Artificial Intelligence Powered Chatbot for Mental Healthcare based on Sentiment Analysis

Publisher: IEEE Cite This PDF

Ansh Mehta ; Sukhada Virkar ; Jay Khatri ; Rhutuja Thakur ; Ashwini Dalvi **All Authors**

Abstract:
The importance of good mental health has been realised in recent times especially after the advent of the pandemic which has further worsened the mental health conditions. An individual facing mental disorders or having a tough time mentally often faces social stigma. Hence Chatbots can potentially prove to be a powerful tool for such people who are afraid of stigmatisation. This paper has been divided into two parts. The first part covers a detailed literature survey of the work that has been carried out in the field of medical chatbots and the algorithms that have been used in the proposed approach. The second part covers the implementation of the proposed system and the results obtained followed by the conclusion and future scope. Bidirectional LSTM model was used for performing sentiment analysis on user's text and an accuracy of 80.88% was obtained. The chatbot functions on the basis of a neural network model that provides a minimal loss.

Published in: 2022 5th International Conference on Advances in Science and Technology (ICAST)

Date of Conference: 02-03 December 2022

DOI: 10.1109/ICAST55766.2022.10039548

Date Added to IEEE *Xplore*: 13 February 2023

Publisher: IEEE

ISBN Information:

Conference Location: Mumbai, India

SECTION I.
Introduction

Personal health monitoring has become very important for every individual and nowadays personal health not only includes physical health but also mental health. Mental health is something that has come into picture in recent times and people have started to realise the importance of being in a good mental space. However, several surveys suggest that people facing mental health issues find it difficult to open up about their issues and the resources to assist them are quite limited. Another survey conducted by Deloitte suggested that emotional well being apps will see a growth of 31.9% in 2022 as compared to 2021. In order to create a source where people can let out their emotions and be open about their issues without the fear of judgement, a chatbot was the most feasible option and in order to understand the users' emotions, sentiment analysis was implemented and further was combined with contextualization in order to provide more personalised responses to the users.

SECTION II.
Related Work

A systematic review of medical chatbot techniques was described in [1] and almost all the methodologies that can be followed to build a medical chatbot were covered. The author has reviewed 27 different medical chatbots and has mentioned the components like algorithm for the model, dataset and pros and cons for each of these chatbots. A variety of methodologies like RNN, Ensemble Learning, SVM, NLP, Knowledge Graph, etc. have been adopted for chatbots serving a wide range of applications. The PRISMA checklist methodology has been used for framing the paper and the author explains the methodology. This paper serves to be an efficient guide as to what purposes have been served by chatbots and what more could be done in the field. It also gives a detailed insight into the algorithms that can be used for building a specific chatbot.

[2] surveyed the sentiment analysis techniques performed on twitter data. The paper gave a detailed understanding of the concept of sentiment analysis and covered basics like the difference between opinions, views, beliefs and sentiments. Furthermore, the preprocessing of data was explained and then feature extraction like determining word frequencies, parts of speech, etc. was carried out. Next, Naive Bayes and SVM were the two classification techniques covered. The author explained the two approaches for sentiment analysis and further compared the performance of different methods on different datasets. Further the various levels of sentiment analysis were discussed and the evaluation metrics for the sentiment classification were explained. Likita [3] is a medical chatbot that has been developed to improve Healthcare delivery in Africa. The objectives of this chatbot included diagnosing common diseases like Malaria and tuberculosis, Offering medical advice, booking appointments and reminders. It was based on 3 pillars of AI namely, Natural Language Processing, knowledge management and Machine learning. It aimed at incorporating speech recognition to provide a voice to voice service and sentiment analysis for tracking user's behaviour. A methodical literature review of a medical chatbot from the perspective of behavior change was conducted in [4]. Functionality of a medical chatbot and its components like database interaction, response, etc. were covered. Further it explained the Theory of planned behaviour and its conceptual model. Next, it covered the detailed stage wise explanation of the transtheoretical models and chatbots. It also explains the concepts that go behind the behaviour change perspective like Perceived behavioural control, subjective norms, etc. [5] discussed a medical chatbot that offers services like disease prediction, general health related queries, age based medicine dosage, medicine details on medicine name and an Online API based service to process voice messages. The proposed system analysed SVM, Naive Bayes and word order similarity and the results show that SVM gave an accuracy of 94.6%.

