Ameliorating Depression through Deep Learning Conversational Agent: A Novel Approach to Mental Health Intervention

Akash Pawar[1], Saurabh Parkar[2], Rithik Rai[3], Abhijay Walia[4], Dilip Kale[5]

[1,2,3,4]*Student, Computer Engineering, MCT's Rajiv Gandhi Institute of Technology, Mumbai University*
[5]*Professor, Dept. of Computer Engineering, MCT's Rajiv Gandhi Institute of Technology, Mumbai University*

---***---

Abstract - *Depression is a prominent cause of disability worldwide, impacting millions of individuals. Despite the availability of many treatment alternatives, many individuals still may not obtain proper care owing to societal stigma, financial restraints, or a shortage of mental health experts. To address these issues, we present a novel approach to depression intervention that leverages deep learning techniques to construct a conversational agent capable of alleviating depressive symptoms. Our proposed method incorporates a combination of natural language processing, sentiment analysis, and transfer learning to create a chatbot that understands the user's emotions and responds in a compassionate and supportive manner. The system is designed to identify the user's cognitive distortions and provide responses that ameliorate the user's mental circumstances to help them reframe their negative views. We assessed the system's performance through a peer-based study by implementing the chatbot on Discord, where many people, notably young people, can utilise the bot. The study comprised participants with mild to severe depression. The results demonstrate that the chatbot substantially lowered depressive symptoms and enhanced the overall mood in the majority of individuals. The users indicated excellent satisfaction with the system and enjoyed the non-judgmental and empathic attitude. Overall, our study illustrates the promise of deep learning-based conversational agents as a scalable and accessible method to relieve depressive symptoms. Further research is needed to explore the long-term effectiveness of the system and its effects on clinical outcomes.*

Key Words: Depression Amelioration, Conversational Agent, Deep Learning, Transfer Learning, NLP, Sentiment Analysis.

1.INTRODUCTION

Depression is a common mental health illness that can have significant consequences on an individual's life. While different treatment alternatives are available, the accessibility and effectiveness of these interventions remain a concern, particularly for people living in distant or disadvantaged locations. Conversational agents, or chatbots, have emerged as a promising way to provide mental health care and deliver evidence-based interventions. Deep learning-based chatbots have shown special promise due to their capacity to learn from massive volumes of data and deliver more tailored and contextually appropriate responses. Transfer learning, a technique that allows the model to exploit pre-existing information and adapt to new tasks, significantly boosts the effectiveness of these chatbots. In this research, we introduce a deep learning chatbot that leverages transfer learning to relieve depression symptoms. We fine-tune a pre-trained binary classification transformer to classify the text as normal or depressed. If the input is classified as depressed, the chatbot uses another pre-trained multiclass classification transformer to accurately choose a response for the provided input. The chatbot leverages the pre-trained Blenderbot transformer pipeline to conduct informal, non-depressed conversation inputs. We analyse the efficacy of our chatbot through user research, including interviews with persons with depression, and report significant reductions in depressive symptoms and user satisfaction. This chatbot is deployed on Discord, a familiar application for many, and simulates the sense of conversing with a friend. Our method illustrates the promise of deep learning-based chatbots in providing accessible mental health care to individuals with depression, particularly those who may not have access to traditional mental health services. The use of transfer learning further boosts the chatbot's performance and may be relevant to various mental health illnesses and healthcare fields.

1.1 Transformers

1) RoBERTa: It stands for 'Robustly Optimised BERT Pretraining Approach' and was used for both the binary and multiclass classification tasks in the chatbot. Precisely, RobertaForSequenceClassification was used, which is a RoBERTa Model transformer with a sequence classification/regression head on top (a linear layer on top of the pooled output), e.g., for GLUE tasks.

2) Blenderbot: Specifically, the Blenderbot-400M-distill model was used for conditional generation to conduct casual conversations. This helps when the user just wants to have a casual conversation and wants normal responses instead of therapeutic ones.

1.2 Deployment Platform

The deployment platform for the chatbot was chosen to be Discord. Not only is it incredibly popular with the youth, it also has significant features such as accessibility, anonymity, 24/7 availability, personalization, scalability & lower cost.

It also opens up a plethora of possibilities such as text-to-speech due to its inbuilt functions and supported APIs.

