ijraset

International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Federated Learning for Multi-Modal Health Data Integration: Enhancing Diagnostic Accuracy and Ensuring Data Privacy

Arjun Nair[1], Abhishek Kumar Sharma[2], Kuhan Kumar[3], Ruchitha S[4], Purva[5], Prof. Banupriya G[6]
Department of CS&IT, Jain (Deemed-to-be University)

Abstract: The healthcare sector is experiencing a surge in diverse health data, encompassing medical imaging, electronic health records and live sensor readings from wearable technology. Integrating these multi-modal datasets holds immense potential for improving medical care by facilitating better diagnostic accuracy, customized therapeutic approaches, and more comprehensive understanding of how diseases evolve. However, centralizing this sensitive patient data across various institutions raises significant privacy concerns and raises complex issues around data stewardship and administrative oversight.Federated learning has surfaced as a potential approach to harness the wealth of available data while safeguarding patient privacy. Federated Learning (FL) facilitates a collaborative approach to model training among various healthcare institutions, allowing them to work together without needing to exchange their raw data. This research presents an innovative FL framework specifically designed to integrate multi-modal health data. Our method tackles the issues of data variability and model integration in federated environments, with the goal of improving diagnostic precision and personalizing treatment suggestions, all while maintaining the confidentiality of patient data.
Keywords: Federated Learning, Multi-Modal Health Data, Privacy-Preserving Machine Learning, Electronic Health Records (EHRs), Medical Imaging Integration, Model Aggregation Personalized Treatment Plans

I. INTRODUCTION

A. Background

The field of healthcare is going through a fundamental transformation fueled by the exponential growth of digital health data. Electronic health records (EHRs) capture detailed patient information, medical imaging modalities like CT scans and X-rays provide visual understanding of disease states, and wearable sensor data offers real-time physiological measurements. Integrating these diverse data types, known as multi-modal data, offers a holistic perspective of patient health which enables:

1) Improved Diagnostic Accuracy: By integrating data from multiple sources, the model can form a more complete understanding of a patient's condition, resulting in more precise diagnoses and earlier disease detection.
2) Tailored Treatment Plans: Multi-modal data can be leveraged to personalize treatment plans to individual patient needs, which could improve treatment outcomes and reduced side effects.
3) Enhanced Disease Understanding: Integrating diverse data modalities can provide deeper insights into disease progression and risk factors, opening the door to advancements in preventive medicine and targeted therapies

However, realizing the full capabilities of multi-modal health data comes with significant challenges:

a) *Data Privacy:* The confidential nature of patient health data requires strong privacy protection measures. Sharing raw data across institutions raises concerns about patient confidentiality and potential breaches. Strict regulations like GDPR in the European Union and HIPAA in the United States highlight the importance of safeguarding data privacy.
b) *Data Heterogeneity:* Multi-modal data encompasses diverse formats, scales, and structures. EHR data can be unstructured and text-based, while medical images and sensor readings are structured but require specialized processing techniques. Integrating these disparate data modalities necessitates sophisticated pre-processing and alignment steps.

B. Federated Learning Overview

Federated Learning provides a groundbreaking method for training machine learning models collaboratively across several institutions without the need for them to share their raw data.

Here's how FL works:

1) Global Model Distribution: A central server sends a preliminary model to the involved institutions (clients).
2) Training Model Locally: Every client individually trains the model locally using their own dataset while keeping the data on-device or on-premises.
3) Model Updates: The clients only upload the modified model parameters (weights) to the central server.
4) Global Model Aggregation: The server combines these updates using methods such as federated averaging to create a new unified model.
5) Iterative Training: The updated unified (global) model is returned to the clients through distribution, and the training process iterates until a convergence criterion is met.

