Developing a Federated Learning-Based System for Personalized MentalHealth Assessment and Prediction

Shrutika Malve*, Mahesh Vaswani*, Abhay Shanbhag*, Prof. Sarang Joshi

SCTR's Pune Institute of Computer Technology

---***---

Abstract

In mental health care, depression presents a significant challenge, especially given the sensitivities around personal health data and the need for tailored therapies. This project introduces an innovative Federated Machine Learning (FML) framework designed specifically to analyze and predict depression while preserving data privacy and enhancing model accuracy.

Traditional depression analysis methods strug- gle with maintaining patient confidentiality and ensuring diverse datasets. Centralized data col- lection, while thorough, poses substantial privacy risks and can deter participation, particularly in mental health contexts. Moreover, these methods often fail to capture the diverse manifestations of depression across different demographics, leading to less effective predictive models and interventions. The FML framework addresses these concerns by processing data locally on individual users' devices, safeguarding personal information and ensuring that sensitive data remains on the user's device. This approach enhances privacy and incorporates diverse and heterogeneous data, resulting in a model that is inclusive and representative of various depression manifestations. The primary goal is to create a robust, scalable model capable of accurately predicting depression and providing insights for personalized treatment strategies. By leveraging FML, the project aims to set a new standard in mental health care analytics, offering a privacy-conscious, scalable solution that can handle diverse data sources. This approach has the potential to revolutionize mental health care, providing a deeper under- standing of depression and paving the way formore effective, personalized care solutions.

1 Introduction

Within the domain of mental health care, the formidable challenge of addressing depression is magnified by the intricate interplay of personal health data sensitivity and the demand for tailored therapeutic strategies. This re- search project, titled "Developing a Federated Learning-Based System for Personalized Mental Health Assessment and Prediction," introduces a groundbreaking initiative – a Federated Machine Learning (FML) framework meticulously crafted for the analysis and prediction of depression. The primary focus is on fortifying data privacy and refining the precision of predictive models which will be an extension to work in Korkmaz et al., 2022

Conventional methods employed in depression analysis Prabhudesai et al., 2021 , Minkowski et al., 2021 confront a dual dilemma: the imperative to uphold patient data confidentiality and the necessity to ensure the of datasets. Centralized data collection, though comprehensive, poses substantial privacy risks, potentially deterring participation, particularly within the sensitive realm of mental health care. Moreover, such centralized approaches as given by Zhang, 2022 often fall short in accommodating the di- verse manifestations of depression across various demographics, leading to less effective predictive models and interventions. The pro- posed FML framework addresses these intricate challenges by empowering local data processing on individual users' devices, a strategic move aimed at safeguarding personal information. This methodology ensures that sensitive data remains within the user's device, markedly enhancing privacy. Beyond its privacy-centric advantages, this approach allows for the assimilation of diverse and heterogeneous data, culminating in a model that is both inclusive and representative of the myriad manifestations of depression. The overarching objective of this project is to craft a robust, scalable model proficient in accurately predicting depression, offering insights crucial for informing personalized treatment strategies. Leveraging the capabilities of FML, the project aspires to establish a new standard in mental health care analytics

– A solution characterized by privacy conscious- ness, scalability, and adaptability to diverse data sources. This innovative approach holds the potential to revolutionize the landscape of mental health care, promising a more nuanced understanding of depression and heralding a new era of effective, personalized care solutions. Lai et al., 2021

2 Related Work

2.1 Federated Machine Learning in Mental Health Analysis

In Konečný et al., 2017 insights into the fundamentals of federated learning, emphasizing strategies to enhance communication efficiency in distributed machine learning. It lays a foundation for understanding how federated learning principles can be applied to mental health analysis.

2.2 Privacy-Preserving Machine Learning Techniques

In McMahan et al., 2023 , Pranto and Al Asad, 2021 and Korkmaz et al., 2022 authors explore communication-efficient learning of deep net-works from decentralized data, offering valuable techniques for privacy-preserving machine learning. The principles presented in this pa- per contribute to the privacy-conscious design of the proposed Federated Machine Learning framework.

