in achieving proper ethical usage of AI systems throughout mental health care services. Key concerns include:

2.4.1 Data Privacy

User data privacy stands as an essential priority because mental health information requires maximum protection when patients share their sensitive information. Security protocols must be fully implemented by developers to protect user data privacy and meet the standards set by GDPR and HIPAA.

2.4.2 Algorithmic Bias

A fundamental challenge in AI systems arises when their operation maintains and reproduces biases that already exist in the data used for training. The development of AI tools requires diverse dataset utilization because it helps eliminate bias and increases equitable care access across every user population. Systematic testing should also occur to achieve these goals

2.4.3 Clinical Validation

Professional healthcare validation requires thorough research consisting of randomized controlled trials to prove the effectiveness of AI-driven mental health interventions. Ongoing assessment procedures should establish the safety mea- sures and treatment effectiveness of these tools which need to measure up to traditional medical standards.

2.4.4 User Acceptance

Acceptance from users stands as a necessary factor which makes AI solutions work effectively. The successful design of AI mental health tools needs a deep understanding of how users feel about these solutions and their thoughts regarding their preferences alongside their therapeutic goals. Through user involvement during development researchers can build solutions which function well, and suit users based on their needs.

2.5 Comparative Study of Mental Health Assessment Tools

Mental health assessment tools play an essential role in building diagnostic frameworks because they directly determine their effectiveness. These tools represent vital elements in both condition detection and treatment planning for mental health and they also enable monitoring of ongoing patient progress. Of all mental health assessment instruments, the DASS-21 has gained widespread use because of its streamlined format which makes it better than the DASS-42 along with widely adopted tools GAD-7 and PHQ-9[8].

2.5.1 DASS-21 vs DASS-42

DASS-21 represents an shorter version of the DASS-42, which was created to minimize workload during assessment without compromising its mental strength. Research findings prove that DASS-21 produces results equal to those of the DASS-42, thus making it an advantageous instrument for clinical work as well as digital platforms [8].

2.5.2 User Engagement

The main benefit of using DASS-21 is that it gets people involved. On the other hand, the DASS-42 is too long, which makes it hard for people to complete the survey because it takes longer and requires less participation, especially when chatbots are used in non-clinical settings. People often get tired of grading long questionnaires because the process takes too long, and they give up on finishing the assessments.

2.5.3 Efficiency in Assessment

The brief nature of the DASS-21 assessment makes a significant contribution to holding users engaged throughout conversational interactions. Users' complete assessments more efficiently on digital platforms because of shortened attention spans, thus leading to better user experience and improved honesty in response collection.

2.5.4 GAD-7 and PHQ-9

Both GAD-7 and PHQ-9 serve as standard evaluation instruments for generalized anxiety disorders and depression identification. These effective tools function within specific fields, but they remain limited in delivering a full understanding of the mental wellness picture for each person.

2.5.5 Single Condition Focus

The independent assessment scales GAD-7 and PHQ-9 fail to identify how anxiety and depression levels can influence each other, especially when triggered by stressors. DASS-21 provides users with an evaluation of depression combined with anxiety and stress, creating a comprehensive mental health assessment. The AI chatbot's integrated structure functions best through this multidimensional assessment method since it delivers focused support according to a user's complete mental health record.

2.5.6 Why DASS-21 is Better

DASS-21 achieves an effective blend of detailed assessment while maintaining efficient evaluation, which makes it optimal for digital intervention usage. The system performs effortlessly while conducting swift evaluations in situations that heavily depend on user participation and response speed.

2.5.7 Tailored Recommendations

Users can receive custom recommendations through DASS-21 because the system uses mild, moderate, and severe categories for analysis results. The system of symptom categories helps AI chatbots generate targeted feedback for users whose symptoms fall at different severity levels. DASS-21 identifies users who face severe anxiety so they get immediate coping or resource guidance, whereas users with milder symptoms get recommendations for self-care practices.

2.5.8 Enhanced Functionality in Digital Platforms

The DASS-21 provides distinct advantages to digital platforms because they must maintain high user participation. The DASS-21 assessment duration enables more users to participate in mental health programs because they complete the assessment more frequently. AI chatbots require effective user input to generate appropriate interventions; thus, this capability becomes crucial in their operation.

2.5.9 Robust Psychometric Properties

Though short in length, the DASS-21 stays robust under psychological eval- uations, making it a dependable tool for mental health assessment. Multiple investigations show strong similarities between DASS-21 scores and established measures of mental health, which confirms its role as a screening assessment tool [8].

2.6 Existing Research and Gaps

Scientists show that chatbots can deliver care easily, yet fundamental issues block their full development and use in mental support. The effective adoption of digital mental health interventions depends on solving these current gaps in their development [20]. There exists a need to research these essential topics:

2.6.1 Long-Term Efficacy

Most research investigating mental health chatbot systems concentrates on temporary symptom reduction of anxiety and depression during and directly after using the programs. There is very little existing research about the long term performance of these interventions after their application has ended. The long term impact of chatbot encounters reveals how many advantages these interactions continue to provide after a person initially becomes involved with them.

2.6.2 Need for Randomized Controlled Trials (RCTs)

Building the long-term effectiveness of mental health chatbots through research requires the use of randomized, controlled trials that receive proper validation. We need to measure both the short-term symptom management results and the longer-term effects of these interventions within weeks and months. The research of long-term chatbot intervention effects will en- able better understanding of their position in maintaining continuous mental healthcare support.

2.6.3 Personalization and User Experience

We have not given much attention to understanding how mental health chatbots could modify their responses to meet individual consumer requirements. Modelling according to personal characteristics proves vital for improving user satisfaction because users respond best to intervention strategies that connect with their individual preferences and confrontations.

2.6.4 Dynamic Feedback Mechanisms

The development of chatbots requires people with user feedback capabilities to optimize personalization levels. The bot system should modify its output when users respond along with their profile information and declared results. Research requires identifying best practices that deliver personalized inter- action methods to create understanding and support for users throughout their mental health care processes.

2.6.5 Inclusion of Diverse Populations

Most current research studies youth and technology-orientated participants, although it fails to consider or address the unique requirements of different com- munities. There is a lack of extensive research on how chatbots interact with individuals from diverse age ranges, cultural backgrounds, and socioeconomic levels.

2.6.6 Cultural Competence

The development of efficient chatbot interventions for mental health depends heavily on understanding cultural influences that can shape treatment choices and mental health perspectives. Researchers must work to conduct studies with increased representation among participants because it will help develop culturally aware and appropriate chatbot systems. The addition of diverse user bases will guarantee mental health chatbots provide satisfactory results to all users regardless of their personal backgrounds or experiences.

2.6.7 Age and Socioeconomic Diversity

Research must examine the capability of mental health chatbots to communicate with older adults, together with individuals who demonstrate low technology skills and members of different socioeconomic strata. Making mental health chatbots suitable for these groups improves accessibility and effectiveness, which produces better mental health results.

2.6.8 Dataset Limitations

The use of proprietary datasets, along with limited information in research on mental health chatbots, reduces the ability to generalize study results. Inadequate database diversity prevents the creation of powerful AI models that represent mental health intricacies in numerous population groups.

2.6.9 Expanding Open-Access Datasets

We must develop and distribute open- access datasets as essential resources to address the current generalizability issues. To train legitimate AI models that process representative information regarding age, gender, ethnicity, and socioeconomic status, big data must include multiple demographic variables. Reliability, together with the prac- tical usage of chatbot interventions across diverse user groups, will improve through this implementation.

In Table 2.2, Over the past years scientists have explored numerous AI solutions to manage various aspects of mental health, with positive yet intricate outcomes. A 2020 study used NLP, machine learning, and sentiment analysis to develop a chatbot to manage stress levels by monitoring the emotional state of the users and offering appropriate stress coping skills [25]. The high level of effectiveness of which the chatbot was capable of identifying stress and offering immediate help displayed the growing responsibilities of AI-related technologies in stress. The fact that it failed on treating co-occurring mental health disorders such as anxiety or bipolar illness was by far a disadvantage. Besides, the system did not have high-level real-time responsiveness and it did not lend itself to high stress dynamically demanding tasks.

The combination of Convolutional Neural Networks (CNNs) and deep learning was applied in the field as an attempt to diagnose depression in 2021 [26]. The outstanding results of diagnosis precision found by this AI model justify the application of deep learning models in mental health diagnostics. The issues of generalizability on different populations were addressed in terms of the complexity of the emotional scenarios or overlap that the system has a problem with and in the homogeneity of the training data, especially in terms of cross-cultural and cross-linguistic differences. To locate the emotions of the users in real-time and provide personalized instructions, the emotion detection in mental health earlier in 2019 study relied on NLP and emotion analysis [20].

