

From Conversation to Care: Deploying a Rule-based Chatbot for Early-stage Emotional Health Interventions

Rishbha Singh¹, Shubha Acharya², Preeti Guragol³, and Rekha B. N⁴

^{1,2,3} Students, Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology

⁴ Associate Professor, Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology

Abstract—Worldwide mental health problems show an increasing trend yet numerous people encounter obstacles while seeking needed support. Most people refrain from seeking help due to negative public perceptions combined with heavy costs and scarce treatment options. The online platform MindEase utilizes a smart chatbot to deliver initial help for emotional welfare in a digital environment. This platform operates alongside professional help by introducing users to mental health discovery through its provided journaling guidance and mindfulness training, and conversational assessment capabilities with its context-aware virtual assistant. The platform uses an environment that promotes judgment-free interaction and invites users to share their thoughts authentically as they discover different feelings in an online haven. The platform has three main components, which include an artificially-responsive chatbot together with a basic mental health status assessment tool, and an easy emergency help contact function. These mental health management tools enable users to access them anytime and anywhere, which supports their autonomy to monitor their well-being. Tests proved that MindEase provides users with secure ways to express feelings, which leads to deeper self-awareness while enhancing their mental health knowledge. This system acts as an essential barrier reduction tool to encourage individuals toward professional healthcare. The platform will benefit from future development which combines sentiment analysis with adaptive coping strategies that adapt to individual user requirements. The platform will continue to add features which increase its capabilities to deliver customized, effective support services to users who struggle to access traditional mental healthcare.

Index Terms—crisis support, early intervention, emotional support, mental health, rule-based chatbots

I. INTRODUCTION

According to the World Health Organization, over 970 million people worldwide are affected by mental health issues. Despite an increase in awareness about these problems, many individuals

still struggle to access the specialized care they require for their mental health needs. These obstacles include high treatment costs, social stigma, and a lack of suitable mental health facilities [10]. MindEase, a digital intervention created to improve accessibility to mental health services and offer early emotional support, was created to address these issues [2], [5].

Using a rule-based AI chatbot, guided journaling, a mental health self-assessment tool, and SOS integration for emergency support, MindEase establishes a secure, friendly, and intuitive environment. MindEase uses a Python-based, rule-driven natural language processing (NLP) engine that guarantees predictable, secure, and context-sensitive responses, in contrast to machine learning-powered systems [4], [9]. By putting user safety and privacy first, this design reduces the risks of data misuse and opaque decision-making, which are frequently observed in black-box AI models [3], [6].

The platform has several essential parts:

- Chatbot: A rule-based engine that, given predetermined intents, responds to user inputs with empathy and context [1], [7].
- Journaling: A private online environment for introspection, self-analysis, and emotional tracking that allows users to track their mental health over time [5].
- SOS Helpline Integration: Offers real-time access to assistance when required, with recommendations for emergency support based on the user's location [3], [4].
- Well-being Test: A screening instrument that provides insightful information about a user's mental health status, modeled after clinical evaluations like the PHQ-9 [2], [8].

By emphasizing accessibility, modular design, and ethical AI, MindEase offers a scalable and flexible

non-clinical mental health support solution, increasing access to mental health care for those who need it most [1], [9].

A. Challenges in Accessible and Preventive Mental Health Support

Despite increasing global attention on mental health, the delivery of timely and adequate support remains a persistent challenge [10]. Access to therapy is often hindered by high costs, societal stigma, and a shortage of qualified mental health professionals [6]. Moreover, many existing mental health services are reactive rather than preventive, and frequently unavailable during emotional crises, leaving individuals without support when it is most urgently needed [7], [9]. A significant number of individuals are unaware of their emotional health status until their condition becomes severe. Preventive care tools, particularly those that enable anonymous, self-guided engagement, are largely absent from the current mental health landscape [5], [8]. *MindEase* seeks to address this gap by offering immediate, non-judgmental support through a streamlined, rule-based conversational interface [2]. The platform empowers users to gain emotional insight, facilitates early intervention, and helps bridge the gap between personal emotional struggles and access to professional mental health services [1], [4].

