Therap.AI: A Computational agent for Mental Health Aid

Madiha Mansoori
Department of Computer Engineering
NMIMS University
Mumbai,India
madiha.mansoori2001@gmail.com

Hrishil Maliwal
Department of Computer Engineering
NMIMS University
Mumbai, India
hvmaliwal@gmail.com

Sharvil Kotian
Department of Computer Engineering
NMIMS University
Mumbai, India
sharvilkotian99@gmail.com

Hersh Kenkre
Department of Computer Engineering
NMIMS University
Mumbai, India
hersh.kenkre31@gmail.com

Ishani Saha
Department of Computer Engineering
NMIMS University
Mumbai, India
ishani.saha@nmims.edu

***Abstract*—In recent times the rate of mental health disorders among the youth has spiked due to increased competitiveness which affects the mental well-being of students. An unhealthy mental state not only affects the daily life of an individual but is also the reason for increased self-harm and suicide rates. In developing nations, the ratio of mental care professionals to mental health patients is far too less for all to receive care. A solution to this problem is internet-delivered cognitive therapy (iCBT). The objective of this paper is to shed light on various techniques that can deliver iCBT to a patient in a comfortable manner. Since iCBT can be delivered from home, it tackles the challenge of societal stigma. Different existing approaches and solutions being implemented like various types of chatbots(SERMO, EMMA) and social robots(Ryan Bot)are analyzed and compared in this paper. We also analyze different types of existing datasets(NHS Mental Health Dataset, CounselChat Dataset, ISEAR Dataset) used to train various models(Convolutional Neural Networks, Recursive Neural Networks, Hierarchical Attention Network, Transformers). Word Count Per Session, Sentiment Analysis and Emotion Analysis were some of the evaluation metrics analyzed.**

***Index Terms*—Artificial Intelligence, Machine Learning, Mental Health, Conversational Agents, Natural Language Processing, Sentiment Analysis, Cognitive Behavior Therapy, Transformers, Neural Networks**

I. Introduction

Mental Health refers to the overall emotional and psychological well-being which affects the daily living, relationships, and physical health of an individual. Mental health is responsible for an individual's response to stress, ability to empathize with others, maintain interpersonal relations, and balance life activities. Taking care of one's mental health can improve one's capacity to appreciate life. To do so, you need to find the right balance between your regular lives, obligations, and efforts to improve your psychological resilience.

The current medical system is ill-equipped to assist patients in distress. According to statistics, 76-85% of individuals in low-middle-income nations lack adequate resources or qualified health care providers. The national health services are unable to meet the increasing demand for resources. In developed nations, there are 9 mental health providers per 100,000 individuals, while in low-income countries, there are as few as 1 per 100,000. Apart from the unavailability of therapists/psychiatrists, treatment sessions are expensive, making it difficult for the ordinary person to afford them. An hour of private counselling might cost anywhere from $70-150$. Couples treatment, marriage counselling, depression therapy, and other types of therapy can cost anywhere from $90-200$. Many people between the ages of 15 and 29 suffer from stress and anxiety. Mental illness affects approximately 20 million college students. In the United States, more than half of college students experience symptoms of depression, anxiety, and stress. Even though there is a definite need for clinical services, up to 75% of college students do not get adequate mental health care. *Table 1* depicts how around 450 million people worldwide suffer from some sort of mental disease.

Stigma, a lack of resources, and an unbalanced income-expenditure ratio all contribute to youth's lack of access to mental health care. Various researchers have come up with methods that leverage technology to make mental healthcare more accessible in order to tackle these difficulties. chatbots, backed by various psychological approaches and computational techniques, is a typical strategy that addresses most of the challenges and eliminates the need for a middleman while offering a sense of comfort to the user. This paper will discuss some of the strategies that have been used and have yielded positive outcomes as well as the technique that we have used. The rest of the paper is structured as follows: Section II talks about the literature review which includes certain existing bots, some datasets, existing approaches and Models and some evaluation Matrices. Our implementation is in Section III; and Conclusion of the paper is given in Section IV.

TABLE I
CLASSIFICATION OF DISORDERS BY PERCENTAGE OF POPULATION.

Disorders by Population		
Disorders	Percentage with Disorder	Population
Any mental disorder	10.7%	792million
Depression	3.4%	264 million
Anxiety disorder	3.8%	284 million
Bipolar disorder	0.6%	4.6 million
Eating disorders	0.2%	16 million
Schizophrenia	0.3%	20 million
Any substance use or mental disorder	13%	970 million
Alcohol disorder	1.4%	107million
Drug use disorder	0.9%	71 million

II. LITERATURE SURVEY

Mental health disorders continue to act as major health and social burden on individuals from various countries over the globe. Several methods to streamline the process of availing therapy are being discussed where various researchers and articles study Mobile application-based delivery of psycho-social interventions as a method that may assist in reducing the weight that mental health practitioners have to carry. According to studies conducted by WHO 29% of the 15,000 available mobile apps catering to mental health are focused on treatment techniques. These technological aids are used to bridge the wide ratio of Therapists to patients. Other techniques also include the usage of Social Robots which employ AI techniques to provide assistance in mental health problems.