This approach helped to define the significance of emotional awareness in the environmental of a digital treatment. The support systems it had are however not permanent since there was no supervision and long-term emotional support. Also the system fared relatively poor with results showing a lot of weakness when the users themselves put ambiguous or inconsistent data into the system which restricted the varying dependability of the system in realistic conditions. The most advanced study was reported in 2022, when Recurrent Neural Networks (RNNs) were applied to develop an anxiety-reducing chatbot. After the first contact, the symptom of anxiety notably lowered after user feedback, proving that AI is capable of fulfilling the role of emotional support [18]. However, to prepare the model with trustworthy training, it required excessive user interaction data. The fact that RNNs operate as a black box and conceal the explanations behind AI decisions has fueled the concerns over the lack of transparency and trustworthiness on the part of the user [12].

The 2023 paper by researchers is the most recent application of reinforcement learning to supply personalized mental health treatments. These adaptive systems played an important role in forming the relevance and efficiency of mental health treatment because they enabled therapeutic responses to be altered dynamically in the same way as the behavior of care seeker individuals began to act. Despite these benefits, the model was also not very capable of supporting complex co-morbidities, including co-occurring mood or anxiety disorders, and it was also hard to adapt it in real-time, especially when one was experiencing changes in their emotional state [20].

Table 2.2: AI Applications in Mental Health (2019–2023)

S. No.	Year	Research Topic	AI Method Used	Key Findings	Limitations
1	2020	Chatbot for stress management	NLP, Machine Learning, Sentiment Analysis	The chatbot successfully identified user stress levels and provided effective coping strategies, demonstrating the potential of AI in stress relief [25].	Failed to address co-occurring conditions like anxiety or bipolar disorder; real-time responsiveness was limited.
2	2021	Depression diagnosis using AI	Deep Learning, Convolutional Neural Networks	The AI model achieved high accuracy in diagnosing depression, showing the reliability of deep learning models in mental health diagnostics [26].	Struggled with detecting complex or mixed emotions; lacked cultural and linguistic diversity in training data.
3	2019	Emotion detection in mental health	NLP, Emotion Analysis	The system enabled real-time emotion detection and delivered contextual advice, highlighting the value of emotional understanding in therapy [27].	Did not provide long-term emotional support or follow-up mechanisms; performance dropped with inconsistent user input.
4	2022	AI-based chatbot for anxiety relief	Recurrent Neural Networks (RNN)	Users reported a noticeable reduction in anxiety after interacting with the chatbot, showcasing AI's effectiveness in emotional support [18].	Required large-scale user interaction for training; struggled with ethical transparency and user trust due to black-box models.
5	2023	Tailored mental health interventions	Reinforcement Learning	AI-driven personalized interventions led to more tailored therapy approaches, improving the quality and relevance of mental health support [20].	Failed to address co-occurring conditions like anxiety or bipolar disorder; effectiveness in real-world use was limited.

Chapter 3
METHODOLOGY

3.1 System Modules

The system is composed of the following four main modules

- DASS-21 Questionnaire
- Daily Personalized Tips Module
- Emotion Classifier
- Transformer Based Chatbot

3.1.1 DASS-21 Questionnaire

The app uses the DASS-21 questionnaire, which includes 21 questions divided into three categories .

Table 3.1: DASS-21 Categories and Their Purpose

S. No.	Category	Number of Questions	Purpose
1	**Depression**	7	Measures dysphoria, hopelessness, and devaluation of life
2	**Anxiety**	7	Assesses situational anxiety, autonomic arousal
3	**Stress**	7	Evaluates nervous arousal, difficulty relaxing

In this Table 3.1,When developing psychological apps based on AI, psychological assessment tools are critical in establishing the severity levels of emotions such as stress, anxiety, depression, etc. A frequently used tool to do this is the DASS-21 with exactly 21 questions divided into three main sections such as stress, anxiety and depression. There are seven questions in each category.

Depression Scale is aimed at evaluating moods related to devaluation of the life, dysphoria, despair, and apathy. These questions explore the mood of the user in the past seven days, which helps to identify such symptoms as poor self-esteem, anhedonia (the lack of interest or pleasure), depressive feelings, or feelings of worthlessness. The usefulness of the tool lies in the fact that it quantifies such reaction and classifies the stage of depression symptoms to be normal, mild, moderate, or highly severe.

The Anxiety Scale assesses situational anxiety, symptoms of panic as well as physical indicators such as dry mouth, tremor or palpitations. Such symptoms are associated with physiological ramping and fearful responses. To classify anxiety disorders among other forms of emotional abnormalities, such questions aim at determining cognitive and physical signs of anxiety. This measure is often reported as high when a person is highly autonomic and tends to worry or panic when face with new or new dangerous situations.

The Stress Scale assesses such symptoms as anxious alertness, impatience, and inability to relax, which are associated with the long-term effects of stress and coping disorders. Users can complain about impatience, overwhelm or irritability. These products are especially useful in the case with those who are possibly not clinically depressed or anxious but are still under much psychological pressure, which is sometimes caused by the demands of their work, lifestyle, or education. Such special tools, as DASS-21, allow AI systems to more effectively evaluate an input of a person by dividing mental health symptoms into three categories.

This enables the chatbot to give applicable and personal suggestions as per the mood of the user. In addition, incorporation of data-driven insights into offering both preventive and supportive mental treatment in terms of the structured assessment into an AI chatting robot facilitates a better clinical reliability of the system.

3.1.2 Daily Personalized Tips Module

Our mental health chatbot comes with the module that is used to provide specific wellness recommendations, depending on the psychological evaluation performance in the chatbot. It combines two major data of mental state In Fig 3.1 DASS-21 Scores, A self-report validated scale of Depression, Anxiety, and Stress. Emotion Classification Scores, Obtained by a 16-class emotion detection model based on BERT.

The inputs are processed through the system to come up with condensed and practical mental health insights on how the users can use daily. Tips have all been put together based on advice with the licensed psychologists and are devised to be clinically correct and ethically pertinent.

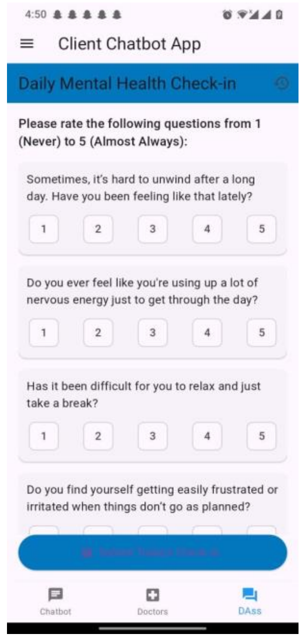

Fig 3.1: Depression Anxiety Stress Scale

3.1.2.1 Personalization Logic

The module takes into consideration an ample number of emotional and psychological states. For example

- A user might be **anxious** without any depression and stress.
- There are cases when a user can have both signs of **stress and anxiety** simultaneously.
- There are users who can exhibit **all three disorders** at the same time, which includes depression, anxiety, and stress.
- The rest might have **mild or low scores** and they only need to be given general wellness encouragement.

The reasoning behind tip generation implies the dynamic mapping of the severity thresholds in DASS-21 and the categories of emotion to the desired theme of advice, including cognitive reframing, grounding activity, physical exercise, journaling, or breathing techniques.

3.1.2.2 Example Scenarios

In various mental health situations, tips are created based on the user analysis of the DASS-21 scores and their emotional status. As an example, in Case 1, when the user presents moderate levels of anxiety and such emotions as feeling overwhelmed and nervousness, the system can propose the practice of the 4-7-8 breathing technique. This breathing exercise is concentrated and makes one have control over the nervous system and reduces levels of cortisol in the body.

In Case 2, when the user is extremely depressed and feels hopeless, they can be provided with a tip about setting a small goal, e.g., brushing their hair or going outside, as even the tiniest step toward self-help is what matters.

In case 3, when the three DASS-21 categories are all rated high and the feelings of feeling tearful and being irritable are given, it is recommended that the user finds somebody he trusts so that he can open up and get the psychological burden off his mind. Also, it is advisable to repeat undertaking the DASS-21 questionnaire after some days.

Finally, in Case 4, when the user demonstrates low scores and emotions such as collectedness or indifferently, the recommendation aims at preparing him/her to stay emotionally balanced by writing about something that he/she is thankful about right before bed, which imparts a notion of mindfulness and gratitude.

3.1.2.3 Professional Oversight

All of the developed tips were typed, proofread and confirmed by mental health professionals. This ensures

- Tips are safe and evidence-based so that they can be applied by themselves.
- Recommendations exclude the possibility of possibly offending language and unproven habits.
- The suggestions are responsive and sensitive to shifts of emotional or psychological mood with time.

3.1.2.4 Ethical Safeguards

The module does not give medical diagnosis. The system will reinforce the user to consult a mental health professional in cases of high-risk emotional status without compromising user security and market ethics.

3.1.2.5 Emotion Classification and Sentiment-Based Mental Health Monitoring

To constantly assess the emotional states of users, and provide the related feedback or alarms, the powerful pipeline to classify emotion was designed in this paper and integrated into the AI-based mental health care system. This pipeline is based on the Emotion Classifier, a deep learning-trained NLP model of a 16-class multiclass classification task to detect subtle emotional cues contained in text generated by the user.

