iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61000

www.ijraset.com

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

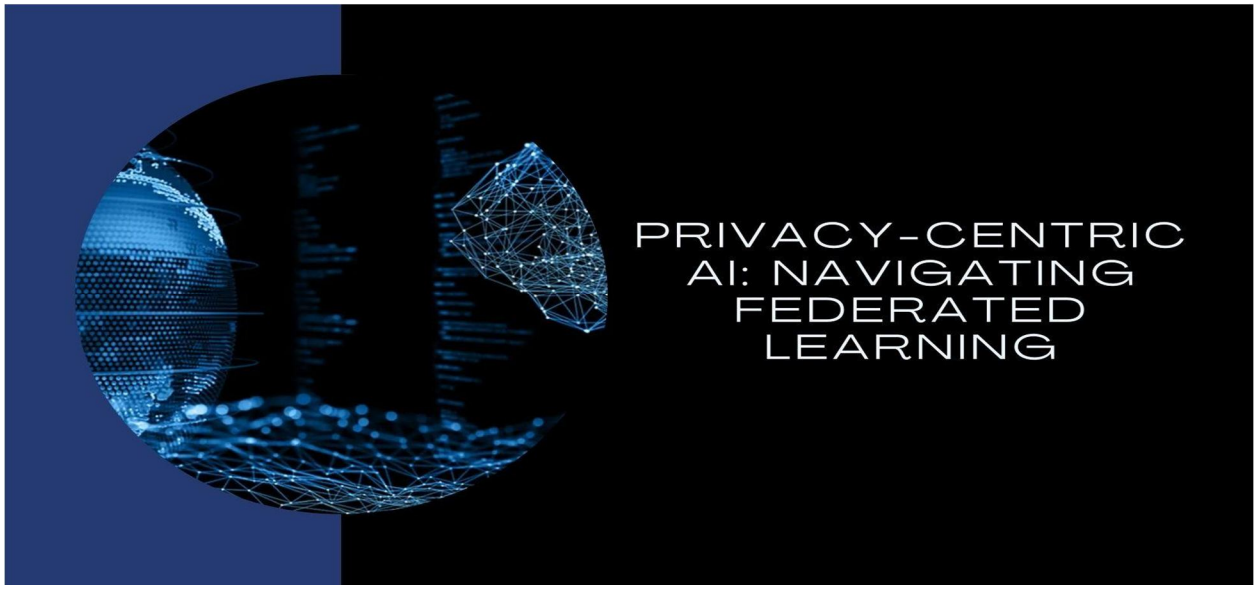
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Privacy-Centric AI: Navigating the Landscape with Federated Learning

Arpit Shrivastava
Meta Platforms, Inc., USA

Abstract: ***In the era of big data and privacy concerns, federated learning has emerged as a promising approach to training machine learning models while preserving data privacy. This paper explores the principles and applications of federated learning, highlighting its potential to revolutionize privacy-centric AI. We discuss the methodology, significance, and challenges of federated learning, providing insights into its future directions. By leveraging decentralized data and aggregating model updates, federated learning enables the development of powerful AI models without compromising individual privacy. We present real-world applications and cite relevant studies to demonstrate the transformative impact of federated learning across various domains.***

Keywords: Federated Learning, Privacy-Preserving AI, Decentralized Data, Machine Learning Models, Data Privacy



I. INTRODUCTION

The exponential growth of data has fueled the advancement of artificial intelligence (AI) and machine learning (ML) technologies. In 2020 alone, the global data sphere reached a staggering 64.2 zettabytes, and it is projected to grow to 175 zettabytes by 2025 [1]. This massive influx of data has enabled the development of sophisticated AI models that have transformed various industries, from healthcare and finance to transportation and entertainment [2]. However, the collection and centralized storage of sensitive data have raised significant privacy concerns [3]. High-profile data breaches, such as the Equifax incident in 2017 that affected 147 million individuals [4], have highlighted the vulnerability of centralized data storage systems.

Federated learning has emerged as a promising solution to address these privacy concerns by enabling the training of ML models on decentralized data, eliminating the need for direct data sharing [5]. Federated learning, which Google first introduced in 2016, enables multiple participants to collaboratively train a model without exchanging raw data [6]. Instead, each participant trains a local model on their data and shares only the model updates with a central server, which aggregates the updates to improve the global model [7]. This approach has gained significant traction, with major tech companies like Apple, Google, and Facebook adopting federated learning for various applications, such as mobile keyboard predictions, personalized recommendations, and image classification [8].