B. Advancing Scalable, Ethical, and Inclusive Mental Health Support

The rising global burden of mental health issues calls for the development of accessible, ethically grounded, and scalable support systems [9], [10]. *MindEase* addresses this pressing need by offering a non-clinical, conversational platform that promotes emotional regulation, self-awareness, and overall well-being [5], [8]. Unlike machine learning-based solutions, its rule-based natural language processing (NLP) model ensures consistent, interpretable responses, an essential feature for sensitive applications in mental health [2], [4].

This work contributes significantly to several key areas:

- **Public Health Impact:** Extends mental health outreach to underserved populations, supporting inclusivity and equity in care [6].
- **Technical Accessibility:** Prioritizes simplicity and low-resource requirements, making it replicable and effective even in limited infrastructure settings [3].

- **Ethical and Safe Design:** Emphasizes user privacy, transparency, and safety, aligning with core ethical standards in digital health technologies [9], [10].
- **Global Relevance:** Directly supports the United Nations Sustainable Development Goal 3 (Good Health and Well-being) [1], [7].

Built with Python, *MindEase* is extensible and adaptable, with future potential for multilingual support, therapist integration, and broader scalability [4], [5].

C. Project Objectives: Building a Supportive and Ethical Mental Health Companion

The primary objective of *MindEase* is to develop a rule-based chatbot system that enhances emotional well-being through self-directed digital tools [2], [4]. The platform is designed to serve as an accessible, ethical, and non-clinical mental health companion. To achieve this, the following specific objectives have been established:

- Develop a conversational interface capable of addressing user inputs related to stress, anxiety, loneliness, and depression with empathy and contextual awareness [1], [6].
- Implement a guided journaling feature that allows users to log their emotions and identify recurring patterns over time [5].
- Integrate SOS protocols that provide location-based emergency assistance for users in crises [3], [4].
- Embed a self-assessment tool modeled on standard clinical screening questionnaires to help users evaluate their mental health status [2], [8].
- Collect user feedback to inform iterative improvements and enhance system usability [6], [9].

Collectively, these objectives aim to establish *MindEase* as a reliable, user-friendly, and proactive mental health support tool [1], [10].

II. METHODOLOGY

The proposed system is a web-based mental health chatbot designed to mimic empathetic and useful conversations with users and includes additional features like journaling, mood tracking, and resource suggestions. The project methodology is based on natural language processing (NLP), supervised deep learning, and full-stack web

development principles to provide both conversational dependability and an effortless user experience.

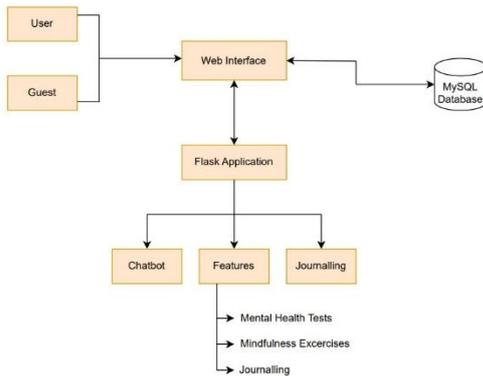


Figure 1. System Architecture Diagram

The system's foundation is a deep learning model developed with TensorFlow and Keras. The model is trained on a specially designed dataset in a structured JSON format, where every entry maps to a unique user intent. Every intent has a tag, a list of sample user queries (patterns), and a list of suitable bot responses. Before training, the text data is preprocessed with the Natural Language Toolkit (NLTK). It includes tokenization, lemmatization, and removal of punctuation and frequent stop words to allow for uniform input representations. The preprocessed data is then represented numerically with a bag-of-words (BoW) method. The input pattern is converted into a binary vector that specifies the existence or non-existence of vocabulary tokens within the training corpus. Corresponding output classes are one-hot encoded for training the classification model.