A. Chatbot

A chatbot is an Artificial intelligence (AI) software that can mimic a natural language discussion with a user through the means of messaging apps, websites, mobile apps, or the telephone. A human user enters text on one side, while an AI bot responds on the other, attempting to act and impersonate a human. Chat bots are created in such a way that they convince the user that they are chatting with a real human through text. By today's standards, only a small proportion of chat bots can pass the Turing test, confirming their humanity. Natural language processing is used by chat bots to read and understand what the user has written. These chatbots are trained on suitable datasets which are designed to fulfill the goal of the system.
chatbots have transformed how technology may be utilised to help with mental health difficulties, and they are a convenient method that allows users to communicate their worries, feel heard, and find closure.

1) EMMA: EMMA stands for Emotion-Aware Mental Health Agent. EMMA[1] is a chatbot which incorporates Ecological momentary interventions(EMI) that is delivering appropriate just-in-time interventions. These interventions provided by EMMA are relevant micro-activities for mental wellness which are applied in an empathetic manner.It also keeps a record of the geo-locations and other detailed activities within the application to retrieve contextual information of the user from the phone.The movement threshold is set to 10 meters and the data is captured periodically every minute in the foreground and background.
The feasibility of automated just-in-time interventions by EMMA was tested in two week long human subject experiment with 39 participants.This study evaluated the design of EMMA, the machine learning models implemented for the automation of mood detection and the effectiveness in influencing the user's mental health. In order to put the retrieved data into to use a prediction engine is developed for the translation of the sensor data.The prediction engine made use of classification and regression models. A range of different classifiers such as Logistic Regression, Ridge, AdaBoost, Bagging, Random Forest, and Gaussian Processes were implemented as binary classifiers for valence i.e negative or positive and arousal i.e low or high. In order to model arousal and valence on a continuous scale Regression models were tried out such as several versions of linear regression like Ridge, Lasso, Elastic Net, Bayesian Ridge, Support Vector Regression, Gradient Boosting and robust to outlier methods like RANSAC, Theil-Sen, and Huber. Since individuals usually had varying baselines and they oscillated near those values a personalised regression was modelled which worked for such personal patterns where for every individual the average of valence (vb) and arousal (ab) was calculated.Finally the personalized model with Random Forest regression for valence prediction and AdaBoost regression for arousal prediction were used as they gave best results.The major drawback of EMMA was that the users over the time found the chatbot's feedbacks and messages repetitive and predictable.For the agent to be more effective the responses and the interventions should cater to specific moods and take into consideration the context while they are being assigned.

2) SERMO: SERMO [2] is a German chatbot which uses a Cognitive Behavioural Therapy (CBT) coupled with Natural Language Processing(NLP) based approach. Its distinctive feature was that it integrated emotion analysis methods to automatically determine emotions from the inputs provided by the user. Unlike, other chatbots which follow a decision tree based system, SERMO relied on semantic and syntactic similarities between the user inputs and the stored inputs from training data and not just on strict patterns. Its architecture consists of an Recurrent Neural Network(RNN) enabled with Gated Recurrent Units(GRU) cells having an attention mechanism. The chatbot was developed using the

'Syn.Bot' framework. The functionalities included the chatbot recognizing natural language user input, reminding the user about therapy appointments, providing an emotion slider and acting as a mood diary for the user. The emotion slider feature asked the user for their current mood and in scenarios of negative emotions, questioned the trigger for the emotion and suggested solutions to deal with the problem. Additionally it suggests a variety of activities to users like mindfulness exercises, leisure activities and relaxation techniques that can be practised outside the environment of the chatbot as well. These activities are found to be increasingly effective within CBT. The app was tested on 21 people, of which 9 were already undergoing treatment in a clinic. Upon usage and evaluation, SERMO was found to be stimulating for patients who experienced difficulty in expressing themselves in face-to-face interactions, they added that the app was understandable, easy to learn and was helpful in dealing with their emotions. The only major drawback of the chatbot was that it could only recognize a limited range of emotions, which could be overcome by training on more data.

B. Social Robots

Social robots an artificial intelligence operated systems that is capable of performing complex interactions with humans. Social bots are believed to take over many basic jobs such as customer services, although as of now, social bots are finding places in people's homes as they can be designed to have personalities and quirks thereby enriching the human-robot interaction (HRI) experience. Some famous social bots are hitchBOT, Kismet, Tico. These robots are used for tutoring, providing customer engagement and provided companionship. Companionship bots provide emotional support like the robot, 'Ryan' created in [3].

