

# Mental Health Chatbot

Publisher: IEEE

Cite This

PDF

Harini V K ; Bhavadharini R M ; Charvi Upreti ; Libin Alex

All Authors

1  
Cites in  
Paper

426  
Full  
Text Views



Abstract

Document Sections

I. Introduction

II. Literature Review

III. Methodology

IV. Results and Discussions

V. Conclusion

Authors

Figures

References

Citations

Keywords

Metrics

More Like This

**Abstract:**

As per a 2023 publication, addressing mental health problems in India holds immense significance due to the scale of human value impact involved, particularly given the country's population size. It is crucial to recognize that mental health issues affect a significant portion of the population and can lead to severe consequences if left unaddressed. With this paper, we aim to contribute to the advancement of mental healthcare provision by developing a task-oriented, closed-domain chatbot. This chatbot will serve as a valuable tool for individuals seeking support and information regarding mental health issues. The paper compares existing models which use various embeddings and transformers and focuses on the development of a task-oriented, closed-domain chatbot using a retrieval-based approach with text data. The maximum accuracy of 91.48% was achieved in the LSTM model.

**Published in:** 2024 IEEE Students Conference on Engineering and Systems (SCES)

<b>Date of Conference:</b> 21-23 June 2024	<b>DOI:</b> 10.1109/SCES61914.2024.10652390
<b>Date Added to IEEE Xplore:</b> 03 September 2024	<b>Publisher:</b> IEEE
<b>► ISBN Information:</b>	<b>Conference Location:</b> Prayagraj, India

## SECTION I. Introduction

Addressing mental health problems in India holds immense significance due to the scale of human value impact involved, particularly given the country's population size. It is crucial to recognize that mental health issues affect a significant portion of the population and can lead to severe consequences if left unaddressed. [1]. Chatbots have become useful resources for mental health support in recent years, by providing immediate access to information [2], self-monitoring capabilities [3], and highly personalized remote health monitoring [4]. The insights garnered from surveys involving mental health professionals suggest that these chatbots are perceived not as replacements but as complements to professional services, with the ability to enhance the quality of care provided to clients [3] [5]. However, it is essential to acknowledge the challenges they bring, including the risk of misdiagnosis, limitations in medical knowledge, and the complex legal responsibilities that may arise in cases of inaccurate diagnoses [6]. Prior research has revealed the positive impact of AI applications on mental health, such as reducing stigma, increasing user comfort, and offering cost-effective access to support [6] [7]. Notably, individuals seeking mental health support have expressed favourable opinions regarding chatbots [8]. Furthermore, users have indicated specific features they desire in chatbots, including positive personality traits and regular check-ins [9].

Among features the “thoughts diary” was found to be particularly beneficial [10]. Importantly, chatbots have been found to contribute to the reduction of depressive symptoms, marking a significant advancement in potential to address mental health concerns [11]– [13]. This paper investigates the potential of chatbots to address mental-health crisis. We explore the current available technologies and introduce our own chatbot solution for immediate and accessible mental health support, to promote mental health awareness and aid in the decrease of mental health issues.

PDF

Help

In this paper we have worked on a retrieval-based, text based, task oriented and closed domain chatbot. The dataset used “Mental Health Conversational Data” was sourced from Kaggle, a Json file containing tags, patterns, and corresponding responses. For the preprocessing purpose we have utilized libraries from TensorFlow, stemming and stop words. We use Natural Language Processing for training the model. Natural language processing (NLP) applies artificial intelligence (AI) to assist chatbots in interpreting and analyzing natural human language when interacting with clients. Chatbots are able to comprehend the context of the communication instead of using the available data only for question and answering.

## SECTION II.

# Literature Review

Authors in [2] explored the impact of task-oriented chatbots in healthcare, highlighting potential dangers of over-reliance and self-diagnosis leading to misdiagnoses and ethical dilemmas. In [3], chatbot acceptance in young adults' mental health was investigated through surveys, literature review, and counselor interviews, suggesting positive outcomes and the importance of user-centered design. Study [4] aimed to develop a mental health chatbot for remote monitoring and personalized behavioral activation, evolving through research, support group input, and expert refinements. In [5], part of the ChatPal project, 100 mental healthcare professionals were surveyed to understand their perspectives on using chatbots for maintaining mental health and well-being. The review [6] outlines existing AI applications, including therapy bots like Woebot, Tess, Replika, and Paro. While these technologies reduce stigma and offer cost-effective accessibility, they lack personalization, emotional awareness, and empathy compared to trained professionals. The proposed work in [7] presents an AI chatbot created to provide therapeutic dialogues and assess users' mental health using the K10 scale. The chatbot predicts user emotions from text input using a Logistic Regression (LR) Model and modifies responses accordingly. It also uses Kessler's Psychological Distress scale to evaluate mental health conditions. Paper [8] is a scoping review on patient perceptions of mental health chatbots, conducted following PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines.