[6] described a basic medical chatbot that gave health related information on the basis of the data stored in the sql database. The process included tokenizing the inputs, assigning them weights using TF-IDF, using N-gram method to extract the keywords and retrieve and display the information on the basis of the similarity computed with keywords using cosine similarity. The medical chatbot developed in [7] involved obtaining the user inputs, finding a class with needed probabilities based on the model used and picking a response at random from the class. The chatbot was a contextual chatbot and used technologies like Tensorflow, TFLearn, NLTK, etc. Medbot [8] was another chatbot developed for delivering telemedicine during the pandemic. It was a stateless chatbot and its backend infrastructure relied on Firebase and Google cloud platform. Dialog Flow API was used to build an automated conversational chat system that understands natural language to converse with users. A futuristic approach involving chatbot for medical assistance has been discussed in [9]. The features of this chatbot included, age based dosage details, solving queries related to medicines, symptom based disease prediction. The chatbot was easily upgradable and integrable and had cross platform compatibility. Psychiatric guidance for mental health through chatbot by analysis of emotional conversation and pointer network model based sentence generation has been described in [10]. The system kept a track of the user's changing emotions through various emotional intelligence techniques such as multi-modal emotion recognition which considers facial expressions, intonation and content of conversation. Unlike most other systems that work on a template or a set of rules, this chatbot generated sentences by psychiatric intervention and machine intelligence. A scoping review of the various features offered by mental health chatbots has been conducted in [11]. The guidelines given by the PRISMA extension have been followed whileconducting this review and the results and findings havebeen presented in a systematic manner. The percentage ofchatbots relying on a particular model or following a rulebased approach has been mentioned. The study acts as aguide to select the approach and the model to be used whiledesigning a mental healthcare chatbot. The effect of negation on sentiment analysis has been explained in [12]. The proposed architecture involved tokenizing the sentence and determining the dependency tree structure and the universal dependency for the tokens. Further, the polarity score is evaluated and lastly the polarity of the sentence and the polarity of the review determined.

SECTION III.
Proposed Methodology

A. Techniques Used

1) Sentiment Analysis

Sentiment analysis essentially involves extracting or identifying views, beliefs or emotions from a text. The opinions identified from the text are often categorised as positive or negative or neutral. This process is also often regarded as subjectivity analysis or opinion extraction. The preprocessing steps that precede the analysis algorithm can be word frequencies, parts of speech and tags identification, negation identification, position of terms, etc. Further any classification algorithm can be used to classify the sentences.

2) Negation Handling

When the author performs sentiment analysis, negations are very important constraints which need to be handled as they affect the overall meaning of the sentence. They include words such as not, no, couldn't and words with prefixes un-, dis- and so on.

There are different types of negations - Syntactical, Morphological, Double, Implicit - that the model should deal with. The syntactic negation includes all those negations which change the meaning of the words completely. Morphological negation includes the use of prefixes and suffixes to form negation. The words with prefixes un-,in-, dis-,il- and the words ending with suffixes like -less are categorised as morphological negations. Implicit negation includes a particular phrase which has a negative impact by use of some adjectives, adverbs or quantifiers. Double negation is the morphological negation that is present within a syntactical negation.

3) Bidirectional LSTM

LSTM or Long Short Term Memory networks are used for analysis of Sequential Data. LSTM Networks were further developed such that they could remem-ber and predict the given data in both forward and reverse directions, i.e. the input flows in both the directions. This improves the network's performance and accuracy as they nowknow what's the future, present and past. However, one needs complete information or data regarding the input sequence. In the use case discussed, the whole input is given at onceand hence, Bidirectional-LSTMs can be used for predictingthe sentiments. The image Fig 1. shows how the input xt-1 is first considered and the output is calculated using the forward and the backward layers (which deal with xT which is the last word)

In comparison with regular LSTM networks, Bidirectional LSTM networks are more effective. However, there is some amount of delay while using Bidirectional LSTMs when compared to the unidirectional ones. In Bidirectional LSTM, The forward and backward passes over the unfolded network are done in same way as a regular LSTM network with the exception that the hidden states need to be unfolded for each time steps. An extra step is required at the start and the endof the data points.