2. LITERATURE REVIEW

A. Literature Review

1) *"The Woebot Trial: A Randomized Controlled Trial of an Automated Conversational Agent for Depression" by Fitzpatrick et al. (2017)*[1]*:* This study evaluated the effectiveness of an automated chatbot named Woebot in reducing symptoms of depression. The chatbot provided cognitive-behavioral therapy and was evaluated through a randomized controlled trial. The study reported significant improvements in depressive symptoms and user satisfaction.

2) *"Using Chatbots for Mental Health: A Systematic Literature Review" by Vaidyam et al. (2019)*[2]*:* This paper provides a systematic literature review of the use of chatbots for mental health support. The review identified 14 studies that evaluated the effectiveness of chatbots in providing mental health support, particularly for depression and anxiety. The paper highlights the potential of chatbots in providing accessible, personalized, and cost-effective mental health support.

3) *"RoBERTa: A Robustly Optimized BERT Pretraining Approach" by Yinhan Liu et al. (2019)*[3]*:* The paper begins by introducing the limitations of the previous pre-training approach, BERT (Bidirectional Encoder Representations from Transformers), and how RoBERTa overcomes these limitations by fine-tuning several hyperparameters in the pre-training process. The authors then describe the pre-training corpus, which was a combination of books, articles, and web pages, resulting in a total of 160GB of text data. The pre-training objective was a masked language modeling task, where the model learned to predict the masked words within a sentence. RoBERTa achieved state-of-the-art results on a range of NLP tasks, including natural language inference, question answering, and sentiment analysis. The authors conducted experiments to show that RoBERTa outperformed BERT on several benchmarks, including the Stanford Question Answering Dataset (SQuAD) and the General Language Understanding Evaluation (GLUE) benchmark. Additionally, RoBERTa was shown to be robust to variations in the training data and hyperparameters, making it a more reliable and flexible model.

4) *"Recipes for building an open-domain chatbot" by Stephen Roller et al. (2020)*[4]*:* The authors evaluate the chatbot's performance on several metrics, including perplexity, human evaluation, and engagement. The chatbot is shown to achieve state-of-the-art performance on several metrics, outperforming existing chatbots on the Persona-Chat dataset. The paper's methodology and results demonstrate the effectiveness of the approach, which has since become a widely used method in the field. The paper has become a go-to resource for researchers and practitioners, and its impact can be seen in the many state-of-the-art chatbots developed using the techniques and methodologies described in the paper.

B. Problems in Existing Systems

1) *Lack of Personalization*: Many mental health chatbots provide a one-size-fits-all solution to mental health problems. They do not take into consideration the individual variances in personalities, experiences, and mental health situations of the users. This can lead to erroneous or ineffective advice.

2) *Data Privacy*: The use of mental health chatbots necessitates the exchange of sensitive personal data, which might be a worry for users who are anxious about their privacy.

3) *Inaccurate Responses*: Mental health chatbots might deliver erroneous or even hazardous advice if they are not properly developed or trained. This can lead to bad results for users.

4) *Limited Scope*: Most mental health chatbots are meant to target a certain set of mental health issues, such as anxiety or depression. They may not be equipped to address more sophisticated mental health difficulties.

5) *Stigma*: Despite the growing acceptance of mental health chatbots, there is still a stigma attached to mental health issues, which may prevent some individuals from seeking help through chatbots, making anonymity an important feature.

6) Lack of available data: Mental health datasets are not as widely available as other healthcare datasets, such as those for diabetes or heart disease, due to the sensitive and personal nature of mental health data, which requires strict ethical considerations and privacy protections. There are also concerns about the quality and consistency of the data collected, as mental health conditions can be difficult to diagnose and may have varying symptoms across individuals.

C. Findings

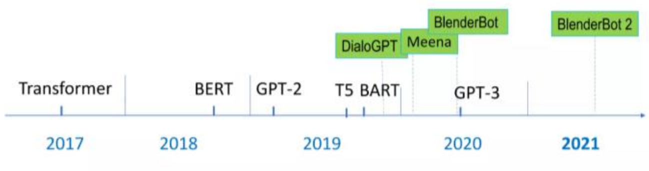
1) RoBERTa was used as it provides state-of-the art results and outperforms the other transformers on GLUE benchmark results. While it gives better results than the previous state-of-the-art transformers used for text classification, it also has a better speed-to-accuracy ratio, making it the best choice.