In [9], desired chatbot features include positive personality traits, regular check-ins, and emotional state tracking, with challenges including NLP issues and responding to diverse human emotions. Study [10] analyzes user logs from ChatPal, noting high return rates for the “thoughts diary” feature, indicating potential mental well-being benefits, though assessing overall effectiveness is challenging due to user abandonment. The efficacy of XiaoE, a cognitive behavioral therapy (CBT)-based mental health chatbot, in lowering symptoms of depression in young individuals during the COVID-19 pandemic was examined in the study reported in paper [11]. The Uses and Gratifications Theory (UGT) along with Technology Acceptance Model (TAM) were employed in the study [12] to ascertain the motivations behind the use of digital technology connected to depression. The investigation described in [13] attempted to assess the effectiveness of Vitalk, a chatbot focused on mental health, in alleviating feelings of depression, anxiety, and stress. An analysis was conducted on real-world data from users who successfully finished the initial stage of the Vitalk program. In article [14], a Twitter-based machine learning prototype for early mental illness detection is introduced. When it comes to digital mental health services, intelligent cognitive assistants (ICAs) play a crucial role in supporting behavior change, especially for disorders like depression, anxiety, and stress (SAD). The technical review in paper [15] aimed to categorize current technology approaches and trends with text-based ICAs for SAD. The paper [16] investigates the feasibility of using a text-based chatbot named “Viki” to initiate a mental health assessment among employees. Viki, a fully automated chatbot, assessed the susceptibility of employees to mental health issues.

Digital mental health technologies, and chatbots in particular, have gained popularity since the COVID-19 pandemic [17]. A multilingual chatbot named ChatPal is created with the intention of enhancing mental health. Symptoms overlap between disorders, and their severity varies among individuals sharing the same diagnosis, leading to prolonged, ineffective treatments and increased healthcare costs. The authors of paper [18] propose precision psychiatry, akin to precision medicine, offers a solution by considering individual variations in biological, ecological, and behavioral variables for individualized therapy suggestions. The research works in [19]–[21] presents an overview of mental health chatbots, examining various mental health disorders. Authors in [22] developed an AI software to forecast progress in mental health issues during internet-delivered Cognitive Behavioral Therapy (iCBT) for depression and anxiety. Research in [23]

validated the use of chatbots for mental health assessments, finding that chatbot assessments required more effort and time compared to established methods.

A study in [24] assessed adolescents' responses to a mental health chatbot, with limited scope and potential influence from decision fatigue. An exploratory investigation in [25] created a prototype chatbot to aid school nurses in delivering preventive mental health support, laying the groundwork for understanding young individuals' needs. The paper in [26] presents an integrated chatbot offering empathetic responses for individuals with mental health concerns, using Seq2Seq architecture with BiLSTM. Research in [27] explores the utilization and perceived reliability of chatbots for self-care, highlighting the impact of chatbot human-likeness on trust formation. Article [28] reports on a controlled study assessing the efficiency of the Elomia chatbot in alleviating inclinations toward depression, anxiety, and adverse emotional impacts, using cognitive-behavioral psychotherapy. The study in [29] evaluated anxiety and depression chatbot apps on Android and iOS, identifying 11 apps as therapeutic substitutes for anxiety and depression concerns. Research in [30] addresses the development of empathetic chatbots based on Cognitive Behavioral Therapy (CBT) to support mental health patients. Researchers in [31] examined user-generated content to understand people's experiences with the digital mental health app Wysa, which blends chatbot-enabled conversations with research-backed CBT techniques.

From the extensive study of the research articles, it is understood that various AI enabled chatbots are playing a vital role in addressing mental health disorders. Above literature survey presented a number of existing chatbots that can assist the mentally distressed people.