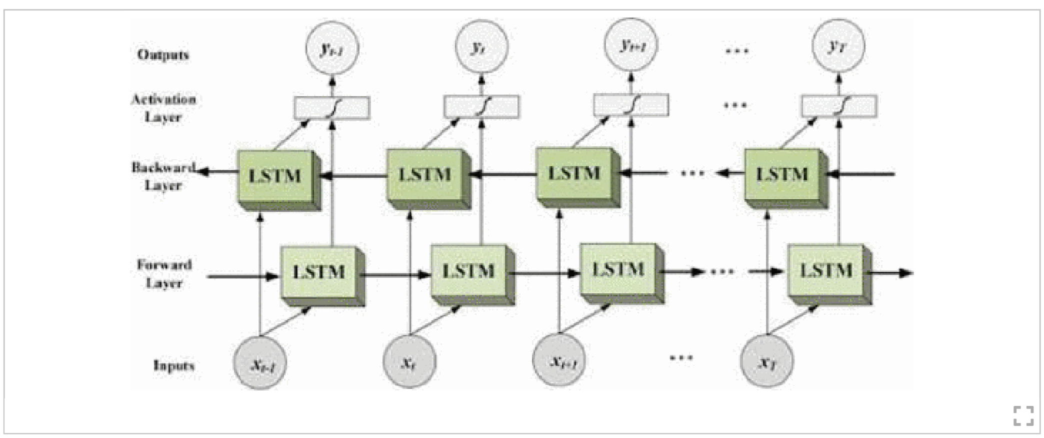

Fig. 1.
Bidirectional LSTM

B. Sentiment Analysis Model

1) Dataset Description

The dataset used by the author was “Sentiment 140 dataset with 1.6 Million tweets” (scraped from Twitter) which was publicly available on Kaggle. It had around 16,00,000 Tweets that were labelled according to the text present in the tweets. The target column includes labels 0, 2 and 4, where 0 meaning negative sentiment, 2 meaning neutral sentiment and 4 meaning Positive sentiment for the given tweet text. The Bidirectional LSTM model was trained and tested on this particular dataset and then employed for the purpose of Sentiment Analysis.

2) Data Preprocessing

The Sentiment 140 dataset tweets were preprocessed with the first step being tokenization using NLTK library's Tweet Tokenizer which is specially built for generating small tweet tokens. After tokenization, data cleaning was performed on the tweet text as it contains a large number of informal words and abbreviations. Further, these cleaned tokens were embedded using GloVe Embeddings which stands for Global Vectors for Word Representation. GloVe Embeddings are trained on 6 billion tokens with approximately 400,000 different words. Thus, a 'word to index' mapping was used to convert the cleaned tokens into their corresponding vector representations and a special value was assigned to unknown words denoted by 'unk'.

The author checks for negations and for punctuation marks to invert the meaning of syntactic words like couldn't, shouldn't. When the model was tested for double negation sentences like, "My mental health is not bad", the author obtained positive results to determine the correct sentiment of the user and further, implicit negation was also handled.

The processed tweet data is split into training and testing sets with test size as 0.2. Further, the embedded tokens are to be passed to the BiDirectional LSTM model. Data padding was performed to improve the accuracy and the resulting vector was fed into the word embedding layer in order to get the representation vector representing each word.

3) Bidirectional LSTM Model

The Sequential model first consists of an embedding layer which uses GloVe embeddings and then the input is further sent to the Mod which comprises 2 layers of Bidirectional LSTMs and a dense layer. The dense layer has one perceptron which tells us the sentiment of the given input in the range of (0, 1). If the output is near less than 0.1, it means that the sentiment of the given input is extremely negative (indeed the person is depressed). Further, if the model outputs a number between 0.1 and 0.3, it means borderline anxiety. If the output is between 0.3 to 0.5, it means the user is sad. Any output greater than 0.5 is considered neutral or happy. In this way, we can classify the given inputs into various segments and then give appropriate responses, which is discussed in the chatbot model section.