1) Ryan Bot: Ryan Bot [3] is a social robot created to deliver internet delivered CBT(iCBT) to senior citizens that are over the age of sixty Ryan uses a dialogue management system was called, *'Program-R'* to deliver therapy sessions. The verbal input of the user is converted into a string and is sent to Program-R to decipher and to generate a proper response. A unique feature of Program-R is that, unlike other chatbots, Program-R asked questions and started the conversation, rather than waiting for the user's initiation. The questions that Ryan asks are from 'Saint Louis University Mental Status Examination' (SLUMS), Patient Health Questionnaire item 9 (PHQ-9) and Geriatric Depression Scale (GDS). Artificial intelligence mark up language(AIML) was used to develop seven hour long iCBT dialogues. Ryan has an expressive digital animated face which is used to accurately depict visual speech and communicate smoothly with the user. This also comforts the user as they feel that they are conversing with a real human/therapist and not a lifeless robot. There is also a touch screen panel on Ryan's chest that is used to receive user input or provide some audio-visual effects in the form of images and videos. A remote control software called Wizard of Oz (WoZ), this can be used by the researchers to intervene, if necessary. This allowed them to control the flow of the therapy session without breaking the feeling of autonomy of the robot. The therapy was spread across seven sessions, first session was a brief introduction and allowed the user to get familiarized with Ryan. Sessions two and three were to study, how thoughts of the user affected their mood and mental health. Sessions four and five were to study, how activities affected the user's mood and mental health.The final two sessions, i.e. six and seven were to study, how people around the user affected their mood and mental health. Ryan bot demonstrated positive results with 3 out of 4 subjects showing an improvement in their mood. The improvement in their mood was gradual and increased at a faster pace as the therapy sessions progress. The reliability of the results was observed from the Cronbach score of the study responses.

C. Data-sets

1) NHS Mental Health Data-set: NHS Digital manages the Mental Health Data Set, which pulls together data collected as part of routine patient care by a variety of secondary care mental health services providers. It's a secondary uses dataset that was designed to keep track of anyone who is considered to have a mental illness and is receiving specialized adult mental health care in an NHS-funded institution, either partially or entirely.
[4] uses the NHS Mental Health dataset to train its SVM based classifier. The dataset was divided into two groups for training and testing. The cleaning and pre-processing steps were done by Multivariate Imputation by Chain Equations(MICE) which is a flexible method for tackling missing values.The original dataset with 31 features was imputed by the feature selector SVM RFE which eliminated 18 irrelevant and low-correlation features. Additionally a feature is added leaving the algorithm with 10 important features to train on.

2) Counsel-Chat Data-set: Counsel Chat is a professional network that connects counsellors with people who could be potential patients. The service allows therapists to react to questions posed by patients, and users can up-vote the responses that they believe are the most appropriate. These counselors are verified therapists, which enhances the reliability of responses. The forum features 31 topics, with several responses for the most of them. The themes vary from intimacy and relationships to self-harm and military traumas. The data set was prepared by scraping data from the website and comprises of ten distinct attributes like question text, topic, answer text, and upvote count.
R. Crasto et al. leveraged the Counsel chat dataset to diversify its Dialo-Generative pre-trained (DialoGPT) model model. Because the DialoGPT model is a conversational model that must be trained on a conversational training set, they construct columns for question type, question, ailment, and answer to provide additional context for fine-tuning the model's text production [5].

3) ISEAR Data-set: The International Survey on Emotion Antecedents and Reactions (ISEAR) dataset was compiled by a broad group of psychologists over the course of several years in the 1990s. The data set was developed by reporting students' both psychologists and non-psychologists moods across seven primary emotion categories, (joy, fear, anger, sadness, disgust, shame, and guilt). It has 7666 responses from a total of over 1000 people. The dataset usually contains phrases and is pre-processed, with operations such as tokenization, word vector generation, embedding, and so on.
F. Patel et al. proposed a chatbot which is trained in a manner where the users input is fed into the chatbot and then classified into one of the seven predetermined emotions [6].

4) Web Scraped data: Over the years, conversations around mental health have significantly increased on several social media sites like 'Reddit', 'Twitter' and other discussion forums. Several researcher have explored the possibility of gathering real world data from these discussion by using web-scraping tools like 'Beautiful-soup' and 'Twitter-API'.
Researchers in [7] created and released the Reddit Mental Health Dataset which includes over 800,000 posts made over two years in several groups like 'r/schizophrenia', 'r/SuicideWatch', 'r/Depression', 'r/HealthAnxiety'. This dataset covered concerns which arose during the pandemic and was helpful in understanding trends. *S. Nagargoje et al.* used the twitter sentiment analysis tool to analyse and learn the most frequent words used by depressed people [8] . DialoGPT [9] a pre-trained Transformer model by Microsoft was trained on a dataset created by scraping reddit data containing over 147M exchanges. This Scraped data is pre-processed by employing NLP techniques before it is used for training models.