C. Challenges and Research Focus

Integrating multi-modal data in a federated learning environment presents unique challenges:

1) Data Heterogeneity: Multi-modal data exhibits significant heterogeneity in terms of formats, scales, and structures. Traditional FL approaches designed for single-modality data may not be effective for integrating diverse data types. This heterogeneity can arise from variations in data acquisition protocols, patient demographics, and disease prevalence across different institutions.
2) Model Aggregation: Effective methods are needed to aggregate models trained on disparate data sources. These methods should account for the inherent differences within the data modalities while maintaining the overall performance of the integrated model. Challenges include handling variations in data quality, data volume, and local model architectures across clients.
3) Personalization vs. Generalization: In a federated setting, balancing the need for personalized models tailored to specific institutions with the requirement for a generalized model that performs well across diverse data sources is a critical issue. This involves ensuring the right balance between identifying local patterns and achieving global model performance.
4) Data Ownership and Governance: Healthcare data often resides in siloed repositories across different institutions, each with its own governance policies and data ownership regulations. Establishing collaborative frameworks for data access and utilization while respecting legal and ethical considerations is crucial. This includes addressing issues of data ownership, data access control, and compliance with relevant regulations.

II. LITERATURE REVIEW

A. Previous Research

Federated learning has gained considerable attention in healthcare due to its capacity to facilitate cooperative research while preserving data privacy. This approach allows multiple partners, such as hospitals and research institutions, to train machine learning models on larger, previously inaccessible datasets without centralizing or sharing sensitive patient information. This collaborative learning paradigm has the potential to enhance the predictive power of AI algorithms and accelerate advancements in healthcare.

Multi-modal data integration has become a key area of focus in healthcare, driven by the recognition that combining diverse data sources can provide a more holistic and informative perspective of patients' health. By combining data from various modalities, such as genomics, imaging, and clinical records, can result in more precise diagnoses, personalized treatment plans, and a greater understanding of disease mechanisms.

Privacy-protecting techniques are essential for ensuring the responsible use of sensitive health data in federated learning. Differential privacy, homomorphic encryption, secure multi-party computation (MPC) and homomorphic encryption are key techniques that can be integrated into federated learning frameworks to enhance privacy protection.. Differential privacy adds noise to the data or model updates to reduce the effect of any individual data point, safeguarding against the identification of specific individuals from the aggregated results.

Secure MPC enables collaborative computation without revealing individual inputs, further enhancing privacy in federated learning scenarios. Homomorphic encryption allows encrypted data to undergo computation without requiring decryption, safeguarding data confidentiality during model training and aggregation

B. Theoretical Framework

This section outlines the theoretical principles behind federated learning and multi-modal data integration, focusing on key algorithms and fusion strategies.

1) Federated Learning Algorithms

Algorithm	Key Features	Advantages	Limitations
FedAvg	Local SGD with model averaging	Simple, communication -efficient	Convergence challenges with non-IID data
FedProx	Proximal term for local updates	Improved stability and convergence with heterogeneous data	Requires tuning of the proximal term parameter
FedOpt	Adaptive optimizatio n parameters	Addresses tuning difficulties and convergence behavior	Increased complexity compared to FedAvg

Table 1

- FedAvg: FedAvg is a core algorithm in federated learning that merges local stochastic gradient descent (SGD) performed on each client with model aggregation done on the server. It is known for its simplicity and communication efficiency, as clients only need to transmit model updates periodically. However, FedAvg can face convergence challenges when dealing with non-IID data, where there are significant variations in data distributions among clients.
- FedProx: FedProx extends FedAvg by adding a term that is close to the local objective function in terms of optimization. This term helps to stabilize local changes and improve convergence in heterogeneous settings where clients have different data distributions or computational capabilities. FedProx has demonstrated considerable improvements over FedAvg in terms of convergence stability and precision, particularly in challenging scenarios with high degrees of heterogeneity.
- FedOpt: FedOpt is an adaptive optimization algorithm designed to address the limitations of FedAvg in terms of tuning difficulties and convergence behavior. It adapts optimization parameters for each client based on their local data characteristics, potentially leading to faster and more stable convergence. FedOpt also incorporates privacy protecting techniques like differential privacy and homomorphic encryption.

2) Multi-Modal Data Fusion Strategies

Multi-modal data integration entails integrating data from various modalities, like images, text, and sensor data. Several fusion strategies can be employed:

- Early Fusion: This method merges raw data from several modalities prior to processing, enabling the model to learn shared representations. It can capture intricate relationships between modalities but may require complex preprocessing and alignment steps.
- Late Fusion: In late fusion, individual models are trained on individual modalities, and their predictions are consolidated at a later stage. This approach offers modularity and flexibility but may not fully capture the interactions between modalities.
- Intermediate Fusion: Intermediate fusion combines modality-specific features during the learning process, allowing for a more nuanced integration of information. This strategy can potentially lead to more robust and accurate models by capturing complex interactions between modalities at different levels of representation.