The author trained the model on the dataset and found out that the best accuracy was 80.88% with batch size as 512 and number of epochs as 10. The author then continued with the model for predicting the sentiment of the input from the user and used it as an input parameter for the chatbot model.

C. Chatbot Model

The chatbot model consisted of a simple ANN which would help us in finding the context of the conversation. For instance, if the user is worried regarding the test in which he might fail or the toxic work environment in his office, the responses should be different for each of them. Hence, the authors made a separate model for predicting the tag which indicated the category of input. This system has a few major categories which include: "greetings", "goodbye", "thanks", "about", "anxiety", "depression", "stress", "sad", "happy", "relationship", "workplace", "eating disorders", "childhood fears", "friends" and "family", "Depression Relief", "Stress Relief" and "Anxiety relief". An intents.json file is built to store the patterns and responses corresponding to each of the above mentioned tags so as to identify the context category referred to by the user during conversation with the chatbot. If the model can't categorise the input into these major categories, a general response is generated by the model and the input is categorised as "Others".

The model here only consists of 1 input and 2 hidden layers. All the layers have a ReLu Activation function and are densely connected. This model is trained on the intents.json file mentioned previously which consists of the possible input patterns for a particular tag. For instance: "tag": "greetings", "patterns":"Hi", "Hello", "Heyy", "Are you there?". Once the model is trained, it predicts the possible category of the input. Once the input is categorised, sentiment analysis is performed on the input and according to both the predictions, appropriate outputs are given to the user.

D. Chatbot Workflow

To design a contextual chatbot, it is important for the model to understand the context about which the end user wants to interact, to make the system more realistic. These were the steps followed by the user to implement the AI powered chatbot:

- Generating Intents: Create Intents.json file which contains tags with corresponding patterns and responses. Every intent consists of a tag, which contains the name of the intent and a pattern, which are the sentence patterns for the neural networks. To design a contextual chatbot, the author has classified the user's input patterns into different categories - friends, family, relationship, workplace and a few more so that the chatbot can provide more contextual response.

- Transforming intents to a PyTorch Model: The Neural Network model is trained on these intentsto identify the tag i.e the context related to which the user is conversing. Whenever the user provides an input to the chatbot, the input is processed and the NN model matches the input with the patterns from the intents file and the appropriate tag is predicted as the context from the intents.json file.

- Chatbot algorithm:

1) Input Preprocessing

Once the user enters a message into the chat, the input is tokenized, cleaned and preprocessed using Snowball Stemmer to send it as an input to the Chatbot NN model that categorises the input into a specific tag to determine the context based on the patterns described in the intents while training.

2) Identifying the Context

The contextual response to be given to the user is generated in 2 stages:

- Stage 1: The output from the Chatbot NN model is one of the tags such as 'workplace', 'relationship', 'anxiety relief' etc based on the input message pattern. If the tag is predicted with an accuracy <0.7 i.e the chatbot cannot predict the context accurately, then the chatbot provides a 'I do not understand' response to the user.

- Stage 2: Now, once the context tag is identified, there are 2 possible conversational paths basedon type of tag- First, general information tag which gives information related to anxiety, stress relief. Second, contextual tags which provide a more curated response by understanding theemotions of the user for a particular context suchas workplace, relationship and so on.

  - If the tag predicted is a general information related tag, a response is given to the userbased on the selecting a random response from a list of responses corresponding to the predicted tag from the intents file.

  - Else if the tag predicted by the Chatbot model is a contextual tag then a more curated re-sponse is given to the user by understanding the emotions of the user in that particular context.

3) Categorization into Sentiment Thresholds

For contextual tag, sentiment value is obtained and message is categorised into a threshold. Following are the thresholds for the generated sentiment value.

a) Sentiment value less than or equal to 0.1 is considered as extremely severe negative sentiment and the user might be feeling depressed and supportive resources are provided for help.

b) Sentiment value greater than 0.1 and less thanor equal to 0.3 is considered moderately severe negative sentiment and it is termed as borderlineanxiety.

c) For sentiment value greater than 0.3 and less than or equal to 0.5, the user is termed as sad andnot much severe in terms of negative sentiment.

d) Sentiment value greater than 0.5 is termed as positive sentiment and an encouraging response is provided to the user.