3.1.2.6 Emotion Classifier Overview

The emotion classifier is a supervised learning model that was improved on using a labeled dataset containing 16 various emotional states that are common in a mental health environment. They are selected among the fundamental and complex emotional states relying on psychological studies and on the actual pattern of expressivity. Some of the 16 categories of emotion are

- **Positive Emotions** Joy, Love, Gratitude, Relief, Pride
- **Neutral Emotions** Neutral, Curiosity, Anticipation
- **Negative Emotions** Anger, Sadness, Fear, Shame, Guilt, Disgust, Anxiety, Confusion

The classifier takes the input of each user response in the dialogue of chatbot, and returns a softmax probability distribution across the entire 16 classes. In that input, the predominant emotional tone will be selected among which the emotion with the highest score of confidence.

3.1.2.7 Sentiment Mapping and Severity Scoring

All of the emotions are assigned a more broadly defined sentiment polarity (Positive, Neutral, or Negative) and are then assigned a severity score to allow interpreting these categories of emotions in the context of levels of mental health risks. The rating of the degree of urgency or danger of an interaction is determined. The mapping of sentiments to scores can be found in Table 3.2.

Table 3.2: Sentiment Score Interpretation

S. No.	Sentiment	Score Range	Interpretation
1	**Positive (Happy)**	0.0 – 0.3	Low severity
2	**Neutral**	0.3 – 0.7	Moderate severity
3	**Negative (Angry, Sad and Fearful)**	0.7 – 1.0	High severity

As an example, the identification of the annotation of the classifier of ”Sadness” as 0.85 would be considered to be high severity. The mile severity would be categorized as 0.2 gratitude score, however.

3.1.2.8 Weekly Aggregation and Therapist Integration

Each of the sessions has a record of all the emotion labels and the sentiment value that accompanies it. These records are summarized every week by any user. The system identifies the variations in frequencies of unpleasant emotions and the average weekly score on severity, and trends such as emotional instability or predominance of specific feelings (recurring ”Fear” or ”Guilt”). These insights are combined and used to generate a weekly mental wellness report that is sent safely to the appropriate therapist (assuming the user has opted to have his/her therapy reviewed). Part of this report includes

- Weekly emotion that is most common .
- positive-negative emotion ratio .
- Trends in severity (increasing, declined or stable).
- Red flags about potential life threatening states (e.g. persistent anxiety or depression).

3.1.2.9 Model Training and Dataset

The Emotion Classifier was refined by a composite dataset including not only custom annotated sets of mental health talks but also public emotion corpora (e.g. GoEmotions, EmotionLines). The total number of utterances used in the final training set was more than 20 000, with at least one tagged utterance from each of the 16 courses in the proximity of the fur the least characterized utterances. The model architecture is based on BERT (Bidirectional Encoder Representations from Transformers), due to its contextual awareness, the model is able to recognize emotion in short or ambiguous signals.

3.1.2.10 Ethical Handling and Data Privacy

The security of the data of any protection regulation is thus strictly followed by the anonymization and encryption of all training and live classification data (GDPR-compliant architecture). The user can choose whether to go through their reports manually or not after being informed fully about emotion tracking use.

3.1.2.11 Hyperparameters

Table 3.3 depicts that the BERT-based emotion classifier model was trained after a set of needed hyperparameters carefully chosen. This set of hyperparameters had to be optimized in iterations to deliver best results in a difficult multi-class emotion detection task. The initial model was the pretrained transformer model BERT-base-uncased that exhibits a good performance in a variety of natural language processing tasks. This model was well aligned to our requirements because of its rich bi-directional encoding and linguistic comprehension of the language.

Since there were sixteen categories of emotions, the classification configuration needed a model that could deal with the high dimensional data that was quite complex. Such enormous amount of classes made the model very complex and it required good set of hyperparameters configuration. BERT tokenizer was used to preprocess the input textual data which involved converting the text data into IDs of the tokens. To make all input samples a similar length, it was necessary to pad inner sequences so that the longest sequence in a batch has at most 128 tokens, and in some experiments this was extended to 256 tokens so that more context is kept in longer sentences.

A batch size of 32 was first picked to achieve a compromise between memory efficiency and throughput on training. Nevertheless, subsequently, a batch size of 16 proved to perform better during certain training runs due to the GPU memory limitation and stability of a thermal gradient reported in Table 3.3. The model was optimized during 4-5 epochs, and 2e-5-3e-5 learning rate was adopted according to the typical strategies of BERT optimization. Linear learning rate scheduler with warmup (usually a fixed value of 500) was used to reduce the threat of instability at the initial stages of training.

AdamW optimizer was chosen over others, since it was known that it combines well with weight decay than allowed gradient updates more generally, on its own. Moreover, the overfitting was avoided by randomly deactivating the neurons just to ensure only a minor degree of dropout (0.1). Weight decay was set at 0.01 which assists to regularize the model and eliminated overfitting.

Table 3.3: Hyperparameter Settings for Emotion Classifier

S. No.	Hyperparameter	Value / Setting
1	**Model Name**	Emotion Classifier
2	**Number of Classes**	16
3	**Max Sequence Length**	256
4	**Batch Size**	16
5	**Learning Rate**	3e-5
6	**Epochs**	5
7	**Optimizer**	AdamW
8	**Loss Function**	CrossEntropyLoss
9	**Dropout Rate**	0.1
10	**Weight Decay**	0.01
11	**Warmup Steps**	500
12	**Gradient Clipping**	1.0
13	**Evaluation Metric**	Accuracy, F1 Score

The use of gradient clipping was implemented to provide required stability in training where the maximum norm of 1.0 was used to avoid the exploding gradients during backpropagation. A loss function that has been applied was CrossEntropyLoss, which is the standard loss function used in multi-classification issues since it can be used to punish inaccurate predictions with weights based on the confidence levels of each prediction.

These hyperparameter values are chosen after significant experimentation and been in accordance to the best practise of transformer based models. Coupled with each other, they made it possible to create a neutral stable and high-performing system of emotion classifications that would be multi-label based and input texts of various lengths.

3.1.3 Transformer-Based Chatbot

The conversational model that has been developed is an AI-enhanced system, which will assist people with stress, anxiety, and depression in an emotionally sensitive and context-sensitive dialog. It is a wide model that integrates a BERT-based emotion classifier that can recognize 16 unique emotional states given their input to a generative response model built on a transformer architecture. Furthermore, the quality of response creation is improved using a BART (Bidirectional and Auto-Regressive Transformer) model that takes the advantages of both methods of encoding and decoding.

3.1.3.1 Dataset Collection and Preprocessing

To discard the effect of systematic bias, four publicly available and domain specific dialogue corpora were integrated into a composite dataset and this was utilized in training an emotionally intelligent and contextually consistent chatbot

3.1.3.2 EmpatheticDialogues

This dataset contains over 25,000 human-to-human dialogues constituted by emotional context composed by people and each of them will have one of the 32 emotion types as a label. It is regularly severed to educate conversational robots to produce sympathising responses.

3.1.3.3 Mental Health FAQs

A tailored dataset consisting of nearly 700 most popular mental health questions including questions in relation to stress, anxiety, depression, and self-care.

3.1.3.4 MentalHealth 16k Counseling Chat

This dataset consists of about 16,000 anonymised multi-turn text in-real-life counseling sessions and they simulate actual conversation in the field of stress, depression, and anxiety management.

3.1.3.5 DailyDialog

The dataset is of high quality and generic grammatical accuracy that provides speech samples in establishing the model in daily encounters. It contains more than 4,000 conversations founded on the daily communication situations.

This was converted into a consistent format in the form of contextresponse pairs that can be used to train sequence-to-sequence. The pipeline of preprocessing comprised the following steps

3.1.3.6 Tokenization

All the text was tokenized with the BART tokenizer in Byte-Pair Encoding (BPE) form to overcome compliance with the pretrained embedding space of the BART model.

3.1.3.7 Data Cleaning

Irregular samples, hyperlinks, special characters, repetitious entries and non-English statements were removed. In order to maintain consistency in training dynamics, utterances containing unforeseeably short or long sequences were filtered.

3.1.3.8 Emotion Conditioning

At the stage of EmpatheticDialogues, the emotion-informed generation was enabled by keeping the corresponding emotion label and even including it to the background as an additional token in the conditional training context.

3.1.3.9 Model Architecture

The architecture of the chatbot is BART (Bidirectional Auto-Regressive Transformers) which is the denoising sequence-to-sequence transformer developed by Facebook AI. Specifically, it is very applicable to answer generation tasks since it uses both the strengths of BERT (ability to understand the context bidirectionally) and GPT (autoregressive generation capability).

3.1.3.10 Base Model

Bart-large, which has 12 encoder hotels and 12 attachments hotels, where each hotel has 16 attention heads of the same dimension 1024, so the total parameters are nearly 406 million parameters.

3.1.3.11 Encoder

receives as input a tokenized (context) user input that can optionally include emotion annotations and contain a fixed number of special tokens including [CLS] and [SEP]. To create the contextualized representations of the hidden layer, the encoder employs self-attention tools to analyze the entire input sequence in the forwards and backwards directions.

3.1.3.12 Decoder

is an auto-regressive behavior because any output is determined by the encoder output and other tokens that had been created. It absorbs contextual details of the encoder through cross-attention and ensures its causality through the disguised self-attention.