Fig -1: Comparison of BERT and recent improvements[5]

	BERT	RoBERTa	DistilBERT	XLNet
Size (millions)	**Base**: 110 **Large**: 340	**Base**: 110 **Large**: 340	**Base**: 66	**Base**: ~110 **Large**: ~340
Training Time	**Base**: 8 x V100 x 12 days* **Large:** 64 TPU Chips x 4 days (or 280 x V100 x 1 days*)	**Large:** 1024 x V100 x 1 day; 4-5 times more than BERT.	**Base**: 8 x V100 x 3.5 days; 4 times less than BERT.	**Large:** 512 TPU Chips x 2.5 days; 5 times more than BERT.
Performance	Outperforms state-of-the-art in Oct 2018	2-20% improvement over BERT	3% degradation from BERT	2-15% improvement over BERT
Data	16 GB BERT data (Books Corpus + Wikipedia). 3.3 Billion words.	160 GB (16 GB BERT data + 144 GB additional)	16 GB BERT data. 3.3 Billion words.	**Base**: 16 GB BERT data **Large**: 113 GB (16 GB BERT data + 97 GB additional). 33 Billion words.
Method	BERT (Bidirectional Transformer with MLM and NSP)	BERT without NSP**	BERT Distillation	Bidirectional Transformer with Permutation based modeling

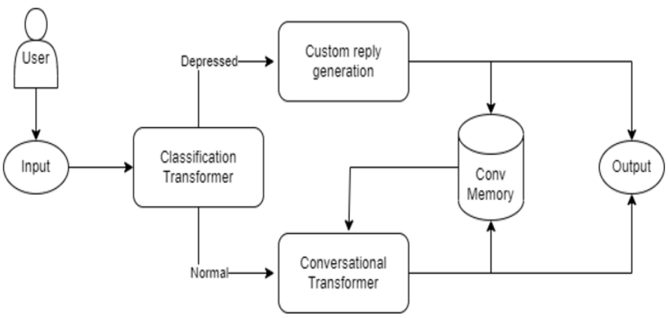
2) *Blenderbot* was chosen as the model for casual conversations as it yielded better results both in answering a single prompt as well as remembering context through conversation history.

Fig -2: Development of Open-domain Chatbots[6]



3. METHODOLOGY

Fig -3: System base workflow



1) *Classification Transformer:* The first step in the system's workflow is passing the user input through the binary classification transformer to classify it as depressed or normal. The model used here is the Roberta-base model. This model was finetuned using the Suicide and Depression Detection Dataset from Kaggle. The dataset has 7730 different Reddit posts classified as depressed or normal. The train and test sets were made with a 75–25% distribution. The trained model yielded an accuracy score of 98.5% on the test set. The hyperparameters used while fine-tuning and the results are shown in Table-1.

Table -1: Hyperparameters and Final Epoch Results

Peak Learning Rate	10^{-5}	Batch size	16
Epochs	3	Training Loss	0.036
Learning Rate Scheduler	Cosine	Evaluation Loss	0.064
Warmup Ratio	0.1	Accuracy	0.9849
Evaluation Steps	99	F1 Score	0.9846

2) *Custom Reply Generation:* This is the second step if the user input is classified as depressed. Here, the input is passed through a multiclass classification transformer to predict the type of mental distress the user has, for example, anxiety, suicidal thoughts, depression, etc., and provide the correct response. The model used here is again the Roberta-base model. The dataset used here is a modified version of the Mental Health Conversational Data Dataset from Kaggle. Due to the extremely small size of the dataset, only the training loss was recorded at 98.2%. The hyperparameters were the same except for the evaluation steps, which were adjusted for the small dataset size. Since the dataset was small, traditional machine learning algorithms were also used as benchmarks. The results of traditional machine learning algorithms when compared with Roberta are shown in Table-2.

Table -2: Roberta vs Traditional ML methods

Methods	Training set Accuracy
Logistic Regression	59.05%
Multinomial Naive Bayes	67.67%
Linear Support Vector Classifier	90.95%
RoBERTa	98.17%

3) *Conversational Transformer:* This is the second step if the user input is classified as normal by the binary classification transformer. Here, we use Blenderbot for casual conversation. Every pair of input and output is stored up to a maximum of three pairs. These pairs are passed through the conversational transformer before every input so as to retain the context of the conversation.

