Federated Machine Learning in Anti-Financial Crime Processes

Frequently Asked Questions

DECEMBER 2020

FinRegLab's investigation into the use of artificial intelligence (AI) in financial services provides the basis for our AI FAQs. This resource is designed to provide financial services stakeholders with accessible information about the technological, market, and policy issues related to the use of these advanced analytical techniques in the service of a vibrant and inclusive financial marketplace. This edition of our AI FAQs focuses on the potential for using federated machine learning to improve the efficiency and inclusiveness of processes to monitor and prevent illicit financial activity. These FAQs have been created to support FinRegLab's Research Proposal to the Central Bank of the Future Conference that was hosted by the Federal Reserve Bank of San Francisco and the University of Michigan's Center on Finance, Law & Policy.[1]

At its core, federated learning inverts how we normally assume machine learning models work, and that inverted architecture can help users navigate requirements related to data privacy, security, and protection, as well as operational constraints, that make data sharing unusually difficult, risky, or costly. Instead of aggregating a large volume of data on which to train an algorithm in a central location, federated learning sends the learning algorithm to the data. Repeating this process across each participating institution enables a networked, hub-and-spoke system for developing a model that reflects insights from a larger and more varied data set than any one participant in the system can access on its own. The coordinating hub will then share back to participants model weights and parameters from that model that reflect those broader insights. In effect, federated machine learning systems have the potential to improve current efforts to identify illicit financial activity by enabling shared learning *without* sharing data.

In the context of financial crimes compliance, this aspect of federated learning poses an elegant and potentially transformative technological solution to problems that have limited our ability to prevent misuse of the financial system. In the aftermath of the terrorist attacks on September 11th,

FinRegLab is a nonprofit research organization founded on the premise that independent, rigorous research is a primary ingredient in developing market norms and policy solutions that will enable responsible innovation and a more inclusive financial system.

the U.S. government expanded the anti-money laundering (AML) and sanctions regime under the Bank Secrecy Act (BSA). These efforts heightened obligations on firms to identify and report activities that may involve illicit financial activity. To support more proactive identification of complex and quickly evolving risks across the marketplace, Section 314(b) of the Patriot Act contemplated information sharing among firms in cases of suspected laundering or terror financing.[2] However, in practice little information sharing occurs under this provision due primarily to concerns related to privacy, information security, and competition, as well as legal uncertainty about this provision's application.[3]

Federated learning might be used in a number of ways to address concerns about financial crimes compliance. It might be applied to different anti-financial crime (AFC) processes,[4] such as customer onboarding or transaction monitoring. Federated learning can also be built using different configurations of participants. For example, the federation might encompass firms and government actors within a single jurisdiction or across multiple jurisdictions. Similarly, a vendor could play the role of hub, or that role may be taken up by a central bank or law enforcement agency.

For this document and in keeping with our proposal to the Central Bank of the Future conference, FinRegLab assumes that the hub role will be played by a central bank that as part of its mandate promotes responsible adoption of advanced technologies and actively supports financial inclusion. Taking on this role can give central banks a pivotal lever to achieve two important goals:

- » Enhancing confidence on the part of investors, firms, and international bodies in the ability of domestic financial institutions and their customers to participate safely and responsibly in the global financial system.
- » Improving domestic access to the international financial system, especially in areas critical for growth and development like foreign direct investment, remittance transfers, and other money services businesses.

Federated machine learning has the potential to enhance stakeholders' confidence in our collective ability to identify correctly individuals and transactions related to illicit financial activity. This improvement has the potential to translate into process changes and cost savings for financial institutions and their regulators, as well as encourage the creation of a more vibrant and inclusive financial system.

These FAQs are designed to explain and enrich the concepts presented in our research proposal. The following questions are answered in this edition:

- » **What is federated machine learning? How is it different from other forms of machine learning?**
- » **How are federated models developed?**
- » **What are the key features of current approaches to financial crimes compliance?**
- » **What are the potential benefits of using federated machine learning to identify risks related to financial crimes?**

» **What are the potential confidentiality and privacy implications of using federated machine learning in anti-financial crime processes?**

» **What is the basis for believing that using federated machine learning to identify illicit financial activity has the potential to expand access to the financial system?**

» **What aspects of using federated learning to improve financial crimes compliance processes warrant further research?**

FinRegLab will continue to explore opportunities to assess federated learning's potential for transforming the efficiency and inclusiveness of financial crimes compliance. We will update and expand these FAQs as that work progresses.

What is federated machine learning? How is it different from other forms of machine learning?