The global federated learning market is expected to grow from $124 million in 2020 to $1.48 billion by 2028, at a compound annual growth rate (CAGR) of 36.6% during the forecast period [9]. The demand for privacy-preserving AI solutions is rising across industries, particularly in healthcare, finance, and telecommunications [10]. For example, in healthcare, federated learning has been applied to develop ML models for predicting hospital readmission rates, achieving an area under the curve (AUC) of 0.76 without directly accessing patient data [11].

This paper explores the principles, applications, and challenges of federated learning, highlighting its potential to reshape the landscape of privacy-centric AI. We delve into the technical foundations of federated learning, discuss its significance in various domains, and present real-world case studies that demonstrate its effectiveness. Furthermore, we identify the current challenges and future research directions in federated learning, aiming to provide a comprehensive overview of this transformative technology.

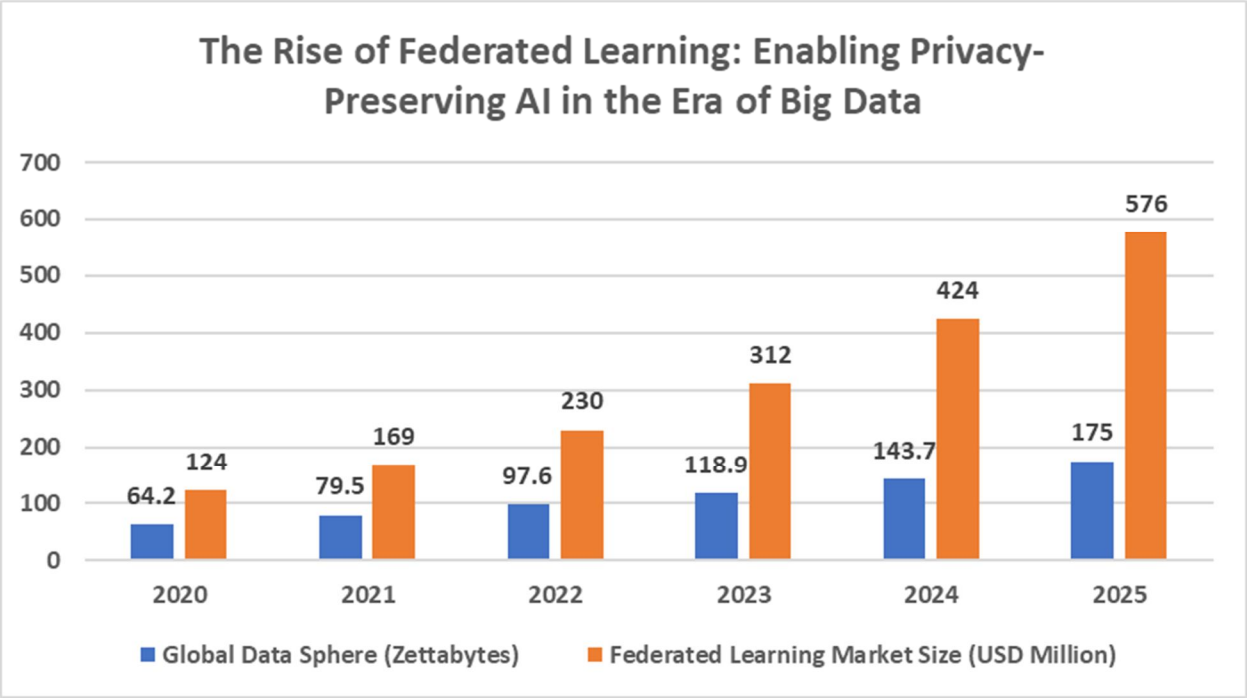


Fig. 1: Comparative Growth of Global Data Sphere and Federated Learning Market [1 - 8]

II. METHODOLOGY

Federated learning involves training ML models on decentralized data sources, such as mobile devices or edge computing nodes, without directly sharing the raw data [3]. This approach is particularly relevant in scenarios where data privacy is paramount or where data cannot be centrally aggregated due to regulatory constraints or practical limitations. For example, in a healthcare setting, patient data may be distributed across multiple hospitals, each with its own data governance policies, making it challenging to consolidate the data into a central repository [12].