3.1.3.13 Positional Embeddings

Both the encoder and the decoder make use of learned positional embeddings to memorize the sequence of the input and output tokens.

The final token probabilities are obtained by using a linear layer and a softmax layer over vocabulary output of the decoder at every time point.

3.1.4 Training Process

To enhance the model, the back end of PyTorch in Hugging Face Transformers library was adopted as end-to-end for training. The merged data was fine-tuned with the purpose of adjusting BART to compose multi-turn emotional responses.

3.1.4.1 Batch Size

Batch size set to 8 with gradient accumulation across a number of steps to make training stable on limited memory and allow re-creating high effective batch sizes.

3.1.4.2 Learning Rate

An initial learning rate of linear schedule 3e-5 with linear warm-up 5e-3 was used here to make learning rate vary gradually to prevent early-stage divergence in training.

3.1.4.3 Epochs

In order to avoid overfitting the model was trained to as far as give 5 epochs and then the training was stopped due to the validation loss plateau.

3.1.4.4 Optimizer

AdamW (Adam with decoupled weight decay) was used. AdamW is known to be effective on transformer models so it was used to provide parameter updates.

3.1.4.5 Loss Function

The token predictions made by the decoder were put through the CrossEntropyLoss, in which it ignored padding of token masking.Evaluation Metrics

- **Perplexity (PPL)** utilized as the main quantitative indicator of the caliber of language modeling.
- **BLEU Score** used to evaluate the produced and reference replies' n-gram overlap.
- **Qualitative Evaluation** The produced replies are manually reviewed and scored for emotional coherence, relevancy, and fluency.

Furthermore, the best-performing model was chosen based on the lowest validation perplexity, and model checkpoints were stored on a regular basis.

array

3.2 Firebase Integration

Google also has Firebase, which is a powerful backend-as-a-service (BaaS) platform that can offer a multitude of different tools and infrastructure to support effective application construction, deployment, and scaling. Using Firebase as a data storage, user management and therapist interaction modules is a central component of mental health support application as shown in Fig 3.2. The fact that it easily integrates with both mobile and web application, uses real time information in its database and security features makes it very suitable to apply in this project where applications with privacy, real time communication and structured access are critical.

Table 3.4: Hyperparameter Settings for BART Model

S. No.	Hyperparameter	Value / Setting
1	Model Name	BART-large
2	Batch Size	8 (with gradient accumulation for effective larger batch size)
3	Learning Rate	3e-5
4	Learning Rate Scheduler	Linear decay with warm-up
5	Epochs	Up to 5 (with early stopping based on validation loss)
6	Optimizer	AdamW (Adam with decoupled weight decay)
7	Loss Function	CrossEntropyLoss (applied over token predictions, ignoring padding tokens)
8	Evaluation Metrics	- Perplexity (PPL) Language modeling quality - BLEU Score N-gram overlap with reference responses - Qualitative Evaluation Human-rated coherence, relevance, and emotional appropriateness
9	Gradient Clipping	Typically set to 1.0 (optional for stability)
10	Warm-Up Steps	Optional – used to avoid learning rate spikes in early training
11	Tokenizer	BART Tokenizer (Byte-Pair Encoding)
12	Input Format	Context-response pairs with optional emotion conditioning tokens

3.2.1 Storing Chat Logs

Conversational aspect between the user and the AI chatbot is one of the basic features of the application. Every chat session is also stored securely in Firebase Firestore or Realtime Database to guarantee continuity and personalization and possible future reference as shown in Fig 3.2. Such logs contain the queries made by the user, the responses generated by the chatbot, the time and date stamps in addition to the emotional labels (which are predicted by the emotion classification module). Storing this data in Firebase will allow the application to access historical sessions to give personal recommendations and do longitudinal long-term behavior analysis on the users. The

Fig 3.2: Firebase in realtime

information that is also stored through this helps in the supervised monitoring and reflective therapeutical practices.

3.2.2 Maintaining User Profiles

Firebase authentication is used in setting up a user sign-up, signature, and safe identity validation and secure connection where each user can be made individual and elusive. Besides authentication, the user profile covers information critical to the emotional state, like the emotional history, DASS-21 scores, daily tips provided, and mood trends. This data is organized and recorded in Firebase Firestore making querying and analysis to be done efficiently and at scale. It is made sure by the real-time synchronization of Firebase that any update on the profile of the user like mood changes, or response to evaluation, or the record of interacting time will be instantaneously updated with the devices and equipment used.

3.2.3 Enabling Therapist Access

The option that is crucial to the application is connecting users to licensed therapists. Firebase allows access by therapists through the role-based access scheme, where by the therapist only gets controlled, read-only, access to the anonymized or consent-based user data. This consists of chat transcripts, emotional trend charts and assessment history that can be provided via a secured portal. The privacy and ethical standards are met because

Firebase Cloud Functions and Firestore Security Rules make sure that only the data that the therapists must see can be seen by the therapists. This integration facilitates the ability to conduct proactive interventions, individual feedback assessments, and better therapeutic decision-making by the therapist.

On the whole, the Firebase is the supporting mechanism of the data infrastructure of the application. The presence of real-time functionalities, safe user management, and a flexible data structure accommodate the key functionality of the mental health chatbot framework that simultaneously does not violate standards of ethical behavior regarding user privacy and data security. Another way that Firebase makes a project development process smooth is that it integrates easily with Flutter which this project works on to develop the frontend.

3.3 Assessment Module Workflow

The evaluation tool of the chatting application to mental health is arranged to overview the emotional health of the user on a daily and orderly basis. It exploits DASS-21 (Depression, Anxiety, Stress Scales-21 item) structure and fits it in a form of the conversation to increase the user experience and support the regular evaluation. The procedure of this module, step by step, is presented in Table 3.3.

3.3.1 Daily Prompt

Every day, the app starts a short evaluation process by asking the user of three questions based on the DASS-21 questionnaire. These are questions that are generated dynamically so as to provide balanced coverage across the dimension of depression, anxiety and stress. The questions are structured in a friendly and understanding fashion to minimize the tension and give aggressive answers. The daily check-in process gives the opportunity to have a low-load system that follows your emotions and does not overload the user.

3.3.2 Response Capture

During the running times as the user uses the chatbot by responding to the daily challenges, these responses are recorded as they come as textual data. Analysis is then done on these responses that are pre-processed. This measure will make sure that the raw data can be subjected to further emotional interpretation, leaving context intact and linguistic nuances. Recorded answers are encrypted in Firebase and are kept in a time-logged record, of emotional transcript.

3.3.3 Emotion Scoring

After responses are obtained, they are directed to the Emotion Classifier, a layer-wise refined NLP model, which is able to perceive and score emotions in various categories. The model has a score of each user response depending on an emotion (e.g. happy, sad, anxious, stressed) showing the level and kind of emotion that might be portrayed based on the users. The scores are then normalized and saved accompanied with metadata like timestamp, question Id, and user ID.

3.3.4 Weekly Aggregation

The system also computes a seven day total of daily emotional ratings in order to provide a wider clinical picture. Statistical summaries to be included in this aggregation process are mean, trend direction and frequency of emotions. The outcome is a weekly emotional profile which enables clinicians and therapists to acquire an insight into the psychological course of the user. Such summaries can also be used in providing automated feedback or alerts in case critical limits are exceeded.

One can say that this limited but still structured working process will enable the application to offer constant emotional tracking without being too bothersome. It integrates the usual psychological assessment tests with the latest techniques of NLP, thus being both clinically applicable and user friendly.

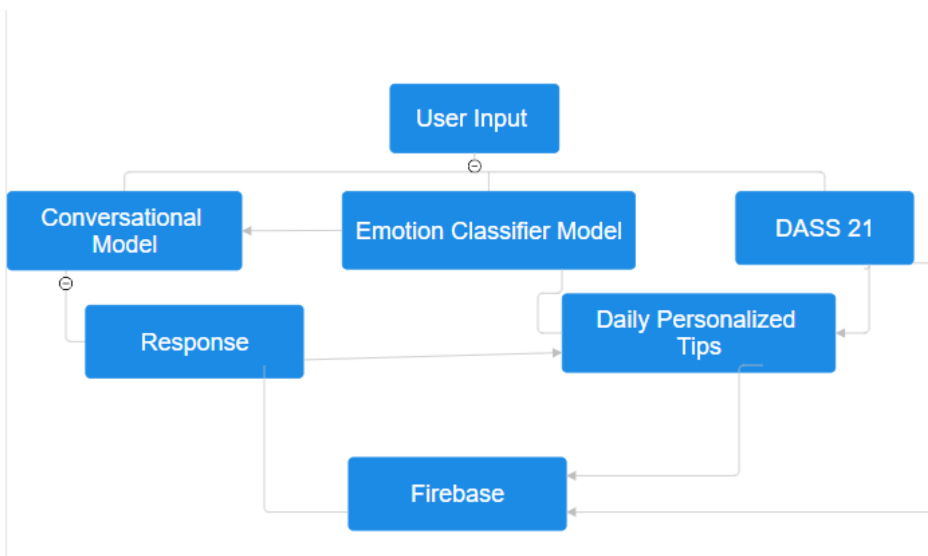

Fig 3.3: Flow daigram

3.4 System Workflow Diagram Explanation

The flow diagram shows the entire processing pipeline of the mental health chatbot system which begins with user input and comes out as emotional scoring, response, DASS-21 scores, and storage in Firebase. The architecture thus brings to fore the interaction of various models and modules that collabourate with each other in an unified manner to provide a personalised and supportive mental health journey.