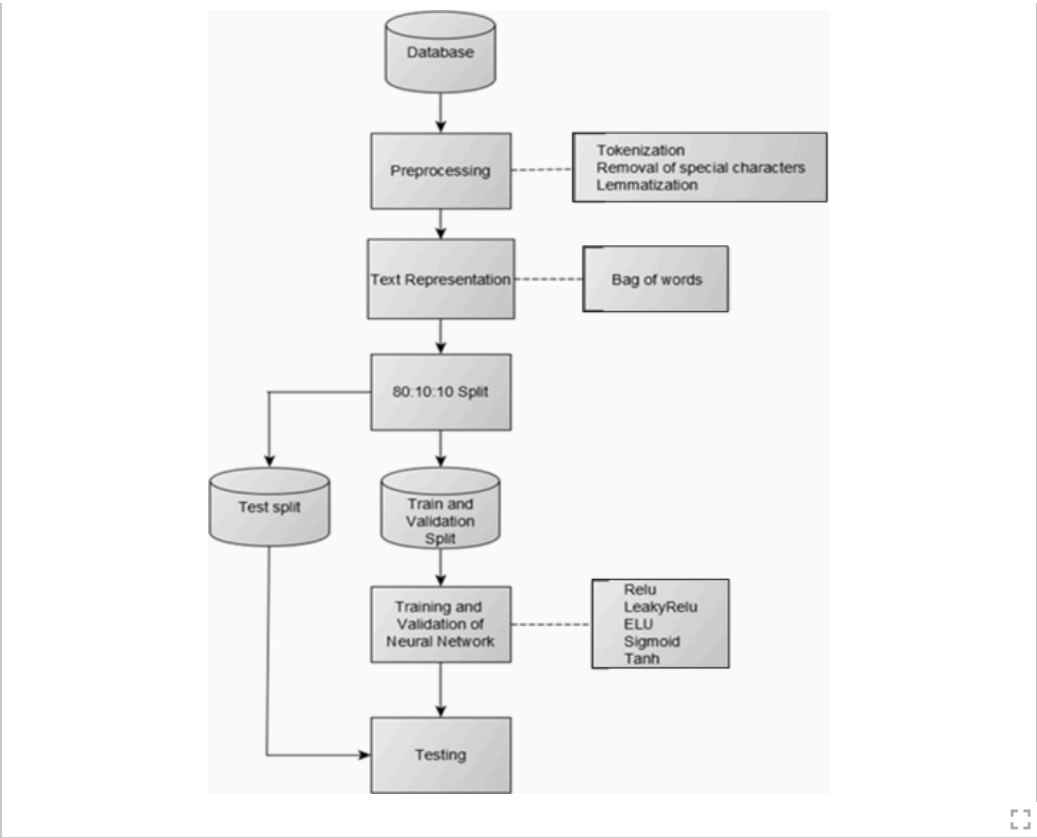
## SECTION III. Methodology

The dataset used is sourced from Kaggle. The results are from models trained on a mental health dataset. To begin, we preprocess the data by tokenizing sentences into words and removing special characters such as exclamation marks and hashtags. Subsequently, lemmatization is applied to the words to reduce them to their base form. We have designed a Neural Network architecture, comprising three fully connected (linear) layers and two batch normalization layers. In order to stabilize and accelerate the training process, Batch normalization is employed to by normalizing the activations of each layer. Additionally, activation functions are applied to the output of the batch-normalized hidden layers. These activation functions introduce non-linearity, facilitating the learning of complex patterns within the data.

In the pursuit of identifying the optimal configuration, we experiment with different activation functions for the hidden layers, including ReLU, Leaky ReLU, ELU, Tanh, and Sigmoid. The model is trained using the PyTorch machine learning library, where Cross Entropy Loss is utilized to measure the disparity between predicted probabilities and actual labels. The Adam optimizer is employed to minimize this loss. The same neural network was used for Bag-of-words, Word2Vec and Roberta Model. To conduct our research, we partitioned the dataset into training, validation, and test sets, maintaining a consistent 80:10:10 ratio for each individual run.

### A. Bag of Words

The bag-of-words model is a method of text classification based on fixed length vectors where the occurrence of each word is used as a feature for training a classifier. In this paper, we're using these word counts as features to train a classifier [32]. The block diagram of Bag-of-Word is as shown in Fig 1 and the parameters used are shown in Table I.



**Fig 1.**  
Bag of words block diagram

**Table I.** Parameters taken for bag of words

Parameters	Values
Batch size	16
Learning rate of Adam optimizer	0.001
Hidden size	10
Input size	295
Output size	80

**B. Word2Vec**

In our experimentation, we explored and evaluated the utilization of pre-trained word embeddings for our model. The Word2Vec model, developed by Google, represents words as vectors of integers. Specifically, we employed the pre-trained Word2Vec model provided by Google, known as “word2vec-google-news-300,” which features a dimensionality of 300 [33].

