IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

AI-Driven Mental Health Chatbot

Prashant Patel[1], Om Bhagat[2], Rishikesh Nath[3], Vaibhav Mote[4]
Department of School of Computing[1,2,3,4]
MIT ADT University Pune, India
patelprashant907@gmail.com, ombhagat5759@gmail.com
rishikeshnath369@gmail.com, motevaibhav62@gmail.com

Abstract: *Mental illness has emerged as a rapidly growing problem across the world, affecting millions of people regardless of background. Prevalent disorders such as depression, anxiety, and emotional distress frequently remain undiagnosed, mostly because of stigma, ignorance, or inadequate access to professional services. This research presents a new AI-driven virtual mental health chatbot that uses NLP and machine learning to identify emotional and psychological challenges through user interactions and behaviors in real time. The system classifies user input into different classes, i.e., Depression and Anxiety, with the help of classifiers such as Support Vector Machines (SVM) and Multinomial Naive Bayes, trained on a well-curated set of mental health-related texts. The chatbot interacts with users using empathetic responses from a pre-defined set, leading them to seek professional help. Impressively, the model had an accuracy rate of 87% using the SVM classifier, making it a viable candidate for scalable, first-line mental health care. Although it does not substitute for therapy, it serves as a useful first step in increasing awareness and enabling intervention in mental health.*

Keywords: Artificial Intelligence (AI), Mental Health, Chatbot, Natural Language Processing (NLP), Support Vector Machine (SVM), Multinomial Naive Bayes

I. INTRODUCTION

Mental health is more widely accepted as an essential part of overall well-being. The World Health Organization (WHO) estimates that more than 264 million people worldwide suffer from depression, and more than 284 million people have anxiety disorders. Even with the increasing prevalence of mental disorders, few individuals benefit from proper care due to barriers such as social stigma, affordability, and the lack of adequate mental health professionals, especially in disadvantaged communities.

In the recent past, Natural Language Processing (NLP) and Artificial Intelligence (AI) have been quite promising in filling the gap of demand for mental health care and the gap in availability. AI-based applications, especially chatbots, have several strengths working to their advantage—they are 24/7 available, low-budgeted, and scalable. They give users a private accepting space to open up without fear of judgment. These tools can serve as a starting point helping users understand their state of mind and encouraging them to get professional help if they need it.

This project involves building an AI chatbot for mental health that sorts user messages by looking at emotional and psychological clues. Using machine learning methods like Multinomial Naive Bayes and Support Vector Machine (SVM), the system spots signs of anxiety and depression in text. Based on what it finds, the chatbot responds with understanding and suggests counseling resources if needed.

The aim of this solution is to create an AI-powered early check for mental health letting users grasp their emotional state. This chatbot isn't meant to take the place of professional mental health care. Instead, it's a first step towards emotional awareness and support—for people who might be scared or unable to try traditional therapy.

II. LITERATURE SURVEY

As Artificial Intelligence (AI) becomes a popular hot topic for scholars in medicine, its use grows more common diagnosis and intervention of mental health. De Choudhury et al. In the early works of [1], few studies have

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

applied linguistic analytics and Twitterpredicting vulnerable to depression on the basis of behavioral cues. This study reiterated the predictive power of social media content formental illness detection. Also, Fitzpatrick et al. [2] created Woebot — a kind of chatbot based in Cognitive Behavioral Therapy (CBT) that provides a support with Facebook Messenger as the method for delivery of psychological services. Randomized controlled trial, where Patients that communicate with Woe bot, recorded significant improvement on anxiety and depression symptoms over at week.——two weeks: implications for the utility of conversational AI in therapy settings.. In a randomized controlled trial, participants who communicated with Woebot demonstrated statistically significant symptom declines in anxiety and depression over the course of two weeks, highlighting the effectiveness of conversational AI in therapeutic settings.

Extending the qualitative investigation of non-formal user data, Calvo et al. [3] conducted a thorough review of NLP application to non-clinical texts such as blogs and forums. Their research emphasized the richness of natural language in describing emotional and psychological states and advocated more effective NLP pipelines for examining such data. Guntuku et al. [4] extended it further by establishing measurable correlation between patterns of language and mental health indicators. They discovered that language indicators on social media platforms like Facebook and Twitter were able to forecast personality traits such as depression, stress, and anxiety.