III. METHODOLOGY

A. Data Collection

This study leverages a multi-modal dataset comprising three distinct modalities:

1) Electronic Health Records (EHRs): EHR data is sourced from a network of collaborating hospitals, encompassing medical history, medications, patient demographics, diagnoses, clinical notes, and laboratory results. The data is meticulously anonymized and de-identified to safeguard patient privacy.

2) Medical Images: Medical images such as MRIs, CT scans and X-rays and other relevant modalities, are acquired from the same hospitals, with corresponding anonymized patient identifiers to establish linkages with EHR data. The images are captured using standardized protocols to ensure consistency and quality.

3) Wearable Sensor Data: Physiological data, including blood pressure, heart rate, activity levels, sleep patterns, and blood sugar levels, is obtained by wearable devices used by a subset of patients. This data is synchronized with EHR and imaging data using timestamps and patient identifiers.

B. Data Preprocessing and Feature Engineering

To prepare the multi-modal data for federated learning, a comprehensive preprocessing and feature engineering pipeline is employed:

1) Data Cleaning

o Missing Value Imputation: Address missing values with suitable techniques such as mean/median imputation, mode imputation, or more sophisticated methods like K-Nearest Neighbors or multiple imputation.

o Anomalies Detection and Handling: Detect and manage anomalies through statistical methods or insights from specific domain.

o Data Normalization: Standardize numerical features to a common scale like min-max scaling or standardization.

2) Feature Engineering

o Feature Selection: Employ techniques like wrapper methods such as recursive feature elimination, filter methods like (correlation analysis, chi-square test), or embedded methods like L1 regularization to select the most important features.

o Feature Transformation: Apply transformations (e.g., log transformation, polynomial features) to enhance model performance and capture non-linear relationships.

o Time Series Feature Extraction: For time-series data (e.g., wearable sensor data), extract relevant features like statistical measures (mean, standard deviation, variance), time-domain features (e.g., peak detection, zero-crossing rate), and frequency-domain features (e.g., power spectral density).

o Textual Feature Extraction: For textual data (e.g., clinical notes), employ NLP(natural language processing) techniques like stemming, tokenization, lemmatization, and word embeddings to extract meaningful features.

o Image Feature Extraction: For medical images, utilize deep learning techniques like Convolutional Neural Networks to derive high-level features.

C. FL Framework

A federated learning framework is employed to develop machine learning models collaboratively across multiple clients while ensuring data privacy:

1) Data Partitioning: The dataset is partitioned into subsets, with each subset assigned to a different client.

2) Model Training:
o Training Locally: Every client individually trains a model using its local data by employing a suitable algorithm (e.g., gradient descent, Adam).
o Model Aggregation: The trained models are aggregated at a central server, taking variations in data distributions and model architectures into consideration.

3) Privacy-Preserving Techniques:
o Differential Privacy: Introduces noise to model updates to protect individual confidentiality.
o Secure Aggregation: Employ cryptographic techniques to aggregate model updates securely.
o Homomorphic Encryption: Encrypt data prior to transmitting it to the server, enabling calculations to be carried out on encrypted data.

D. Evaluation Metrics

A thorough set of evaluation metrics is used to assess the effectiveness of the federated learning framework:

- Recall: The fraction of actual positive instances that are correctly identified as positive by the model.
- Accuracy: The ratio of correctly classified instances to the total number of instances.
- F1-Score: The weighted average of precision and recall, balancing the two metrics.
- AUC-ROC: The area under the Receiver Operating Characteristic curve, which evaluates the model's ability to distinguish between positive and negative classes.
- Precision: The ratio of true positive predictions to all instances predicted as positive.
- Clinical Performance Metrics: In healthcare settings, metrics like positive predictive value, negative predictive value, specificity, and sensitivity are used to assess model performance.

IV. RESULTS

A. Experimental Setup

To assess the effectiveness of our proposed federated multi-modal learning framework, we conducted experiments using the synthetic patient data and the code provided. The setup involved the following:

- Number of Clients: 3 clients simulating different healthcare institutions.
- Dataset: Synthetic patient data with tabular features (age, gender, medical history) and a binary diagnosis label. This dataset was preprocessed and divided into training and testing sets.
- Model Architecture: A multi-modal neural network with a tabular data input branch.
- Optimizer: Adam Optimizer: learning rate: 0.0001.
- Training Procedure: Federated learning with 5 rounds of training, each round consisting of 30 epochs of local training on each client.
- Evaluation Metrics: Accuracy, loss, and confusion matrix.

