AI Chatbots In Therapy For Depression

Ayden Nelson, Jenny Epstein, Chris Sanchez, Janine Ortiz, Daniel Chelling
SCI 250-06: Grand Challenge Initiative

Abstract

Our challenge was harnessing AI to be used in depression therapy. First step was researching how to train an AI model, and what data we should train it on. We collected data from outside sources, specifically Kaggle and GitHub, and manually cleaned the data in excel before training our model on that data. After finding and formatting the data, we would train an AI model, then implement it into a UI so that we can test it's conversational behavior. With this AI, it would allow therapeutic services to be accessible to all individuals.

Introduction

Because of the hardships caused by COVID-19, there has been a spike in mental illness throughout the population. Despite the effects COVID-19 may medically have, other mental issues can also arise during this time.(Pfefferbaum and North, 2020). With the new rise of mental health disorders, we have decided to create an AI therapy chatbot that can create original responses. With this chatbot, we hope to be able to make a therapy setting less intimidating. Some of the knowledge gaps we encountered during this project was the fact that AI can be a challenging subject, and coding an AI from scratch is rather difficult.

Literature Cited

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 mental health. 2017;4(2):e19.

Pfefferbaum, Betty, and Carol S. North. "Mental Health and the Covid-19 Pandemic." *New England Journal of Medicine*, vol. 383, no. 6, 2020, pp. 510–12. *Crossref*, doi:10.1056/nejmp2008017.

Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. n.d. "Language Models Are Unsupervised Multitask Learners," 24.

D'Alfonso, Simon, Olga Santesteban-Echarri, Simon Rice, Greg Wadley, Reeva Lederman, Christopher Miles, John Gleeson, and Mario Alvarez-Jimenez. 2017. "Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health." *Frontiers in Psychology* 8. https://doi.org/10.3389/fpsyg.2017.00796.

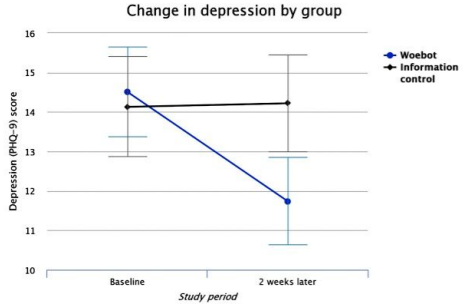

***Figure 1**: Change in mean depression after 2 weeks using a Therapy chatbot*

```
sess = gpt2.start_tf_sess()

gpt2.finetune(sess,
              dataset=file_name,
              checkpoint_dir='/content/drive/My Drive/checkpoint',
              model_name='774M',
              steps=3000,
              restore_from='fresh',
              optimizer='adafactor',
              learning_rate=0.0001,
              run_name='therapyrun',
              print_every=10,
              sample_every=500,
              save_every=1000,
              overwrite=False
              )
```

***Figure 2**: Code for training AI Model, given the parameters*

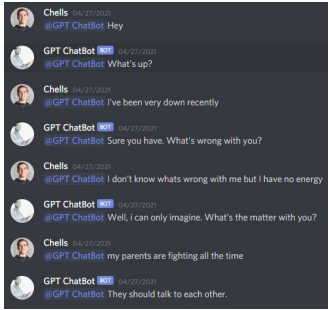

***Figure 3:** UI implementation of AI into discord*

Acknowledgements

- Dr. Bingjie Zhang
- Dr. Kenjiro Quides
- Dr. Brenna Gormally
- Dr. Alison Ramirez

Methods

1. Identified the challenge of applying AI to therapy in a chatbot format
2. Researched topics under AI, User Interfaces, and Depression therapy
3. Gathered, cleaned, and formatted data from external sources
4. Taking the formatted data, and training our AI model
5. Integrate the AI into a discord bot, allowing us to interact with the AI
6. Test the responses of the AI by exchanging messages and observing the response
7. From the observed behavior, continue training on existing or new data

Results

- We found that our AI is able to mimic the trained behavior
- The data we use is important, it heavily affects how the AI responds
- The main impact overall is making an accessible alternative to therapy
- Recognizing the limitations of online learning, and self taught work
- Future improvements can be applied, no where near a finished product
- Different data sets, as well as different size models can affect results.

Conclusion

Based on our progress, a person must look over certain number of responses and identifying if the responses are appropriate. We hope to continue to reassessing our model with our current data by making appropriate changes to enhance our model. We also hope to find and clean new data to train our models with to ensure that we have a generally constant response and accuracy provided by our models.