**Table II.** Parameters taken for Word2Vec

Parameters	Values
Batch size	10
Learning rate of Adam optimizer	0.001
Hidden size	10
Input size	300
Output size	80

PDF  
Help

C. Roberta Model

Roberta, short for ‘Robustly Optimized BERT Approach,’ stands as a transformer-based language model renowned for its utilization of self-attention mechanisms to process input sequences and generate contextualized word representations within a sentence. Roberta serves as the cornerstone for generating sentence embeddings within this framework. These embeddings are subsequently fed into a Neural Network tasked with intent classification, thereby predicting user intentions based on the Roberta-encoded representations of the sentences for response retrieval. The block diagram of Word2Vec and Roberta is depicted in Fig 2 and the parameter used to train the model are shown in Table II and Table III.

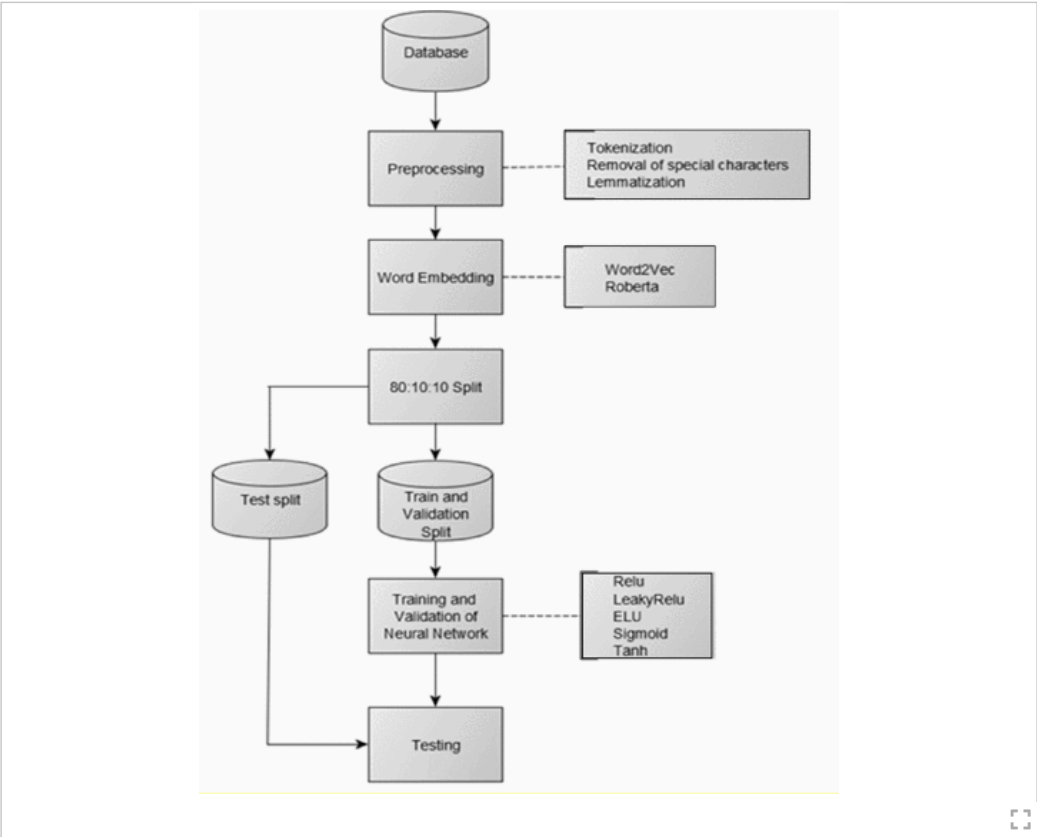


Fig 2.  
Word2Vec and roberta block diagram

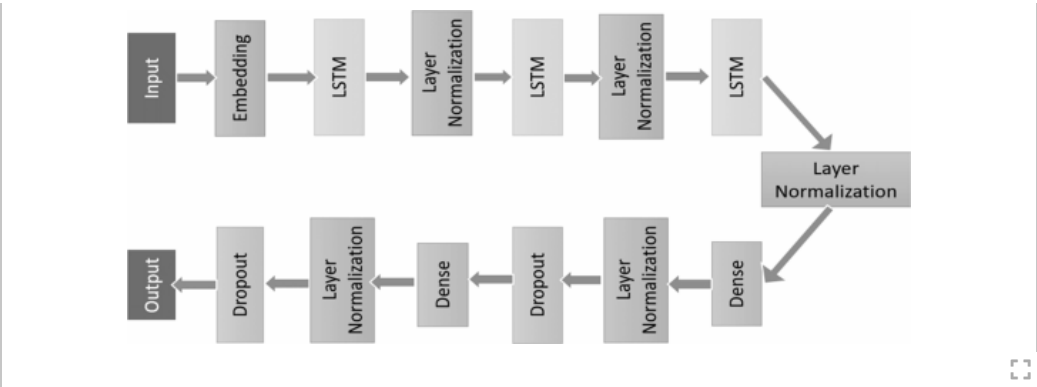
Table III. Parameters taken for RoBERTa

Parameters	Values
Batch size	10
Learning rate of Adam optimizer	0.001
Hidden size	10
Input size	768
Output size	80

D. LSTM

LSTM (Long short-term memory) was used to calculate the accuracy of the chatbot model. LSTM is a recurrent neural network (RNN) architecture widely used in Deep Learning as shown in Fig.3. We tried various activation functions for the model and found the accuracy in each case.