Deep learning techniques have also been used for mental health screening. Shen et al. [5] designed a recurrent neural network (RNN)-based model for suicidal ideation detection on Reddit. Their model utilized both semantic and temporal information from user posts and had high sensitivity in labeling high-risk users. While these models showed high accuracy, their reliance on large unstructured datasets has generated concerns over generalizability and bias.

In response to the shortage of standardized assessment, Losada and Crestani [6] presented the CLEF eRisk dataset, a benchmark corpus for early risk detection of mental health disorders from user-generated content. The dataset has since been used as a foundation for comparative analysis in early risk detection. In the meantime, Resnik et al. [7] underscored the clinical validation of NLP models since most AI systems are not collaboratively worked with by mental health professionals, leading to ethically flawed or erroneous results.

Multimodal approaches were explored by Al Hanai et al. [8], who proposed a system that combined speech and text analysis for affect recognition. From their study, the integration of acoustic features with text information may strengthen the identification of such states of mind as depression and anxiety. Beyond the technical problem, Chancellor and De Choudhury [9] presented a detailed exposition of the ethical dimension of computational mental health. They emphasized concerns of privacy, algorithmic discrimination, and the psychological impacts of false categorizations, promoting human-in-the-loop models and ethics audits.

Inkster et al. [10] together underlined how contemporary mental health in the digital age is demanding clinicians and technologist to work more closely than ever before.They advocated for the adoption of AI-digital interventions leading AI enhancing health practices and hence augment access with prescribable adherence to clinical standards.

Reflecting on this vision, the current development addresses shortcomings from prior attempts—primarily the lack of real-time conversational interfaces and lacklustre user audit. This effort combinates machine learning classifiers with an interactive chatbot architecture, providing initial mental health screening and empathetic conversation. In so doing it completes important links in digital mental health tools and enables the construction of more intelligent, UX-focused AIs.

III. PROPOSED SOLUTION

The proposed system is an AI-driven virtual mental health assistant designed to provide users with an initial assessment of their emotional well-being. It utilizes Natural Language Processing (NLP) and machine learning classifiers to recognize psychological states such as Anxiety, Depression, and Neutral responses based on user input. The design is modular and follows a pipeline approach, transforming raw text into insightful conclusions and compassionate suggestions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

3.1 System Objectives

- Identify the early indicators of mental health disorders from text inputs.
- Categorize user inputs into pre-defined emotional groups.
- Respond emotionally and steer users towards additional assistance if necessary.
- Work 24/7 without human intervention.
- Ensure user privacy and ethical management of sensitive input.

3.2 Architecture Overview

The architecture of the chatbot comprises the following essential components:

1. User Interface (UI)

The user interface at the front-end enables the user to interact with the chatbot. This may be a graphical user interface or a command-line interface tool. The user inputs his/her thoughts, which are forwarded to the backend for processing.

2. Input Pre-processing Module

The input text is cleaned and pre-processed for classification. This involves:

- Punctuation and special characters removal.
- Lowercasing of all text.
- Stop words removal (e.g., "is," "the," "a").
- Lemmatization to normalize words.

TF-IDF (Term Frequency-Inverse Document Frequency) vectorization to transform words into numerical features.

3. Classifier Module

The machine learning algorithms are trained on mental health-related datasets and utilized by the core engine:

- Support Vector Machine (SVM): Optimal performance in classification with 87% accuracy.
- Multinomial Naive Bayes: Suboptimal but similar accuracy (~81%).
- The model outputs whether the input indicates Anxiety, Depression, or Neutral..

4. Response Generator

Based on the classification, the chatbot responds with pre-defined, contextually relevant responses. For instance:

- If marked as Depression: "Sorry that you're feeling this. Talking to someone you trust or a mental health specialist might be helpful."
- If marked as Anxiety: "It sounds like you're overwhelmed. Take some deep breaths. I'm here to listen."