4) Generating a Chatbot Response

An empathetic, random response is chosen from the available responses for that particular context and the responsesare framed using Mental health support resources to portray human like emotions and presented as chatbot's response to the user's input.

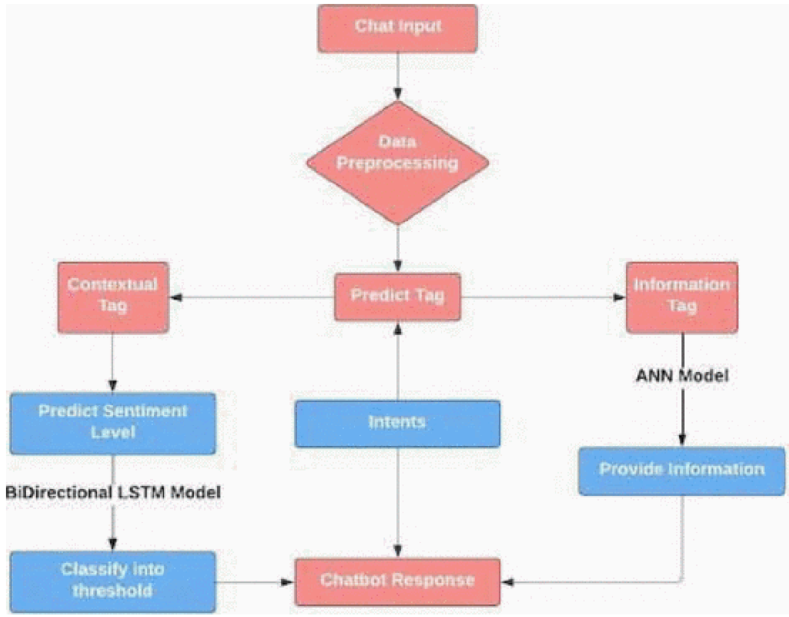

Fig. 2.
Chatbot workflow diagram

SECTION IV.
Results and Discussion

A. Sentiment Analysis Model Performance

On training the Bidirectional LSTM model on Sentiment 140 tweet dataset, an accuracy of 80.88% was obtained.