2.3 Towards Reliable and Empathetic Depression-Diagnosis-Oriented Chats

This paper Lan et al., 2024 proposes a novel framework for using chatbots in preliminary depression diagnosis through interactive conversations with potential patients. It acknowledges the complexity of blending task-oriented and empathetic chit-chat elements in diagnosis- related dialogues, which necessitates both professional expertise and empathy. Traditional dialogue frameworks often prioritize single optimization goals, which may not effectively cater to the multifaceted nature of depression diagnosis conversations. Similar work is carried out at Gupta et al., 2022

To address these challenges, the paper Ogamba et al., 2023 introduces an innovative ontology definition and generation framework specifically tailored for depression diagnosis dialogues. This framework aims to combine the re- liability of task-oriented conversations with the empathetic appeal of chit-chat, thereby enhancing the effectiveness of chatbots in supporting individuals with depression symptoms.

2.4 Scalability and Adaptability in Federated Learning

Addressing the scalability and adaptability aspects of federated learning, Concone et al., 2022 explores the concept of distributed learning. The methodologies discussed provide a basis for creating a scalable and adaptable Federated Machine Learning framework capable of handling diverse data sources in mental healthanalytics.

By incorporating insights from these key studies, the literature survey establishes a foundation for the development of the proposed Federated Learning-Based System for Personalized Mental Health Assessment and Prediction.

3 Methodology

3.1 Server-side Implementation:

The server, built with Python and using Flask for client communication, managed the federated learning process. Initially, it created the global model with TensorFlow/Keras, tailored to the specific needs of mental health data, such as multi-class classification. While clients handled core data preprocessing, the server took care of additional tasks like label encoding and one-hot encoding for the target variable. During training, the server split the data, distributed initial model weights, collected updates from clients, and aggregated them to form a new global model. To ensure privacy, optional security measures like differential privacy and SSLencryption were implemented.

The Keras library was used to define a multi- class classification model for mental health analysis. The model architecture consisted of asequential stack of layers:

- **Input Layer:** This layer received the in- put data with a shape based on the number of features in the training data
- **Hidden Layer:** A Dense layer with 100 neurons and ReLU activation was used as the first hidden layer. L1 regularization (kernel-regularizer = l1-reg) was applied to this layer to potentially prevent overfitting.
- **Output Layer:** A Dense layer with 5 neurons and softmax activation was employed as the output layer, suitable for multi-class classification with five mental health classes (normal, stress, anxiety, depression, loneliness).
- **Dropout:** A Dropout layer with a rate of

0.25 was included to randomly drop out a portion of neurons during training, further helping to prevent overfitting.

This model architecture was then employed within the federated learning framework for collaborative training on the distributed mentalhealth dataset.

3.2 Client-side Implementation:

Client devices, also running Flask in Python, played a crucial role in the federated learning process. They received the global model from the server and trained it locally on their own mental health data, ensuring that the raw data never left their device. Secure communication with the server was maintained through SSL encryption. Additionally, to further protect privacy, differential privacy techniques, such as adding noise to updates, could be optionally implemented. After training, clients sent the updated model weights back to the server, contributing to the global model without compromising their data.

3.3 Federated Learning Architecture:

The federated learning setup included a central server and multiple client devices. The server managed the training process by distributing initial model weights, collecting updates from the clients, and aggregating them using methods like federated averaging. This approach combined the insights gained by clients from their local data without exposing the data it- self, enabling collaboration while maintainingprivacy.

3.4 Model Training and Evaluation:

We developed a multi-class classification neural network model using TensorFlow/Keras on the server, specifically tailored to a mental health dataset that included conditions like depression and anxiety. The training process followed a federated learning approach: the server sent initial model weights to the clients, who then trained the model locally on their own data. After training, the clients sent the updated weights back to the server, which aggregated these up- dates using methods like federated averaging to create an improved global model. We assessed the model's performance on a separate test set, using metrics such as accuracy, precision, recall, and F1-score for each mental health condition.