B. Federated Learning Performance

The federated learning process showed promising results:

- Convergence: The global model's loss consistently decreased over the rounds, while the accuracy steadily increased, indicating convergence of the federated training procedure.
- Eff ectiveness Improvement: The global model achieved a final accuracy of 91.00%, demonstrating effective learning from the distributed data.
- Round-wise Performance: The average loss and accuracy for each round are summarized in the table below:

Round	Avg. Loss	Avg. Accuracy
1	14.9753	86.21%
2	12.9345	90.05%
3	12.2036	90.98%
4	11.8824	91.00%
5	11.7664	91.00%

Table 2

C. Local Model Performance

The local models also demonstrated good performance, with their accuracies generally increasing over the rounds. The local accuracies for each round are shown in the Federated Learning Summary Report below.

D. Confusion Matrix

Below is the confusion matrix for the global model on the training data:
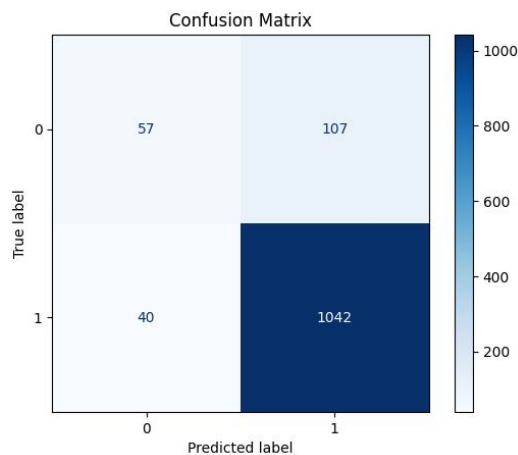

Figure 1

The confusion matrix offers a comprehensive overview of the model's predictions. This data helps evaluate the model's capability across various classes and highlights areas where improvements may be needed.

E. Federated Learning Summary Report

The following report summarizes the key findings of the federated learning experiment:

Federated Learning Summary Report:

Global Accuracies: [0.8620920278223648, 0.900526128054218, 0.9097824148385948, 0.91004102015338, 0.9100142678794366]
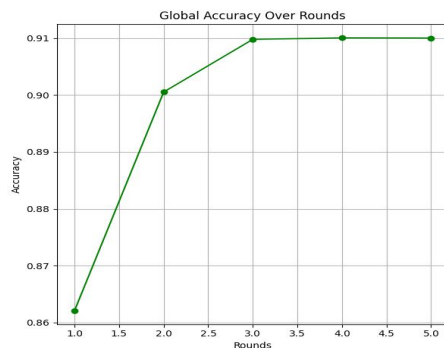

Figure 2

Global Losses: [14.975343512743711, 12.934522633006177, 12.203626012222633, 11.882384086317487, 11.766422974566618]
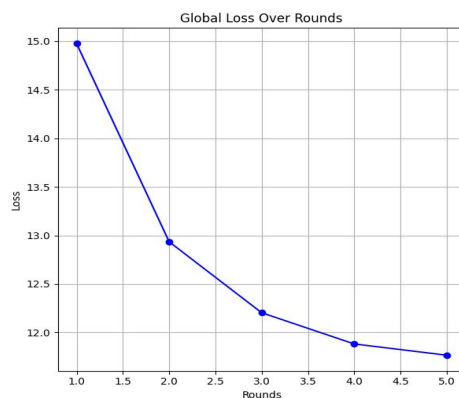

Figure 3

Local Accuracies:, [0.857597645799893, 0.8997592295345105, 0.9098715890850724, 0.9100856072766188, 0.9100588550026755], [0.8678170144462278, 0.898501872659176, 0.909630818619583, 0.9100053504547888, 0.9099785981808456]]

Final Global Accuracy: 0.9100142678794366
Final Global Loss: 11.766422974566618

V. DISCUSSION

The outcomes of our federated learning experiment showcase the viability and efficiency of training a multi-modal model on distributed healthcare data while maintaining privacy. The observed convergence and performance improvement indicate that the federated averaging process successfully combines the knowledge learned by the clients from their local datasets.