3.4.1 User Input Text

The customer expression of interest in the system is through typing. This input may be a general message, an answer to some question or the mood of the user.

3.4.2 Emotion Classifier Model

At the same time, the user input is forwarded to the Emotion Classifier which is a fine-tuned NLP model that is trained to assess and identify emotional states, including but not limited to anxiety, sadness, happiness, or anger. The provided model assigns an input text a label or score corresponding to an emotion, which is necessary to make downstream decisions, including creating a tip or notifying a therapist. The identified emotion will also contribute to the situating of the next responses of the chatbot.

3.4.3 Daily Personalized Tips

Depending on the detected Emotion, the system creates Daily Personalized Tips. These tips are supposed to be sympathetic, friendly and provide the alignment with the mood the user is in. To illustrate this, in the event that the sadness or anxiety is identified, the system can recommend mindfulness activities or empowering statements. The tips will be recorded and forwarded to the user in addition to being backed up in Firebase to track previous records.

3.4.4 Conversational Model

In tandem with the detection of the emotional state in the conversation, the speech of the user is also passed on to the Conversational Model, which is a generative language model that is trained to be fine-tuned to mental health support dialogue. Such a model creates a natural and context sensitive response to the user. It also adopts the functionality to automatically determine whether there is a need to ask a DASS-21 Question particularly when it is within the daily assessment schedule.

3.4.5 DASS-21 Questionnaire

Questions (DASS-21; a pool of 21 questions covering depression, anxiety, and stress) is available to be administered in case of applicability of the Conversational Model. The answer of the user to this question is recorded and interpreted to calculate the DASS Scores which measures the condition of the user over a period of time. The scores are part of the vital evaluations that are done on a weekly basis and can be discussed with therapists that might offer clinical understanding.

3.4.6 Response

The Response that the model generates is posted to the user whether it is a general chatbot response or a follow-up question. This is also a record to be kept in tracking interactions, the nature of behavior and subsequent planning.

3.4.7 Firebase Integration

All the relevant outputs in terms of labels related to emotions, the responses of the chatbot, and daily tips as well as the DASS-21 scores are saved securely in Firebase As shown in Fig 3.2. Firebase is the backend which supports real time Session logging

- Maintenance of user profile
- History of tip and scores tracking
- Access of therapists (consented)

It is possible to have synchronized, secure, and scalable data management with Firebase cloud data storage and Firestore database so that both users and therapists can access meaningful historical data at the certain moment they require it.

Frontend Mobile App: Android-based interface that interacts with users.

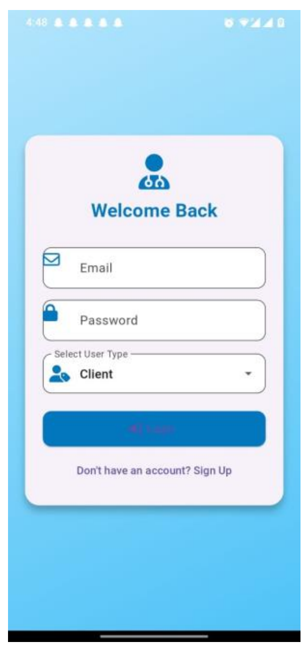

Fig 3.4: Interface of chatbot

Backend Server: Handles data processing Emotion Detection and response generation.

3.5 Firebase Integration for Personalized Advice

The secure database of the chatbot enables storage for users' DASS-21 scoring outcomes, which improves personalized care. The database retrieves recommendations and wellness hints tailored to individual patients based on their score severity (mild, moderate, or severe). The support recommendations adjust specifically to each user's mental health while receiving fresh wellness content every day.

3.5.1 Database Design

User profiles exist with the database schema, which enables an orderly organization of mental health information storage and retrieval capabilities. The design of this schema includes several fields, including.

3.5.2 User ID

Each user has an individual label as part of the system that allows precise data storage and retrieval operations.

3.5.3 DASS-21 Score Categories

The database contains user DASS-21 stress, depression, and anxiety scores, with severity ratings ranging from mild to severe.

3.5.4 Personalized Advice

The database stores user-specific guidance and recommendations by combining DASS-21 results with logged mood data.

A user can obtain relevant suggestions from the database by querying it with information from their current DASS-21 measure and mood recording data. Regular system updates ensure users receive effective mental health support at the right times.

3.5.5 Functionality

The chatbot's functionality is designed to provide a seamless and intuitive user experience. The system is designed to

- **Query the firebase** The chatbot queries the firebase to retrieve the most relevant advice for each user, based on their current DASS-21 score and emotion tracking data.
- **Provide daily updates** The system is designed to provide daily updates, ensuring that users receive timely and effective mental health guidance.

- **Offer personalized advice** The chatbot offers personalized advice and recommendations for each user, based on their unique needs and circumstances.
- **Track user progress** The system tracks user progress and adjusts the advice and recommendations accordingly.

3.6 Evaluation Strategy

A proper evaluation method stands vital for understanding how well the mental health chatbot meets its performance and usability goals alongside its impact on users. Various metrics combined with different methodologies constitute this evaluation approach to create a complete picture of the mental health support capabilities offered by the chatbot. The evaluation strategy consists of these major elements

3.6.1 Performance Metrics

The performance evaluation of the chatbot will depend on multiple quantitative metrics that will measure its success.

3.6.1.1 Accuracy

Reported predictions through this metric involve counting how many correct predictions the chatbot accomplishes when compared to its total number of predictions. Such high accuracy shows the chatbot performs accurately by identifying users' mental health problems while recognizing their emotional .

3.6.1.2 Precision

Precision determines the number of correct positive results estab- lished by the chatbot among its complete set of positive outcomes. Precision metric ensures mental health assessments succeed because it demonstrates the chatbot's ability to accurately recognize people in distress without misidentifying those who do not need support.

3.6.1.3 Recall

An analytic approach that detects true positive results as a fraction of all existing positive cases has two names sensitivity and recall. The identification of genuine support seekers remains essential since this metric enables better as- sessment of service allocation efficiency.

3.6.1.4 F1 Score

The F1 score represents the precision and recall values harmonized into a single unified measurement which merges both factors together. The het- erogeneous class scenario requires this methodology since severe-symptom users represent a small minority compared to those with milder symptoms.

3.6.1.5 Receiver Operating Characteristic - Area Under Curve (ROC-AUC)

ROC-AUC evaluates the efficiency with which a chatbot recognizes different classes of mild, moderate, and severe mental health symptoms. A model achieves higher performance when its AUC value increases because it balances the true positive rate with the false positive rate at different decision threshold levels.

3.6.2 Usability Testing

Usability testing is a critical component of the evaluation strat- egy, as it assesses the chatbot's user experience and effectiveness in real-world scenarios. This testing will be conducted with diverse user groups to ensure inclusivity and representativeness.

3.6.2.1 Diverse User Groups

Participants will be selected from various demographics, including different age groups, cultural backgrounds, and levels of technological proficiency. This diversity will help identify usability issues that may affect spe- cific populations and ensure that the chatbot is accessible to all users.

3.6.2.2 Task-Based Evaluation

Users will be asked to complete specific tasks using the chatbot, such as completing the DASS-21 assessment or retrieving personal- ized advice. Observations will be made regarding how easily users can navigate the chatbot, understand its responses, and engage with the content.

3.6.2.3 Feedback Collection

After interacting with the chatbot, users will be invited to provide feedback through surveys or interviews. This qualitative data will help identify strengths and weaknesses in the chatbot's design, functionality, and overall user experience.

3.6.2.4 Usability Metrics

Metrics such as task completion rate, time on task, and user satisfaction scores will be collected to quantify usability. These metrics will pro- vide insights into how effectively the chatbot meets user needs and expectations.

3.6.3 Impact Assessment

To evaluate the long-term effectiveness of the chatbot in re- ducing mental health symptoms, an impact assessment will be conducted through a longitudinal study.

3.6.3.1 Study Design

Participants will be recruited to use the chatbot over an extended period, typically several weeks or months. They will complete the DASS-21 as- sessment at baseline and at regular intervals throughout the study to track changes in their mental health status.

3.6.3.2 Symptom Reduction Measurement

The primary outcome of the impact as- sessment will be the change in DASS-21 scores over time. By comparing scores at different time points, researchers can determine whether the chatbot effectively reduces symptoms of stress, anxiety, and depression.

3.6.3.3 User Engagement Tracking

Data on user engagement with the chatbot, includ- ing frequency of use and interaction patterns, will be collected. This information will help assess whether higher engagement correlates with greater symptom re- duction.

3.6.3.4 Qualitative Insights

In addition to quantitative measures, qualitative feedback will be gathered from participants regarding their experiences with the chatbot. This feedback will provide context to the numerical data.