D. Approaches

1) Cognitive behavioral therapy: Cognitive behavioral therapy (CBT) is a treatment approach that aims to help you identify and explore the ways your emotions and thoughts can affect your actions this helps the patient recognize negative or unhelpful thought and behavior patterns. Once the patient notices these patterns, they can begin learning to reframe their thoughts in a more positive and helpful way. CBT is based on the concept that your thoughts, emotions, and actions are all interconnected. In other words, a person's mental state has an impact on their actions and influences their behavior to a great extent. CBT has aided patients with a variety of issues and provided them with tools to deal with mental illnesses and conditions such as depression, eating disorders, post-traumatic stress disorder (PTSD), bipolar disorder, anxiety disorders, including panic and phobia, obsessive-compulsive disorder (OCD), schizophrenia, and substance abuse. CBT employs techniques such as setting smart goals,guided discovery and questions, journaling to keep a track,self monitoring ,indulging in positive activities,Dialectical behavior therapy (DBT), multi-modal therapy, Cognitive restructuring, and Rational emotive behavior therapy. CBT is one of the most researched types of therapy, owing to the fact that the treatment is focused on highly specific goals and the process of measuring the results is relatively convenient.
Researchers have studied the usage of VR delivered CBT approaches coupled with Explainable Artificial Intelligence (XAI) techniques to tackle Speech Anxiety, Agoraphobia and Driving Anxiety. One of the principles that CBT is based on, involve Exposure therapy where the patient is exposed to certain situations which might act as a Trigger [10], [11]. Computational methods are expected to face lesser reluctance as patients are more willing to face VR based stimuli in contrast to actually facing the trigger in the real world setting. XAI techniques were chosen as they provide easier understanding from psychologists who need to interpret how various treatment variables will interact with each other within the black-box.These methods showed positive results which were cost- effective and improved the existing CBT treatments

2) Acceptance and commitment therapy: Acceptance and commitment therapy (ACT) teaches mindfulness skills to help individuals live and behave in ways consistent with personal values while developing psychological flexibility while being in contact with ones inner emotions, and thoughts. ACT relies of 4 core values.
1.Developing creative hopelessness involves exploring past attempts at solving or getting away from those difficulties bringing an individual to therapy. Through recognition of the workability or lack of workability of these attempts, ACT creates opportunity for individuals to act in a manner more consistent with what is most important to them.
2.Acceptance of one's emotional experience to further learn the range of human emotions.
3.Taking action to commit to making changes in move in a positive direction.
4.Laying clear definitions of what's important in life to the individual.
ACT has enjoyed a steady rise in interest as an alternative therapeutic intervention to CBT. Currently ACT is being delivered through websites which help users by suggesting different exercises and methods to accept the issue being faced. These exercises often help with emotional and moral support and can be the first step to a better mental well being[12].

E. Models

1) Convolutional Neural Networks: CNN stands for Convolutional Neural Networks, it is a class of deep neural networks which are largely used in the field of image processing and computer vision. Applications commonly apply CNN algorithms for analyzing visual imagery, image classification, facial recognition, object detection. However Convolutional Neural Networks have also been proven to be effective in dealing with tasks problems associated with NLP and where they are implemented for carrying out processes like tokenizations, emotion detection, sentence classification, text classification, sentiment analysis, text summarization and

machine translation. Therefore, there is considerable amount of scope for implementation of CNN in the development of chatbots. In Image processing the images are represented in the form of array of pixel or float values similarly the text can be stored in the form of array of vectors, where every word will be mapped to specific vector, these vectors can then be used to analyse the patterns and sequences.
[6] The tested the usability of a CNN based architecture in their chatbot where non cyclical connections were made between the nodes of Multi layer perceptrons. When the text data was passed to the model, a patter in the form of a N-gram word expression was detected at every convolution layer. The model was able to achieve an accuracy of 75% in 15 epochs but was proved to have very high computational costs and execution times.

2) Recurrent Neural Networks: Recurrent neural networks (RNN) are a class of neural networks which are based on Feed-forward network and are useful for modeling and predicting sequence data. Feed-forward neural networks are unidirectional networks i.e the data only moves from the input layer, through the hidden layers and producing results in the output layer. RNN works in the form of a loop. The decision making process follows a cycle where the hidden layer considers the current input and also what it has learned from the past data collected from previously received inputs. RNN can be very precise in predicting what will happen next, based on the information it has received but takes more time for execution than HAN and CNN. [6] Since, recurrent neural networks have the ability to efficiently deal with various input and output types they are implemented in a diverse range of applications, namely Sentiment Classification, Image Captioning, Language Translation etc.

3) Hierarchical Attention Networks: Hierarchical Attention Networks (HANs) are a self-attention-based architecture that are similar to RNNs, can capture linguistic relationships over long sequences while remaining fast to train. On a number of classification tasks, HANs can achieve accuracy that exceeds the current state-of-the-art while training twice as fast. Hierarchical Attention Networks use stacked recurrent neural networks on the word level, followed by an attention model, to extract important words for the sentence's meaning and aggregate the representation of those informative words to form a sentence vector. The same procedure is then applied to the derived sentence vectors, resulting in a vector that represents the meaning of the given document, which can then be used for text classification.
[6] After the selection of a HAN based model over a CNN and RNN model which was trained on the ISEAR dataset. This model performed at par with the RNN model and led the researchers to believe that it could outperform the RNN model if a larger training dataset was available.