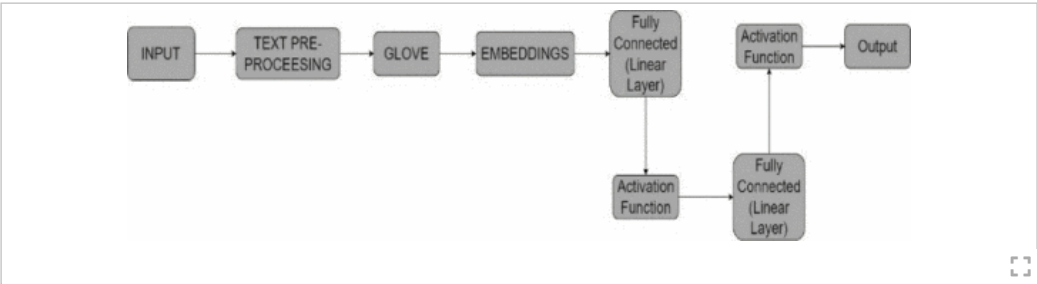
PDF  
Help



**Fig 3.**  
LSTM architecture

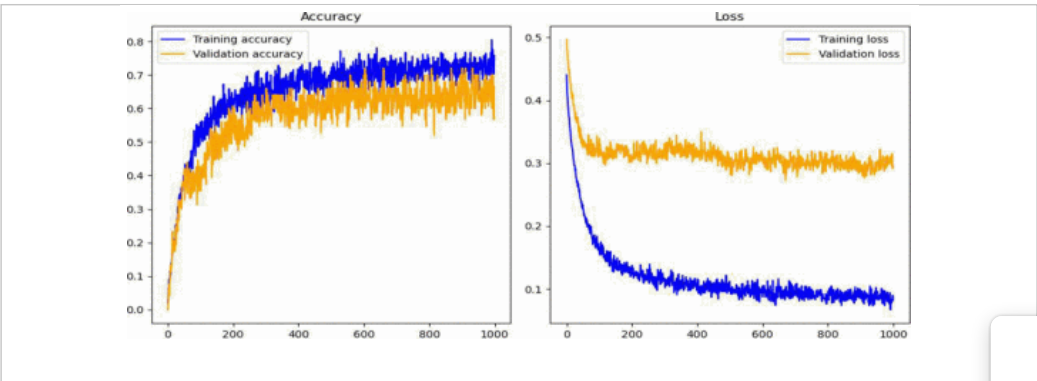
**E. Glove+CNN**

We calculated the accuracy of the chatbot model using Glove preprocessing and CNN(Convolutional Neural Network). After preprocessing, we use glove embedding. The GloVe word embedding method shown in [Fig 4.](#) is based on co-occurrence matrix to determine the relationship between words. The co-occurrence matrix indicates how frequently a given word pair appears together. It focusses on the conditional probability for the words that occur together in the given context.