Evaluation Metrics: We evaluated the model's performance on a held-out test set using variousmetrics:

- **Test Accuracy:** The model achieved a test accuracy of 0.8014, indicating good generalizability to unseen data.
- **Loss:** The final training loss converged to 1.4052, suggesting reasonable learningprogress.
- **Overall AUC (macro-averaged):** The model achieved a macro-averaged Overall AUC of 0.95, demonstrating good discrimination between the mental health classes.
- **Per-Class Performance:** The model's performance also varied across different mental health classes. Precision, recall, and F1-score metrics revealed an averageprecision of 0.70, an average recall of 0.80,and an average F1-score of 0.73.

By employing these evaluation metrics, we comprehensively assessed the performance of the federated learning model for mental healthanalysis.

4 ALGORITHM

This research uses a federated learning approach to analyze mental health, aiming to distinguish between normal, stress, anxiety, depression, and loneliness. The client-server architecture ensures privacy-preserving training of a multi-class classification model with TensorFlow/Keras. We also use a separate questionnaire to collect user data for analysis. Additionally, as an optional feature, a chatbot can be integrated to offer users a platform for mental health inquiries.

4.1 Model Creation and Distribution

The central server begins by creating the initialglobal model using the TensorFlow/Keras Sequential API. This model architecture is specifically designed for the mental health data under investigation, as described in section 1 of the previous response. Once established, the model architecture and its corresponding weights are securely distributed to participating client de- vices through well-defined APIs. Secure communication protocols are employed to ensure the integrity and confidentiality of the transmit- ted information.

4.2 Local Training and Prediction

Upon receiving the model architecture and weights, client devices can reconstruct the model locally without compromising the privacy of their raw data. This localized model is then leveraged to predict mental health labels (normal, stress, anxiety, depression, loneliness) for the client's own data.

4.3 Federated Learning Loop

Following the local prediction phase, a feder- ated learning loop is established. This loop fa- cilitates the iterative improvement of the global model by incorporating knowledge gained from each participating client. Here's the breakdown of this loop:

- **Client-side Retraining and Data Collec- tion:** Based on the predicted label, user responses from the separate questionnaire, and optionally, interactions with the chat- bot (if integrated), clients can perform lo-cal retraining on their model using their own data. User responses and chatbot interactions, when applicable, contribute to enriching the client-side data used for re-training.
- **Updated Weight Transmission:** After retraining, the client transmits the updated model weights back to the server. This transmission should occur through securechannels to ensure data integrity.

4.4 Model Aggregation

The server aggregates the model updates received from multiple clients using techniques like federated averaging. Federated averaging allows the server to combine the knowledge gained by individual clients while preserving privacy, as it operates on the updated model weights without accessing the raw client data. The aggregated weights are then used to update the global model on the server. This iterative process of local training, weight transmission, and aggregation enables the global model to continuously improve by incorporating insights from all participating devices, user responses from the separate questionnaire, and optionally, interactions with the chatbot.

4.5 Recommendation Engine Integration

The progressively improved global model, re- siding on the server, can then be utilized for further analysis or integrated into a recommendation engine. This engine can leverage the predicted mental health label, user responses from the questionnaire, and chatbot interactions (if applicable) to deliver even more relevant con- tent. For instance, the top eight most relevant questions (e.g., for self-assessment or intervention) could be displayed based on the user's predicted state, their responses in the questionnaire, and the information gleaned from the chatbot conversation (if available).

4.6 Chatbot Functionality

The chatbot, as an optional feature, provides users with a safe space to ask questions and receive informative, contextually relevant responses about mental health in a descriptive manner. It is designed to avoid collecting any personally identifiable information (PII) from the user, ensuring privacy Oreščanin et al., 2024. Interactions with the chatbot can con- tribute valuable data for client-side retraining, potentially enhancing the model with additional insights.