4) Transformers: A transformer is a deep learning architecture that differentiates the importance of each portion of the input sequence based on the principle of attention. In Natural Language Processing and Computer Vision, transformer models are frequently employed. Transformers, like RNNs, are designed to handle sequential input, but they have one key difference: they do not necessarily process the data in an orderly manner, instead relying on the attention mechanism to grasp the context of each word. This offers considerable advantages to transformers over RNNs since it allows for better parallelization, which benefits in training on larger datasets. A feed-forward encoder-decoder structure makes up the transformer architecture. The encoder layer processes the input layer by layer, while the decoder does the same thing in the exact opposite direction. The encoder layer encodes the inputs while the self-attention mechanism assigns weights to each input based on their inter-relevance.
Carebot [5] employs a pre-trained transformer model: the DialoGPT model developed by Microsoft researchers, that operates entirely on the principle of self-awareness. The DialoGPT model is a conversational model that may be tuned based on data taken from a Reddit discussion thread. This model was chosen for Carebot because it has a higher BLEU score than other state-of-the-art models. The Transformer model's attention mechanism calculates the significance of a set of data based on particular keys and queries, with the outcome being a weighted sum of values assigned by a compatibility function. On the counselchat data-set, the transformer model was fine-tuned and trained. Owing to parallelization, training takes significantly longer but eventually produces higher outcomes.

5) Mental state identification from text: This model distributes the text into emotional labels such as; Happy, Joy, Shame, Anger, Disgust, Sadness, Guilt, and Fear.Based on these labels chat data is used to identify the mental state of a user , whether he or she is normal, stressed or depressed.Post classification of the emotions from the chat data an algorithm is applied to calculate the negative percentage and positive percentage. Finally predefined ranges of negative percentages will conclude the mental state of the user, there are five classed namely normal, slightly stressed, highly stressed, slightly depressed and highly depressed.

F. Evaluation metrics

1) Word count per session: This technique evaluates a user's involvement in the therapy process by evaluating the number of words used by the user to answer a particular question. The idea behind this evaluation metric is that a person suffering through any mental health disorder won't choose to talk to a chat bot about their issues. So, whenever a new user will start conversing to a chat bot they will use less number of words to answer any particular question. Across various sessions, if the word count increases then that could imply that the user's condition is improving. If the the word count decreases or remains stagnant, we can assume that the user's situation is remaining the same, if not getting worse.

To calculate this average response length, the user answers were tokenized and and average length of tokens per session were calculated. In [3], average increase in word count was seen in almost all users.

2) Sentiment Analysis: Sentiment analysis is used to analysis and detect the positive or negative sentiment of the user through their answers. Sentiment analysis can have multiple levels, most commonly they are divided in five, very negative, negative, neutral, positive and very positive. In [3], Stanford CoreNLP was used to perform sentiment analysis. They only used three levels of sentiments, negative, neutral and positive, because it was seen that very negative and very positive were not seen very frequently and can be integrated in negative and positive respectively. The idea behind this technique is that a user with a mental disorder will tend to be more negative or use more negative words while talking to a chat bot. As the sessions proceed, if the user's sentiment analysis gives an average positive result, it would imply that the therapy is working and the user is getting better. We can say that the therapy is working as user is getting more optimistic and positive in life.

3) Emotion Analysis: Emotion analysis is the technique used to analyse the mood of the user, in terms of percentage positivity or negativity. First all sentences are analysed and assigned a mood. Then using an algorithm, percentage positivity or negativity is calculated. In [6], seven moods were considered, 'joy', 'happy', 'anger', 'disgust', 'shame', 'fear' and 'guilt'. Out of these seven emotions, 'happy' and 'joy' are considered to be positive, whereas rest all are considered to be negative. After classifying all sentences, an algorithm calculates the percentage negativity. This variable is used to determine the state of the user. Five types are considered, normal, highly stressed, slightly stressed, highly depressed and slightly depressed. If percentage is is below 20, we classify the user as normal. If percentage negativity is between 20-40, we classify the user as slightly stressed. If percentage negativity is between 40-60, we classify the user as highly stressed. If percentage negativity is between 60-75, we classify the user as slightly depressed. Lastly, if percentage negativity is above 75, we classify the user as highly depressed. This metric is used to diagnose the user across various sessions. If the percentage negativity keeps on increasing for a particular user, we can assume that the therapy is not working and some other method of treatment is required. Opposite to that, if percentage negativity keeps on decreasing, we can assume that the therapy is working and the user is getting better.

III. Implementation

We have implemented a chatbot system which replies after evaluating the users input coupled with the baseline mental state which is determined by using the WHO-5[13] and PHQ-9[14] questionnaires which are the standard questionnaires used by psychologists. The chatbot then recommends micro-activities which could be useful in reducing the stress, anxiety or other negative feelings that the user might be feeling.

