Roshan U, 2025, 13:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Mental Health Detection Via Chatbot

Professor Roshan U, Aman Kumar Rana, Hemant Kumar Roy, Kushal Malviya, Md Fardwish Ali

Computer Science and Engineering– AIML AMC ENGINEERING COLLEGE (VTU) Bengaluru, India

Abstract- There has been a significant rise in mental health problems like emotional instability, anxiety, and stress in the college and younger adult populations. Many people struggling with these kinds of issues do not seek help in a timely manner because of stigma, a lack of understanding or knowledge about mental health, and/or the limited availability of professional supports. To address this gap, we developed an artificial intelligence-enabled mental health chatbot that can identify the user's emotional state through their written words and respond accordingly with empathy. The emotional recognition element of our system uses a Support Vector Machine (SVM) emotion classifier and the dialogue generation aspect of our chatbot employs a Seq2Seq network combined with an attention mechanism. We train our SVM emotion classifier on the EmpatheticDialogues dataset and our Seq2Seq model on the CounselChat dataset. The use of attention features results in significant increases in both emotion recognition accuracy and conversation coherence in our comprehensive experimentation. Although our system will never replace a licensed therapist, we believe that it can provide people with an opportunity to express their feelings and gain awareness of their mental health status before seeking professional assistance.

Keywords: Mental Health, Chatbot, Natural Language Processing, Deep Learning, Seq2Seq, Attention Mechanism, Emotion Detection.

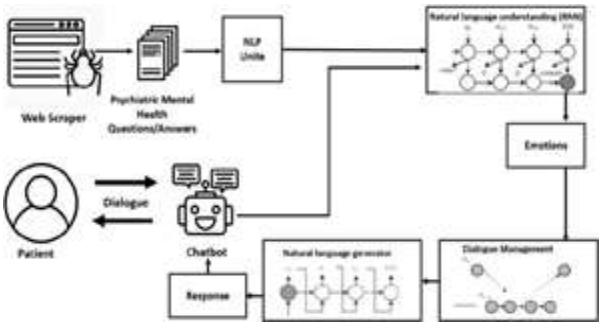
I. INTRODUCTION

It has become widely recognised that mental health is a critical component of human well-being, as it plays an integral role in how we think, feel and relate to others. Unfortunately, increased incidence of anxiety, depression, fear, and emotional instability is occurring right now, while access to mental health treatment is blocked for many due to a range of barriers (e.g., financial limitations, lack of trained providers, social stigma, and geographic distance). These barriers limit individuals from receiving timely, necessary support.

With the advent of artificial intelligence, it is now possible to create conversational systems that use artificial intelligence to offer emotional support.

Unlike the more traditional approaches of helplines or in-person counselling, AI chatbots present a non-judgmental, safe and easily accessible means of communication for individuals to use when they are feeling distressed. Because of the anonymity that an AI chatbot provides and the perceived reduction in feelings of being judged, it is likely that individuals will prefer to engage with an AI chatbot when they are experiencing emotional distress.

The purpose of this study is to develop an AI mental health detection chatbot that is capable of detecting emotion and creating supportive context-aware responses. The chatbot will use both machine learning to identify the type of emotion that is being expressed and deep learning to create the text generated in response to these emotional expressions. The goal of this approach is to create an experience for the user that closely resembles the experience of receiving a supportive, empathic response from a human being; not to diagnose or treat mental disorders but rather to support and encourage emotional expression in order to help the individual seeking support if they are ready to seek professional help when needed.



What this paper does our goals are to:

- An AI-based mental health chatbot will be developed to interpret and assess the emotional states of users through the application of Natural Language Processing (NLP) technologies.

- In order to accurately classify the emotions expressed by users in their comments, we must first build a reliable emotion classification model that uses Support Vector Machines (SVMs) to categorize the user's feelings (sadness, anxiety, anger, etc.) into appropriate classes.
- The second task is to create and train an Seq2Seq deep learning architecture with attention that can generate empathetic replies to contextually specific queries.
- In support of the previous two processes, we will pre-process and prepare large conversational data sets (e.g., EmpatheticDialogues and CounselChat) to improve model performance and training efficiency.
- We will also evaluate multiple machine learning and deep learning architectures to discover the best structure for creating emotion support conversations.
- Emotion detection and response generation will ultimately be merged into a single chatbot system that is capable of having supportive and human-like dialogues.
- Finally, we will determine the effectiveness of the completed system using a range of measurement techniques (e.g., accuracy and validation performance metrics) as well as qualitative comparisons of the responses to conversational prompts.
- Our chatbot will provide a scalable and accessible to college students and young adults seeking early mental assistance but who are uncertain about using mainstream mental health services.