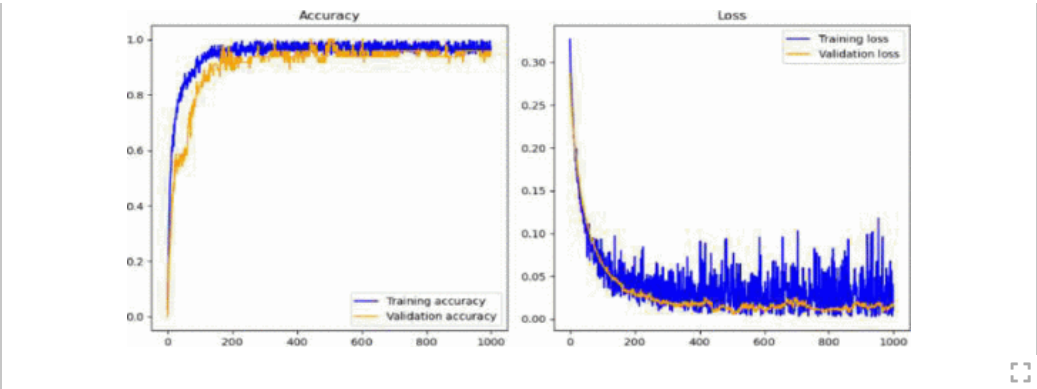


**Fig 4.**  
Glove+CNN architecture

**SECTION IV.**  
**Results and Discussions**

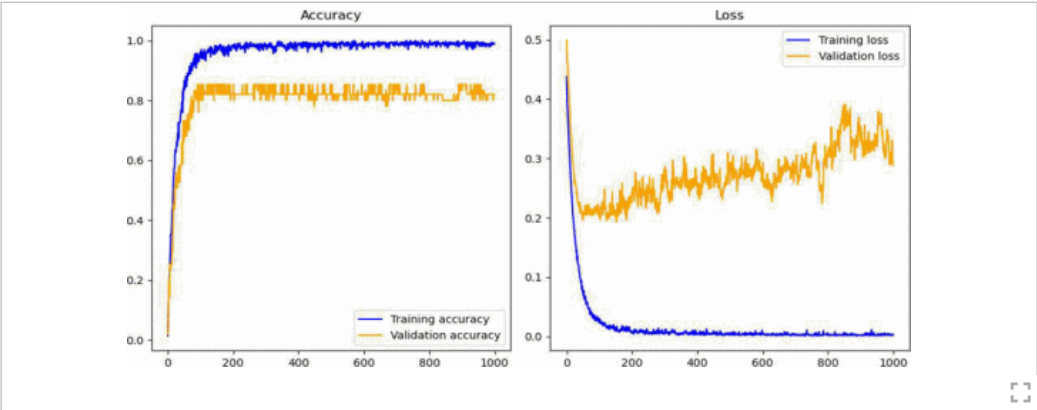


**Fig 5.**  
Training and validation accuracy and loss graph for Word2Vec with ReLU activation over 1000 epochs

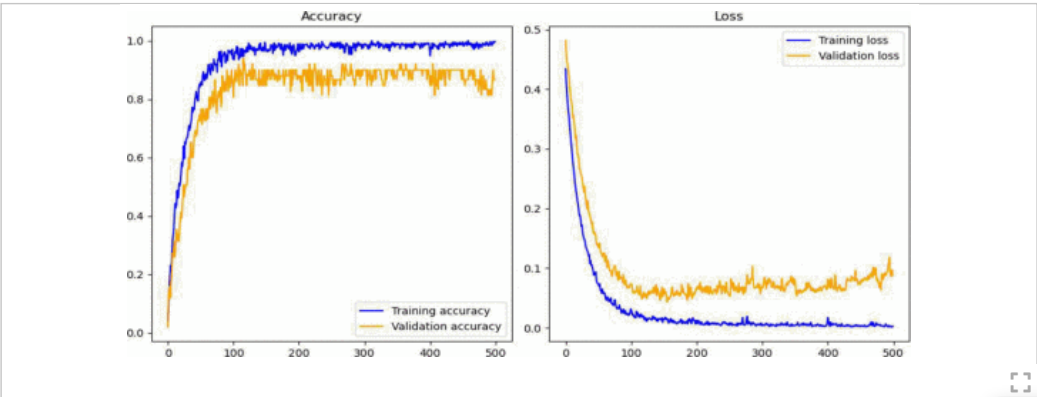


**Fig 6.** Training and validation accuracy and loss graph for bag of words with tanh activation over 1000 epochs.

The Word2Vec implementation achieves its highest accuracy using ReLu activation, as indicated in [Fig.5](#) and [Table IV](#), with performance metrics of 75.76% for training, 66.66% for validation, and 70.28% for testing. The Bag of words implementation achieves its highest accuracy using Tanh activation, as indicated in [Fig.6](#) and [Table V](#), with performance metrics of 97.30% for training, 95.83% for validation, and 91.38% for testing.