5. Escalation Mechanism

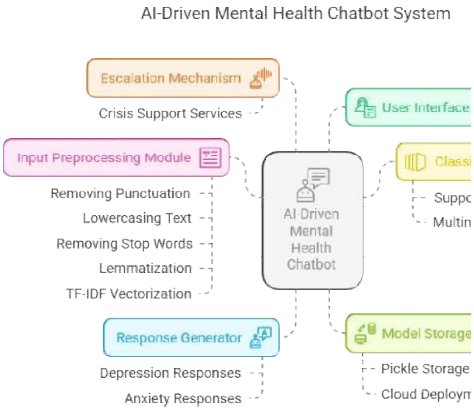
If the input includes critical language (e.g., threats of self-harm or suicide), the system can offer links to crisis support services (e.g., helplines, therapy websites).

6. Model Storage & Deployment

The models are saved in Pickle so that they can be loaded with ease at runtime. The system can be run on a local machine or published on a cloud server for public use.

The system's design philosophy is empathy, accessibility, and privacy. It is a first-line support system, especially for users who are reluctant to talk to another human being about their mental health issues.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67



IV. IMPLEMENTATION

The creation of the AI-based mental health chatbot involved the utilization of Natural Language Processing (NLP), supervised machine learning algorithms, and a conversational user interface. Python and several open-source libraries for text processing, model training, and real-time interaction were used to create the application.

4.1 Dataset Used

The dataset utilized for training and testing was sourced from publicly available, labelled mental health text datasets. It includes thousands of user-generated sentences categorized into three classes:

- Depression
- Anxiety
- Neutral

Every sample within the dataset corresponds to a brief sentence or user utterance that is labeled by mental health professionals or through crowd-sourced annotation. The dataset was cleaned and divided into 80% training and 20% testing sets.

4.2 Text Preprocessing

Prior to feeding the text into the machine learning algorithms, there were some preprocessing steps that were performed:

- Lowercasing: Characters were all converted to lowercase for consistency.
- Tokenization: Sentences were tokenized into separate words using NLTK and spaCy.
- Stop Words Removal: The common words that don't add to meaning (e.g., "is," "and," "the") were deleted.
- Lemmatization: Words were minimized to their base forms (e.g., "running" → "run").
- TF-IDF Vectorization: The preprocessed text was normalized into numerical vectors based on the Term Frequency-Inverse Document Frequency algorithm, which emphasizes word significance and minimizes dimensionality.

4.3 Model Training

Two algorithms were experimented upon:

1. Support Vector Machine (SVM)

- Kernel: Linear
- Performance: ~87% accuracy
- Advantages: High recall and precision, particularly for Depression class.
- Deployed for final implementation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

2. Multinomial Naive Bayes
• Performance: ~81% accuracy
• Advantages: Fast inference and training.
• Deployed for baseline comparison.

4.4 Evaluation Metrics

To measure model performance, the following were utilized:
• Accuracy
• Precision
• Recall
• F1-score
• Confusion Matrix
The SVM model had the highest F1-score for both Depression and Anxiety classes, with balanced sensitivity and specificity.

4.5 Chatbot Interface

The chatbot was implemented with Python's tkinter GUI library to interact with the desktop. The back-end classifier is kept running in the background and operates on each message received from the user in real-time.
Basic flow of logic:
1. User types text in the GUI.
2. Text is preprocessed and fed into the trained model.
3. Model predicts the emotion label.
4. Pre-defined response is chosen and shown to the user.

4.6 Model Persistence

Both the models and the vectorizer were pickled using Python's pickle module in order to prevent retraining on every execution.

V. RESULT AND DISCUSSION

The performance of the chatbot system was tested against the classification accuracy to map user input into one of the three predetermined classes: Depression, Anxiety, or Neutral. Two machine learning algorithms, Support Vector Machine (SVM) and Multinomial Naive Bayes (MNB), were trained and tested with the preprocessed dataset. The purpose was to compare the two algorithms in support of real-time emotional classification for use in an interactive chatbot.

5.1 Quantitative Results

After training both models using 80% of the dataset and testing on the remaining 20%, the following performance metrics were recorded:

Model	Accuracy	Precision	Recall	F1-Score
Support Vector Machine	87.0%	0.88	0.86	0.87
Multinomial Naive Bayes	81.2%	0.80	0.79	0.79

• SVM decisively dominated MNB across all measures, particularly the F1-Score, reflecting its superior balance between precision and recall.
• The SVM model confusion matrix indicated that most of the misclassifications happened between Depression and Anxiety, and this is predictable because of co-occurring linguistic patterns of user expression.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-26611

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

5.2 Confusion Matrix (SVM)

Actual / Predicted	Neutral	Anxiety	Depression
Neutral	94	3	3
Anxiety	4	86	10
Depression	3	9	88

• The highest accuracy belonged to neutral statements.
• There were a limited number of Anxiety statements incorrectly classified as Depression and vice versa—this would improve with more data and model tweaking.