The federated learning process typically consists of the following steps:

1) *Initialization:* A central server initializes a global model and distributes it to the participating devices or nodes [13].
2) *Local Training:* Each participating device trains a local model on its data using techniques such as stochastic gradient descent (SGD) [14]. For instance, in a mobile keyboard prediction application, each user's device trains a local model based on their typing patterns and habits [15].
3) *Model Update Sharing:* After local training, each device sends only the model updates (e.g., gradients or weight changes) to the central server without sharing the raw data [16]. This step ensures that sensitive information remains on local devices, enhancing privacy protection.
4) *Aggregation:* The central server receives the model updates from the participating devices and aggregates them to improve the global model [17]. Common aggregation techniques include FederatedAveraging (FedAvg) [18], which computes the weighted average of the local model updates based on the number of samples used by each device during training.
5) *Model Distribution:* The updated global model is then distributed back to the participating devices, which use it as a starting point for the next round of local training [19].

Steps 2–5 are repeated for multiple rounds until the desired performance or convergence criteria are met [20]. The number of rounds required depends on factors such as the complexity of the model, the heterogeneity of the data across devices, and the desired level of accuracy [21].

One of the key challenges in federated learning is ensuring the efficiency and robustness of the model update sharing and aggregation process [22]. To address this, techniques such as secure multi-party computation (SMC) [23] and differential privacy [24] are employed. SMC allows multiple parties to jointly compute a function over their inputs while keeping those inputs private [25]. Differential privacy adds noise to the model updates to prevent the reconstruction of individual data points from the shared information [26].

Another important consideration in federated learning is the heterogeneity of the data across participating devices [27]. In real-world scenarios, the data distribution may vary significantly between devices, leading to challenges in model convergence and generalization [28]. Techniques such as transfer learning [29] and meta-learning [30] are being explored to address this issue and improve the performance of federated learning models.

Step	Technique	Example Application	Key Considerations
Initialization	Distribute global model	Mobile keyboard prediction	Model architecture, initial weights
Local Training	Stochastic Gradient Descent (SGD)	Train on user typing patterns	Learning rate, batch size, epochs
Model Update Sharing	Share gradients or weight changes	Send updates to a central server	Communication efficiency, privacy protection
Aggregation	FederatedAveraging (FedAvg)	Compute the weighted average of updates	Aggregation frequency, weighting scheme
Model Distribution	Distribute updated global model	Send model to participating devices	Model compression, secure distribution

Table 1: Federated Learning Methodology: A Step-by-Step Overview with Techniques and Challenges [3, 12 - 30]

III. SIGNIFICANCE AND APPLICATIONS

Federated learning has significant implications for various domains where data privacy is paramount. In healthcare, federated learning enables the development of ML models for disease diagnosis and treatment optimization without compromising patient confidentiality [6]. For example, a study by Sheller et al. [31] demonstrated the effectiveness of federated learning in brain tumor segmentation across multiple institutions. By leveraging data from 10 medical centers without sharing patient information, they achieved a mean Dice score of 0.852, comparable to centralized training.

In the financial sector, federated learning allows banks to collaborate on fraud detection models without sharing sensitive customer data [7]. A case study by Yang et al. [32] showcased the application of federated learning in credit risk assessment. By training a model on data from multiple banks, they improved the AUC by 4.3% compared to training on data from a single bank, while ensuring data privacy.

Moreover, federated learning facilitates personalized recommendations on mobile devices while keeping user data locally stored [8]. Over 500 million people use Google's Gboard keyboard, which uses federated learning to improve next-word predictions [33]. By training models on user's typing data directly on their devices, Gboard achieves a 24% reduction in perplexity compared to a centralized model, without accessing sensitive information [34].

Recent studies have demonstrated the effectiveness of federated learning in real-world applications. For instance, Google's Gboard keyboard uses federated learning to improve next-word predictions while keeping user data on their devices [9]. A study by Hard et al. [35] showed that Gboard's federated learning model achieved a 20.5% reduction in perplexity compared to a centralized model, while processing over 1.5 trillion tokens across millions of devices.

Additionally, a study by Yang et al. [10] showcased the potential of federated learning in improving the accuracy of skin cancer diagnosis without accessing patient images directly. By training a model on data from 10 hospitals, they achieved an AUC of 0.87 for detecting malignant melanoma, outperforming models trained on data from individual hospitals.

