Narrative Analysis: Results Analysis and Discussion

Nolan Jones, Jordan duBarry, Joseph Henshaw

11/29/2024

In today's world, narratives regarding mental health are dissolving. No longer does agitation or depression mean that someone requires extreme treatment, and more people than ever have access to resources for their mental health. With all this progress aside, mental health resources are not perfect. Due to the nature of therapy and the complexity of one's mind, mental health issues may require long periods of observation when compared to other illnesses [1].

This diagnostic process can absolutely be improved though. Emotional analysis and voice analysis AI provides an alternative approach for doctors, allowing them to use tools like journaling or emotional surveys to help quickly and accurately diagnose someone without requiring them to personally interact with every journal entry or GAD-7 score [2]. That being said, patient data is a difficult dataset to use, as it needs to comply with HIPPA as well as other consumer privacy laws like the GLBA [3]. This puts limitations on the accuracy of emotional analysis AI, since training datasets are composed of specific cases patients have consented to use or artificial cases [4].

To work around this limitation, we have worked to introduce federated learning to narrative analysis. By continuously updating the global model with changes to local models based on individual user input, it should lead to a more accurate global model as well as faster training times. Since the user data is all contained locally, it is also in compliance with necessary data privacy laws [5].

This paragraph will be a description of our framework including but not limited to aspects of our pipeline, datasets used, experimental design, and other prerequisite information for our results. Our project uses the Flwr federated learning framework in conjunction with the Hugging Face version of Google Research's BERT base uncased language model. The model will be trained using PyTorch and on Google Research's Go Emotions dataset. The dataset samples over 58,000 comments from Reddit which are annotated as consisting of 27 different emotional categories, achieving an F1-score of 0.46. Our pipeline uses Hugging Face's auto tokenizer.

At the moment, we do not have any experimental results as we are working on setting up our experiment environment. The data we will be looking at will be the model convergence over several iterations of federated learning environments. In other words, we are attempting to see whether the models will converge to generate a useful analysis of control data between iterations of the model. In furtherance of this goal, we will run three experiments with the same dataset, partitioned randomly into 50 segments using the Flwr Dataset package's IidPartitioner function.

The Flower framework further creates 50 virtual clients, each of which train their own client-specific models. Once this is complete, the smaller models are used to create a population-wide model for the clients. We have had significant difficulty in initializing our experiments using Chameleon Cloud resources as we have so far been unable to satisfy the required dependencies to run our experiments on CUDA-enabled machines.

When it comes to user data, especially medical data, privacy is a major point of concern. Federated learning keeps data private by training local models and sending changes in the model up to a central model rather than data. Our approach also implements differential privacy which adds a layer of noise to model changes so that individual data contributions cannot be discerned. Keeping data decentralized reduces the risk of data breaches because there is no centralized dataset for attackers to target. Also, there is no personal data that is transferred which prevents it from being maliciously intercepted by foreign parties.

References:

[1] U.S. Department of Health and Human Services. (n.d.). *Depression*. National Institute of Mental Health. https://www.nimh.nih.gov/health/publications/depression

[2] Place, S., Blanch-Hartigan, D., Rubin, C., Gorrostieta, C., Mead, C., Kane, J., Marx, B. P., Feast, J., Deckersbach, T., Pentland, A. "Sandy," Nierenberg, A., & Azarbayejani, A. (2017). Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. *Journal of Medical Internet Research*, *19*(3). https://doi.org/10.2196/jmir.6678

[3] Murray, C. (2024, June 3). *U.S. Data Privacy Protection Laws: A Comprehensive Guide*. Forbes. https://www.forbes.com/sites/conormurray/2023/04/21/us-data-privacy-protection-laws-a-comprehensive-guide/

[4] (OCR), O. for C. R. (2024, August 22). *Methods for de-identification of phi*. HHS.gov. https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

[5] Yin, X., Zhu, Y., & Hu, J. (2021). A comprehensive survey of privacy-preserving federated learning. *ACM Computing Surveys*, *54*(6), 1–36. https://doi.org/10.1145/3460427