Federated machine learning is a distributed form of machine learning that uses decentralized model training and development processes. In federated learning, a learning or classification algorithm trains on data held by node entities—such as financial institutions—that are participating in the federated learning process. Each node provides access for that algorithm to training data held on its servers and uploads updates to the algorithm to a coordinating server.[5] By contrast, non-federated forms of machine learning typically train on data held on a central server, which can be costly, impractical, or risky to acquire and aggregate in some contexts given confidentiality requirements and competitive considerations.[6]

In leaving training data at its origin, federated machine learning enables access to a "significantly wider range of data than what any single organization possesses in house"[7] by using an architecture that is designed to avoid compromising the privacy, information security, and competitive interests of each firm that provides training data.[8] Among the most important aspects of this approach is its ability to increase the number of "true positives"—examples of the behavior that the model is designed to detect—in the available training data.[9] In practice, a financial institution may only be able to train a risk identification model on 10 examples of activity that represent a particular type of illicit financial activity, whereas a federated model for identifying the same kind of risk may be trained on an aggregated set of 500 examples of that activity across all the participating node institutions, and each participating firm will be able to use the insights derived from the richer training data in the federated model.

Federated machine learning has potential application in other areas where improving access to certain kinds of data for training models is critical to model performance, such as healthcare, autonomous vehicles,[10] retail marketing personalization, and in other financial services applications such as fraud.[11]

Further Reading

Jakub Konecny, H. Brendan McMahan, Daniel Ramage, Peter Richtárik, Federated Optimization: Distributed Machine Learning for On-Device Intelligence (October 11, 2016), available at: https://arxiv.org/abs/1610.02527

Jakub Konecny, H. Brendan McMahan, Felix X. Yu, Ananda Theertha Suresh & Dave Bacon, Federated Learning: Strategies for Improving Communication Efficiency, Google (October 30, 2017), available at: https://arxiv.org/abs/1610.05492.

Heiko Ludwig, Nathalie Baracaldo, Gegi Thomas, Yi Zhou, Ali Anwar, Shashank Rajamoni, Yuya Ong, Jayaram Radharkrishnan, Ashish Verma, Mathieu Sinn, Mark Purcell, Ambrish Rawat, Tran Minh, Naoise Holohan, Supriyo Chakraborty, Shalisha Whitherspoon, Dean Steuer, Laura Wynter, Hifaz Hassan, Sean Laguna, Mikhail Yurochkin, Mayank Agarwal, Ebube Chuba, Annie Abay, IBM Federated Learning: An Enterprise Framework, IBM (July 22, 2020), available at: https://arxiv.org/pdf/2007.10987.pdf.

Brendan McMahan, Daniel Ramage, Federated Learning: Collaborative Machine Learning Without Centralized Training Data, Google (April 6, 2017), available at: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

Daniel Nelson, What is Federated Learning?, Unite.AI (August 23, 2020), available at: https://www.unite.ai/what-is-federated-learning/.

Nicola Rieke, What is Federated Machine Learning?, Nvidia (October 13, 2019), available at: https://blogs.nvidia.com/blog/2019/10/13/what-is-federated-learning/.

Related FAQs

How are federated models developed?

What are the potential benefits of using federated machine learning to identify risks related to financial crimes?

How are federated models developed?

Federated machine learning can be used to develop models using the following steps:[12]

- **Initial model development:** One hub institution—a central bank in our proposal—develops a learning algorithm that is designed to identify activities and patterns that point to potential illicit financial activity. The central hub trains that algorithm on a preliminary data set to create a model that will detect trends, anomalies, or to make predictions about potential risk levels.
- **Model shared to nodes:** That preliminary model or the learning algorithm is then shared with institutions participating as nodes—which in an AFC context may be financial institutions, law enforcement, or regulatory agencies.
- **Model training in nodes:** Each institution participating as a node will then train a copy of the model on their own institutional transaction data. With potentially hundreds of participants, each model copy is re-trained, reflecting new parameters and weights based on the training data available at each participating node.
- **Re-trained models returned to hub:** Each participant transmits to the hub either a version of the retrained model back or detailed information on the updated parameters and weights, without sharing any of their data. This could happen periodically or on a set schedule independent of other participants.
- **Aggregation by the hub:** The hub server aggregates and analyzes the revised model parameters received from the nodes and updates the central model based on this information.
- **Updated model shared back to nodes:** The hub then shares the revised model back to participants. The model now reflects insights derived from analysis of all the participants' data. Alternatively, the hub could instead share revised weights and parameters for each participant to use in their own individual risk identification models. The model weights and parameters would need appropriate security to ensure illicit actors cannot evade the updated processes.

A short video describing this process can be viewed here.

Further Reading

Robert Carlsson, Privacy-Preserved Federated Learning: A Survey of Applicable Machine Learning Algorithms in a Federated Environment, Uppsala University (October 2020), available at: https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1497137&dswid=-8830.

Zhiyuan Chen, Le Dinh Van Khoa, Ee Na Teoh, Amril Nazir, Ettikan Kandasamy Karuppiah, Kim Sim Lam, Knowledge and Information Systems (February 2018), available at: https://link.springer.com/article/10.1007/s10115-017-1144-z.