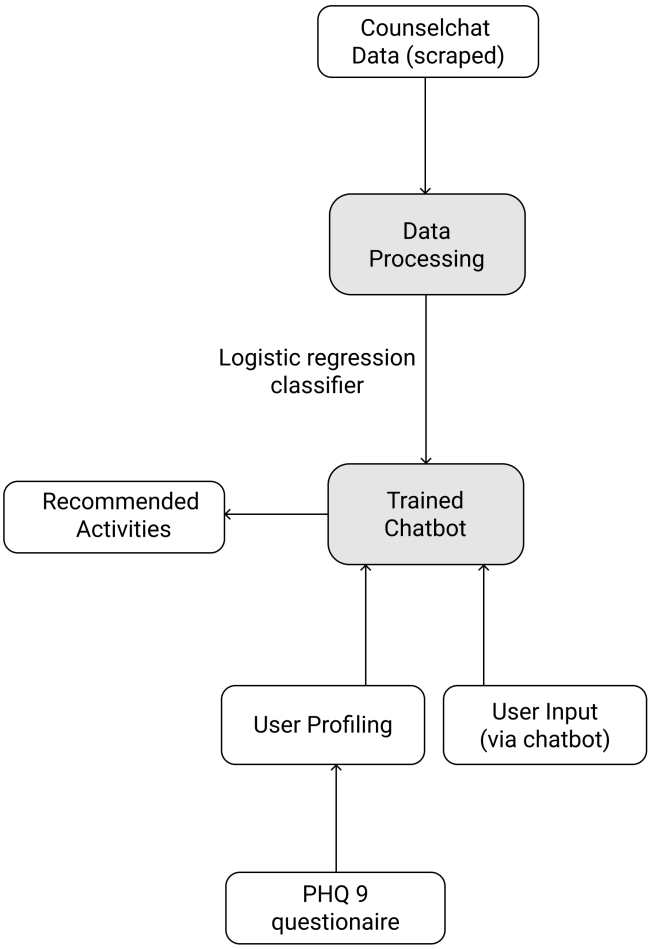


Fig. 1. Proposed Methodology.

A. Dataset

We intend on training our model on a dataset that is created by scraping the counselchat website which contains thousands of responses from qualified professionals ranging over 30 different topics. It contains responses on topics like relationships, intimacy and depression.

1) Data Preprocessing: The dataset contains several attributes like, 'questionID', 'questionTitle', 'questionText','questionLink', 'topic', 'therapistInfo', 'therapistURL', 'answerText', 'upvotes', 'views', 'split'. Redundant attributes are dropped and only 'questionText' and 'answerText' are used for further processing. The Natural Language Toolkit (NLTK) library is used for further

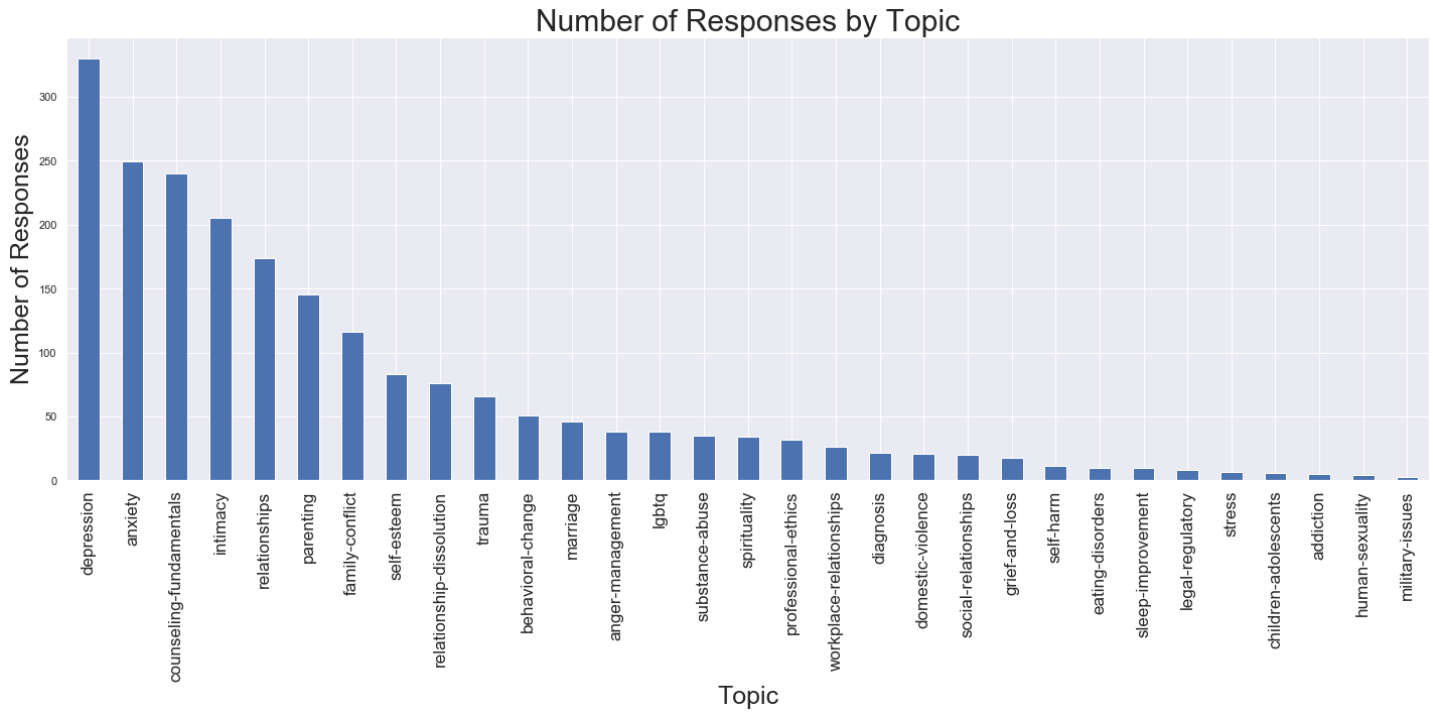


Fig. 2. Data Exploration.

processing. Stopwords and punctuations are removed before using the data for training the model.