This federated learning approach offers an effective solution for collaborative training on a distributed dataset while safeguarding user privacy by keeping raw data on client devices. Through the client-server architecture, secure communication protocols, a separate questionnaire, and the optional chatbot, this method supports the development of a robust, privacy-preserving model for mental health analysis. It also provides users with a potential support sys- tem through the chatbot, enhancing the overall utility and accessibility of the system.

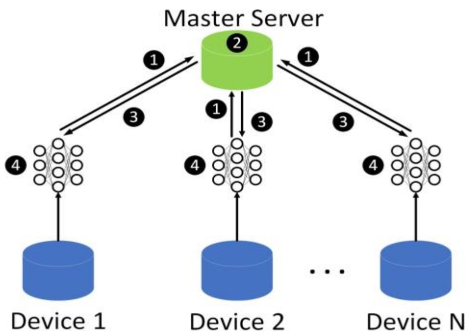

Figure 1: Architecture of Federated Machine Learning

5 CONCLUSION AND FUTURESCOPE

5.1 Conclusion

This research explores the application of federated learning in mental health analysis, focusing on distinguishing between normal, stress, anxiety, depression, and loneliness. The pro- posed approach utilizes a client-server architecture and employs TensorFlow/Keras for privacy-preserving training of the model. Additionally, a separate questionnaire gathers user data for analysis, and optionally, a chatbot canbe integrated to offer a platform for user inquiries. The federated learning architecture ensures user privacy by keeping raw data on client devices. The model is trained iteratively through the transmission of updated model weights from clients to the server for aggregation, without disclosing the underlying individual data. This collaborative approach facilitates the development of a robust model by incorporating insights from a distributed dataset. The key findings of this research include:

- The feasibility of employing federated learning for mental health analysis whilepreserving user privacy.
- The effectiveness of the separate questionnaire in enriching the client-side data used for model training.
- The potential benefits of a chatbot in pro- viding user support and enhancing the model's understanding through user inter-actions (if integrated).

5.2 Future Scope

This research opens doors for further exploration in the domain of federated learning for mental health analysis. Here are some promising avenues for future investigation:

- **Incorporation of additional data sources:** Explore the integration of sensor data (e.g., wearables) or anonymized electronic health records to potentially improve the model's accuracy.
- **Differential privacy advancements:** Explore novel techniques for differential privacy to further enhance user privacy while maintaining model performance.
- **Explainable AI for mental health:** Develop interpretable models to understand the factors influencing the model's predictions, potentially leading to more targetedinterventions.
- **Large-scale federated learning deployments:** Evaluate the scalability of the pro- posed approach in real-world scenarios with a larger number of participants.

6 Outcome/ results of research (screenshots of work done)

Development of a Mental Health Diagnosis System: Successfully implemented a system capable of diagnosing mental health disorders based on user responses to a questionnaire.

Integration of Machine Learning Models: Integrated machine learning models using Keras and TensorFlow to predict mental healthdisorders with high accuracy.

Creation of a Web Interface: Developed a user-friendly web interface using Flask, HTML, and CSS, allowing users to interact with the diagnosis system conveniently.

Enhanced User Experience: Implemented features such as clear questionnaire prompts and responsive design to enhance user experi-ence and accessibility.

Application of Modern Engineering Tools: Utilized Python, Flask, Keras, Tensor- Flow, HTML, and CSS to create a robust and efficient mental health diagnosis system.

Successful Demonstration of Federated Learning: Demonstrated the feasibility of federated learning techniques for enhancing model performance while preserving data privacy.

Recommendation Engine Integration: Integrated a recommendation engine based on prediction results to provide personalized recommendations for mental health supportresources.

Incorporation of Chatbot Functional- ity: Implemented chatbot functionality to provide immediate support and guidance to users based on their responses and diagnosis results.

References

Federico Concone, Cedric Ferdico, Giuseppe Lo Re, and Marco Morana. 2022. A federated learning approach for distributed human activity recognition. In *2022 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 269–274. Arti Gupta, Chanchal Maurya, Khyati Dhere, and Vijay Kumar Chaurasiya. 2022. Wellness detection using clustered federated learning. In *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, pages 1–5. Jakub Konečný, H. Brendan McMahan, Fe- lix X. Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. 2017. Federated learning: Strategies for improving communica-tion efficiency.