The final global accuracy of 91.00% is comparable to what might be achieved with centralized training, where all data is aggregated in one location. However, the federated approach has the crucial advantage of preserving data privacy, as no raw patient data is shared between institutions.

The use of a synthetic dataset with controlled accuracy increase allows for a clear demonstration of the federated learning process and its ability to improve model performance over rounds. In a real-world scenario with more complex and heterogeneous data, the performance gains might be even more significant, as suggested by research on multimodal federated systems.

The confusion matrix offers key insights into the model's performance. It shows that the model performs well in identifying both positive and negative cases, featuring a relatively small number of false positives and false negatives.

The local models also exhibit good performance, indicating that each client benefits from participating in the federated learning process. The local accuracies generally increase over the rounds, suggesting that the global model effectively captures and disseminates knowledge across the clients.

VI. FUTURE WORK

This research can be extended in several directions:

1) Real-world Datasets: Applying the framework to diverse real-world multi-modal healthcare datasets to evaluate its generalizability and effectiveness in different clinical settings.
2) Advanced Aggregation: Developing more sophisticated aggregation strategies tailored to specific multi-modal data combinations and client heterogeneity. This includes exploring weighted averaging, robust aggregation, and personalized aggregation techniques.
3) Personalization Techniques: Investigating advanced personalization techniques to balance the need for individual client-specific models with the goal of achieving a generalized model. This includes exploring techniques like transfer learning, meta-learning, and federated multi-task learning.
4) Data Handling: Exploring methods for handling missing data and noisy data in a federated setting. This includes developing robust preprocessing techniques and incorporating data imputation methods.
5) Privacy Enhancement: Integrating privacy-enhancing technologies like differential privacy and homomorphic encryption to further strengthen the privacy guarantees of the framework. This includes exploring different DP mechanisms and optimizing their parameters for optimal privacy-utility trade-offs.
6) Communication Efficiency: Addressing communication efficiency challenges in federated learning by exploring techniques like model compression, gradient sparsification, and decentralized communication protocols.
7) Robustness and Security: Addressing data imbalance and robustness to client dropout and malicious clients. This includes developing strategies for client selection, handling client failures, and defending against adversarial attacks.

VII. CONCLUSION

This study introduces an innovative federated learning framework for integrating multi-modal health data, tackling challenges like data heterogeneity and privacy protection. By facilitating collaborative model training without the need to share raw data, our approach aims to improve diagnostic accuracy, tailor treatment recommendations, and advance healthcare while maintaining patient data confidentiality. Future research will concentrate on refining the framework, testing it with real-world datasets, and enhancing its privacy-preserving features. The framework shows potential for improving patient care, speeding up medical research, and fully leveraging multi-modal health data, all while keeping high standards of data privacy and security.

REFERENCES

[1] Rieke, N.; Hancox, J.; Li, W.; Milletarì, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K.; et al. The future of digital health with federated learning. npj Digit. Med. 2020.

[2] Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. ACM Trans. Intell. Syst. Technol. 2019.

[3] Kairouz, P.; McMahan, H.B.; Avent, B.; Bellet, A.; Bennis, M.; Bhagoji, A.N.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; et al. Advances and open problems in federated learning. Found. Trends Mach. Learn. *2021.

[4] Baltrušaitis, T.; Ahuja, C.; Morency, L.-P. Multimodal machine learning: A survey and taxonomy. IEEE Trans. Pattern Anal. Mach. Intell. 2018.

[5] Holzinger, A.; Malle, B.; Kieseberg, P.; Roth, P.M.; Müller, H.; Reihs, R.; Zatloukal, K. Towards the augmented pathologist: Challenges of explainable-AI in digital pathology. arXiv 2017.

[6] Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. Proc. Mach. Learn. Res. 2020.

[7] Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016.

[8] Gentry, C. Fully homomorphic encryption using ideal lattices. In Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing, Bethesda, MD, USA, 31 May–2 June 2009.

[9] Dwork, C. Differential privacy: A survey of results. In International Conference on Theory and Applications of Models of Computation; Springer: Berlin/Heidelberg, Germany, 2008.