Peter Kairouz, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, Rafael G.L. D'Oliveira, Salim El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaid Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara Javidi, Gauri Joshi, Mikhail Khodak, Jakub Konecny, Aleksandra Korolova, Farinaz Koushanfar, Sanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, Rasmus Pagh, Mariana Raykova, Hang Qi, Daniel Ramage, Ramesh Raskar, Dawn Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu, Sen Zhao, Advances and Open Problems in Federated Learning (December 10, 2019), available at: https://arxiv.org/pdf/1912.04977.pdf.

LexisNexis, True Cost of Financial Crime Compliance Global Report (April 7, 2020), available at: https://risk.lexisnexis.com/insights-resources/research/true-cost-of-financial-crime-compliance-study-global-report.

Qiang Yang, Yang Liu, Tianjian Chen, Yongxin Tong, Federated Machine Learning: Concept and Applications, ACM Transactions on Intelligent Systems and Technology Journal (February 13, 2019), available at: https://arxiv.org/pdf/1902.04885.pdf.

Toyotaro Suzumura, Yi Zhou, Nathalie Baracaldo, Guangnan Ye, Keith Houck, Ryo Kawahara, Ali Anwar, Lucia Larise Stavarache, Yuji Watanabe, Pablo Loyola, Daniel Klyashtorny, Heiko Ludwig, and Kumar Bhaskaran, Towards Federated Graph Learning for Collaborative Financial Crimes Detection, IBM TJ Watson Research Center (October 2, 2019), available at: https://arxiv.org/pdf/1909.12946.pdf.

Related FAQs

How are AI models and machine learning models developed?

What forms of AI and machine learning are most commonly used in financial services? How do they work?

What are the potential confidentiality and privacy implications of using federated machine learning in anti-financial crime processes?

What are the key features of current approaches to financial crimes compliance?

The potential for federated machine learning to create more effective and inclusive AFC processes is particularly important because substantial efforts by both financial institutions and regulators to improve compliance have produced uneven results . The current system is characterized by:

- **High pressure and escalating costs:** Firms have experienced intense and sustained regulatory oversight of their efforts to detect and report risks of illicit activity since September 11th. The core risk management expectations have been reinforced by subsequent waves of banking regulation.[13] Both enforcement actions and the compliance costs that firms incur to forestall those actions are measured in billions.[14] To help manage this level of regulatory oversight, financial institutions spend $1.28 trillion annually to combat financial crime around the world,[15] including $181 billion on AFC compliance.[16] In 2020, those compliance costs increased by an estimated 33 percent.[17]

- **Weak results:** Notwithstanding substantial investment and levels of activity to secure the financial system, less than 1 percent of an estimated $2 trillion laundered annually through the world financial system is actually caught.[18] This means that—in addition to spending on compliance—companies, individuals, and governments lose an estimated $1.45 trillion annually to financial crime as victims of fraud, bribery or corruption, money laundering, theft, cybercrime, and issues related to slave labor or human trafficking.[19] A large part of the problem is that financial institutions develop models that are limited to their own information or rely on vendor-provided models that may provide marginally better insight on patterns of illicit activity that affect multiple firms or markets. Further, law enforcement does not generally provide feedback about which reported transactions are confirmed instances of illicit activity and which are not. As a result, risk identification models produced by firms and vendors are developed without access to deep or current information about true positives.

» **Little coordination:** The structure of BSA/AML oversight is designed primarily to detect and prosecute bad actors, rather than to position firms and their regulators to prevent and disrupt illicit financial activity.[20] In some cases, this means that domestic requirements prevent sharing information about potential risks among operating subsidiaries of global firms. Competitive interests and market structures may also undermine each individual firm's incentives to share information about emerging risk patterns among peers. Further, as noted above, firms have little access to training data from government sources that provides critical information about which people or activities flagged as risks were confirmed to be illicit financial activity. This absence of "true positives" hampers efforts by financial institutions and vendors to develop effective risk identification and improve their detection efforts as the nature of financial crime changes.[21]

» **Financial exclusion:** Though unintended, financial exclusion has also become a hallmark of the global BSA/AML regime. Firms often cite compliance costs related to these regulations as key factors in decisions about whether and how they serve developing countries. In many instances, the potential for significant losses, sanctions, and reputational damage related to AFC problems have skewed the risk-reward ratio that informs what regions, products, and customers become the focal point of a financial institution's strategy. This is particularly acute for firms providing correspondent banking services in emerging markets where data availability and systems limitations make AFC processes more time consuming and expensive, as well as potentially less effective, in identifying illicit activity. Risks related to providing intermediation to domestic financial institutions has result in a 20 percent decrease in correspondent banking relationships over the last seven years. Reduction in correspondent banking has a particularly acute effect on financial inclusion, because these activities include remittance transfers and foreign direct investment.[22] Restricting these activities can stunt development, growth, and asset formation, as well as cutting off individuals from immediately accessible financial resources from family members living abroad. It may also push people and businesses to use less transparent and unregulated financial networks.