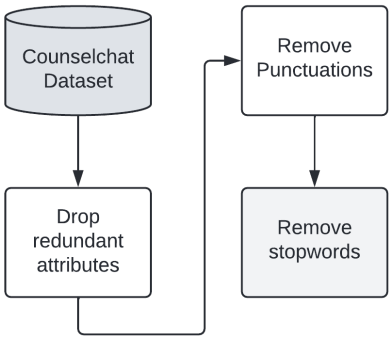
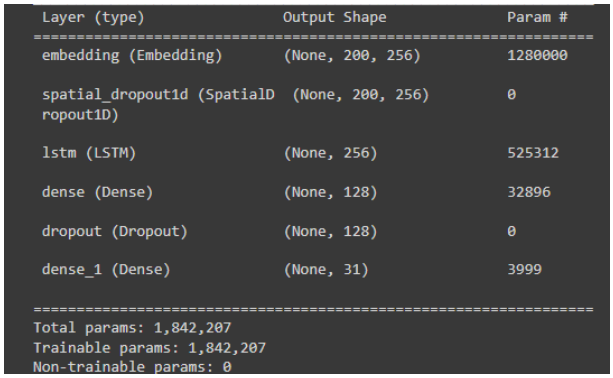


Fig. 3. Data Preprocessing.

B. Model

The model aims to use a Logistic Regression Classifier coupled with an LSTM to generate responses due to certain superior features like accepting the entire input sequence and its ability to retain data from it's memory mechanism. The Logistic Regression Classifier architecture offers ease of implementation, interpretation, and is very efficient to train. It is very fast at classifying unknown records, in this case the input sequences from the user.



Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 200, 256)	1280000
spatial_dropout1d (SpatialDropout1D)	(None, 200, 256)	0
lstm (LSTM)	(None, 256)	525312
dense (Dense)	(None, 128)	32896
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 31)	3999

Total params: 1,842,207
Trainable params: 1,842,207
Non-trainable params: 0

Fig. 4. Long Short Term Memory model.

IV. Results Discussion

The Logistic Regression shows a testing accuracy of 85.58%. Further metrics used for the evaluation are the F1 score, Precision, Recall and the confusion matrix.

To use classification metrics, we had to convert our testing data into a different NumPy format, NumPy array, to read. Precision is a good measure when classification accuracy is not a good indicator of your model performance, when your class distribution is imbalanced. Recall is defined as the fraction of samples from a class which are correctly predicted by the model. One popular metric which combines precision and recall is called F1-score. Support is the number of samples of the true response that lie in that class.

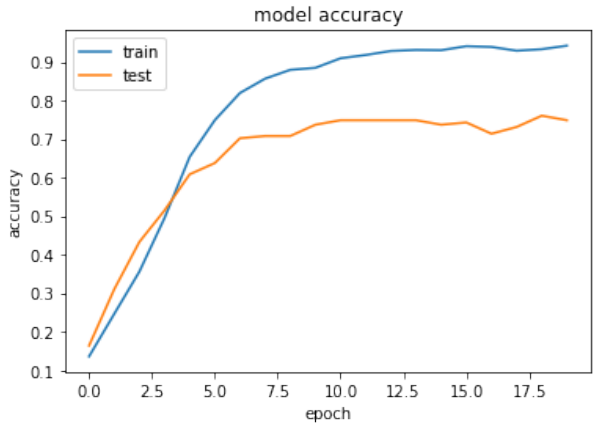


Fig. 5. Training and Validation Accuracy.

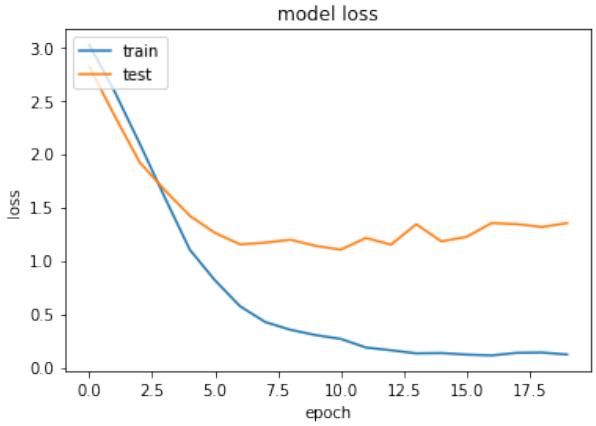


Fig. 6. Training and Validation Loss.

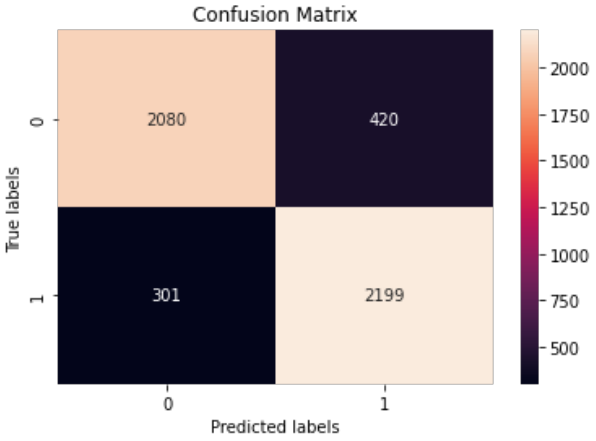


Fig. 7. Confusion Matrix.