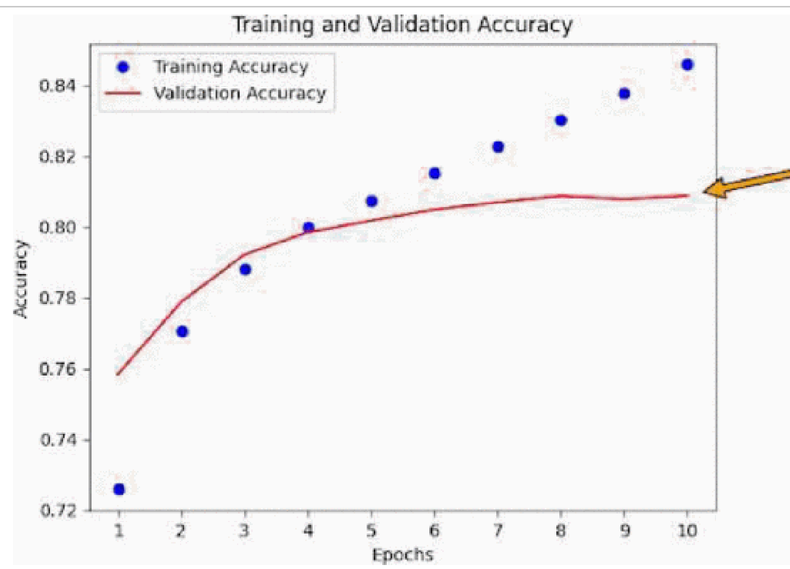

Fig. 3.
Graph for model's training and validation accuracy

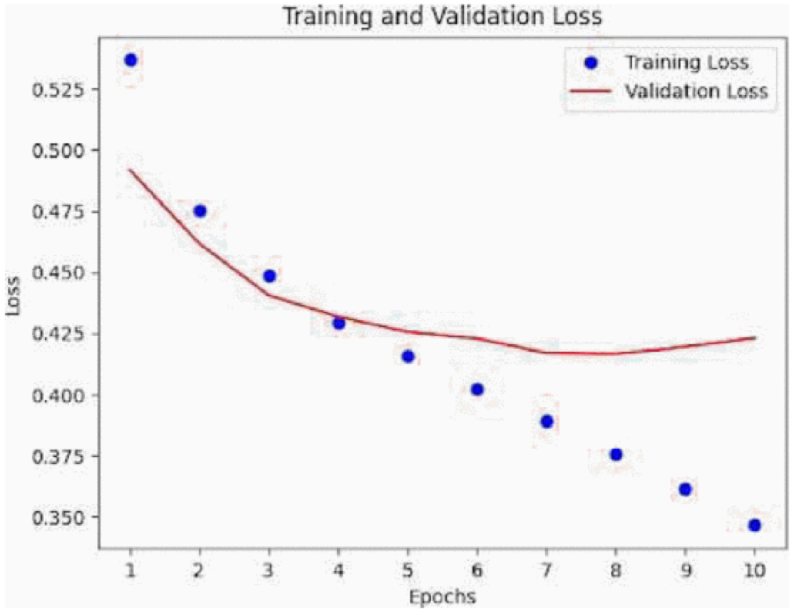

Fig. 4.
Graph for model's training and validation loss

Negative Sentiment Prediction Results

```
print(predict_sentiment(" I have been through a lot of things which have affected me"))
0.068

print(predict_sentiment(" I am feeling a little alone lately"))
0.417

print(predict_sentiment(" I wish I could open up about it and talk to someone"))
0.092
```

Fig. 5.
Negative sentiments

In the first example, the input sentence has extreme negative sentiment probably leaning towards depression and thus, the model has appropriately identified the severe negative andsad emotion by predicting the sentiment value as 0.068. Similarly, in the second example, the input sentence has a somewhat negative and unhappy tone but not too severe, and thus the model has given the sentiment value appropriately as 0.417. The third example sentence has a very dejected feeling associated with it, portraying melancholy and loneliness and hence, it is categorised as extreme emotion by the model by giving a sentiment value of 0.092.

Positive Sentiment Prediction Results

```
print(predict_sentiment(" I'm quite happy with my family , they support me when I am down"))
0.745

print(predict_sentiment(" I am so glad to have such supportive friends"))
0.997

print(predict_sentiment(" I received a promotion yesterday!"))
0.893
```


Fig. 6.
Positive sentiments

In the first example, the input sentence has a happy and positive sentiment and since the word “quite” is present, the model has appropriately identified the moderately happy emotion by predicting the sentiment value as 0.745. However, in the second example, the input sentence has a very positive and cheerful emotion since it has words such as “so” before “glad” which emphasises the feeling, the model has predicted the extreme positive sentiment value appropriately as 0.997. The third example sentence has a positive and happy tone and thus, the model has given a sentiment value of 0.893.

B. Chatbot Model Performance

The chatbot implementation includes training a neural network model which provides a minimal loss of 0.0185 and further the chat is loaded which appropriately identifies the appropriate context tag for user's message and randomly selects an output from the instructional intents framework if the text message has a probability greater than 0-7 in terms of being related to a tag.

C. Conversational Responses Results from the Chatbot

On conversing with the chatbot to get mental health support, the following interaction was obtained based on sentiment analysis and intent being identified using Chatbot Neural network model: The Chatbot identifies the context anxiety referred to by the user and offers help by conversing further to know about the user's feelings.