Further Reading

Raghad Al-Shabandar, Gaye Lightbody, Fiona Browne, Jun Liu, Haiying Wang, Huiru Zheng, The Application of Artificial Intelligence in Financial Compliance Management, AIAM (October 2019), available at: https://dl.acm.org/doi/abs/10.1145/3358331.3358339

Esman Kurum, RegTech Solutions and AML Compliance: What Future for Financial Crime?, Journal of Financial Crime (May 22, 2020), available at: https://www.emerald.com/insight/content/doi/10.1108/JFC-04-2020-0051/full/html.

Andrea Scripa Els, Artificial Intelligence as a Digital Privacy Protector, Harvard Journal of Law and Technology (2017), available at: https://jolt.law.harvard.edu/assets/articlePDFs/v31/31HarvJLTech217.pdf.

Juan Zarate and Chip Poncy, Designing a New AML System, The Clearing House (2016), available at: https://www.theclearinghouse.org/banking-perspectives/2016/2016-q3-banking-perspectives/articles/a-new-aml-system.

Related FAQs

What are the potential benefits of using federated machine learning to identify risks related to financial crimes?

What are the potential confidentiality and privacy implications of using federated machine learning in anti-financial crime processes?

What is the basis for believing that using federated machine learning to identify illicit financial activity has the potential to expand access to the financial system?

What are the potential benefits of using federated machine learning to identify risks related to financial crimes?

Federated machine learning can be operationalized in a number of ways to support risk identification in various components of AFC programs. Its potential benefits include:

- » **Improving accuracy of risk identification models:** Federated learning may improve the accuracy of risk identification by expanding the data available for model training and sharing the resulting insights without requiring sharing or aggregation of the underlying data from participating firms or government agencies.[23] There are two primary mechanisms at work. First, federated learning can expand insights available to individual firms by effectively allowing access to training data from peers as they independently develop and operate their own risk identification models or work with vendors to tailor such models for their business and operating processes. This can improve each participating firm's ability to detect rapidly evolving, market-wide patterns that suggest illicit activity. Second, federated learning can provide a way for models designed to identify potential risks of financial crime to incorporate insights using richer and more timely information from law enforcement about whether previously suspected individuals or transactions proved to be risky that agencies are not permitted or willing to share directly.
- » **Respecting privacy, information security, and data protection requirements:** Data used to build risk identification models and screen individual applications and transactions are generally subject to privacy and confidentiality requirements. They may also be subject to localization requirements in a growing number of domestic laws and regulations that prescribe the handling of data created in a particular country or concerning that country's citizens. Collectively, these data privacy, security, and protection requirements have limited the ability of AFC processes at individual firms to benefit from the kind of data sharing that the Patriot Act contemplated. However, federated learning presents a compelling potential solution—instead of amassing large volumes of data from various sources, the algorithm, rather than the data, travels in this approach and enables sharing of insights among participants without exposing data that is legally protected or privileged or competitively sensitive.
- » **Enhancing agility of detection efforts:** The financial system is comprised of tens of thousands of financial institutions conducting transactions that move trillions of dollars

around the globe every day. One major payments network estimated that it handled 500 million transactions a day in 2018.[24] The diversity of financial institutions, intermediaries, products, and jurisdictions, as well as the speed at which transactions occur in a digital marketplace, place a premium on being able to use the available information to detect illicit activity. This requires not only access to broad, deep, and diverse data, but also the creation of a structure that allows timely sharing of insight and learning about how risks are evolving. Federated learning models not only provide a dynamic mechanism for shared learning about common threats, but also have the capacity to efficiently add participating entities as nodes.

» **Better resource allocation:** Improving the accuracy of risk identification models can help firms and government agencies alike focus investigative resources on individuals and transactions that present real financial crime risk. One provider of BSA/AML risk identification models that uses federated learning suggests that its technology reduces false positives in screening from roughly 95 percent to 12 percent.[25] The cost of false positives can be measured in the substantial resources dedicated by firms and regulators to manually review and investigate each transaction flagged as potentially suspicious. Accordingly, significant decreases in false positives, if realized at scale, present a substantial opportunity to re-allocate AFC resources, enabling more efficient pursuit of real risks of illicit financial activity and potentially reducing costs related to operating AFC programs.

As discussed more fully in a separate question, these benefits if realized at scale may create opportunities to expand access to the financial system.

Further Reading

Alon Kaufman, How Privacy-Enhanced Technologies Can Make Financial Crime Compliance More Effective, ABA (June 11, 2020), available at: https://bankingjournal.aba.com/2020/06/how-privacy-enhanced-technologies-can-make-financial-crime-compliance-more-effective/

Josh Lake, Federated Learning: Is it Really Better for your Privacy and Security?, Comparitech (October 25, 2019), available at: https://www.comparitech.com/blog/information-security/federated-learning/.

Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith, Federated Learning: Challenges, Methods, and Future Directions (Aug. 2019), available at: https://arxiv.org/pdf/1908.07873.pdf.

Parviz Peiravi, Gary Shiffman, Juan Zarate, and Nikhil Deshpande, Fighting Financial Crime Effectively through Federated Machine Learning, Intel (November 18, 2020), available at: https://www.brighttalk.com/webcast/10773/448012/fighting-financial-crime-effectively-through-federated-machine-learning.

Gary Shiffman, Juan Zarate, Nikhil Deshpande, Raghuram Yeluri, Parviz Peiravi, A 21st Century Solution for Combating Money Laundering and the Financing of Terrorism, Consilient (November 6 2020), available at: https://consilient.com/whitepaper/federated-learning-through-revolutionary-technology/.

Suzumura (2019).

What are the potential confidentiality and privacy implications of using federated machine learning in anti-financial crime processes?

Both legal and competitive factors constrain information sharing for AFC processes. For example, individual countries are developing standards that define when and how their citizens can control the use of their digital data and may require that the individual be given notice of and consent to a specific use of their data.[26] In the same vein and often with the intent of reinforcing national privacy rules, as well as national economic interests, data localization or residency laws require that data about a nation's citizens or residents be collected, processed, and/or stored inside the country, and may require demonstration that local requirements have been satisfied before data can be transferred internationally.[27] India, China, and Russia have moved in this direction, as have many African nations.

Federated learning is a particularly promising potential solution for improving AFC processes precisely because it enables the benefits of machine learning—better pattern identification and enhanced predictive power—without requiring more data sharing or compromising compliance with privacy and information security expectations, data localization laws, or law enforcement confidentiality considerations.

However, several questions remain to determine whether its potential can be realized at scale. For example, implementation of a federated system that delivers on its promises in this regard will likely require well designed model architecture, appropriate systems (including encryption), and effective governance of the participants and outputs of the federated learning system. In addition, the performance of federated learning in delivering these privacy benefits is still subject to evaluation. Substandard encryption or failure to implement a system update in a timely way in one of the participating nodes could expose sensitive information.[28] It may also be possible to reverse engineer portions of training data sets from the updated models.[29] Finally, appropriate oversight systems and governance mechanisms need to be in place to detect and respond to attempts by outsiders or one of the participating nodes to prevent data poisoning or attempts to manipulate the model being shared across the federated system.

Further Reading

Nicholas Carlini, Chang Liu, Úlfar Erlingsson, Jernej Kos, and Dawn Song, The Secret Sharer: Measuring Unintended Neural Network Memorization and Extracting Secrets (July 16, 2019), available at: https://arxiv.org/abs/1802.08232.

Yang Feng, Xue Yang, Weijun Fang, Shu-Tao Xia, Xiaohu Tang, Jun Shao, Tao Xiong, A Practical Privacy-Preserving Method in Federated Deep Learning (August 6, 2020), available at: https://arxiv.org/pdf/2002.09843.pdf.

Daniel Gutierrez, How You Can Use Federated Learning for Security & Privacy, Open Data Science (May 25, 2020), available at: https://opendatascience.com/how-you-can-use-federated-learning-for-security-privacy/.

Josh Lake, Federated Learning: Is it Really Better for your Privacy and Security?, Comparitech (October 25, 2019), available at: https://www.comparitech.com/blog/information-security/federated-learning/.

Sheng Shen, Tianqing Zhu, Di Wu, Wei Wang, Wanlei Zhou, From Distributed Machine Learning to Federated Learning: In the View of Data Privacy and Security (September 23, 2020), available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.6002.

Stacey Truex, Nathalie Baracaldo, Ali Anwar, Thomas Steinke, Heiko Ludwig, Rui Zhang, Yi Zhou, A Hybrid Approach to Privacy-Preserving Federated Learning (August 14, 2019), available at: https://arxiv.org/pdf/1812.03224.pdf.

Related FAQs

What are the key features of current approaches to financial crime compliance?

What are the potential benefits of using federated machine learning to identify risks related to financial crime?

What aspects of using federated learning to improve financial crime compliance processes warrant further research?

What is the basis for believing that using federated machine learning to identify illicit financial activities has the potential to expand access to the financial system?

Federated machine learning has the potential to confer on all relevant stakeholders an essential component of an effective BSA/AML regime: enhanced confidence in our collective ability to identify correctly individuals and transactions related to illicit financial activity. This confidence has the potential to translate into opportunities to improve processes, focus resources, and reduce costs, as well as to create a more vibrant and inclusive financial system.

At the firm level, substantial improvements in the accuracy of models that screen individuals and transactions for indications of financial crimes risk may affect not just the allocation of resources within AFC programs, but also reduce their overall operating costs. Over time, this may affect the strategic decisions that firms make about which markets to serve by improving their perception of potential risks and rewards.

Several market mechanisms may reinforce the incentives to expand access to the financial system that come with these altered firm-level dynamics. For instance, domestic incumbents may be able to expand their operations because international investors and firms, including correspondent banks, are more willing to do business with them. New entrants may also seek to develop products and services designed to serve those who had previously been cut off by risk identification models that identified certain populations or activities as having higher levels of risk than perhaps is warranted.

Federated learning may be a particularly powerful bridge for establishing access in high-risk corridors excluded from the financial system. Areas in active conflict or newly emerging from conflict present extraordinary risks of terrorism and financial crime due to the breakdown of civil society. However, the inability of the global financial system to establish connectivity in these regions can heighten national security and financial crimes risks and complicate recovery from conflict. In these areas, federated learning may provide a mechanism for creating a trusted mechanism for funding essential development and supporting legitimate financial activities.

Further Reading

The Decline in Access to Correspondent Banking Services in Emerging Markets: Trends, Impacts, and Solutions, The World Bank (2018), available at: http://pubdocs.worldbank.org/en/786671524166274491/TheDeclineReportlow.pdf.

Scott B. MacDonald, Is There a New Normal for De-Risking in the Caribbean?, CSIS (October 2019), available at: https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/191016_There_a_New_Normal_for_De-risking_in_the_Caribbean_v1.pdf.

Brendon Machado, Federated Machine Learning for Loan Risk Prediction, InfoQ (September 9, 2020), available at: https://www.infoq.com/articles/federated-machine-learning/.

Suzumura (2019).

Shirish Wadivkar, De-Risking in Correspondent Banking: Existential Challenge or Catalyst for Change?, Standard Chartered (2019), available at: https://www.sc.com/en/feature/de-risking-in-correspondent-banking-existential-challenge-or-catalyst-for-change/.

Related FAQs

What are the key features of current approaches to financial crime compliance?

What aspects of using federated learning to improve financial crime compliance processes warrant further research?

What aspects of using federated learning to improve financial crimes compliance processes warrant further research?

The use of federated machine learning within financial services may be novel, but the technology is being successfully used in areas like health care where significant privacy requirements and other regulations apply to holders of data that is necessary for analytical purposes. For example, federated learning is used to help identify clinically similar patients and predict mortality, hospitalization from cardiac events, and ICU stay length.[30] This suggests there is ongoing work by researchers and practitioners that can inform the design and implementation of federated learning in financial services.

Nevertheless, further research is needed to validate the benefits of using federated learning to improve AFC risk identification processes, to assess the attendant risks, and to evaluate its compatibility with applicable laws and regulations. The primary areas where research is needed include:

- Developing an appropriate framework for evaluating the performance of federated learning technology in the context of financial crime risks, including how well it reduces improper denials of access that result from false positives;[31]
- Assessing the performance of federated machine learning models in reliably protecting sensitive information while enabling sharing of model parameters and other data insights, including their robustness to various forms of data manipulation[32] and the efficacy of efforts to prevent data leaks to or reverse engineering of training data by participating firms;[33]
- Measuring how the use of federated machine learning affects individuals and businesses that are excluded from the financial system by current risk management decisions (using metrics such as income, geography, and other variables); and
- Identifying policy options for central banks and other financial regulators to support the use of this technology generally and in ways that improve financial inclusion.

Further Reading

Abhishek Bhowmick, John Duchi, Julien Freudiger, Gaurav Kapoor, Ryan Rogers, Protection Against Reconstruction and Its Applications in Private Federated Learning (June 3, 2019), available at: https://arxiv.org/abs/1812.00984.

Kairouz (2019).

Li (2019).

Machado (2020).

Endnotes

[1] FinRegLab's prior FAQs on AI in financial services provide an overview of key concepts in machine learning and discuss the particular importance of model transparency and explainability and consider the implications of using of machine learning for credit underwriting.

[2] 31 U.S.C. 5311. This section enables financial institutions to share information among themselves in the case of suspected money laundering or terror financing. Given that notification is provided to FinCEN, these acts of sharing customer data are generally allowed under a safe harbor from liability of violating other laws including the Right to Financial Privacy Act and the Gramm-Leach-Bliley Act.

[3] Joe Mont, Data Sharing, AI May be Antidote to Failing AML Efforts, Compliance Week, Jan 17, 2018. https://www.complianceweek.com/data-sharing-ai-may-be-antidote-to-failing-aml-efforts/2411.article.
Jonah Anderson, Jeremy Kuester, John Wagner, Rebecca Copcutt, and John Timmons, AML Information Sharing in a Technology-Enabled and Privacy-Conscious World, White & Case, (January 31, 2019), available at: https://www.whitecase.com/publications/alert/aml-information-sharing-technology-enabled-and-privacy-conscious-world.

What is federated machine learning? How is it different from other forms of machine learning?

[4] AFC as used in this document reflects the umbrella term increasingly used by practitioners around the world to refer to activities to implement requirements of the Bank Secrecy Act and similar laws in other jurisdictions to prevent, monitor, detect, and report financial crime. These include requirements related to anti-money laundering (AML), countering the financing of terrorism (CFT), and sanctions monitoring, as well as customer identification or know-your-customer (KYC) requirements. KYC programs typically include customer identification programs (CIP) to conduct customer due diligence (CDD) and enhanced due diligence (EDD) as needed. See: What is Financial Crime?, International Compliance Association, available at: https://www.int-comp.org/careers/your-career-in-financial-crime-prevention/what-is-financial-crime/.

[5] Jakub Konecny, H. Brendan McMahan, Daniel Ramage, Federated Optimization: Distributed Machine Learning for On-Device Intelligence (October 11, 2016), available at: https://arxiv.org/pdf/1610.02527.pdf; Nicola Rieke, What Is Federated Machine Learning?, Nvidia (October 13, 2019), available at: https://blogs.nvidia.com/blog/2019/10/13/what-is-federated-learning/; Daniel Nelson, What Is Federated Learning?, Unite.AI (August 23, 2020), available at: https://www.unite.ai/what-is-federated-learning/.

[6] Konecny, *et al.* (2018); Qiang Yang, Yang Liu, Tianjian Chen, Yongxin Tong, Federated Machine Learning: Concept and Applications, ACM Transactions on Intelligent Systems and Technology Journal (February 13, 2019), available at: https://arxiv.org/pdf/1902.04885.pdf.

[7] Rieke (2019).

[8] Qiang Yang, *et al.* (2019).

[9] Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletarì, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N. Galtier, Bennett A. Landman, Klaus Maier-Hein, Sébastien Ourselin, Micah Sheller, Ronald M. Summers, Andrew Trask, Daguang Xu, Maximilian Baust, M. Jorge Cardoso, The Future of Digital Health with Federated Learning, Nature (September 14, 2020), available at: https://www.nature.com/articles/s41746-020-00323-1.

[10] Tian Li, Federated Learning: Challenges, Methods, and Future Directions, Carnegie Mellon University (November 12, 2019), available at: https://blog.ml.cmu.edu/2019/11/12/federated-learning-challenges-methods-and-future-directions/.

[11] Qiang Yang (2019).

How are federated models developed?

[12] Brendan McMahan, Federated Learning: Collaborative Machine Learning without Centralized Training Data, Google (April 6, 2017), available at: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html; Nelson, (2020), available at: https://www.unite.ai/what-is-federated-learning/; Rieke (2019).

What are the key features of current approaches to financial crimes compliance?

[13] For example, in 2014, the prudential regulators' heightened standards for the largest and complex banks ratcheted up expectations in AFC and other regulatory disciplines for the quality, scope, and resourcing of risk management, and compliance, and governance activities. OCC Finalizes Its Heightened Standards for Large Financial Institutions, Office of the Comptroller of the Currency (September 2, 2014), available at: https://www.occ.treas.gov/news-issuances/news-releases/2014/nr-occ-2014-117.html.

[14] Juan Zarate and Chip Poncy, Designing a New AML System, The Clearing House (2016), available at: https://www.theclearinghouse.org/banking-perspectives/2016/2016-q3-banking-perspectives/articles/a-new-aml-system.

[15] Revealing the True Cost of Financial Crime: 2018 Survey Report, Refinitiv (May 2018), available at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/reports/true-cost-of-financial-crime-global-focus.pdf.

[16] Financial Services Firms Spend $180.9 Billion on Financial Crime Compliance, According to LexisNexis Risk Solutions Global Study, LexisNexis (April 7, 2020), available at: https://risk.lexisnexis.com/about-us/press-room/press-release/20200407-fcc-global-study.

[17] LexisNexis Risk Solutions Study Reveals Financial Crime Compliance Costs Increased 33 percent in 2020 at Financial Institutions in the United States and Canada, LexisNexis (October 21, 2020), available at: https://risk.lexisnexis.com/about-us/press-room/press-release/20201021-tcof-study-us-canada

[18] Estimating Illicit Financial Flows Resulting from Drug Trafficking and Other Transnational Organized Crimes, United Nations Office on Drugs & Crime (October 2011), available at: https://www.unodc.org/documents/data-and-analysis/Studies/Illicit_financial_flows_2011_web.pdf.

[19] Revealing the True Cost of Financial Crime: 2018 Survey Report, Refinitiv (May 2018), available at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/reports/true-cost-of-financial-crime-global-focus.pdf.

[20] Zarate (2016).

[21] Gary M. Shiffman, Giant Oak's Comments Regarding FinCEN's Proposed AML Program Effectiveness Rule, Giant Oak (November 12, 2020), available at: https://resources.giantoak.com/anprm (containing text of comment letter submitted to FinCEN in response to Advanced Notice of Proposed Rulemaking on November 12, 2020).

[22] Given that remittances and foreign direct investment are about five times higher than official development assistance (around $150 billion in 2018), the development impact of restricting cash transfers when correspondent banking relationships are severed could be drastic. Dilip Ratha, Remittances on Track to Become the Largest Source of External Financing in Developing Countries, World Bank Blogs (April 8, 2019), available at: https://blogs.worldbank.org/peoplemove/remittances-track-become-largest-source-external-financingdeveloping-countries; Development Aid Drops in 2018, Especially to Neediest Countries, Organisation for Economic Co-operation and Development (October 2019), available at: https://www.oecd.org/development/development-aid-drops-in-2018-especially-to-neediest-countries.htm.

What are the potential benefits of using federated machine learning to identify risks related to financial crimes?

[23] Toyotaro Suzumura, Yi Zhou, Nathalie Baracaldo, Guangnan Ye, Keith Houck, Ryo Kawahara, Ali Anwar, Lucia Larise Stavarache, Yuji Watanabe, Pablo Loyola, Daniel Klyashtorny, Heiko Ludwig, and Kumar Bhaskaran, Towards Federated Graph Learning for Collaborative Financial Crimes Detection, IBM TJ Watson Research Center, (October 2, 2019), available at: https://arxiv.org/pdf/1909.12946.pdf) (discussing efforts to demonstrate proof-of-concept of federated machine learning at the United Kingdom Financial Conduct Authority week-long Global Anti-Money Laundering and Financial Crime TechSprint and noting the federated models outperformed standard models by 20 percent); *see also* Li et al. (2019).

[24] Visa, Annual Report: 2018 (2018), available at: https://s1.q4cdn.com/050606653/files/doc_financials/annual/2018/Visa-2018-Annual-Report-FINAL.pdf

[25] 'False positive' is an error in a binary classification (either-or statement). In AFC, an application or proposed transaction may be flagged for further review because the individual's identity cannot be confirmed in a given database or because the transaction is over a specific threshold or involves a high-risk jurisdiction or counter-party. A false positive in the context of AFC risk identification occurs when a permissible activity is inappropriately flagged as suspicious. PwC, From Source to Surveillance: The Hidden Risk in AML Monitoring System Optimization (September 2010), available at: https://www.pwc.com/us/en/anti-money-laundering/publications/assets/aml-monitoring-system-risks.pdf;
Gary Shiffman, Juan Zarate, Nikhil Deshpande, Raghuram Yeluri, and Parviz Peiravi, Federated Learning Through Revolutionary Technology: A 21st Century Solution for Combating Money Laundering and the Financing of Terrorism, Consilient (November 6, 2020), available at: https://consilient.com/whitepaper/federated-learning-through-revolutionary-technology/).

What are the potential confidentiality and privacy implications of using federated machine learning in anti-financial crime processes?

[26] For example, the European Union's General Data Protection Regulation includes both data protection and privacy requirements, keeping data safe from unauthorized access as well as empowering users to decide how their data can be processed, by whom, and for what purpose. A Guide to GDPR Data Privacy Requirements, Proton Technologies AG, available at: https://gdpr.eu/data-privacy/.

[27] Arindrajit Basu, The Retreat of the Data Localization Brigade: India, Indonesia, and Vietnam, The Diplomat (January 2020), available at: https://thediplomat.com/2020/01/the-retreat-of-the-data-localization-brigade-india-indonesia-and-vietnam/.

[28] Georgios A. Kaissis, Marcus R. Makowski, Daniel Rückert, and Rickmer F. Braren, Secure, Privacy-Preserving and Federated Machine Learning in Medical Imaging, Nature (June 8, 2020), available at: https://www.nature.com/articles/s42256-020-0186-1.

[29] In at least one documented situation, researchers were able to recover private information (in this case a credit card number) from the parameters shared among the federated learning models in the context of a specific learning algorithm that sought to predict the next character typed on a smartphone. Nicholas Carlini, Chang Liu, Úlfar Erlingsson, Jernej Kos, and Dawn Song, The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Network (February 2018), available at: https://arxiv.org/abs/1802.08232.

What aspects of using federated learning to improve financial crimes compliance processes warrant further research?

[30] Rieke (2020).

[31] Shiffman (2020).

[32] Josh Lake, Federated Learning: Is It Really Better for Your Privacy and Security?, Comparitech (October 25, 2019), available at: https://www.comparitech.com/blog/information-security/federated-learning/.

[33] Brendon Machado, Federated Learning for Loan Risk Prediction, InfoQ (September 9, 2020), available at: https://www.infoq.com/articles/federated-machine-learning/.

FinRegLab
900 19th Street NW, Suite 600
Washington DC 20006

Copyright 2020 © FinRegLab, Inc. All Rights Reserved.