Federated Learning - Methods, Applications and beyond

Moritz Heusinger, Christoph Raab, Fabrice Rossi, Frank-Michael Schleif

Federated Learning
Methods, Applications and Beyond

Moritz Heusinger[1], Christoph Raab[1], Fabrice Rossi[2] and Frank-Michael Schleif[1]

1- University of Applied Science Würzburg-Schweinfurt - Department of Computer Science, Würzburg - Germany
2- CEREMADE, University Paris Dauphine PSL, France

Abstract. In recent years the applications of machine learning models have increased rapidly, due to the large amount of available data and technological progress. While some domains like web analysis can benefit from this with only minor restrictions, other fields like medicine with patient data are stronger regulated. In particular *data privacy* plays an important role as recently highlighted by the trustworthy AI initiative of the EU or general privacy regulations in legislation. Another major challenge is, that the required training *data is* often *distributed* in terms of features or samples and unavailable for classical batch learning approaches. In 2016 Google came up with a framework, called *Federated Learning* to solve both of these problems. We provide a brief overview on existing Methods and Applications in the field of vertical and horizontal *Federated Learning*, as well as *Federated Transfer Learning*.

1 Introduction

Federated learning (FL) is a novel concept for learning distributed data, which was first introduced by Google [1, 2, 3] in 2016.

Definition 1 (Federated Learning)**.** *Given a large number of N clients and a particular data analysis task, each client C_i has its own data addressing the task, without direct access to the other clients data. The objective in FL is to learn a predictive model M such that the error on the objective function E is minimized, in a distributed way. In particular various data processing clients are involved. The communication takes place by a distributed protocol where in general a master is identified to aggregate the prediction model. FL has three steps (1) an initial model is distributed to the clients (2) the local model of C_i is trained on its local data by taking the model information of the master into account (3) the master aggregates the local models M_i to the global model M and communicates the global model back to the clients. The steps are visualized in Figure 1.*

FL has gained substantial interest in the machine learning (ML) community with different frameworks implementing the main concept [4], in particular *flower* [5] which is considered very mature. Applications of FL are more and more frequent [6, 7, 8, 9]. The research field is also very active with new communication protocols [1], encryption concepts [10] and particular optimization algorithms [2, 11, 12].

FL has numerous advantages over classical ML. While in classical machine learning training data are submitted to a central instance and a model is learned in batch processing, FL shifts the actual learning to the data source. This allows one to employ the power of distributed client machines, keeps the user data private, and permits to use information that is otherwise inaccessible and spread over different clients.

In this view FL perfectly aligns with recent trends on machine learning on large community data [13] and the increasing set of constraints due to privacy regulations like the GDPR[1], trustworthy AI[2] but also objectives covered in various AI manifests[3].

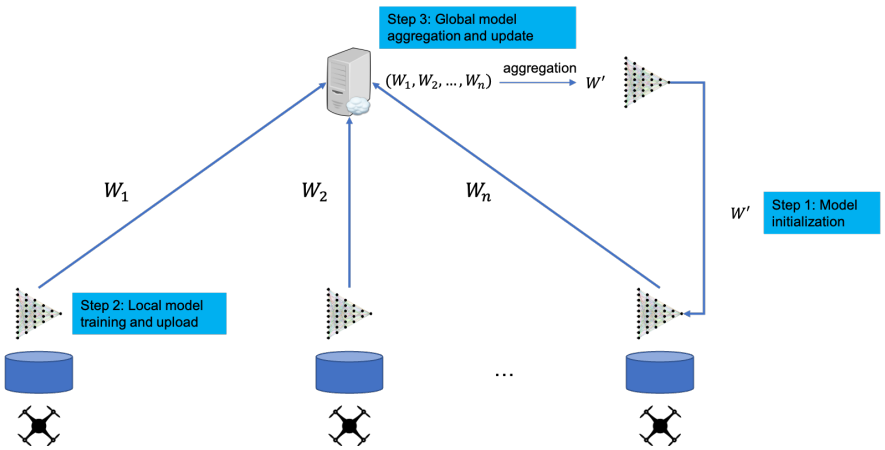


Fig. 1: Diagram of FL. To preserve data privacy, local model gradients are only sent to one trusted server (or a primary coordinator) and not directly to other clients. The local instances are training their own model with the local data and after multiple iterations the gradient is sent to the primary coordinator. The primary coordinator aggregates all local gradients to a central global update, which is used to update the global model. Finally, the global model is sent back to the clients to replace their previous models.

Broadly speaking, Federated Learning methods belong to three different categories, which depend on how comparable are the local data of the clients, as detailed in Sec. 2.

2 Types of federated learning

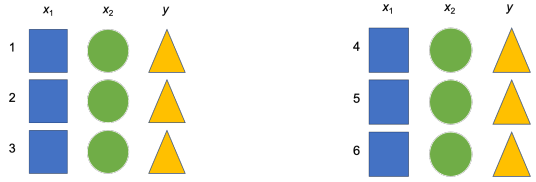
We denote the data held by client C_i as $\mathcal{D}_i$. In FL a data set consists of the features $\mathbf{X}$, the sample id space $\mathbf{i}$ and an optional label space $\mathbf{y}$. The sample space, as well as the feature space, may not be identical over different data

[1] https://gdprinfo.eu/
[2] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[3] https://nouvelles.umontreal.ca/en/article/2018/12/04/developing-ai-in-a-responsible-way/

owners. By this characteristic, FL is categorized into horizontal FL (HFL), vertical FL (VFL), and Federated Transfer Learning (FTL), as detailed below.