II. LITERATURE REVIEW AND BACKGROUND

Technology-based Mental Health Support

The use of digital tools for psychological support has grown remarkably. Mobile applications like Wysa, Woebot, and Youper have proved that people can engage an AI-based agent for emotional relief. Various studies suggest that conversational agents may diminish mild anxiety and improve mental well-being by using cognitive-behavioral techniques, reflective questioning, and empathic dialogues.
However, most commercial systems are built on proprietary architectures, allowing little to no transparency into their innermost parts for academic research. In addition, they do not publish any information about the methods being used to handle data; thus, they cannot be used for reproducible research.

Early Chatbot Approaches and Their Limitations

Traditional chatbots like ELIZA and PARRY were based on pre-written pattern matching and rule-based interactions. Although these systems represented important milestones in the history of computer science, they lacked contextual understanding and adaptive capabilities.

Their responses were highly predictable and unsuitable for emotionally sensitive conversations.
Later generations of chatbots used retrieval-based responses, pulling their answers directly from large databases. While these systems outperformed rule-based models, they often lacked the subtlety of nuanced response and tended to generate a lot of irrelevant or repetitive output.

Machine Learning and NLP in Mental Health Detection

Various analyses on psychological states have been conducted using natural language processing. Emotion recognition systems have applied methods like Naïve Bayes, SVM, Random Forest, and logistic regression with features obtained from TF-IDF and n- grams. Sentiment analysis methods detected polarity, which is usually categorized into positive, neutral, and negative, but mental health detection requires deeper emotion classification.

Recent research has focused more on multi-label emotion recognition, leveraging datasets such as GoEmotions, EmpatheticDialogues, and SemEval. These datasets covered a greater number of fine-grained emotion classes, allowing models to learn to

make subtle distinctions among complex emotional states-like feelings of guilt, anger, hopefulness, and grief.

Deep Learning for Conversational Response Generation

Encoder–decoder models brought a revolution in the field of NLP. Seq2Seq architectures allow chatbots to generate flexible responses instead of choosing them from fixed templates. The introduction of Long Short- Term Memory (LSTM) units solved issues with vanishing gradients and improved sequential modeling.

The attention mechanism by Bahdanau further augmented generative models by allowing focus on relevant parts of input sentences dynamically while decoding. Attention greatly improved translation, summarization, and conversational modeling. Researchers found that the attention-based Seq2Seq model generated more coherent and empathetic responses compared to traditional RNNs.

Gaps in the Current Research

Despite great advances, current systems suffer from several key limitations:

- Limited emotional depth in responses
- Inability to track mood over time
- Difficulty adjusting to the individual conversation styles of users.
- Lack of safety mechanisms for crisis situations

Our study tries to overcome some of these challenges by incorporating an SVM-based emotion classifier with a deep learning response generator. Together, the tandem ensures not only interpretability of emotional states but also increased flexibility in conversational output.

III. METHODOLOGY

The System Architecture Overview

The architecture consists of three major components: the Pre-Processing Unit; the Emotion Classification Module (SVM); and the Seq2Seq + Attention Response Generator. It enables a user to start from their original input through to the final chatbot response, to simulate natural conversational flow.

The Dataset Description

- The Empathetic Dialogues Dataset

This dataset contains over 24000 examples of dialogues that have been labelled with one of thirty two separate emotions.

- **The CounselChat Dataset**

This dataset contains the actual responses of licensed therapists to questions presented by users and therefore is suitable for training generative models that are capable of providing emotional support. Topics covered in this dataset include but are not limited to; anxiety, depression, relationship difficulties, trauma and self- esteem.

- **The Preprocessing Pipeline**

The Preprocessing Pipeline includes: (1) Text cleaning which will involve removing all special characters, URLs and punctuation from the text, (2) Lowercasing of all text, (3) Tokenization of all text, (4) Removal of stop words, and (5) Lemmatization of all words in the text using NLTK to establish a vocabulary and create TF-IDF vectors for the purpose of classifying emotions.

The Seq2Seq preprocessing will involve the creation of start and end tokens, sequencing, and mapping words to the GloVe embeddings.