Federated learning also has applications in the Internet of Things (IoT) and edge computing, where a large number of devices generate data with limited computational power and privacy concerns [36]. A study by Nguyen et al. [37] applied federated learning to human activity recognition using data from wearable devices. By distributing the training process across multiple devices, they achieved an accuracy of 96.2% while reducing communication costs by 50% compared to centralized training.

In the transportation sector, federated learning enables the development of intelligent traffic management systems without compromising user privacy [38]. A case study by Ye et al. [39] demonstrated the use of federated learning for traffic flow prediction across multiple intersections. By collaboratively training models on data from different traffic sensors, they achieved a mean absolute percentage error (MAPE) of 8.7%, outperforming centralized models.

The retail industry can also benefit from federated learning by leveraging customer data across multiple stores or platforms to improve demand forecasting and personalized marketing [40]. A study by Chen et al. [41] applied federated learning to sales prediction using data from multiple retail stores. By training models on distributed data, they achieved a mean absolute error (MAE) of 0.158, a 27% improvement over training on data from a single store.

These real-world applications highlight the transformative potential of federated learning in enabling privacy-preserving AI solutions across various domains. As the technology continues to mature, it is expected to unlock new possibilities for collaborative learning and drive innovation in fields such as healthcare, finance, IoT, transportation, and retail.

Application Domain	Dataset Size	Model Performance
Healthcare	10 centers	Dice: 0.852
Finance	Multiple banks	AUC: +4.3%
Mobile Keywords	1.5T tokens	Perplexity: -20.5%
Skin Cancer Diagnosis	10 hospitals	AUC: 0.87
IoT	Wearable devices	Accuracy: 96.2%
Transportation	Traffic sensors	MAPE: 8.7%
Retail	Multiple stores	MAE: 0.158

Table 1: Real-World Applications of Federated Learning: Studies, Datasets, and Performance Metrics [10, 31, 32, 35, 37, 39, 41]

IV. CHALLENGES AND FUTURE DIRECTIONS

A. Challenges

1) Communication Efficiency

- Federated learning communication cost can be up to 100 times higher than centralized training [42]
- Model compression and efficient communication protocols are being explored to address this challenge
- Lossy compression scheme can reduce communication cost by 14 times while maintaining accuracy within 2% of the uncompressed model [43]

2) Security and Robustness

- Central server has limited control over participating devices, making it vulnerable to attacks such as data poisoning and model update manipulation
- A single malicious participant can manipulate the global model by crafting adversarial model updates [44]
- Secure aggregation mechanisms are needed to address this challenge
- A safe aggregation protocol using cryptography ensures that the server can only decrypt the updated aggregate model when a sufficient number of devices have participated [45]

3) Data Heterogeneity

- Non-IID data distribution across devices can lead to slower convergence and reduced model accuracy
- Accuracy of a federated learning model can degrade by up to 55% when trained on non-IID data compared to IID data [46]
- Adaptive algorithms that account for different data types are being investigated to improve generalization and convergence
- Clustered federated learning method groups devices with similar data distributions and performs local training within each cluster, improving accuracy by 10% compared to standard federated learning on non-IID data [47]

B. Future Directions

1) Model Compression and Efficient Communication Protocols

- A sparse ternary compression scheme can reduce communication costs by 28 times while maintaining model accuracy within 1% of the uncompressed model [48]

2) Secure aggregation methods and Differential Privacy Mechanisms

- Differentially private federated learning framework protects user privacy by adding noise to model updates, achieving an ε-differential privacy guarantee of ε = 1 with almost no effect on model accuracy [49]

3) Transfer Learning and Meta-learning

- Techniques are being investigated to improve the performance of federated learning on non-IID data.
- The federated transfer learning framework uses pre-trained models to improve federated learning performance on non-IID data, leading to a 5% increase in accuracy compared to regular federated learning [51]

4) Integration with other Privacy-preserving Techniques

- Research is needed on how federated learning can be used with other privacy-protecting techniques like secure multi-party computation and homomorphic encryption [52]

5) Application to Domains beyond Mobile Devices

- Federated learning in healthcare, finance, and IoT presents new challenges and opportunities for future research [53]