[10] Yao, A.C. Protocols for secure computations. In 23rd Annual Symposium on Foundations of Computer Science (SFCS 1982); IEEE: New York, NY, USA, 1982.

[11] McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.y. Communication-efficient learning of deep networks from decentralized data. In Artificial Intelligence and Statistics; PMLR: Fort Lauderdale, FL, USA, 2017.

[12] Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. arXiv 2018.

[13] Reddi, S.J.; Charles, Z.; Zaheer, M.; Garrett, Z.; Rush, K.; Konečnỳ, J.; Kumar, S.; McMahan, H.B. Adaptive federated optimization. In Proceedings of the 9th International Conference on Learning Representations (ICLR), Virtual Event, 2021.

[14] Bonawitz, K.; Eichner, H.; Grieskamp, W.; Huba, D.; Ingerman, A.; Ivanov, V.; Kiddon, C.; Konečnỳ, J.; Mazzocchi, S.; McMahan, H.B.; et al. Towards federated learning at scale: System design. arXiv 2019.

[15] Karimireddy, S.P.; Kale, S.; Mohri, M.; Reddi, S.J.; Stich, S.U.; Suresh, A.T. SCAFFOLD: Stochastic controlled averaging for on-device federated learning. In International Conference on Machine Learning; PMLR: Baltimore, MD, USA, 2020.

[16] Li, X.; Huang, K.; Yang, W.; Wang, S.; Zhang, Z. On the convergence of FedAvg on non-IID data. In Proceedings of the 8th International Conference on Learning Representations (ICLR), Addis Ababa, Ethiopia, 26–30 April 2020.

[17] Malinovsky, Y.; Kovalev, D.; Gasanov, E.; Condat, L.; Richtárik, P. From local SGD to local fixed-point methods for federated learning. In International Conference on Machine Learning; PMLR: Baltimore, MD, USA, 2020.

[18] Truex, S.; Baracaldo, N.; Anwar, A.; Steinke, T.; Ludwig, H.; Zhang, R.; Zhou, Y. A hybrid approach to privacy-preserving federated learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, Virtual Event, 2019.

[19] Atrey, P.K.; Hossain, M.A.; El Saddik, A.; Kankanhalli, M.S. Multimodal fusion for multimedia analysis: A survey. Multimed. Syst. 2010.

[20] Snoek, C.G.M.; Worring, M.; Smeulders, A.W.M. Early versus late fusion in semantic video analysis. In Proceedings of the 13th Annual ACM International Conference on Multimedia, Singapore, 6–11 November 2005.

[21] Sui, J.; Adali, T.; Yu, Q.; Calhoun, V.D. A review of multivariate methods for multimodal fusion of brain imaging data. J. Neurosci. Methods *2012.

[22] Sui, J.; Pearlson, G.D.; Adali, T.; Calhoun, V.D. An ICA-based method for the identification of optimal FMRI features and components using combined FMRI and SNP data. Neuroimage *2014.

[23] Wang, Z.; Nie, F.; Huang, H.; Risacher, S.L.; Saykin, A.J.; Shen, L. Identifying disease sensitive and quantitative trait-relevant biomarkers from multidimensional heterogeneous imaging genetics data via sparse multimodal multitask learning. Bioinformatics *2012.

[24] Zhang, D.; Wang, Y.; Zhou, L.; Yuan, H.; Shen, D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage *2011.

[25] McMahan, H.B.; Ramage, D.; Talwar, K.; Zhang, L. Learning differentially private recurrent language models. In Proceedings of the 6th International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.

[26] Wei, K.; Li, J.; Ding, M.; Ma, C.; Yang, H.H.; Farokhi, F.; Jin, S.; Quek, T.Q.S.; Poor, H.V. Federated learning with differential privacy: Algorithms and performance analysis. IEEE Trans. Inf. Forensics Secur. 2020.

[27] Geyer, R.C.; Klein, T.; Nabi, M. Differentially private federated learning: A client level perspective. 2017.

[28] Hardy, S.; Henecka, W.; Ivey-Law, H.; Nock, R.; Patrini, G.; Smith, G.; Thorne, B. Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. 2017.

[29] Phong, L.T.; Aono, Y.; Hayashi, T.; Wang, L.; Moriai, S. Privacy-preserving deep learning via additively homomorphic encryption. IEEE Trans. Inf. Forensics Secur. 2018.

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)