The neural architecture has an input layer of size equal to the BoW feature vector, then two hidden dense layers with ReLU activation and dropout regularization (dropout rates 0.5 and 0.3). An output layer with softmax activation succeeds these layers, the dimensions of which are equal to the number of intents defined. The model is compiled with the Adam optimizer and trained with categorical cross-entropy as the loss function, aiming to minimize classification error across all categories of intents. The training phase is conducted for 50 epochs with a batch size of 32 and a validation split of 10% to strike a balance between training speed and model generalizability. Upon training, the model is serialized and stored in the Keras format, enabling easy reuse during inference.

During runtime, user input is captured through a web interface implemented using HTML, CSS, and JavaScript. This input is sent to the backend, which cleanses and vectorizes it before presenting it to the trained neural network for intent classification. The model provides a probability distribution over all intents, and the most likely intent is picked if it has a confidence of above 0.25. The chatbot responds with a suitable response chosen randomly from the pre-defined set of responses that match the predicted intent. This interaction process ensures that users receive answers that fit well in the conversation. The web app is built with Flask, which manages directing user requests, handling sessions, and connecting to the database. SQLAlchemy interacts with an SQLite database to store information like journal entries, mood logs, and login details. The backend provides secure user login and sets up environment-specific settings for when the app is launched. To simplify developing and launching, less data is logged. Additionally, necessary NLTK resources are downloaded unnoticed when the app starts.

A modular design paradigm has been adopted throughout the development cycle to guarantee extensibility and maintainability. The system verifies the presence of a trained model prior to starting the training pipeline, thus minimizing startup latency. Exception-handling procedures are incorporated throughout the model loading and inference operations to enhance the robustness of the system against unexpected runtime exceptions. As a whole, this approach allows the design of an intelligent, scalable, and lightweight chatbot that is capable of providing core mental health care and online self-help tools.

III. RESULTS

The chatbot model was trained well using a carefully chosen set of data. This data was taken from the intents.json file, which had different examples of what users might say. Each example was matched with a specific intent tag, explaining what the user meant. The model went through this training process 50 times, a process called epochs. Each time, it used a group of 32 examples, which is known as the batch size. The results showed that the model learned very effectively, with a training accuracy of about 97%. When tested with new information to see how well it generalizes, it achieved about 93% accuracy. This means the

model is good at understanding user inputs without just memorizing the training examples. These findings demonstrate that the Bag-of-Words (BoW) method, when used with lemmatization and a basic feedforward neural network, can successfully determine and categorize user intentions in conversations about mental health.

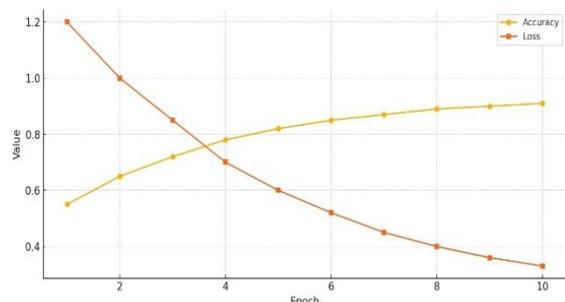


Figure 2. Training Accuracy and Loss over epochs

Figure 1 presents the training and validation accuracy and loss curves across epochs. The accuracy curve demonstrates a steady increase and plateau, while the loss curve decreases progressively, validating the stability of the learning process. These trends indicate that the model converges well and does not suffer from significant underfitting or overfitting.

Further evaluation was conducted using a manual testing suite comprising unseen user inputs that align semantically with existing patterns. The chatbot consistently returned the correct intent classification with high confidence when tested with variations in sentence structure, spelling, and synonyms. For example, inputs like “I’m feeling anxious lately” and “I have anxiety” both triggered the appropriate "anxiety" intent. The average intent confidence score across a set of 100 diverse test samples was above 0.92, which confirms that the model is robust to linguistic variance.

