Federated Learning in Healthcare is the Future, But the Problems Are Contemporary

Mustafa Y. Topaloglu[1], Elisabeth M. Morrell[1] and Umit Topaloglu[2]a
[1]*Department of Computer Science, Wake Forest University, Winston Salem, NC, 27109, U.S.A.*
[2]*Departments of Cancer Biology and Biostatistics, Wake Forest School of Medicine, Winston Salem, NC, 27157, U.S.A.*

Keywords: Federated Learning, Machine Learning, Privacy, Machine Learning Ethics, Machine Learning Legal Issues.

Abstract: Federated Learning (FL) has originated out of a need to mitigate certain inherent limitations of ML, particularly the capability to train on larger datasets for improved performance, which is typically an unwieldy coordination for an inter-institutional collaboration due to existing patient protection laws and regulations. FL may also play a crucial role in bypassing ML's innate algorithmic discrimination issues via the access of underrepresented groups' data spanning across geographically distributed institutions and the diverse populations. FL inherits many of the difficulties of ML and as such we have discussed two pressing FL challenges, namely: privacy of the model exchange as well as equity and contribution considerations.

1 INTRODUCTION

Machine Learning (ML) is poised to provide an incomparable opportunity to overcome the traditional paradigms of the healthcare (Griffin et al., 2020; Topol, 2019). However, data availability and underrepresentation of minorities in healthcare datasets are traditionally accepted disadvantages to ML research(Obermeyer et al., 2019) and lead to relatively low performance for disproportionately represented ethnic and minority groups due to bias that the model might develop(Gao & Cui, 2020). Correspondingly, the training data from these populations result in distribution discrepancies that are highly susceptible to biases. Problems that arise from data heterogeneity, depth, and breadth are a hindrance to the generalization of ML approaches. Given the data intensive nature of model training, Federated Learning (FL) approaches may provide a novel opportunity for the future of ML applications(Rajendran et al., 2021; Sarma et al., 2021). FL is a collaborative ML training approach illustrated in Figure 1.

FL has recently received a greater emphasis in recent years due to its privacy preserving potential in healthcare despite certain structural issues which necessitate an address. Characteristic of many recent advancements, the Friedman curve indicates that technological advancement has overtaken present human governing capacity, and the only way to bridge the gap in the case of FL is through the introduction of rapid problem identification and prudent regulation of FL and cooperation from the public and private sector respectively(Friedman, 2016).

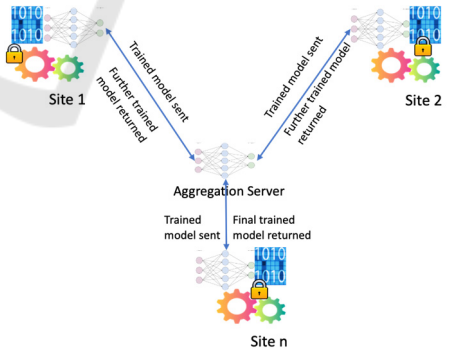

Figure 1: Federated Learning Overview.

Protected Health Information (PHI) is covered by the Health Insurance Portability and Accountability Act (HIPAA). Due to existing risk for FL models as shown by attacks, it would be appropriate to classify the model under the definition of PHI, as there is a

a https://orcid.org/0000-0002-3241-8773

Topaloglu, M., Morrell, E. and Topaloglu, U.
Federated Learning in Healthcare is the Future, But the Problems Are Contemporary.
DOI: 10.5220/0010722600003058
In *Proceedings of the 17th International Conference on Web Information Systems and Technologies (WEBIST 2021)*, pages 593-598
ISBN: 978-989-758-536-4; ISSN: 2184-3252

risk of the model itself containing the PHI of patients. This classification would be beneficial to the overall security of FL on account of the HIPAA Security Rule which outlines specific guidelines for the utilization, employment, and protection of PHI.

2 CHALLENGES

Among the issues, training data variations due to data types and their capture quality make data preparation salient to the success of the endeavor. Some of the most relevant clinical information may not be accessible or be recorded incorrectly in a way that is not representative of the studied population or missing. Moreover, data quality challenges in healthcare are an acknowledged barrier to research in general and ML in particular, however it is not within the scope of this work. Despite the perceived and studied benefits, there are some challenges for wider implementation and acceptance of FL that can be categorized in two: privacy and equity and contribution considerations.