Abdulkadir Korkmaz, Ahmad Alhonainy, and Praveen Rao. 2022. An evaluation of fed- erated learning techniques for secure and privacy-preserving machine learning on med- ical datasets. In *2022 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–7.

Fan Lai, Xiangfeng Zhu, Harsha V. Mad- hyastha, and Mosharaf Chowdhury. 2021. Oort: Efficient federated learning via guided partici- pant selection.

Kunyao Lan, Cong Ming, Binwei Yao, Lu Chen, and Mengyue Wu. 2024. Towards reliable and empathetic depression-diagnosis- oriented chats.

H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2023. Communication-efficient learn- ing of deep networks from decentralized data.

Laura Minkowski, Kristiina Valter Mai, and Dharmendra Gurve. 2021. Feature extraction to identify depression and anxiety based on eeg. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine BiologySociety (EMBC)*, pages 6322–6325.

Maria Ogamba, Joseph Gitonga, Betsy Muri- ithi, John Olukuru, and Joseph Sevilla. 2023. Wellness buddy: An ai mental health chatbot for kenyan university students. In *2023 First International Conference on the Advancements of Artificial Intelligence in African Context (AA- IAC)*, pages 1–6.

Dražen Oreščanin, Tomislav Hlupić, and Boris Vrdoljak. 2024. Managing personal identifi- able information in data lakes. *IEEE Access*, 12:32164–32180.

Siddharth Prabhudesai, Apurva Mhaske, Manvi Parmar, and Sumedha Bhagwat. 2021. Depres-sion detection and analysis using deep learn- ing: Study and comparative analysis. In *2021 10th IEEE International Conference on Com- munication Systems and Network Technologies (CSNT)*, pages 570–574.

Md. Appel Mahmud Pranto and Nafiz Al Asad. 2021. A comprehensive model to monitor men- tal health based on federated learning and deep learning. In *2021 IEEE International Confer- ence on Signal Processing, Information, Com-munication Systems (SPICSCON)*, pages 18– 21.

Ye Zhang. 2022. The intelligent prediction model of college students' mental health basedon cluster analysis algorithm. In *2022 Inter- national Symposium on Advances in Informat- ics, Electronics and Education (ISAIEE)*, pages 653–657.