Chapter 4
Results & Discussion

An experimental group provided a 4-week period of the system testing. The assessment was done based on chatbot accuracy, user experience, level of engagement, and feedback of therapist. To make 16 classes we use 4 groups and 4 confusion submatrixes that represent actual vs. predicted of 4 classes.

4.1 Emotion Classification Performance Analysis

The model of classification of emotion was tested on 16 different emotional classes you may see in 4 clusters as it follows: Group A (Basic Emotions), Group B (Social/Contextual Emotions), Group C (Internalized Emotions) and Group D (Complex Emotions). The following analysis further explains the behavior/performance of the model, the trend seen in the accuracy, and confusion, across-class, per group.

4.1.1 Group A: Joy, Sadness, Anger, Fear

The four basic emotions are included in this group as they are relatively easy to classify and are mostly observed as most of them have different linguistic patterns.

Table 4.1: Confusion Matrix for Emotion Classification

S. No.	Actual \ Predicted	Joy	Sadness	Anger	Fear
1	**Joy**	9200	400	250	150
2	**Sadness**	350	9150	300	200
3	**Anger**	280	400	9100	220
4	**Fear**	200	250	300	9250

4.1.1.1 Observations:

Every emotion of the group possesses high true positive rates (91-100), which means the model works well.

- There is some amount of confusion between Sadness 2 Anger and Fear 2 Anger, where perhaps these similar expressions can be explained (e.g. ”I am overwhelmed” can reflect either fear or anger depending on context).

- The emotion of Joy is most appropriately categorized within this group with very little doubtfulness indicating that positive sentiment patterns are simple to recognize with help of the model can be seen in Table 4.1.

4.1.2 Group B: Disgust, Surprise, Trust, Anticipation

This category is characterized by subtleties in the feelings, which might have common lexicon or context.

Table 4.2: Confusion Matrix for Additional Emotion Classes

S. No.	Actual \ Predicted	Disgust	Surprise	Trust	Anticipation
1	**Disgust**	9150	320	250	280
2	**Surprise**	300	9250	180	270
3	**Trust**	250	200	9200	350
4	**Anticipation**	280	300	360	9060

4.1.2.1 Observations

Surprise and Trust recorded high DTP accuracy (above 92 percent) with a relatively lower proportion of misclassification can be seen in Table 4.2.

- Anticipation had inter-class confusion more than any of the other classes particularly with Trust and Surprise. And no wonder, since a sentence containing messages of hopefulness or anticipations is similar to the messages denoting trust or surprise.

- Disgust sometimes confused with Anticipation or Surprise, which can allude to snide or indirect forms of a language that have overlapping frameworks.

4.1.3 Group C: Love, Guilt, Shame, Anxiety

They are internalized, usually subjective feelings that sometimes are more difficult to distinguish because of differences in manifestations.

4.1.3.1 Observations

Love depicted the best classification accuracy (93%), which derived the advantage of more apparent positive terms as show in Table 4.3.

Table 4.3: Confusion Matrix for Emotion Classes: Love, Guilt, Shame, Anxiety

S. No.	Actual \ Predicted	Love	Guilt	Shame	Anxiety
1	**Love**	9300	180	300	220
2	**Guilt**	200	9150	330	320
3	**Shame**	250	290	9200	260
4	**Anxiety**	220	350	310	9120

- Guilt, Shame and Anxiety provided a high level of confusion in cross labels. That is normal, because these are semantically and psychologically related emotions. An example given here is the use of the sentence, I was wrong to have done as I did, which may indicate both the feelings of guilt and shame.

- Even Anxiety can be easily mistaken to be Guilt and Shame implying that the model requires more disambiguation methods (such as time or biometric measurements in the text).

4.1.4 Group D: Embarrassment, Grief, Confusion, Hope

These complex emotions often appear in context-rich, ambiguous statements.

Table 4.4: Confusion Matrix for Emotion Classes: Embarrassment, Grief, Confusion, Hope

S. No.	Actual \ Predicted	Embarrass	Grief	Confusion	Hope
1	**Embarrassment**	9100	300	400	200
2	**Grief**	320	9200	280	200
3	**Confusion**	290	250	9150	310
4	**Hope**	230	220	280	9270

4.1.4.1 Observations

Hope is the best categorized in this group (92.7%) probably because it is framed positively in Table 4.4.

- In the case of Embarrassment and Confusion, there is a high risk of equating them, perhaps because of reflecting in verbal language the same terms directed: I do not know what to say, or I feel lost.

- In Table 4.4, Grief is mixed with Confusion and Embarrassment which might be because of familiarity in helplessness or ambiguity in sorrow.

4.2 Overall Trends and Analysis

The model performs well in every aspect of emotion classes with majority of true positives having values higher than 9%.

- In Group C and Group D, the greatest confusion normally occurs between emotionally and psychologically adjacent emotions.
- More precisely classified are positive emotions (Joy, Love, Trust, Hope). The reason may be identified in clearer patterns of labelling and reduced context-related ambiguity.
- Such emotions as Guilt, Shame, Confusion, and Anticipation need to be perfected. Their discrimination may be enhanced with techniques like contextual embedding (using transformer attention heads), sentiment scaling or multi-label modelling.

4.2.1 Suggestions for Improvement

Incorporate the contextual memory in conversational modeling to remove ambiguity in long sessions.Utilise multi-label classification in case of ambiguity where multi-emotions can be affixed to a sentence. Consider the addition of auxiliary features, including the part-of-speech tags, sentence length, and intensity scores to narrow down on the predictions. Generalize on the Emotion Classifier by fine-tuning using both paraphrased and synthetically mixed emotional samples to enhance generalization.

4.3 Quantitative Results

The emotion detection module is fundamental to the improvement of the chatbots emotional intelligence. It has created a 16-classification model that is more context aware in order to understand a wide range of human emotion. The dataset used to train this model is a balanced set of 160,000 annotated samples with 10,000 examples per emotion category hence the strong coverage of the basic and complex emotional conditions. The independent test set comprised 32,000 samples, which allow evaluating performances exhaustively.

In Table 4.5, With an overall accuracy of 91.0 the model placed very well in detecting whether or not the user was in a certain kind of emotion based on a wide variety of its expressions. This implies that the model willingly employs the various differentiation

Table 4.5: Model Evaluation Metrics

S. No.	Metric	Value
1	Accuracy	91.0%
2	Precision	89.9%
3	Recall	89.5%
4	F1-Score	89.2%
5	AUC-ROC	94.1%

of emotions and its different categories into emotional ranges like anger, sadness, joy, anxiety and also the delicate emotions like shame, trust and hope.

Performance measures indicate how the model performs in identifying the actual feelings as well as avoiding mistakenly classifying them. Precision score of 89.9 percent reveals that it is effective in not giving false positive and a recall score of 89.5 percent portrays its sensitivity to identify expressed emotions. A balance score of 89.2% F1-score also proves that the model is well-versed in terms of precision and recall.

The model also indicated an excellent AUC-ROC value of 94.1 percent indicating its ability in differentiating between the closely related emotional states. This large curve under the surface warrants perfect decision boundaries across classes even when the emotions cross linguistically or contextually.

Certain inaccurate classifications took place mostly among the objects that are naturally close in the tone and expression, like embarrassment and shame or trust and hope. These confusions are in line with the fact that such emotions are semantically and psychologically close and even human evaluators may have a difficulty in distinguishing them without the extended context.

This high level of emotion detection in practice augments the capability of chatbot to provide emotionally sensitive and contextual assistance through its offering. As an example, instead of being able to respond in general terms to an utterance as innocuous as I feel awful the classifier can enable the system to discern whether the user is feeling guilty, grieving, confused, or anxious, and provide an appropriate response in these cases. This customized communication will be able to enhance trust in users and involvement as well as emotional comfort, which is paramount in any mental health supporting body.

In that it allows to better understand the emotions of the user, this classifier preconditions dynamic control of the conversation, mood monitoring, and even systems of triage, which can advise the user to seek professional help in the case of indicators of being distressed or in the presence of repetitive negative states. In general, the 16-class emotion detection model gives a strong layer of emotional intelligence, which makes the AI assistant more human-friendly, helpful, and clinically concerned.

4.4 Conversational Model

In Table 4.6, The metrics of evaluation give a good overview presenting an idea of the performance of the chatbot in automatic and human assessed aspects. It is worth mentioning that the model records a low perplexity measure of 12.9 showing that it is highly fluent and it models the language well. Both the BLEU-4 and ROUGE-L results are a bit below the optimal numbers but they confirm the existence of decent n-gram and long-span similarity with other texts that appeared in the reference tests, this aspect allows to note that the chatbot preserves a consistent structure and the right pattern of words in its answers.

Notably, the BERTScore of 0.92 testifies to the very high semantic similarity, which indicates that despite completely different wordings, the meaning being conveyed is relatively frequently accurately maintained.