2.1 Privacy

The purpose of Federated Learning is to train machine learning models while preserving the privacy of individual contributors. However, the models are capable of unintentionally revealing sensitive information, therefore the security and privacy of FL has become an area of extensive research (Beaulieu-Jones et al., 2018; Duan et al., 2020; Hitaj et al., 2017; Li et al., 2019). The main methods of protecting training data in the FL process are differential privacy, model encryption, blockchain based computing, and homomorphic encryption.

Differential privacy is performed by randomly perturbing the parameters of the local model with noise (e.g., Gaussian noise, Laplacian noise) before communicating with and incorporating into the global model (Li et al., 2020). Model encryption is accomplished by encrypting the parameters of the global model before they are sent to local data collection for training. Local models are communicated back to the global model with encrypted local gradients (Lu, 2021). Homomorphic encryption is implemented by computing on encrypted models (Kim et al., 2018).

Despite the improvements in security, FL privacy methods continue to prove vulnerable to attacks. These breaches and data leaks fit into two main categories: inference during the learning process, and inference over the output(Truex et al., 2019). In general, the more overfitted the model is to its training data, the more vulnerable it is to an attack(Shokri et al., 2017).

While troubleshooting these flaws in the FL process, new research shows that a variety of privacy features can prove a suitable defense (Beaulieu-Jones et al., 2018; Li et al., 2020; Li et al., 2019; Shokri & Shmatikov, 2015; Truex et al., 2019; Wei et al., 2020).

In some cases, research will stray from the centralized FL system and perform all model training decentralized as an additional security measure. For example, Swarm Learning (SL) builds ML models from local data. These models never leave the host site and instead the learning parameters are shared via blockchain with other local models. After the arrival of new parameters, the local models train with the new parameters until standards for synchronization are met. Each contributing site has locally installed nodes which are responsible for building and synchronizing models and implementing the blockchain (Warnat-Herresthal et al., 2021).

Blockchain provides a secure distributed ledger-based computing framework that has started to show promise in healthcare(Norgeot et al., 2019). Several studies proposed a Blockchain implementation for ML models to benefit security and privacy promised by the Blockchain(Hathaliya et al., 2019; Vyas et al., 2019). Despite the consensus of Blockchain being very secure, its security level is directly correlated with the hashing power an implementation may have. Some of the known attacks on Blockchain include Finney attack, race attack, 51% attack (i.e., majority attack), eclipse attack, Sybil attack, routing attack, Decentralized Autonomous Organization (DAO) attack. Other than the 51% attack, the aforementioned attacks depend on the implementation and may not be very common. In spite of the numerous vulnerabilities blockchain implementations produce, methods to counter and mitigate the risks posed rely on computationally expensive measures, which entails an honest dialogue about the fruitfulness of these models (Aggarwal & Kumar, 2021).

2.2 Equity and Contribution Valuation

A tradeoff for collaborating institutions exists between the privacy of the data and the joint effort of model development(Rieke et al., 2020). For example, since the training data will be decentralized and if there are two institutions developing a model, it would be beneficial to have them coordinate on such tasks as taking measurements. Potential solutions

could include having a restricted access for limited amounts of coordinators in order to compromise, or to simply decline sharing any information not explicitly related to the model. Deciding between lacking the ability to investigate the data or possible privacy issues is a difficult dilemma. Due to the unique and distributed nature of FL, it would be advantageous to have a standard evaluation strategy for purposes such as determining remuneration. The Gini coefficient ("Gini Index," 2008), which determines income inequality could be employed to assess the contribution (figure 2). Even though the Information Gain Function could be considered as an alternative, Raileanu et al. have proved that the Gini coefficient only disagrees by 2% with Information Gain in all cases(Raileanu & Stoffel, 2004). In a similar vein, we utilize two factors to determine the level of participation: model development contribution, and data contribution; formula (1) is for the total contribution and (2) can be used to calculate an individual site's contribution.

Mi= Model development contribution level of *ith* site: Mi ε $\mathbb{R}$ = [0,1]

Di= Data contribution level of *ith* site: Di ε $\mathbb{R}$ = [0,1]

$$Total\ Contribution = \frac{\sum_{i=0}^{n}(M_i + D_i)}{2 * N} \quad (1)$$

$$Contribution\ of\ ith\ node = \frac{M_i + D_i}{\sum M + \sum D} \quad (2)$$

N= Number of participants in a FL framework

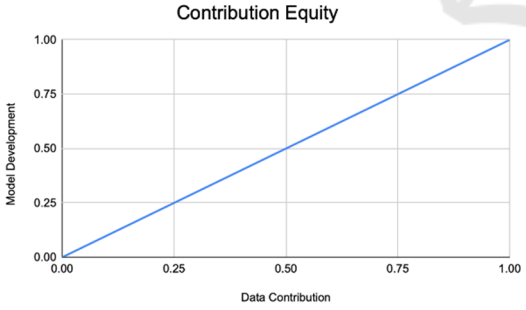

Figure 2: Contribution Equity representation for data and model based on Gini Index.