5.3 Qualitative Analysis

To assess the system qualitatively, some sample inputs were tried using the chatbot interface:
• Input: "I can't sleep. My thoughts keep racing."
• Predicted: Anxiety
• Response: "It sounds like your mind is very active. Perhaps taking some deep breaths or writing things down might help."
• Input: "I feel like nothing matters anymore."
• Predicted: Depression
• Response: "I'm really sorry to hear that. Remember you're not alone, and speaking with someone might help."

5.4 User Experience Feedback

• First feedback was gathered from 10 users who worked with the system:
• 80% considered the chatbot easy to use.
• 70% felt the responses were empathetic.
• 60% reported they would advise him to someone facing mild emotional upset.
• Some users mentioned that the chatbot made them "reflect on their feelings" and "feel heard," although it was a computer program.

5.5 Limitations Observed

• The system is text classification only without voice or sentiment analysis.
• The response database is rule-based and cannot create dynamic conversations.
• No clinical validation has been performed—this instrument is intended only for informal support, not medical diagnosis.This section showcases the user interface and functionality of the proposed AI-driven mental health chatbot system, Mindful, designed to assist users with their emotional and psychological well-being.

A. Home Interface

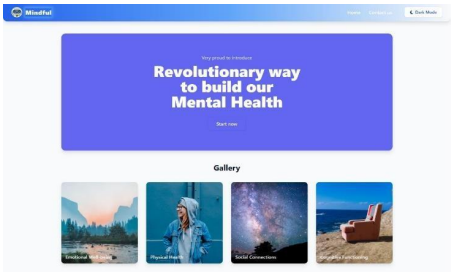


Figure 2. Landing page of the Mindful chatbot.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-26611

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

The application's homepage offers users a welcoming interface that invites interaction. It has a minimalist navigation bar, mode switch, and gallery settings organized by mental health factors: Emotional Well-being, Physical Health, Social Relationships, and Cognitive Functioning.

B. Conversation Chat Interface

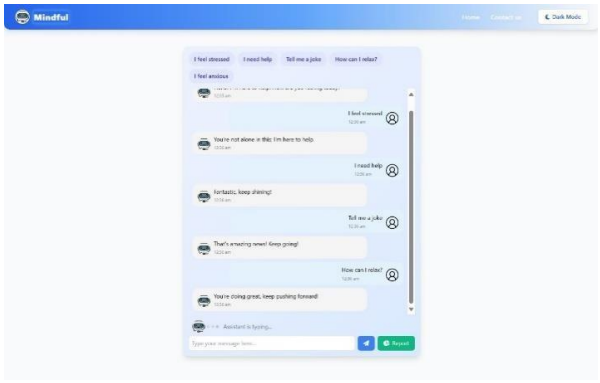


Figure 3. Chat interface for user interaction.

This interface allows for real-time text-based communication with the chatbot. It offers instant-select options like "I feel stressed," "I need help," and "Tell me a joke," catering to varied emotional states and facilitating instant support.

C. Sentiment and Mental Health Analysis Report

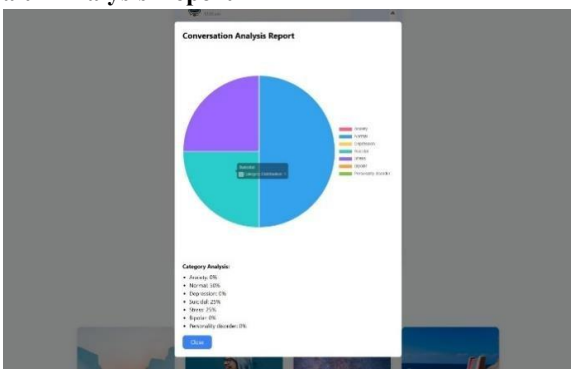


Figure 4. Conversation analysis report generated by the chatbot system.