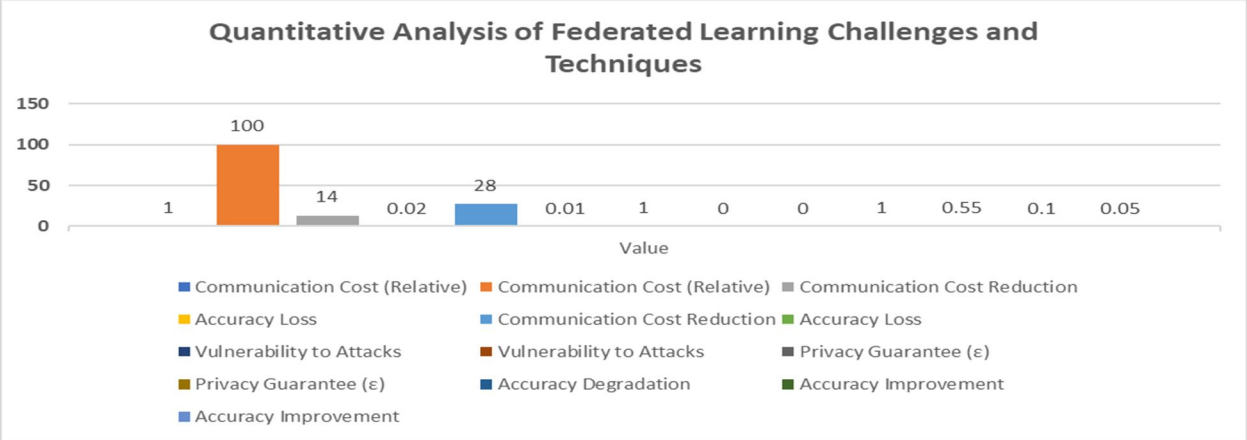


Fig. 2: Evaluating Federated Learning Performance: Challenges, Techniques, and Metrics [42 - 47]

V. CONCLUSION

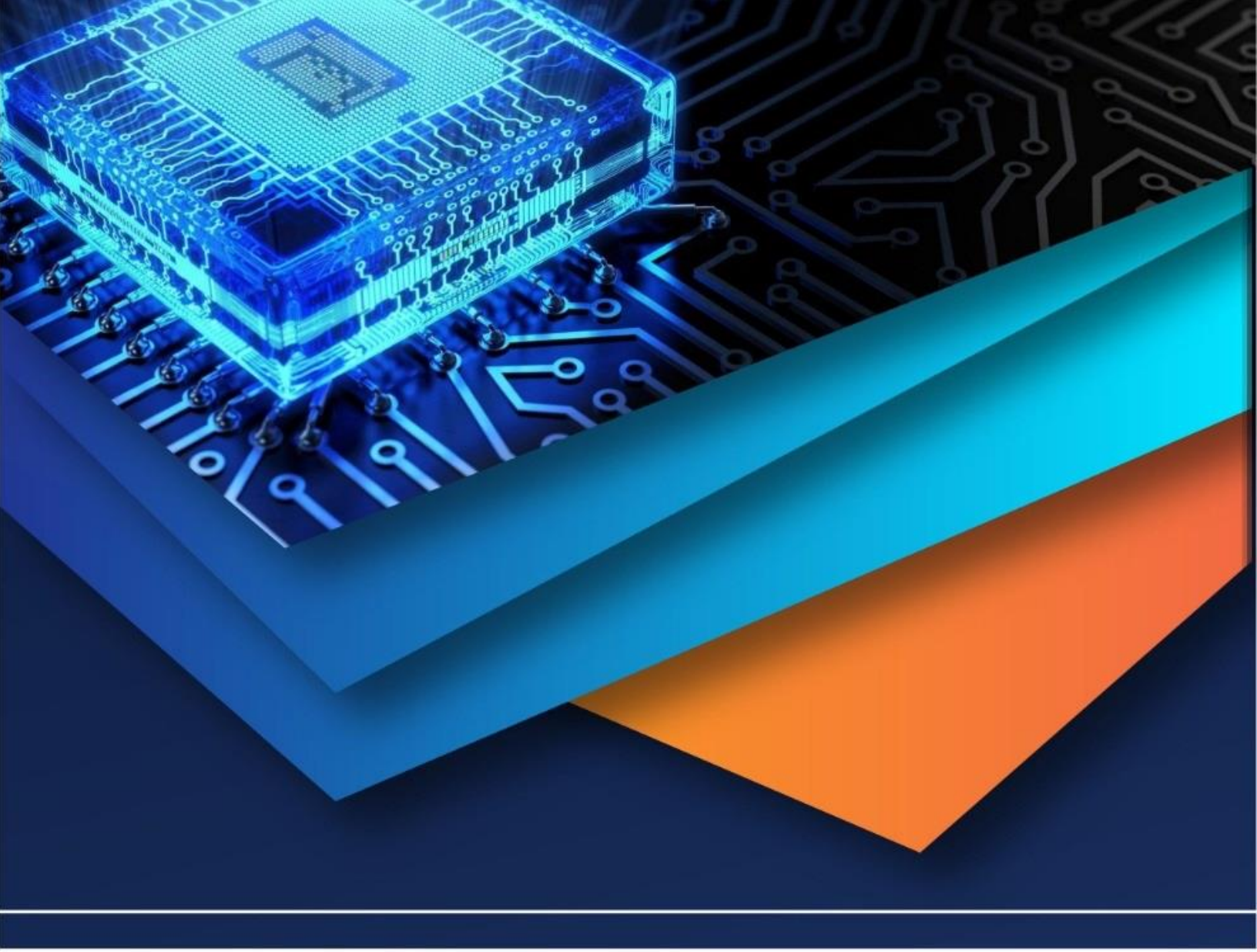
In conclusion, federated learning has emerged as a transformative approach to privacy-centric AI, enabling the development of powerful machine learning models while preserving data privacy. By leveraging decentralized data and aggregating model updates, federated learning has demonstrated significant potential across various domains, including healthcare, finance, mobile applications, IoT, transportation, and retail. Real-world case studies and performance metrics highlight the effectiveness of federated learning in achieving comparable or even superior results to centralized training while ensuring data confidentiality. However, challenges such as communication efficiency, security, robustness, and data heterogeneity need to be addressed to fully realize the potential of federated learning.