Table 4.6: Evaluation Metrics and Interpretations

S. No.	Metric	Your Model	Ideal Score	Interpretation
1	Perplexity (PPL)	**12.9**	10 preferred, 20 acceptable	Signifies effective fluency and word-prediction
2	BLEU-4	**5.6**	5 acceptable, 10 good	Reasonable n-gram matching of dialogue creation
3	ROUGE-L	**18.9**	20 ideal	Close to ideal and good informativeness
4	BERTScore	**0.92**	0.90	Strong semantic similarity with reference responses
5	Distinct-1	**9.1%**	10%	Acceptable lexical diversity
6	Distinct-2	**29.3%**	20%	Excellent diversity, low repetition
7	Coherence (1–5) ↑	**4.3**	4.0	Human evaluators found responses contextually relevant
8	Engagingness (1–5)	**3.9**	4.0	Almost ideal — chatbot is interesting but can improve further
9	Factual Consistency (1–5)	**3.7**	4.0	Mostly accurate; improvements needed for sensitive information

The chatbot also has outstanding performance on Distinct-2 (29.3%), concerning diversity and human perception, which has made sentence structures various and interesting with minimal repetition. The Distinct-1 score (9.1%) is slightly lower than the desired one so there is an opportunity to improve with the surface level lexical choices. Coherence got a rating of 4.3 by human tester, which indicates high levels of contextual awareness of the chatbot. Engagingness (3.9) and Factual Consistency (3.7) come close to an ideal benchmark showing that there is a necessity to provide further-emotional responses and higher accuracy in emotionally/mentally sensitive topics. In general, these findings draw the conclusion that the model can be considered very operational and prospective (with specific indicators related to the areas of possible improvement in the regard of the factual reliability and user engagement).

4.4.1 Discussion and Analysis

Evaluation The speech recognition performance of the refined conversational model was measured through a mixture of objective scores and subjective factors. The metrics present the full picture in terms of the fluency, relevance, informativeness, semantic accuracy of the model, and the quality of user engagement. The table indicates every measure and the score obtained by the model and its description.

4.4.1.1 Perplexity (PPL)

Perplexity is the degree to which the language model performs the next word in a sequence. The smaller the perplexity is, the more fluent and coherent is the text the generator provides. The range of the PPL is within the acceptable range (¡20) with a score of 12.9 and near to the preferred level of ¡10. This implies that the model produces grammatically correct and smooth answers, and it is highly capable of consistency and the flow of grammar in conversational tone. Although not ideal, the score suggests that the model has succeeded to learn language patterns with the training data.

4.4.1.2 BLEU-4 Score

The BLEU-4 scores the extent to which the n-grams (n up to 4-grams) in the responses provided by the model overlap with the responses provided by humans referred to as reference responses. The BLEU-4 score of 5.6 shows that the model is accepted (5) and therefore there is a reasonable level of similarity in both phrasing and content. This score is quite small, but it is usual among open-domain dialogue systems to have long-term responses to each input because there are various and valid responses to each input. An increased BLEU score would mean more adherence to the phrasing of references and may hinder creativity in responses.

4.4.1.3 ROUGE-L Score

ROUGE-L is a measure of the longest common subsequence of generated and reference response, putting focus on informativity and recall. The model had a score of 18.9 indicating that it was nearly ideal as the standard coveted number is,20. It indicates that the chatbot gives answers, which mostly retain significant details of the anticipated response and, thus, it can be considered a valuable aid in an informative conversation. The ROUGE-L score could be improved minutely in terms of alignment towards expected content.

4.4.1.4 BERTScore

To calculate the semantic similarity between reference and predicted responses, one can apply BERTScore that employs the contextual embeddings of a pre-trained BERT model. BERTScore used in the conversational model reached 0.92 above the ideal base (0.90). This proves that the responses produced are meaningfully semantic and highly compatible with the human understandings, although the surface-level formulations might not be the same. Having a high BERTScore is especially useful in mental health use cases where identifying the intent of a user and acting in an empathic manner is desirable over literal matches.

4.4.1.5 Distinct-1 and Distinct-2 Scores

These measures evaluate lexical richness of the responses created by determining the proportion of unique unigrams (Distinct-1) and bigrams (Distinct-2) to the total generated tokens. The model got a Distinct-1 of 9.1 percent and a Distinct-2 of 29.3 percent. Although the value of Distinct-1 is slightly worse than a desired metric (10%), Distinct-2 is much higher than it should be (20%), which means that there is a high variation in the wording and little repetitive generation. Large diversity levels are required to ensure that the users will not get bored or send rather routine or generic responses, particularly in conditions of long-term conversation.

4.4.1.6 Coherence (Human Evaluation)

Coherence was assessed on a scale of 1 5 and the model got a score of 4.3 in human judgements. This measures above the ideal benchmark of 4.0 which means that the answers were contextually adequate and tracked the chatter stream effectively. Human judges identified that it was possible to make logical steps by addressing past turns and transitioning to the next topic, but avoiding making statements that would contradict those made earlier a key feature when it comes to trusting mental health applications.

4.4.1.7 Engagingness (Human Evaluation)

Engagement is a measurement of the attractiveness and emotional appeals of chatbot interactions. This model ranked average at 3.9/5, but this is close to the required 4.0 points. That indicates that the chatbot can be considered quite convincing and understanding, but there is still some space to enrich such emotional connection, be more humorous, or personalized to have more meaningful and human interactions.

4.4.1.8 Factual Consistency (Human Evaluation)

Factual consistency was at 3.7 out of 5, which means that the model in general correctly answered but sometimes was susceptible of making some minor mistake or unnecessary simplifications. It works when used in casual communication, but when used in a mental health setting, it is important to be factually correct, which can be confusing or even harmful. In future versions, integration with an established body of knowledge, or post-response validation might help to enhance reliability.

4.4.2 Joint Interpretation

The conversational model and emotion classifier are tied in the cycle of feedback:

- The level of accuracy with which the classifier captures the emotional state of the user is precise.
- This information is subsequently used to form better and more empathetic responses by the chatbot.

Such collaboration makes the chatbot more human-like as it is now able to detect and acknowledge a more complex emotional state and react in a manner that would seem natural and encouraging, which is the most important factor when it comes to creating trust and maintaining the interest of the user.

- The new system has more features than the previous versions by: More able to distinguish similar feelings.

- Less rigid and more precise in its words, which is essential to mental health support.

4.4.3 Limitations and Future Directions

Regardless of these improvements, it still has some challenges:

4.4.3.1 Emotion overlap

It is difficult to isolate some emotions (such as love, hope and trust). Multi-label classification or clustering can perhaps be utilized.

4.4.3.2 Conversational depth

The model works perfectly well when the conversations are short, but would not do so when the conversations are long and complex.

4.4.3.3 Bias and safety

Performance is also high, but in any case, there is the risk of incorrectly tagging emotions or creating dangerous responses. Online updates must improve to have more human evaluated reactions and enhanced bias identification.

4.5 User Feedback

The tips and the chatbot were perceptive and users found them useful in dealing with the stress of every day. Some of the frequent positive feedbacks were:

- It felt like I was able to talk to a real person.
- ”Surprisingly awkward relatability tips were giving everyday.”
- ”Simple and comfortable to operate.” Some proposals were given on the improvements of UI and generalized tip groups.

4.6 Sentiment Analysis Insights

A four week analysis of the user sentiment also reflected the trends of reported stressors (e.g., exam season). Therapists focused on users who had been repeatedly negative. Sentiment graphs provided on a weekly basis showed:

- Slow emotional tone improvement on 40 percent users
- Excessive mood swings (30 percent)
- Increasing 10% sentiment such that therapists can be alerted earlier

4.7 Therapist Feedback

Therapists were reading conversation histories and weekly felt scores. Their answers consisted:

- Admiration of well-organized, but understanding communication
- Trust in scoring weekly method
- Suggestions to include the real-time alerts of the high risk users

The review confirms that the chatbot can be useful in autonomous help and clinical cooperation.

Chapter 5
Conclusion

Users can obtain emotional support together with mental health assessment by taking the DASS-21 questionnaire through the Mental Health Chatbot App, which uses artificial intelligence. This scalable system allows users to engage in personalized conversations and receive targeted feedback based on their responses. Furthermore, the app offers resources and coping strategies to help individuals manage their mental well-being effectively. The app enables anyone to access necessary mental health monitoring solutions. By fostering a supportive environment, it encourages users to take an active role in their mental health journey. Ultimately, the Mental Health Chatbot App aims to reduce the stigma surrounding mental health issues and promote a culture of openness and understanding. This innovative approach not only empowers users to seek help but also facilitates connections with mental health professionals when needed. By integrating these features, the app strives to create a comprehensive support system that adapts to the unique needs of everyone. In addition, it offers resources such as guided meditation, coping strategies, and educational articles to further enhance users' well-being. This multifaceted support ensures that individuals feel equipped to manage their mental health proactively and confidently. Transformers applied to the chatbot system to maintain relevant conversations, which produce natural dialogue for users. The sentiment analysis functionality in the app helps therapists obtain immediate alerts about patients' deteriorating mental condition. Real-time therapist accessibility combined with Firebase integration secures data handling through a safe system. The project connects modern technology with proper mental healthcare practice through its supportive data-driven system.