Fig -4: Conversation without context

```
Human: Yes i have a cat
Bot:  What kind of cat do you have?
Human: The black kind
Bot:  Do you like the color black? It
is one of the three primary colours
of pigments in painting.
Human: quit
```

Fig -5: Conversation with context

```
Human: Yes i have a cat
Bot:  What kind of cat do you have?
Human: The black kind
Bot:  I love black cats! I have two
myself. Do you have any other pets?
Human: quit
```

4. RESULTS

Fig -6: Casual conversation without depression detection

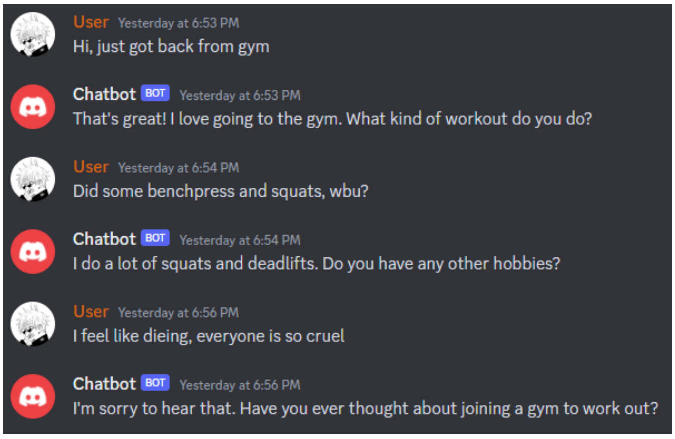
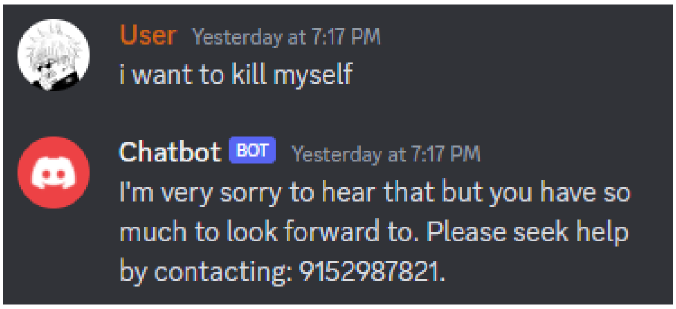


Fig -7: Depression Detection implemented



5. CONCLUSION

To summarize, developing a depression amelioration chatbot that employs both transfer learning and pre-trained transformers is a potential way for enhancing mental health results. This chatbot has the ability to give individuals with customised and accessible help, allowing for early intervention and even preventing symptoms from deteriorating. The chatbot's use of natural language processing and machine learning technologies enables it to give tailored responses that adapt to the user's needs over time. Furthermore, the chatbot is offered on an easy-to-use platform, making it more accessible than a personal therapist. While the chatbot's performance is not equivalent to that of a real-life therapist, more research is needed to improve the chatbot's effectiveness and user experience. These preliminary findings imply that a depression amelioration chatbot is a viable tool for enhancing mental health outcomes and giving persons in need with accessible support.

ACKNOWLEDGEMENT

We wish to express our sincere gratitude to *Dr. Sanjay U. Bokade*, Principal, and *Prof. S. P. Khachane*, Head of Department of Computer Engineering at MCT's Rajiv Gandhi Institute of Technology, for providing us with the opportunity to work on our project, "Depression amelioration Chatbot using NLP and Deep Learning." This project would not have been possible without the guidance and encouragement of our project guide, *Prof. Dilip Kale*, and the valuable insights of our project expert, *Dr. Sharmila Gaikwad*. We would also like to thank our colleagues and friends who helped us complete this project successfully.

REFERENCES

[1] Fitzpatrick KK, Darcy A, Vierhile M. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Ment Health. 2017 Jun 6;4(2):e19. doi: 10.2196/mental.7785. PMID: 28588005; PMCID: PMC5478797.

[2] Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. Can J Psychiatry. 2019 Jul;64(7):456-464. doi: 10.1177/0706743719828977. Epub 2019 Mar 21. PMID: 30897957; PMCID: PMC6610568.

[3] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. ArXiv [Cs.CL]. From http://arxiv.org/abs/1907.11692

[4] Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, Jason Weston (2020). Recipes for building an open-domain chatbot. ArXiv [Cs.CL]. From http://arxiv.org/abs/2004.13637

[5] Khan, S. (2019, September 4). BERT, RoBERTa, DistilBERT, XLNet — which one to use? Retrieved 29 March 2023, from Towards Data Science website: https://towardsdatascience.com/bert-roberta-distilbert-xlnet-which-one-to-use-3d5ab82ba5f8

[6] LianaYe2, (n.d.). The future of_conver_ai[6933]. Retrieved 29 March 2023, from https://www.slideshare.net/LianaYe2/the-future-ofconverai6933