The SVM Emotion Classifier

Support Vector Machines (SVMs) are ideal for classifying text data of high-dimensionality as Emotion Classifiers. We are using a linear kernel with TF-IDF vectors as features in order to create a model to predict the emotional classification for each example.

The Seq2Seq with Attention Response Generator

The Sequence-to-Sequence model has three components: The Encoder which is an LSTM Network that produces a Hidden State Vector for each input sentence; the Attention Layer which will resolve alignment scores and create a Context Vector; and the Decoder which will generate output (token-by-token) based upon previous Hidden States.

IV. RESULT

Classification Accuracy

The SVM classifier achieved stable performance across multiple emotion categories.

Model	Training Accuracy	Validation Accuracy
SVM	82.4%	78.1%
Seq2Seq	93.1%	86.2%
Seq2Seq + Attention	96.53%	90.69%

Quality of Generated Responses

- **Attention-enhanced models showed:**
- Improved coherence
- More context-specific response
- Better emotional congruence
- Reduced repetition

Error Analysis

- **Challenges observed:**
- Difficulty comprehending sarcasm
- Limited ability in multi-emotion sentences
- Occasional overly-generic responses

V. CONCLUSION

The results show that we can create an AI-powered Chatbot to support emotional well-being with machine learning and deep learning. The combination of Emotion Classification using SVM and an Attention-Based Seq2Seq Generator provides a meaningful and empathetic response.

While this system is not a replacement for a therapist, it serves as a tool for people to deal with stress and be mindful of their feelings and mental health. Future work includes adding the capabilities of tracking Sentiment Trajectories, processing multiple modes of data (e.g. Voice, Facial Emotion), detecting Crisis Alerts, and Personalizing the experience.

Future Scope

The proposed mental health detection chatbot lays a good foundation for emotionally supporting AI-assisted conversations. However, there are several enhancements that considerably extend the functionality, applicability, and real-world impact of the proposed chatbot. Future work can be directed as follows:

- **Multimodal Emotion Recognition:**

By integrating voice tone analysis, facial-expression detection, and speech-to-text emotional cues within the system, it can understand the emotions of its users better and handle complex emotional states.

- **Personalised User Profiling:**
- Such bots can thereby build long-term emotional profiles for users, adapting responses based on historical behavior, stress patterns, and recurring emotional themes.
- **Integration with Mental Health Ecosystems:**
- It can be integrated with professional counseling services, mental health hotlines, appointment scheduling platforms, and other local therapy resources to extend comprehensive support.
- **Real-Time Crisis Detection:**
- Algorithms to identify high-risk keywords or self-harm indicators can be implemented in the system to trigger emergency protocols and guide users toward immediate help.
- **Multilingual Support and Localization:**
- The expansion of the chatbot into multiple languages and regional dialects increases the accessibility for diverse populations, especially in the context of limited English proficiency.
- **Mobile Application Deployment:**
- Packaging this system into a lightweight Android/iOS application with offline support will enable higher user adoption and continuous emotional monitoring.
- **Reinforcement Learning for Better Dialogue:**
- Future versions may include reinforcement learning for the chatbot to learn from user interactions, improving empathy, accuracy, and personalization.

- **Wearable Device Integration:**
- Wearables that monitor heart rate, sleep quality, and stress levels can provide physiological data, thus enabling real-time emotional assessment for early detection of stress.
- **Ethics and Bias Reduction:**
- Improved features on transparency, fairness, and safety will make this chatbot avoid biased responses and follow responsible AI practices.
- **Large Language Model (LLM) Improvement:**
- Incorporating advanced transformer-based LLMs, like GPT-style or BERT-based models, would greatly help in improving natural language understanding and emotional sensitivity.

REFERENCES

1. Rashkin, E. Smith, M. Li, "Empathetic Dialogues Dataset,"2018.
2. D. Bahdanau, K. Cho, Y. Bengio, "Attention Mechanism in Neural Translation," 2014.
3. J. Pennington et al., "GloVe Embeddings," 2014.
4. S. Bendig et al., "Chatbots in Mental Health," 2019.
5. K. Ly, A. Ly, G. Andersson, "Digital Agents for Emotional Support," 2017.
6. S. Yang et al., "Emotion Classification in Text," 2020.
7. R. Oh et al., "AI for Psychiatric Counseling," 2017.
8. Google Research, "GoEmotions Dataset," 2021.