2.2.1 Model Development Expertise (*Mi*)

Model development requires understanding the problem domain with adequate subject matter expertise and ability to translate that knowledge into ML models. In classical ML algorithms, this involved extracting features and designing a ML task (e.g., classifier). With Deep Learning, the initial feature engineering can be done by the algorithms. Model development can be broken into the three following subcategories:

Model Development Cost: The Cost here refers to such efforts as the labor behind the creation of the algorithm until initial operating capacity and implementation occur. This is a multifaceted component, which will be able to leverage a certain institution's healthcare informatics and data science expertise if they lack significant Data Contribution.

Model Validation and Benchmarking Cost: Without proper validation and benchmarking, the algorithms cannot be integrated into clinical care[49]. Over the years, the machine learning community has developed several statistical methods to properly evaluate AI algorithms. The "intrinsic uncertainty" in medicine introduces variations in result interpretation, which suggest that model performance criteria should be use case specific vs. using standard scoring metrics[50].

Continuous Model Improvement Cost: The Model itself will see countless iterations and frequent evolution to accommodate new aspects and features. The work put into the model after deployment will be factored in accordingly to reward constant improvements and reflect the reality of Model Development.

3 CONCLUSION

Data and model privacy is essential for any FL implementation in healthcare in order to realize its potential. We have discussed current privacy challenges and corresponding proposals to address those deficiencies. We believe that none of the proposed solutions have sufficient safeguards that is practical to implement. Therefore, further studies and solutions are needed for FL to strive.

Another identified challenge is the contribution assessment and corresponding profit (or responsibility) sharing among the FL participating institutions. Unfortunately, there is no widely accepted models for such collaboration. We have proposed a conceptual model that relies on the Gini coefficient. The model considers the model development attributes that need to be taken into account along with the data that each institutions contributes. There are some proposed models for data contribution but not for the model and data, to our knowledge.

Upon addressing these challenges, we strongly believe that FL will be widely accepted and contribute to the biomedical advancements.

ACKNOWLEDGEMENTS

We would like to thank Drs. Can Bora Unal and Todd Morrell for their invaluable feedback. The work is partially supported by the Cancer Center Support Grant from the National Cancer Institute to the Comprehensive Cancer Center of Wake Forest Baptist Medical Center (P30 CA012197). The authors also acknowledge use of the services and facilities, funded by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (UL1TR001420).

REFERENCES

Aggarwal, S., & Kumar, N. (2021). Chapter Twenty - Attacks on blockchain☆☆Working model. In S. Aggarwal, N. Kumar, & P. Raj (Eds.), *Advances in Computers* (Vol. 121, pp. 399-410). Elsevier. https://doi.org/https://doi.org/10.1016/bs.adcom.2020.08.020

Allen, B., Agarwal, S., Kalpathy-Cramer, J., & Dreyer, K. (2019). Democratizing AI. *Journal of the American College of Radiology*, *16*(7), 961-963. https://doi.org/10.1016/j.jacr.2019.04.023

Azoulay, A. *Towards an Ethics of Artificial Intelligence*.

Beaulieu-Jones, B. K., Yuan, W., Finlayson, S. G., & Wu, Z. S. (2018). Privacy-Preserving Distributed Deep Learning for Clinical Data. *arXiv e-prints*. Retrieved December 01, 2018, from https://ui.adsabs.harvard.edu/abs/2018arXiv181201484B

Bujalkova, M. (2001). Hippocrates and his principles of medical ethics. *Bratisl Lek Listy*, *102*(2), 117-120.

Caribbean, R. o. A. I. T. A. a. t. *https://www.loc.gov/law/help/artificial-intelligence/americas.php#us*. Retrieved February 2021 from

Char, D. S., Abràmoff, M. D., & Feudtner, C. (2020). Identifying Ethical Considerations for Machine Learning Healthcare Applications. *Am J Bioeth*, *20*(11), 7-17. https://doi.org/10.1080/15265161.2020.1819469

Cheng, K., Fan, T., Jin, Y., Liu, Y., Chen, T., & Yang, Q. (2019). SecureBoost: A Lossless Federated Learning Framework. *ArXiv*, *abs/1901.08755*.