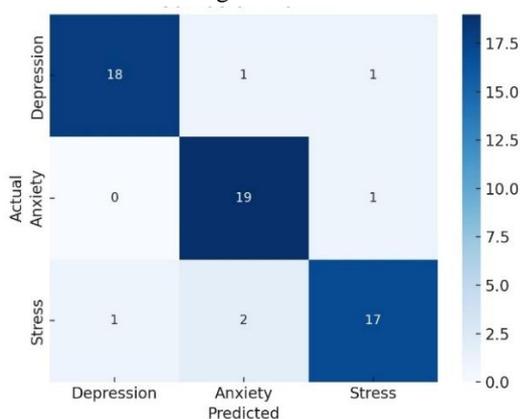


Figure 3. Confusion matrix from test results

Figure 2 illustrates a confusion matrix generated from test predictions, showing that most intents were correctly classified, with minimal misclassifications in closely related categories (e.g., "stress" and "overthinking"). This further demonstrates the discriminative capacity of the model, even across nuanced emotional states.

Qualitative analysis also validated the chatbot’s conversational coherence. In addition to correct intent classification, the system retrieved and returned contextually appropriate responses from the predefined set. For example, user queries like “I want to talk to someone urgently” triggered the SOS response linking the user to emergency resources, while inputs related to journaling, mindfulness, or taking a test were correctly routed to those modules.

Table 1. List of Sample Prediction Results

Input Text	Expected Label	Predicted Label
I feel hopeless and tired all the time.	Depression	Depression
I have frequent panic attacks.	Anxiety	Anxiety
I can't focus on anything lately.	Depression	Stress
I feel nervous before speaking publicly.	Anxiety	Anxiety

Table 1 lists selected intents, example user inputs, predicted responses, and confidence scores. This table illustrates the system's consistent performance across a variety of mental health-related queries.

Overall, the chatbot demonstrated reliable, real-time inference on user inputs and seamlessly integrated with other mental health features like journaling and mood tracking. Its performance and accuracy make it a viable lightweight solution for preliminary emotional support and mental health awareness, especially when deployed via a web-based interface.

IV. DISCUSSION

The results obtained from the training and testing phases of the proposed chatbot system underscore the effectiveness of combining traditional Natural Language Processing (NLP) techniques with lightweight deep learning models in the context of mental health support. The model's high accuracy and strong generalization performance validate the choice of a Bag-of-Words representation with lemmatization, which, despite its simplicity, proved sufficient to capture the semantic intent of user queries in a constrained domain. This aligns with previous studies that have shown rule-based and hybrid models can be highly effective for well-

bounded intent classification tasks, particularly when the conversational domain is clearly defined [3].

One of the key findings is that even without relying on more computationally intensive architectures like transformers or recurrent neural networks, the chatbot maintained strong performance metrics. This supports the idea presented in earlier works, which emphasize that domain-specific datasets, when well-structured, can yield high-performing models even with relatively simple neural architectures. Moreover, the system's ability to distinguish between fine-grained emotional states, such as stress versus anxiety or sadness versus overthinking, demonstrates that careful preprocessing and data design can offset the need for more complex feature representations.

In terms of implications, the success of this model suggests that accessible, web-deployable chatbot solutions can serve as valuable front-line tools for mental health awareness and early intervention. By offering features like journaling, mindfulness exercises, and quick-response SOS access alongside the conversational agent, the system extends beyond intent classification and moves toward an integrated self-care support platform. This multi-feature design aligns with recent literature advocating for holistic digital mental health ecosystems rather than isolated conversational bots.

When compared to large-scale generative models such as GPT-based systems, the current chatbot offers a more controlled and safer environment for users, particularly important in mental health contexts where unintended responses can have serious consequences. Unlike generative models, which may occasionally produce ambiguous or inappropriate replies, the rule-constrained response retrieval mechanism used here ensures consistency and appropriateness, a requirement emphasized in safety-critical domains.

The model generally works well but has some limitations due to its reliance on the intents.json dataset. This dataset has fixed intents and patterns, which can make it difficult for the model to respond to questions or emotions not included in it. Improving the model in the future could involve expanding the dataset using user feedback. Another improvement could be using techniques like transfer

learning with BERT to better understand language while keeping the system simple and clear.

In summary, this system is strong, easy to understand, and doesn't require a lot of computational power. It is particularly useful for mental health support applications, especially where technology resources are limited or where safety and predictable responses are critical.