- Precision=True_Positive/(True_Positive + False_Positive)
- Recall=True_Positive/(True_Positive + False_Negative)
- F1-score=2 * Precision * Recall/(Precision + Recall)

The higher the F1 score the better the model is.

Fig 7. highlights the evaluation metrics of the Logistic Regression classifier implemented.

```
[8] from sklearn.metrics import classification_report
    print(classification_report(y_test,y_pred_test))
```

	precision	recall	f1-score	support
0	0.87	0.83	0.85	2500
1	0.84	0.88	0.86	2500
accuracy			0.86	5000
macro avg	0.86	0.86	0.86	5000
weighted avg	0.86	0.86	0.86	5000

Fig. 8. Confusion Matrix.

V. Conclusion

In this paper, a thorough study has been done on different agents used for mental health aid along with the various datasets that are used in this domain. This paper also proposes a new methodology that uses a combination of two different datasets and employs a Logistic Regression Classifier model and LSTM model to train the chatbot to identify the emotions in the conversations between the system and the user. After analysing and evaluating the model using Classification Metrics and confusion matrix we found that the Logistic Regression Classifier and an LSTM to generate appropriate responses due to certain superior features like accepting the entire input sequence and it ability to retain data from it's memory mechanism.

References

[1] A. Ghandeharioun, D. McDuff, M. Czerwinski, and K. Rowan, "EMMA: An Emotion-Aware Wellbeing Chatbot," 2019 8th Int. Conf. Affect. Comput. Intell. Interact. ACII 2019, pp. 15–21, 2019, doi: 10.1109/ACII.2019.8925455.

[2] K. Denecke, S. Vaaheesan, and A. Arulnathan, "A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test," IEEE Trans. Emerg. Top. Comput., vol. 14, no. 8, 2020, doi: 10.1109/TETC.2020.2974478.

[3] F. DIno, R. Zandie, H. Abdollahi, S. Schoeder, and M. H. Mahoor, "Delivering Cognitive Behavioral Therapy Using A Conversational Social Robot," IEEE Int. Conf. Intell. Robot. Syst., pp. 2089–2095, 2019, doi: 10.1109/IROS40897.2019.8968576.

[4] R. Wang, J. Wang, Y. Liao, and J. Wang, "Supervised machine learning chatbots for perinatal mental healthcare," Proc. - 2020 Int. Conf. Intell. Comput. Human-Computer Interact. ICHCI 2020, pp. 378–383, 2020, doi: 10.1109/ICHCI51889.2020.00086.

[5] R. Crasto, L. Dias, D. Miranda, and D. Kayande, "CareBot: A Mental Health ChatBot," pp. 1–5, 2021, doi: 10.1109/incet51464.2021.9456326.

[6] F. Patel, R. Thakore, I. Nandwani, and S. K. Bharti, "Combating depression in students using an intelligent ChatBot: A cognitive behavioral therapy," 2019 IEEE 16th India Counc. Int. Conf. INDICON 2019 - Symp. Proc., pp. 2019–2022, 2019, doi: 10.1109/INDICON47234.2019.9030346.

[7] Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study. J Med Internet Res. 2020 Oct 12;22(10):e22635. doi: 10.2196/22635. PMID: 32936777; PMCID: PMC7575341.

[8] S. Nagargoje, V. Mamdyal, and R. Tapase, "Chatbot for Depressed People," in UIJRT United International Journal for Research and Technology, 2021, vol. 02, no. 07, pp. 208–211.

[9] Y. Zhang et al., "DIALOGPT : Large-Scale Generative Pre-training for Conversational Response Generation," pp. 270–278, 2020, doi: 10.18653/v1/2020.acl-demos.30.

[10] P. Das Monalisa and S. Kumar, "A Chatbot System For Mental Health Care," pp. 1–5.

[11] F. Åhs, P. Mozelius, and F. Dobslaw, “Artificial intelligence supported cognitive behavioral therapy for treatment of speech anxiety in virtual reality environments,” Proc. Eur. Conf. Impact Artif. Intell. Robot. ECIAIR 2020, no. October, pp. 1–6, 2020, doi: 10.34190/EAIR.20.030.

[12] Brown M, Glendenning A, Hoon A, John A ”Effectiveness of Web-Delivered Acceptance and Commitment Therapy in Relation to Mental Health and Well-Being: A Systematic Review and Meta-Analysis” J Med Internet Res 2016;18(8):e221 URL: https://www.jmir.org/2016/8/e221 DOI: 10.2196/jmir.6200

[13] Topp C, W, Østergaard S, D, Søndergaard S, Bech P: The WHO-5 Well-Being Index: A Systematic Review of the Literature. Psychother Psychosom 2015;84:167-176. doi: 10.1159/000376585

[14] Kroenke, K., Spitzer, R.L. & Williams, J.B.W. ”The PHQ-9.” J GEN INTERN MED 16, 606–613 (2001). https://doi.org/10.1046/j.1525-1497.2001.016009606.x