Cohen Healthcare. *Artificial Intelligence and the Health Insurance Portability and Accountability Act of 1996 (HIPAA),*. Retrieved February, 2021 from

Duan, R., Boland, M. R., Liu, Z., Liu, Y., Chang, H. H., Xu, H., Chen, Y. (2020). Learning from electronic health records across multiple sites: A communication-efficient and privacy-preserving distributed algorithm. *J Am Med Inform Assoc*, *27*(3), 376-385. https://doi.org/10.1093/jamia/ocz199

Enterprise, T. C. f. O. D. (2019). *Sharing and Utilizing Health Data for AI Applications*.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118. https://doi.org/10.1038/nature21056

Evans, B., & Ossorio, P. (2018). The Challenge of Regulating Clinical Decision Support Software After 21(st) Century Cures. *Am J Law Med*, *44*(2-3), 237-251. https://doi.org/10.1177/0098858818789418

FDA. *Digital Health Software Precertification (Pre-Cert) Program*. Retrieved February, 2021 from

Friedman, T. (2016). *Thank You for Being Late: an Optimist's Guide to Thriving in the Age of Accelerations.* Farrar, Straus and Giroux.

Gao, Y., & Cui, Y. (2020). Deep transfer learning for reducing health care disparities arising from biomedical data inequality. *Nature Communications*, *11*(1), 5131. https://doi.org/10.1038/s41467-020-18918-3

Geis, J. R., Brady, A. P., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., Kohli, M. (2019). Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. *Journal of the American College of Radiology*, *16*(11), 1516-1521. https://doi.org/10.1016/j.jacr.2019.07.028

Gini Index. (2008). In *The Concise Encyclopedia of Statistics* (pp. 231-233). Springer New York. https://doi.org/10.1007/978-0-387-32833-1_169

Griffin, A. C., Topaloglu, U., Davis, S., & Chung, A. E. (2020). From Patient Engagement to Precision Oncology: Leveraging Informatics to Advance Cancer Care. *Yearb Med Inform*, *29*(1), 235-242. https://doi.org/10.1055/s-0040-1701983

Guidance, F. C. *https://www.fda.gov/media/109618/download*. Retrieved November from

Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Groups, l.-I. (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, *29*(8), 1836-1842. https://doi.org/10.1093/annonc/mdy166

Hallevy, P. G. (November 17, 2015). *AI v. IP - Criminal Liability for Intellectual Property IP Offenses of Artificial Intelligence AI Entities*.

Hathaliya, J., Sharma, P., Tanwar, S., & Gupta, R. (2019, 13-14 Dec. 2019). Blockchain-Based Remote Patient Monitoring in Healthcare 4.0. 2019 IEEE 9th International Conference on Advanced Computing (IACC),

Hitaj, B., Ateniese, G., & Perez-Cruz, F. (2017). *Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning* Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, Texas, USA.

Institute, F. o. L. (2021). https://futureoflife.org/lethal-autonomous-weapons-pledge/?cn-reloaded=1&cn-reloaded=1

Kim, M., Song, Y., Wang, S., Xia, Y., & Jiang, X. (2018). Secure Logistic Regression Based on Homomorphic Encryption: Design and Evaluation. *JMIR Med Inform*, *6*(2), e19. https://doi.org/10.2196/medinform.8805

Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: a systematic review of discrimination and

fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, *13*(3), 795-848. https://doi.org/10.1007/s40685-020-00134-w

Lee, S. S., Kelley, M., Cho, M. K., Kraft, S. A., James, C., Constantine, M., Magnus, D. (2016). Adrift in the Gray Zone: IRB Perspectives on Research in the Learning Health System. *AJOB Empir Bioeth*, *7*(2), 125-134. https://doi.org/10.1080/23294515.2016.1155674

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated Learning: Challenges, Methods, and Future Directions. *IEEE Signal Processing Magazine*, *37*(3), 50-60. https://doi.org/10.1109/MSP.2020.2975749

Li, W., Milletarì, F., Xu, D., Rieke, N., Hancox, J., Zhu, W., Feng, A. (2019). Privacy-preserving Federated Brain Tumour Segmentation. *arXiv e-prints*. Retrieved October 01, 2019, from https://ui.adsabs.harvard.edu/abs/2019arXiv191000962L

Lu, X. Y. a. Y. F. a. W. F. a. J. S. a. X. T. a. S.-T. X. a. R. (2021). Computation-efficient Deep Model Training for Ciphertext-based Cross-silo Federated Learning. *arXiv:2002.09843*.