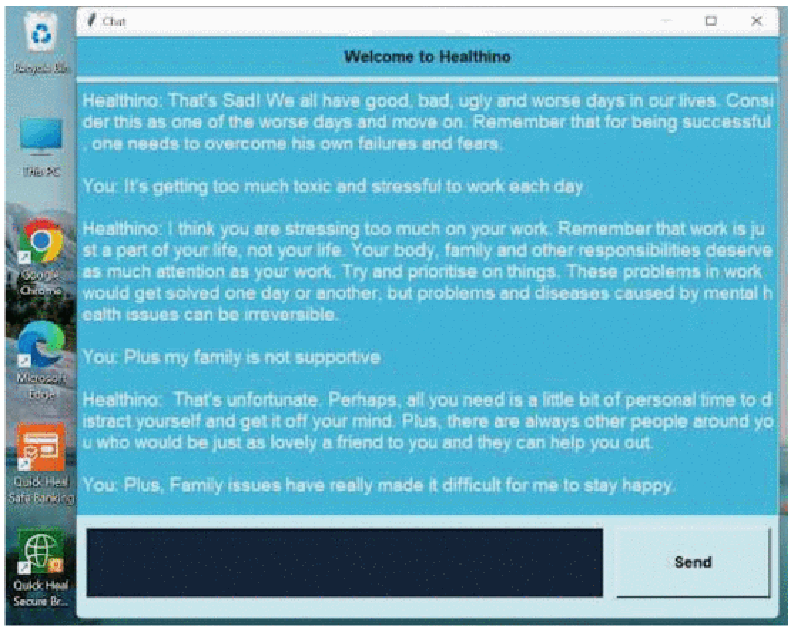

Fig. 7.
Chatbot conversation

PDF
Help

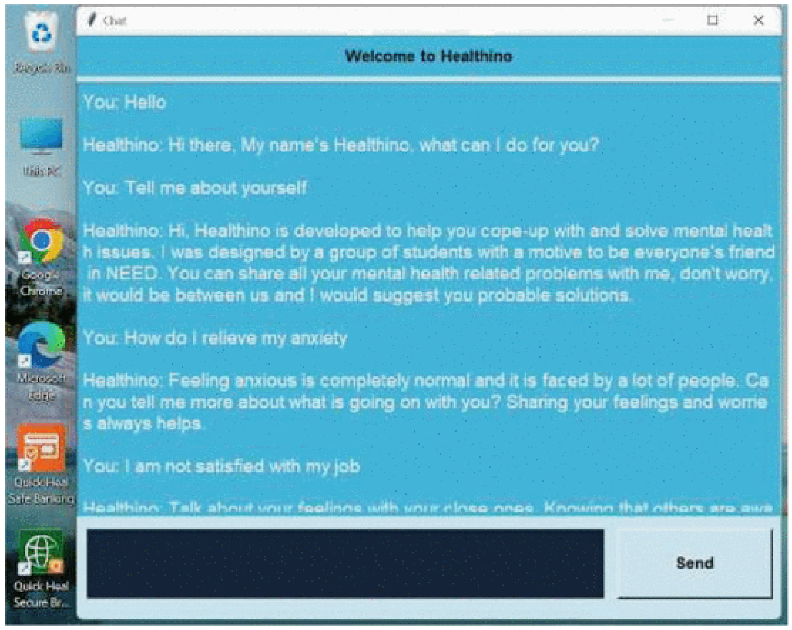

Fig. 8.
Chatbot conversation - workplace and family issues

When the user shares anxiety issues and its reason as workplace stress, the chatbot model categorizes the tag as ‘workplace’ and performs sentiment analysis on the message to determine the emotion severity and provides an appropriate and comforting response to ease the user's anxiety.

When the user explains further that they are facing family problems as well, the new context tag is identified accurately as ‘family’ and based on the severity of the user's sentiment, which is extreme, the appropriate guidance and response is provided to resolve the problem.

SECTION V.
Conclusion and Future Scope

The implementation of a mental healthcare chatbot heavily depends on AI and ML algorithms and training data as discussed in the paper. The chatbot developed serves the purpose of a listener for people who are scared to open up and tries to provide appropriate responses with it's major highlight being context i.e topic wise customized responses along with consideration of emotion severity. However, the chatbot is still in its early stages and as of now works on the basis of the intents laid down by the developer which tends to limit its scope. As part of the future work, the intents can be made more complex or the chatbot can work with generative algorithms to automatically generate appropriate responses. Moreover the accuracy for the sentiment analysis can be improved by considering more thresholds and types of sentiments.

Authors

Figures

References

Citations

Keywords

Metrics