[b]0.4

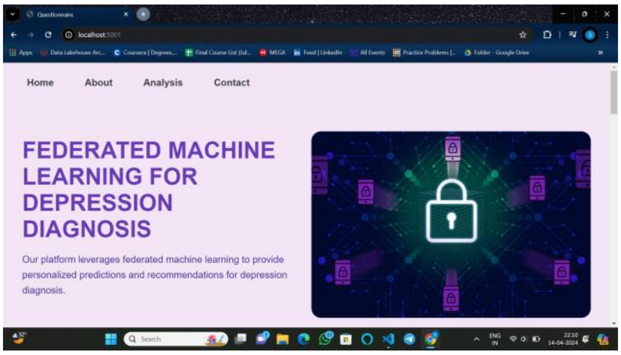

Figure 2: Implementation 1

[b]0.4

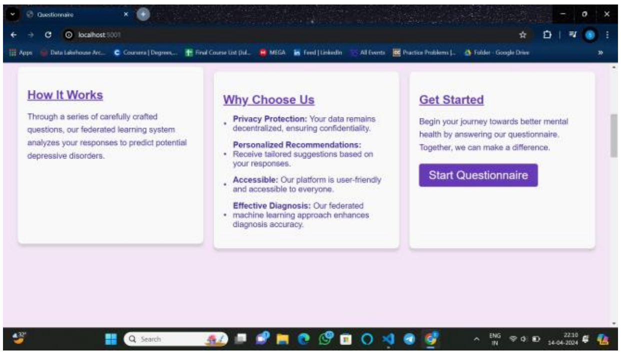

Figure 3: Implementation 2

[b]0.4

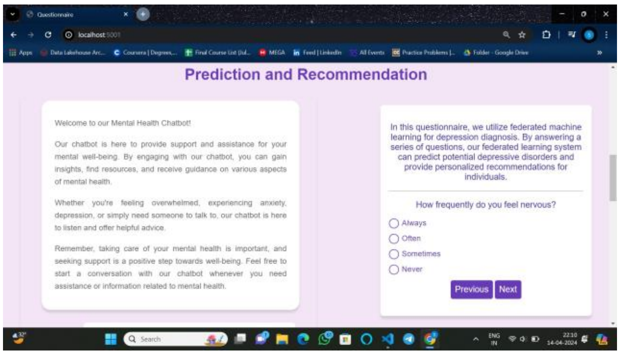

Figure 4: Implementation 3

[b]0.4

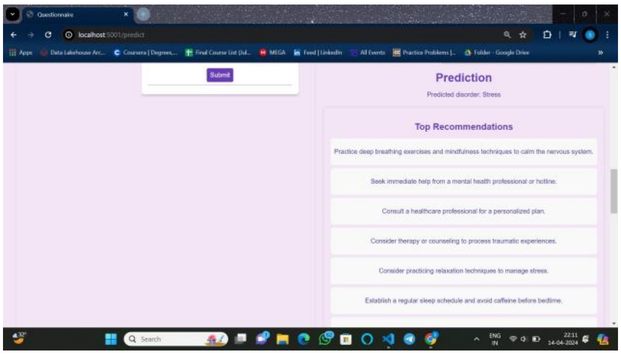

Figure 5: Implementation 4

[b]0.4

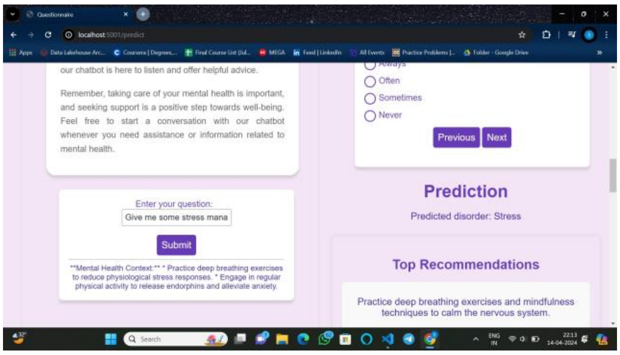

Figure 6: Implementation 5

[b]0.4

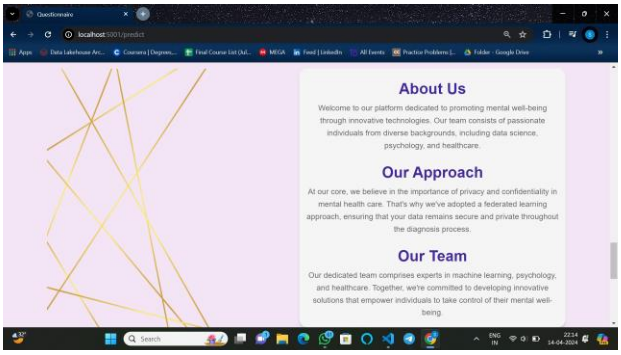

Figure 7: Implementation 6

[b]0.4

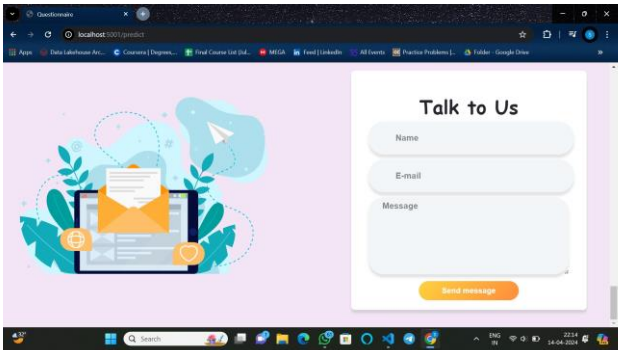

Figure 8: Implementation 7