Markose, A., Krishnan, R., & Ramesh, M. (2016). Medical ethics. *J Pharm Bioallied Sci*, *8*(Suppl 1), S1-S4. https://doi.org/10.4103/0975-7406.191934

Mulshine, M. A major flaw in google's algorithm allegedly tagged two black people's faces with the word 'gorillas'. In BusinessInsider, 2015.

Nicholas Carlini, C. L., Úlfar Erlingsson, Jernej Kos, Dawn Song. (2019). The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks. *arXiv:1802.08232*

Noor, P. (2020). Can we trust AI not to further embed racial bias and prejudice? *BMJ*, *368*, m363. https://doi.org/10.1136/bmj.m363

Norgeot, B., Glicksberg, B. S., & Butte, A. J. (2019). A call for deep-learning healthcare. *Nature Medicine*, *25*(1), 14-15. https://doi.org/10.1038/s41591-018-0320-3

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447-453. https://doi.org/10.1126/science.aax2342

Policy, T. W. H. O. o. S. a. T. (2018). *Summary Of The 2018 White House Summit On Artificial Intelligence For American Industry https://www.hsdl.org/?view&did=811092*.

Price, W. N. (2018). Big data and black-box medical algorithms. *Sci Transl Med*, *10*(471). https://doi.org/10.1126/scitranslmed.aao5333

Raileanu, L. E., & Stoffel, K. (2004). Theoretical Comparison between the Gini Index and Information Gain Criteria. *Annals of Mathematics and Artificial Intelligence*, *41*(1), 77-93. https://doi.org/10.1023/B:AMAI.0000018580.96245.c6

Rajendran, S., Obeid, J. S., Binol, H., D Agostino, R., Foley, K., Zhang, W., . . . Topaloglu, U. (2021). Cloud-Based Federated Learning Implementation Across Medical Centers. *JCO Clin Cancer Inform*, *5*, 1-11. https://doi.org/10.1200/CCI.20.00060

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digit Med*, *3*, 119. https://doi.org/10.1038/s41746-020-00323-1

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, *4*, 100005. https://doi.org/https://doi.org/10.1016/j.jrt.2020.100005

SaMD, F. *https://www.fda.gov/media/122535/download*. Retrieved February from

Sarma, K. V., Harmon, S., Sanford, T., Roth, H. R., Xu, Z., Tetreault, J., Arnold, C. W. (2021). Federated learning improves site performance in multicenter deep learning without data sharing. *J Am Med Inform Assoc*. https://doi.org/10.1093/jamia/ocaa341

Scientist, F. o. A. (2021). *Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems*. Retrieved May 2021 from https://fas.org/sgp/crs/natsec/IF11150.pdf

Shokri, R., & Shmatikov, V. (2015, 29 Sept.-2 Oct. 2015). Privacy-preserving deep learning. 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton),

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017, 22-26 May 2017). Membership Inference Attacks Against Machine Learning Models. 2017 IEEE Symposium on Security and Privacy (SP),

Technology, N. S. a. T. C. C. o. (October 2016). *https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf*. Retrieved February, 2021 from

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44-56. https://doi.org/10.1038/s41591-018-0300-7

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., & Zhou, Y. (2019). A Hybrid Approach to Privacy-Preserving Federated Learning. *Informatik Spektrum*, *42*(5), 356-357. https://doi.org/10.1007/s00287-019-01205-x

Vyas, S., Gupta, M., & Yadav, R. (2019, 4-6 Feb. 2019). Converging Blockchain and Machine Learning for Healthcare. 2019 Amity International Conference on Artificial Intelligence (AICAI),

Warnat-Herresthal, S., Schultze, H., Shastry, K. L., Manamohan, S., Mukherjee, S., Garg, V., . . . Deutsche, C.-O. I. (2021). Swarm Learning for decentralized and confidential clinical machine learning. *Nature*. https://doi.org/10.1038/s41586-021-03583-3

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Poor, H. V. (2020). Federated Learning With Differential Privacy: Algorithms and Performance Analysis. *IEEE Transactions on Information Forensics and Security*, *15*, 3454-3469. https://doi.org/10.1109/TIFS.2020.2988575

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Goldenberg, A. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, *25*(9), 1337-1340. https://doi.org/10.1038/s41591-019-0548-6

Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021). Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers. *ArXiv*, *abs/2105.02117*.