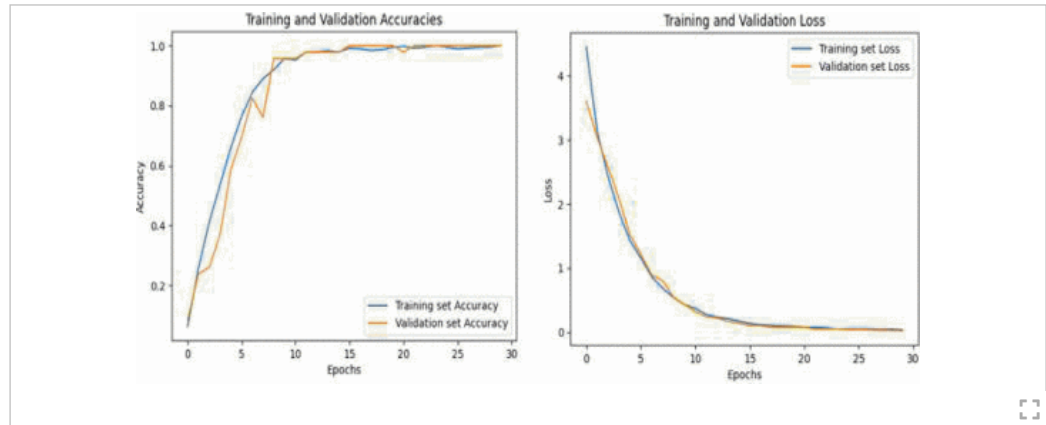
**Fig 7.** Roberta model trained over 1000 epochs



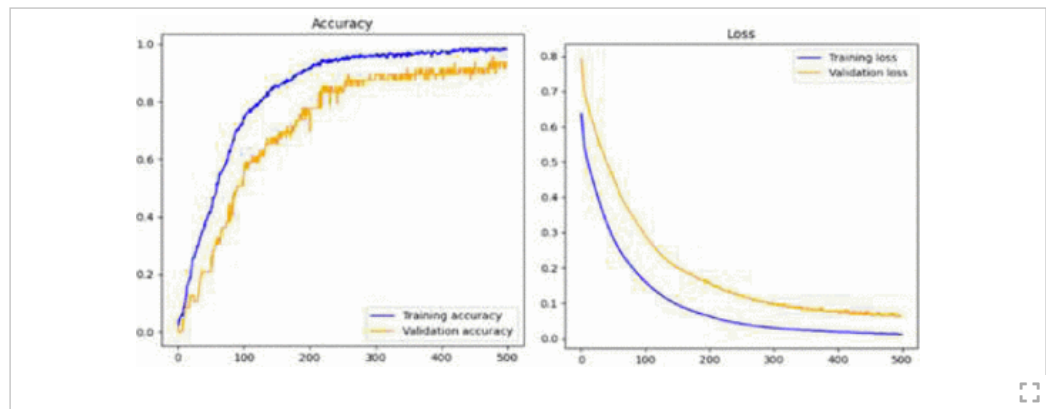
**Fig 8.** Training and validation accuracy and loss graph for roberta with ReLU activation over 500 epochs

During the training of Roberta model over 1000 epochs, it was observed that the validation loss began to increase beyond a certain point, while validation accuracy was similar as seen in [Fig.7](#). Consequently, for the

analysis, the number of epochs was reduced to 500 as shown in [Fig.8](#). The Roberta implementation achieves its highest accuracy using ReLu activation with 500 epochs, as indicated in [Table VI](#) and [Table VII](#), with performance metrics of 99.72% for training, 86.66% for validation, and 87.14% for testing.



**Fig 9.**  
Training and validation accuracy and loss graph for LSTM with tanh activation over 30 epochs



**Fig 10.**  
Training and validation accuracy and loss graph for Glove+CNN with tanh activation over 500 epochs

The LSTM implementation achieves highest accuracy using tanh activation function as shown in [Fig. 9](#) and [Table VIII](#), with performance metrics of 100% for training, 100% for validation and 91.14% for testing. The Glove+CNN implementation achieves its highest accuracy using Tanh activation with 500 epochs, as indicated in [Fig. 10](#) and [Table IX](#), with performance metrics of 98.26% for training, 93.75% for validation, and 79.86% for testing.

**Table IV.** Results of Word2Vec - google-news-300

S. No.	Number of epochs	Activation function	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss	F1 Score	Recall
1.	1000	ReLU	75.7658	0.0820	66.6667	0.2943	70.2857	0.3581	0.7110	0.7021
2.	1000	Leaky ReLU	74.8348	0.0793	72.0000	0.3899	72.2857	0.3965	0.7272	0.7234
3.	1000	Tanh	72.8829	0.0912	70.0000	0.2525	68.2857	0.2867	0.6419	0.6809
4.	1000	Sigmoid	69.6396	0.1034	64.6667	0.3038	58.2857	0.3450	0.5405	0.5745
5.	1000	ELU	75.3453	0.0752	68.6667	0.3312	75.1429	0.3493	0.7364	0.7447