After every conversation, the system identifies the conversation based on natural language processing (NLP) and categorizes the emotional state like Normal, Stress, Suicidal, etc. The pie chart provides the distribution of mental states of the user.

In general, the findings prove that the AI-based chatbot is efficient in initial emotion categorization and response generation. As a non-alternative to therapy, it is a beneficial first-level support system, particularly for users afraid of seeking human contact.

VII. FUTURE WORK

Although the AI-based mental health chatbot demonstrated here is an inspiring initial-stage solution for measuring emotional well- being, many avenues are available for future growth and expansion.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

7.1 Multimodal Support

Future versions of the system could incorporate additional modalities like for speech and facial expression understanding for enhanced emotional comprehension. As multimodal systems have outperformed text-based systems for identifying psychological state in prior studies, this is to be expected given the richer analysis input available to them. When auditory or visual information was added, the chatbot could pick up a sliver of how the user is really feeling and then improved classification accuracy.

7.2 Personalization

The chatbot can be personalised for better user engagement and a more precise emotional analysis. The system can learn from history, adapt to a change in responses that notice subtle changes in emotional trends over time. Feedback to be personalized would help in enhancing the system performance to render individual help and interventions

7.3 Clinical Validation

Following as one of the next major steps for this chatbot in further development Is clinical validation. To validate that the diagnoses and responses by the system are correct and, should be deemed safe for actual use in clients, mental health professionals would have access. Psychiatrists could then fine-tune the system during clinical trials and demonstrate that it works in actual treatment settings, a basis for potential regulatory approval.

7.4 Ethical and Privacy Issues

When AI starts touching mental health, ethics and privacy are beginning. The next time it comes around, new implementations will need to think about user data being protected, anonymized and stored as required by regulations such as GDPR. It is also necessary to satisfy the needs of an non-prejudiced algorithm and make the system fair for all users.

7.5 Integration with Mental Health Services

One possible future evolution would be to integrate the chatbot into existing mental healthcare (Teletherapy sites, crisis hotlines etc.). This would enable the chatbot to serve as a triage system to which should redirect only if necessary, users to the relevant professional intervention. Tackling these will make the system a more excellent mental health aid, better being able to cater to more diverse users.

REFERENCES

[1] S. De Choudhury, M. Gamon, and S. Counts, "Predicting depression via social media," Proceedings of the 7th International Conference on Weblogs and Social Media, 2013, pp. 128–137.
[2] L. Fitzpatrick, K. Darcy, and C. L. O'Reilly, "Woebot: A chatbot for mental health," Journal of Medical Internet Research, vol. 21, no. 7, 2019, e12156. doi: 10.2196/12156.
[3] P. Calvo, S. D'Mello, and J. K. Gratch, "Affective computing and intelligent interaction," Proceedings of the 1st International Conference on Affective Computing and Intelligent Interaction, 2005, pp. 1–10.
[4] P. Guntuku, M. Y. Lin, E. Klinger, et al., "Detecting depression and mental illness in social media: Analyzing large-scale datasets," Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining, 2016, pp. 507–516.
[5] L. Shen, L. J. Ji, and J. K. Y. Lee, "Identifying suicidal ideation in social media posts using recurrent neural networks," Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 1079–1088.
[6] J. Losada and G. Crestani, "CLEF eRisk 2019: Early risk detection on social media," Proceedings of the International Conference of the Cross-Language Evaluation Forum for European Languages, 2019, pp. 1–10.
[7] A. Resnik, R. G. M. L. de Silva, and S. S. Y. Mendelson, "Ethical issues in computational psychiatry," Journal of Psychiatry & Neuroscience, vol. 43, no. 6, 2018, pp. 388–395. doi: 10.1503/jpn.180107.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

[8] R. Al Hanai, T. Desai, and H. Zhu, "Speech-based emotional recognition for mental health," Proceedings of the 20th International Conference on Multimodal Interaction, 2018, pp. 55–64.
[9] E. Chancellor and S. De Choudhury, "Ethical considerations in computational mental health research," Proceedings of the 1st Workshop on Ethics in AI, 2020, pp. 1–12.
[10] C. Inkster, N. S. Sarda, and P. G. Caverly, "Digital mental health tools in psychiatry: A changing landscape," The Lancet Psychiatry, vol. 5, no. 4, 2018