Ongoing research efforts focus on developing efficient communication protocols, secure aggregation mechanisms, differential privacy techniques, and adaptive algorithms to mitigate these challenges. As federated learning continues to evolve, its integration with other privacy-preserving techniques and its application to diverse domains beyond mobile devices present exciting opportunities for future research and innovation.

REFERENCES

[1] D. Reinsel, J. Gantz, and J. Rydning, "The digitization of the world from edge to core," IDC White Paper, 2018.
[2] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," Mobile Networks and Applications, vol. 19, no. 2, pp. 171-209, 2014.
[3] J. Lane, V. Stodden, S. Bender, and H. Nissenbaum, "Privacy, big data, and the public good: Frameworks for engagement," Cambridge University Press, 2014.
[4] U.S. Government Accountability Office, "Data protection: Actions taken by Equifax and federal agencies in response to the 2017 breach," GAO-18-559, 2018.
[5] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, 2016.
[6] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial Intelligence and Statistics, 2017, pp. 1273-1282
[7] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 2, pp. 1-19, 2019.
[8] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, 2020.
[9] MarketsandMarkets, "Federated learning market by application, vertical, and region - Global forecast to 2028," Report, 2021.
[10] W. Y. B. Lim et al., "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 2031-2063, 2020.
[11] W. Li et al., "Privacy-preserving federated brain tumour segmentation," in International Workshop on Machine Learning in Medical Imaging, 2019, pp. 133-141.
[12] W. Y. B. Lim et al., "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 2031-2063, 2020.
[13] I. Sim et al., "Mobile devices and health," The New England Journal of Medicine, vol. 381, no. 10, pp. 956-968, 2019.
[14] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, 2016.
[15] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial Intelligence and Statistics, 2017, pp. 1273-1282.
[16] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, and F. Beaufays, "Applied federated learning: Improving google keyboard query suggestions," arXiv preprint arXiv:1812.02903, 2018.
[17] K. Bonawitz et al., "Towards federated learning at scale: System design," arXiv preprint arXiv:1902.01046, 2019.
[18] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 2, pp. 1-19, 2019.
[19] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial Intelligence and Statistics, 2017, pp. 1273-1282.
[20] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," arXiv preprint arXiv:1610.02527, 2016.
[21] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," arXiv preprint arXiv:1812.06127, 2018.
[22] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," arXiv preprint arXiv:1806.00582, 2018.
[23] C. Fang, Y. Guo, N. Wang, and A. Ju, "Highly efficient federated learning with strong privacy preservation in cloud computing," Computers & Security, vol. 96, p. 101889, 2020.
[24] P. Mohassel and Y. Zhang, "SecureML: A system for scalable privacy-preserving machine learning," in 2017 IEEE Symposium on Security and Privacy (SP), 2017, pp. 19-38.
[25] R. C. Geyer, T. Klein, and M. Nabi, "Differentially private federated learning: A client level perspective," arXiv preprint arXiv:1712.07557, 2017.
[26] P. Mohassel and P. Rindal, "ABY3: A mixed protocol framework for machine learning," in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018, pp. 35-52.
[27] K. Wei et al., "Federated learning with differential privacy: Algorithms and performance analysis," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3454-3469, 2020.
[28] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," arXiv preprint arXiv:1806.00582, 2018.
[29] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-iid data," IEEE transactions on neural networks and learning systems, vol. 31, no. 9, pp. 3400-3413, 2019.
[30] Y. Liu, J. Peng, J. Kang, A. M. Iliyasu, D. Niyato, and A. A. A. El-Latif, "A secure federated learning framework with ternary gradient aggregation in support of IoT applications," IEEE Internet of Things Journal, 2020.
[31] A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning: A meta-learning approach," arXiv preprint arXiv:2002.07948, 2020.
[32] M. J. Sheller et al., "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," Scientific Reports, vol. 10, no. 1, pp. 1-12, 2020.
[33] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Federated learning," Synthesis Lectures on Artificial Intelligence and Machine Learning, vol. 13, no. 3, pp. 1-207, 2019.