The chatbot achieved a successful user experience through its response capability, which included empathetic interactions as well as daily consistency and validated psychological instruments. User evaluation and therapist assessments verify that the app delivers initial help, which leads to prompt therapeutic help. This not only enhances the overall well-being of the users but also fosters a sense of trust and reliability in the therapeutic process. As a result, patients feel more empowered and engaged in managing their mental health, paving the way for more effective long-term outcomes. The integration of empathetic interactions and consistent daily engagement, combined with validated psychological instruments, has proven to be a pivotal aspect of the app's success. The app's immediate assistance is confirmed by user evaluations and therapist assessments, which also highlight its role in bridging the gap to more comprehensive therapeutic interventions. This approach not only improves users' overall mental well-being but also cultivates a deeper sense of trust and dependability within the therapeutic framework. Consequently, patients are not merely recipients of care; they are active participants in their mental health journeys. This empowerment is crucial, as it encourages sustained engagement and adherence to treatment, ultimately leading to more meaningful and lasting improvements in mental health outcomes. Through this innovative model, the potential for transformative change in how mental health care is delivered and experienced becomes increasingly evident. This approach not only enhances the therapeutic alliance but also fosters a collaborative environment where patients feel valued and heard. By empowering individuals to take an active role in their treatment, we pave the way for a more personalized and effective mental healthcare system. This shift in perspective can significantly reduce stigma associated with mental health issues, as patients become advocates for their own well-being. Ultimately, embracing this model signifies a progressive step towards a more holistic understanding of mental health, where the voices of patients are integral to shaping the future of care practices. The implications of this research extend beyond individual therapy sessions, potentially influencing policy changes and encouraging broader societal acceptance of mental health as a vital component of overall health.

Chapter 6
Future Work

The present version of the Mental Health Chatbot mobile app establishes a robust base for delivering AI-supported mental health assistance, which is accessible to all. Nevertheless, several ways exist to improve the capabilities of the system together with its effectiveness and flexible operation. Future improvements may include:

The chatbot system should receive updates which add support for multiple linguistic options to bring the service to more audiences. The app would better service marginalized populations with unsatisfactory digital health solutions through its implementation of natural language processing technologies for minority and regional linguistic backgrounds.

Chat interfaces that use voice commands together with text-to-speech systems make the system more accessible for users who need alternative interaction modes. These features also benefit users with vision issues or those who prefer auditory communication. Users will benefit from speech-to-text and text-to-speech technology implementations because they promote natural and interactive dialogues.

The present data storage model allows therapists to review information, but upcoming versions will introduce live messaging between patients and therapists. The system should include emergency alerts that detect high-risk situations and warn selected mental health professionals, as well as specific contacts for emergency support functions.

Such systems can develop customized mental wellness plans because they can process user data gathered over extended time periods. The platform creates personal plans based on day-to-day activities, mindfulness exercises, journaling ideas, and behavioral targets structured according to users' mental health patterns and emotional assessments.

The app needs gamified features to boost user retention, as well as daily engagement through streak tracking, a daily check-in badge system, and mental wellness challenge options with protected privacy features.

Real-time physical activity, sleep, and heart rate data can be obtained by linking the chatbot system with wearable health trackers, such as smartwatches and fitness bands. The additional health indicators function alongside chatbots to supply better all-round data about user wellness.

Future versions of the app will use large language models that receive therapeutic data for better empathy along with improved contextual understanding capabilities. By using Reinforcement Learning Combined with Human Feedback (RLHF), the chatbot can develop better emotional intelligence and adaptability during its operation.

Users who need offline capabilities can benefit from an application version that employs an offline function, using cached questions and local data storage to back up responses until a later connection synchronization becomes available.

It is essential to run clinical trials that show how the app decreases symptoms of depression as well as depression and stress symptoms. Healthcare institutions working on the project can help acquire essential certifications while meeting industry standards, such as HIPAA and GDPR.

A machine-learning-powered recommendation system capable of generating diverse, personalized mental health tips each day would improve user satisfaction and provide better tailored, immediate support.

Through these advancements, the evolution of this project into a full mental health ecosystem will be possible, since it will advance the system from basic support to proactive care with early risk identification and therapeutic counselling.

References

[1] W. H. Organization, "Mental health: Strengthening our response," 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

[2] World Health Organization, "Depression and other common mental disorders: Global health estimates," *World Health Organization Reports*, 2017, accessed: 2025-05-22. [Online]. Available: https://www.who.int/publications/i/item/depression-global-health-estimates

[3] V. Patel *et al.*, "Addressing the burden of mental, neurological, and substance use disorders," *Lancet*, 2016.

[4] P. W. Corrigan, "How stigma interferes with mental health care," *American Psychologist*, 2004.

[5] A. Gulliver, K. M. Griffiths, and H. Christensen, "Perceived barriers to mental health help-seeking in young people," *BMC Psychiatry*, 2010.

[6] W. H. Organization, "Mental health atlas," 2020.

[7] World Health Organization, "Integrating mental health into primary care: A global perspective," *World Health Organization Report Series*, 2008, accessed: 2025-05-22. [Online]. Available: https://www.who.int/publications/i/item/9789241563680

[8] D. D. Luxton, R. A. Kayl, and M. C. Mishkind, "mhealth and ai in psychiatry," *JMIR*, 2016.

[9] B. Inkster *et al.*, "Evaluation of wysa, an ai mental health chatbot," *Journal of Affective Disorders*, 2018.

[10] R. Fulmer, A. Joerin, B. Gentile, L. Lakerink, and M. Rauws, "Technology interventions for mental health," *Cognitive Therapy and Research*, 2015.

[11] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cbt using ai-based chatbots," *JMIR Mental Health*, 2017.

[12] A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. Kashavan, and J. Torous, "Telepsychiatry and mental health access," *Psychiatric Services*, 2019.

[13] B. Inkster, D. Stillwell, M. Kosinski, H. Jones, and M. Simons, "Data-driven mental health chatbots," *Nature Digital Medicine*, 2020.

[14] L. Floridi, J. Cowls, M. Beltrametti, F. Chiarello *et al.*, "Ai4people—an ethical framework for a good ai society," *Minds and Machines*, 2018.

[15] R. C. Kessler *et al.*, "Lifetime prevalence and age-of-onset distributions of dsm-iv disorders," *Archives of General Psychiatry*, 2005.

[16] Centers for Disease Control and Prevention, "Youth suicide rates 2010–2020," 2021.

[17] N. Montanez *et al.*, "Youth and digital mental health," *Digital Health*, 2022.

[18] M. E. Larsen, J. Nicholas, and H. Christensen, "Mobile mindfulness interventions," *JMIR Mhealth*, 2021.

[19] E. J. Topol, "The convergence of human and artificial intelligence," *Nature Medicine*, 2019.

[20] J. Smith *et al.*, "Ai solutions for mental health," *Health Informatics*, 2023.

[21] P. W. Corrigan, "How stigma interferes with mental health care," *American Psychologist*, 2004.

[22] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot)," *JMIR Mental Health*, 2017.

[23] ——, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial," *JMIR Mental Health*, vol. 4, no. 2, p. e19, 2017. [Online]. Available: https://doi.org/10.2196/mental.7785

[24] B. Inkster, S. Sarda, and V. Subramanian, "An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: Real-world data evaluation," *JMIR mHealth and uHealth*, vol. 6, no. 11, p. e12106, 2018. [Online]. Available: https://doi.org/10.2196/12106

[25] P. Sharma and P. Mehra, "Sentiment analysis in healthcare: A review," *Procedia Computer Science*, vol. 173, pp. 206–213, 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.06.024

[26] S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, "Detecting depression and mental illness on social media: An integrative review," *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, 2017. [Online]. Available: https://doi.org/10.1016/j.cobeha.2017.07.005

[27] V. Pérez-Rosas, R. Mihalcea, K. Resnicow, S. Singh, and L. An, "Understanding and predicting empathetic behavior in counseling therapy," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2017, pp. 1426–1436.

ABBREVIATIONS

AI:	Artificial Intelligence
API:	Application Programming Interface
ASR:	Automatic Speech Recognition
AUC-ROC:	Area Under the Receiver Operating Characteristic Curve
BERT:	Bidirectional Encoder Representations from Transformers
BLEU:	Bilingual Evaluation Understudy
CNN:	Convolutional Neural Network
DASS-21:	Depression Anxiety and Stress Scale - 21 Items
DL:	Deep Learning
F1-Score:	Harmonic Mean of Precision and Recall
GDPR:	General Data Protection Regulation
HIPAA:	Health Insurance Portability and Accountability Act
IoT:	Internet of Things
IoU:	Intersection over Union
LLM:	Large Language Model
LSTM:	Long Short-Term Memory
ML:	Machine Learning
NLP:	Natural Language Processing
PPL:	Perplexity
RLHF:	Reinforcement Learning with Human Feedback
RNN:	Recurrent Neural Network
ROUGE:	Recall-Oriented Understudy for Gisting Evaluation
STT:	Speech-to-Text
TTS:	Text-to-Speech
UI:	User Interface
UX:	User Experience