**Table V.** Results of bag of words

S. No.	Number of epochs	Activation function	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss	F1 Score	Recall
1.	1000	ReLu	94.5312	0.0710	93.7500	0.0194	89.3056	0.0902	0.9021	0.8936
2.	1000	Leaky ReLu	94.0104	0.0698	91.6667	0.0283	87.2222	0.1147	0.8706	0.8723
3.	1000	Tanh	97.3090	0.0102	95.8333	0.0177	91.3889	0.0709	0.9255	0.9149
4.	1000	Sigmoid	96.4410	0.0266	91.6667	0.0341	87.2222	0.0819	0.8735	0.8723
5.	1000	ELU	96.4410	0.0237	95.8333	0.0524	91.3889	0.0841	0.9149	0.9149

**Table VI.** Roberta model - 1000 epochs

S. No.	Number of epochs	Activation function	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss	F1 Score	Recall
1.	1000	ReLu	98.9189	0.0020	82.0000	0.2886	85.1429	0.3131	0.8511	0.8511

**Table VII.** Results of roberta model - 500 epochs

S. No.	Number of epochs	Activation function	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss	F1 Score	Recall
1.	500	ReLu	99.7297	0.0028	86.6667	0.0872	87.1428	0.1927	0.8794	0.8723
2.	500	Leaky ReLu	99.1892	0.0021	81.3333	0.1618	83.1429	0.2052	0.8284	0.8298
3.	500	Tanh	98.1081	0.0054	79.3333	0.2056	81.1429	0.1761	0.8099	0.8085
4.	500	Sigmoid	97.8378	0.0101	81.3333	0.1552	83.1429	0.1902	0.8163	0.8298
5.	500	ELU	98.6486	0.0025	83.3333	0.1579	85.1429	0.1922	0.8582	0.8511

**Table VIII.** Results LSTM model — 30 epochs

S. No.	Activation Function	Final Training Accuracy	Final Training Loss	Final Validation Accuracy	Final Validation Loss	Test Accuracy	Test Loss
1	relu	0.9919	0.0586	0.9783	0.0564	0.89361	0.6145
2	tanh	1.0000	0.0341	1.0000	0.0211	0.91489	0.8077
3	leaky_relu	0.9892	0.0745	0.9565	0.1191	0.91489	0.6692
4	Sigmoid	0.9919	0.0387	1.0000	0.0584	0.91489	0.6640

**Table IX.** Results glove+CNN model — 500 epochs

S. No.	Activation Function	Final Training Accuracy	Final Training Loss	Final Validation Accuracy	Final Validation Loss	Test Accuracy	Test Loss
1	relu	0.9903	0.0041	0.8750	0.2388	0.7778	0.7922
2	Leaky_relu	0.9855	0.0033	0.8750	0.1837	0.7778	0.5347
3	tanh	0.9826	0.0116	0.9375	0.0640	0.7986	0.2289
4	Sigmoid	0.8750	0.0806	0.6111	0.2634	0.6597	0.3532

## SECTION V. Conclusion

In this study, we investigated the potential of chatbots to address mental health crises, exploring various existing technologies and identifying the best chatbot solution for immediate and accessible mental health support. Our approach focused on a retrieval-based, text-based, task-oriented, and closed-domain chatbot, utilizing a dataset sourced from Kaggle and leveraging Natural Language Processing techniques for model training. Our results indicate that chatbot based on LSTM using tanh as activation function, gave the highest accuracy of 91.48%, demonstrating its potential to provide valuable mental health support through an accessible and interactive platform. Moving forward, continued research and development in this area hold promise for advancing mental health awareness, promoting early intervention, and ultimately reducing the burden of mental health issues in society. In conclusion, this research contributes to the growing field of AI-driven mental health support systems, offering a promising avenue for leveraging technology to address mental health challenges and provide support to individuals in need.

PDF

Help

Authors	▼
Figures	▼
References	▼
Citations	▼
Keywords	▼
Metrics	▼

IEEE Personal Account

CHANGE  
USERNAME/PASSWORD

Purchase Details

PAYMENT OPTIONS  
VIEW PURCHASED  
DOCUMENTS

Profile Information

COMMUNICATIONS  
PREFERENCES  
PROFESSION AND  
EDUCATION  
TECHNICAL INTERESTS

Need Help?

US & CANADA: +1 800  
678 4333  
WORLDWIDE: +1 732  
981 0060  
CONTACT & SUPPORT

Follow

[f](#) [@](#) [in](#) [v](#)

[About IEEE Xplore](#) | [Contact Us](#) | [Help](#) | [Accessibility](#) | [Terms of Use](#) | [Nondiscrimination Policy](#) | [IEEE Ethics Reporting](#) [↗](#) | [Sitemap](#) | [IEEE Privacy Policy](#)

A public charity, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2025 IEEE - All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies.

PDF

Help