[34] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated learning for mobile keyboard prediction," arXiv preprint arXiv:1811.03604, 2018.
[35] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, and F. Beaufays, "Applied federated learning: Improving google keyboard query suggestions," arXiv preprint arXiv:1812.02903, 2018.
[36] A. Hard, K. Rao, R. Mathews, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated learning for mobile keyboard prediction," arXiv preprint arXiv:1811.03604, 2018.
[37] W. Y. B. Lim et al., "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 2031-2063, 2020.
[38] D. T. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A. Sadeghi, "DÏoT: A federated self-learning anomaly detection system for IoT," in 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 756-767.
[39] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4177-4186, 2020.
[40] Q. Ye, Y. Zhou, Z. Wang, B. Jiang, and J. Xu, "Federated learning for traffic flow prediction with cross-region data sharing," in 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020, pp. 1-6.
[41] S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated learning for wireless communications: Motivation, opportunities, and challenges," IEEE Communications Magazine, vol. 58, no. 6, pp. 46-51, 2020.
[42] J. Chen, X. Pan, R. Monga, S. Bengio, and R. Jozefowicz, "Revisiting distributed synchronous SGD," arXiv preprint arXiv:1604.00981, 2016.
[43] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, 2016.
[44] S. Caldas, J. Konečný, H. B. McMahan, and A. Talwalkar, "Expanding the reach of federated learning by reducing client resource requirements," arXiv preprint arXiv:1812.07210, 2018.
[45] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning," arXiv preprint arXiv:1807.00459, 2018.
[46] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175-1191.
[47] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," arXiv preprint arXiv:1806.00582, 2018.
[48] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-iid data," IEEE transactions on neural networks and learning systems, vol. 31, no. 9, pp. 3400-3413, 2019.
[49] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Sparse binary compression: Towards distributed deep learning with minimal communication," arXiv preprint arXiv:1805.08768, 2018.
[50] S. Truex, L. Liu, M. E. Gursoy, L. Yu, and W. Wei, "Towards demystifying membership inference attacks," arXiv preprint arXiv:1807.09173, 2018
[51] Y. Liu, Y. Kang, C. Xing, T. Chen, and Q. Yang, "A secure federated transfer learning framework," IEEE Intelligent Systems, vol. 35, no. 4, pp. 70-82, 2020.
[52] J. Chen, X. Pan, R. Monga, S. Bengio, and R. Jozefowicz, "Revisiting distributed synchronous SGD," arXiv preprint arXiv:1604.00981, 2016
[53] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 2, pp. 1-19, 2019.
[54] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, 2020.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY