Federated Learning and Applications in Cybersecurity

Ani Sreekumar
William & Mary

Federated Learning and Applications in Cybersecurity

Ani Sreekumar

asreekumar@wm.edu

Coastal Virginia Commonwealth Cyber Initiative

December 15, 2022

Abstract

Machine learning is a subfield of artificial intelligence that focuses on making predictions about some outcome based on information from a dataset. In cybersecurity, machine learning is often used to improve intrusion detection systems and identify trends in data that could indicate an oncoming cyber attack. Data privacy is an extremely important aspect of cybersecurity, and there are many industries that have more demanding laws to ensure the security of user data. Due to these regulations, machine learning algorithms can not be widely utilized in these industries to improve outcomes and accuracy of predictions. However, federated learning is a recent development in the field of machine learning that allows for the training of a model using decentralized data. Federated learning is a practical solution in cases where a machine learning model needs to be trained with data from different servers, devices, or organizations and the data from one party can not be shared with the other parties. Federated learning is also a form of cybersecurity in itself, as it improves the security of machine learning models in terms of data privacy. This paper explains the concept of federated learning and its specific applications to cybersecurity, with a focus on federated learning's impact on the healthcare industry. Cyber threats to machine learning models as well as recent improvements in federated learning algorithms and their implications in the field of cybersecurity are also discussed.

Federated Learning and Applications in Cybersecurity

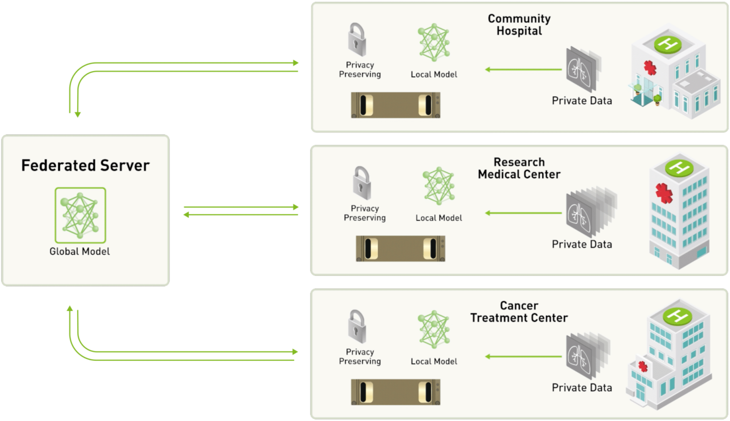
Machine Learning (ML) is a concept that falls under the broader field of Artificial Intelligence (AI) and is focused on making intelligent predictions based on an existing dataset. Applications of machine learning algorithms include prediction systems, object recognition, speech recognition, and computer vision. These applications can be utilized in practically all industries with notable ones including healthcare, defense, and finance. In the field of cybersecurity, ML techniques can be applied to create detection systems and gain insights into patterns of intrusion. These techniques are often utilized in areas of critical infrastructure like power grids and industrial control systems, as well as many other industries (Sarker, 2021).

In order to utilize an ML algorithm, existing data has to be available in order to train the ML model. In the applications of machine learning to cybersecurity, the security and privacy of large datasets is crucial given the necessity of the data to create the model. One method to improve the accuracy of a machine learning model is to train it on a more representative sample of the population that it is making predictions on. As a result, it is usually beneficial for a machine learning model to be trained on as much applicable data as possible. In many industries where machine learning is currently applied, there are concerns with data privacy. In the healthcare industry, patient data is often protected by privacy laws and the sharing of data between organizations is often illegal (Antunes, 2022). This is where the concept of federated learning (FL) can become useful, as it allows for models to be trained with data that is not in a centralized server.

Introduction to Federated Learning

Federated learning is a relatively new subfield in ML that seeks to train a model collaboratively, but without actually exchanging the data between parties. The goal is to address the issue of data privacy and the difficulties associated with training a model using data that is not centralized. A typical machine learning model necessitates the aggregation of all the data that is used to train the model to some central server. Federated learning allows parties to train the model locally using only their own data. This model update is then sent to the central server, which then utilizes all of the local models to create a new global model (Alazab, 2022).

Federated learning was first introduced by Google AI in 2016 as an alternative to distributed learning. The primary difference between federated learning and distributed learning is that only the encrypted parameters of the locally trained model are updated to the global model, rather than the training data being transmitted itself (Alazab, 2022). Both of these aspects help to protect data privacy and security, while also allowing for a greater range of applications given the privacy of the data. There are several advantages to the decentralized nature of data aggregation under federated learning. Since the data of each party is not being shared with the other parties contributing to the global model, federated learning theoretically allows for sensitive data to be used to train a model. This would allow the use of machine learning techniques in a wider range of industries where there is a particular need for stringent data privacy.



The figure above (Roth, 2022) shows an example of how federated learning can utilize data from multiple entities while keeping the data secure, and private from the other parties in the network. Each entity uses their own private data to train a local model, and then the encrypted parameters of the model are uploaded to the federated server, preserving the privacy of each entity.

Use Cases in Healthcare

Machine learning is currently utilized extensively in the healthcare industry, particularly using Electronic Health Records (EHR) data. The healthcare space is also an industry that has many laws protecting sensitive data, like the Health Insurance Portability and Accountability Act (HIPAA) and the California Consumer Privacy Act (CCPA). Under these laws, patient data is subject to certain protections of privacy and must be authorized before use (Vayena, 2018). The sensitive nature of healthcare data and the growing need for large datasets for AI training has led to the start of Data Lakes - or pooled data from multiple institutions. Many countries like the UK and France have created large databases of healthcare data in an effort to promote better data governance (Antunes, 2022). However, large data lakes pose many regulatory and technical challenges. Data privacy laws call for anonymization of sensitive data, and this is particularly difficult to do with patient data. Using an aggregation server, federated learning could allow for institutions to remain unknown to each other if they are entities in the same federated network (Rieke, 2020). A recent example of federated learning efforts in the healthcare industry is called the HealthChain project. The goal of this project is to develop a federated learning system using four different hospitals in France as the entities in the network. The global model uses data from all four hospitals and is used to predict treatment response for breast cancer patients (Rieke, 2020). Using a more diverse array of data will generally yield positive results for the predictive power of a learning algorithm, and federated learning presents a framework to securely use data from a much wider variety of sources.

Although data privacy is a point of concern for machine learning in healthcare, there are still many areas where it is being used in the industry. One application is the development of ML algorithms that help providers diagnose diseases more accurately using data from MRI studies and EHR data (K. Shailaja, 2018). Another common application is using data to predict patient outcomes based on a variety of factors. Wearable devices are also a significant area of research in the machine learning healthcare space (Sabry, 2022). Researchers have trained models on data from wearable devices in order to create a wearable fall-monitoring device for elderly people. The device is capable of identifying the pattern of activity prior to a fall with 99% accuracy (Hussain, 2019). However, the accuracy of this model drops when tested on a different set of data. This is a common problem with machine learning models, especially in healthcare. Federated learning would allow for models like this to be trained on a much larger quantity of data while maintaining data privacy standards.

Potential Security Concerns

As with all systems and models, federated learning models are certainly vulnerable to certain security risks. Although the data from each party that is used to train a global model is private and not accessible by the other parties, in most cases it is still sent to the central server at some point for aggregation into the global model. Thus, this data could be compromised by a cyber attack on the central server (Kairouz, 2019). This is not a unique risk to federated learning, as this issue certainly exists in centralized machine learning models as well. In this case, the advantage of federated learning would be that the risk of such an attack is mitigated because the data is not permanently stored in the central server. This is currently a highly researched topic within the field of federated learning, and much of existing research focuses on hiding the local model updates from the server, while still using them to update the global model (Nguyen, 2022).

Another security concern involved in ML and FL is model poisoning. Model poisoning is a type of cyberattack on machine learning systems where the goal is to change or manipulate the training data in order to negatively affect the accuracy of the model. This is also a very highly researched problem within ML and FL and remains one of the most significant security risks to federated learning. A recent study proposed a model poisoning attack in which the attacker would add fake local model updates to the central server in order to lower the accuracy of the global model (Cao, 2022). The attacker is then able to increase the influence of their fake local models on the global model, and significantly decrease the accuracy of the global model. An example of model poisoning is the attacks against Gmail's spam filters, where Google reported "at least four malicious large-scale attempts to skew our classifier." (Constantin, 2021). Other techniques for model poisoning include introducing random noise into the dataset, or altering the data in some way to skew the relationship between variables. Another advantage that federated learning has over machine learning is that it allows for federated averaging, which can help to reduce the impact of poisoned data in the global model. Federated averaging uses a weighted average of the model updates from each party, as opposed to the average of all of the updates (Bietti, 2022). Federated learning is inherently less vulnerable to model poisoning than machine learning because the data is less centralized. However, as shown in the study mentioned previously there are certainly still methods that hackers may utilize to poison a federated learning model.

Algorithms

Machine learning has introduced many new algorithms to improve predictions like KNN, Random Forest, SVN, Naive Bayes, etc (Dasgupta, 2022). One of the most significant optimization algorithms that federated learning has introduced is called federated stochastic gradient descent (FSGD). This algorithm is a slight variation of ordinary stochastic gradient descent that has been adapted for use in a decentralized setting. One of the primary benefits of this algorithm is that it increases the computational efficiency of the federated learning system (Bietti, 2022). It also helps to reduce the amount of communication between entities in the network and helps to improve the scalability of federated learning systems (Yuan, 2020).

Applications of FL and ML in Cybersecurity

Data Privacy and Security

One of the primary advantages to federated learning from a cybersecurity perspective is data privacy. Federated learning generally improves the data privacy of training data in all areas where machine learning is currently utilized. Centralized data is more vulnerable to all kinds of cyber attacks, and federated learning's decentralized nature helps to reduce this risk. Federated learning also opens doors to industries and use cases that traditional machine learning techniques can not currently be applied to because of data privacy requirements. Instead of sending raw data to some centralized server, federated learning allows for each entity in the network to send the parameters of their local model to the global model, helping to maintain data privacy.

Data security is also an aspect that federated learning improves on. Because the raw data is typically not communicated over the network, attacks that rely on manipulation of the actual data are minimized. Denial of service attacks and SQL injections are common examples of cyber attacks that would be much more difficult to implement in a federated learning system. These types of attacks are more common threats when using centralized data for a machine learning algorithm.

Intrusion Detection

Traditional machine learning techniques have been utilized to identify patterns in intrusions, which helps to identify malicious actors in the network. Thus, attacks that have predictable patterns can be identified using an intrusion detection system. Data regarding patterns of cyber intrusion is very sensitive, and organizations are often hesitant to share this data due to fear of attackers or bad actors gaining access to it. Federated learning helps to improve the accuracy of intrusion detection systems, as it would allow researchers to train their models on larger datasets that are more representative of all patterns of cyber intrusions. This would help to increase the accuracy of such models and improve intrusion detection systems.

A notable use case of federated learning for intrusion detection systems is critical infrastructure like power grids. Machine learning techniques are utilized to determine the best power utilization strategy, but the data that is used to inform these algorithms is often owned and stored by different parties who do not necessarily have interests in sharing data with one another. In fact, sharing data in an industry that is so critical to national infrastructure may even compromise national security. A paper written by Liu and Zhang outlines a federated learning framework that would enable collaborative learning of both power consumption and intrusion detection patterns without leaking data about individual entities. According to the authors, the models constructed from the framework would be privacy-preserving with proper encryption schemes. Intrusion detection systems are widely utilized in all industries, and the application of federated learning to these systems would allow for greater model accuracy and preservation of data privacy (Liu, 2021).

Conclusion

Federated learning presents a powerful alternative to traditional machine learning techniques and further ensures the data privacy of the entities in the network. The decentralized nature of the data aggregation in federated learning systems allows for machine learning algorithms to be applied to use cases in healthcare and defense. This also facilitates use cases within other industries that may not have been able to use these technologies due to privacy concerns. Additionally, the greater availability of datasets to use for training these algorithms will ultimately lead to more accurate models, and thus better patient outcomes in healthcare as well as better intrusion detection systems in cybersecurity. Federated learning has already had a large impact on many industries in just the last six years, and research on the topic will continue to expand. There are still questions about certain frameworks in federated learning, especially regarding data heterogeneity and privacy. Although federated learning certainly has its own security risks and points of vulnerability, the infrastructure and design of federated systems tends to prove more difficult for cyber criminals to penetrate. Ultimately this technique serves as a promising framework for applying machine learning algorithms to more widespread use cases in a more effective, efficient, and secure manner.

References

Alazab, RM, S. P., M, P., Maddikunta, P. K. R., Gadekallu, T. R., & Pham, Q.-V. (2022). Federated Learning for Cybersecurity: Concepts, Challenges, and Future Directions. IEEE Transactions on Industrial Informatics, 18(5), 3501–3509. https://doi.org/10.1109/TII.2021.3119038

Antunes, André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated Learning for Healthcare: Systematic Review and Architecture Proposal. ACM Transactions on Intelligent Systems and Technology, 13(4), 1–23. https://doi.org/10.1145/3501813

Bietti, Wei, C.-Y., Dudík, M., Langford, J., & Wu, Z. S. (2022). Personalization Improves Privacy-Accuracy Tradeoffs in Federated Learning. https://doi.org/10.48550/arxiv.2202.05318

Cao, & Gong, N. Z. (2022). MPAF: Model Poisoning Attacks to Federated Learning based on Fake Clients. arXiv.org.

Constantin. (2021). How data poisoning attacks corrupt machine learning models. CSO (Online).

Dasgupta, Akhtar, Z., & Sen, S. (2022). Machine learning in cybersecurity: a comprehensive survey. Journal of Defense Modeling and Simulation, 19(1), 57–106. https://doi.org/10.1177/1548512920951275

Hussain, Hussain, F., Ehatisham-ul-Haq, M., & Azam, M. A. (2019). Activity-Aware Fall Detection and Recognition Based on Wearable Sensors. IEEE Sensors Journal, 19(12), 4528–4536. https://doi.org/10.1109/JSEN.2019.2898891

K. Shailaja, B. Seetharamulu and M. A. Jabbar (2018). Machine Learning in Healthcare: A Review. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 910-914, doi: 10.1109/ICECA.2018.8474918.

Kairouz, McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., Rouayheb, S. E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., … He, L. (2019). Advances and Open Problems in Federated Learning. https://doi.org/10.48550/arxiv.1912.04977

Liu, Zhang, X., Shen, X., & Sun, H. (2021). A Federated Learning Framework for Smart Grids: Securing Power Traces in Collaborative Learning. https://doi.org/10.48550/arxiv.2103.11870

Nguyen, & Thai, M. T. (2022). Preserving Privacy and Security in Federated Learning. https://doi.org/10.48550/arxiv.2202.03402

Rieke, Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3(1), 119–119. https://doi.org/10.1038/s41746-020-00323-1

Roth, Cheng, Y., Wen, Y., Yang, I., Xu, Z., Hsieh, Y.-T., Kersten, K., Harouni, A., Zhao, C., Lu, K., Zhang, Z., Li, W., Myronenko, A., Yang, D., Yang, S., Rieke, N., Quraini, A., Chen, C., Xu, D., … Feng, A. (2022). NVIDIA FLARE: Federated Learning from Simulation to Real-World. https://doi.org/10.48550/arxiv.2210.13291

Sabry, Eltaras, T., Labda, W., Alzoubi, K., & Malluhi, Q. (2022). Machine Learning for Healthcare Wearable Devices: The Big Picture. Journal of Healthcare Engineering, 2022, 4653923–4653925. https://doi.org/10.1155/2022/4653923

Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x

Vayena, Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. PLoS Medicine, 15(11), e1002689–e1002689. https://doi.org/10.1371/journal.pmed.1002689

Yuan, & Ma, T. (2020). Federated Accelerated Stochastic Gradient Descent. https://doi.org/10.48550/arxiv.2006.08950