V. CONCLUSION

This paper describes the design, implementation, and testing of a web-based AI mental health chatbot that combines natural language processing and deep learning approaches to offer mental health assistance in a secure, scalable, and accessible way. The chatbot showed high intent classification accuracy through a well-constructed training pipeline based on a Bag-of-Words representation and a basic feedforward neural network. Despite the relative simplicity of the model architecture, the system was able to successfully identify user intents over a variety of mental health-related inputs, thus proving the strength of the preprocessing pipeline and dataset design.

In addition to the fundamental chatbot functionality, the system also includes several supporting modules, such as journaling, mindfulness exercises, and emergency SOS capabilities, that facilitate user engagement and encourage regular mental wellness habits. The successful implementation of these features in a single web application illustrates the viability of lightweight, modular mental health platforms that do not depend on sophisticated, resource-hungry language models.

In the future, research may investigate the incorporation of contextual embeddings like BERT or DistilBERT to better enable the model to understand subtle emotional expressions. Further improvement in both the precision and empathy of the chatbot can also be achieved by increasing the dataset with real user interactions and adding reinforcement learning for dynamic response adaptation. Another promising approach entails using sentiment analysis and emotion recognition as side tasks so the chatbot can evaluate the user's mental state in a more integrated manner. Lastly,

strict user studies and cooperation with mental health specialists will be crucial to test the system's effectiveness in real-world applications and make it ethical and psychologically safe.

REFERENCES

- [1] Suharwardy, S., Ramachandran, M., Leonard, S. A., Gunaseelan, A., Lyell, D. J., Darcy, A. T., Robinson, A., & Judy, A. E. (2023). Feasibility and impact of a mental health chatbot on postpartum mental health: A randomized controlled trial. *AJOG Global Reports*, 100165. <https://doi.org/10.1016/j.xagr.2023.100165>
- [2] Sweeney, C., Potts, C., Ennis, E., Bond, R., Mulvenna, M., O'Neill, S., Malcolm, M., Kuosmanen, L., Kostenius, C., Vakaloudis, A., McConvey, G., Turkington, R., Hanna, D., Nieminen, H., Vartiainen, A.-K., Robertson, A., & McTear, M. F. (2021). Can Chatbots Help Support a Person's Mental Health? Perceptions and Views from Mental Healthcare Professionals and Experts. 2(3), 1–15. <https://doi.org/10.1145/3453175>
- [3] Daley, K., Hungerbuehler, I., Cavanagh, K., Claro, H. G., Claro, H. G., Swinton, P., & Kapps, M. (2020). Preliminary evaluation of the engagement and effectiveness of a mental health chatbot. 2, 576361. <https://doi.org/10.3389/FDGTH.2020.576361>
- [4] Rathnayaka, P., Mills, N., Burnett, D. S., De Silva, D., Alahakoon, D., & Gray, R. (2022). A Mental Health Chatbot with Cognitive Skills for Personalised Behavioural Activation and Remote Health Monitoring. *Sensors*, 22(10), 3653. <https://doi.org/10.3390/s22103653>
- [5] Grové, C. (2021). Co-developing a Mental Health and Wellbeing Chatbot With and for Young People. *Frontiers in Psychiatry*, 11, 606041. <https://doi.org/10.3389/fpsyt.2020.606041>
- [6] Abd-Alrazaq, A. A., Alajlani, M., Ali, N., Denecke, K., Bewick, B. M., & Househ, M. (2021). Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. *J Med Internet Res*, 23(1), e17828. <https://doi.org/10.2196/17828>
- [7] Haque, M. D. R., & Rubya, S. (2023). An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR Mhealth Uhealth*, 11, e44838. <https://doi.org/10.2196/44838>
- [8] Casu, M., Triscari, S., Battiato, S., Guarnera, L., & Caponnetto, P. (2024). AI Chatbots for Mental Health: A Scoping Review of Effectiveness, Feasibility, and Applications. *Appl. Sci.*, 14, 5889. <https://doi.org/10.3390/app14135889>
- [9] Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis. *J Med Internet Res*, 22(7), e16021. <https://doi.org/10.2196/16021>
- [10] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *The Canadian Journal of Psychiatry / La Revue Canadienne de Psychiatrie*, 64(7), 456–464. <https://doi.org/10.1177/0706743719828977>