(a) Client A (b) Client B

Fig. 2: Vertical Federated Learning (inspired by [14])



(a) Client A (b) Client B

Fig. 3: Horizontal Federated Learning (inspired by [14])

(a) Client A (b) Client B

Fig. 4: Federated Transfer Learning (inspired by [14])

2.1 Horizontal FL

The scenario of HFL is depicted in Figure 3 and shows that the users are split across various clients, while the feature space is always the same. A typical scenario is given by a global business model which is implemented on smartphones or other IoT devices. Thereby the user generates the same data, like interaction events or shopping activities but may be located in very different geographical regions. HFL is beneficial by calculating a model which employs the information from a large (distributed) user group instead of focusing on a centralized approach with a rather limited amount of training data. In HFL it is common to calculate and upload local gradients calculated from the objective function, which are aggregated by a central master client. The data transmissions can be encrypted to improve the level of privacy using homomorphic encryption [15], differential privacy [16] or secure aggregation as discussed in Sec. 4.

2.2 Vertical FL

Vertical federated learning considers the case where we have multiple different feature sets on a common basis of users. A typical case for this scenario (depicted in Figure 2) is an analysis task where the user is using at least two ways of interaction and is generating data on two channels, like an online store and in a brick and mortar business.

2.3 Transfer learning

As described above, in FL, it is generally assumed that every user provides a sufficient set of labeled data for a model to learn a specified task. In general, it is also assumed that the analyzed data are given in common feature spaces, although potentially split across various clients [9]. If the first assumption does not hold, the model cannot capture the whole data set characteristic leading to poor prediction performance. Further, the provided gradients are biased towards the present data characteristics, which eventually deteriorates the global learning model. If the second assumption does not hold, the entities and the corresponding feature spaces are disjoint. This prevents the model from discovering the feature space characteristics, and therefore, the FL system cannot extract useful information for the specific user model [9].

In the cases just described, methods of Federated Transfer Learning need to be applied. The FTL setup is illustrated in Figure 4. FTL can be seen as the cross-section of horizontal and vertical transfer learning. In general, Transfer Learning is an algorithmic technique to improve a model trained on one data set, by using related information from another data set. The just mentioned data sets are usually denoted as source and target domain. FTL helps to improve the model of user U_i on its data by using data or model information from one (or more) users U_j, where $i \neq j$. Hence, the learning environment of U_i is called target, and U_j is source [17]. Note, that there can be small intersections regarding feature space or sample space as displayed in Figure 4. A model improvement on the target by means of FTL is achieved by creating a shared representation keeping the federation of data. For example via manifold alignment [18] or domain adversarial learning [19]. Alternatively, FTL methods improve the target model leveraging the source model by instance reweighting [20] or by receiving gradients from source [17]. Note that some approaches also consider different data distributions between source and target [20].

3 Frameworks and Algorithms

In FL the communication architecture and protocols are of particular importance. FL has to deal with non i.i.d. data [21] and due to the FL learning concept substantial communication costs can occur. Considering the typical use cases of FL in the field of mobile devices, IoT [22], unmanned vehicles [8] or large distributed server systems additional compression techniques, averaging strategies, and sparsity constraints are applied to obtain real time-efficient sys-

tems [23, 1, 2]. One can also employ quantization approaches [24] or information about the transmitted data to control the communication load [25].

Only recently some FL frameworks have been proposed which simplify the implementation of own FL models. In particular, for HFL the BlockFL framework was proposed in [26] which makes use of a blockchain during the updates of the model parameters. A framework for VFL is provided by SecureBoost [27]. And in the context of Federated Transfer Learning the framework in [28] is suggested. The framework *Flower* [5] scales well to a large number of clients.

The machine learning community has also recently started to design dedicated learning algorithms for FL. One example is given in [11] where hyperparameter learning on distributed systems is considered. Also information theoretic strategies have been proposed [29], matrix factorization techniques [13], spectral clustering [30], particular designed gradient descend techniques [31, 32] or multi-objective solvers [12]. Furthermore, the Learning Vector Quantization (LVQ) concept has been adopted, to fit in a VFL environment by training separate LVQ models locally and using the relevance matrix to update a global model [33].

4 Privacy methods

The key element of all FL approaches is to keep the data on the user side and in particular to avoid any disclosures. Three techniques are most common to ensure this goal and are frequently combined:

1. homomorphic encryption
2. differential privacy
3. (secure) model aggregation

The most common is model aggregation which trains the global model by summarizing the model parameters from all clients to avoid disclosure of original data. Many optimization concepts which are roughly based on a kind of iterative stochastic gradient descent on E perform a natural averaging of various update steps which is also used in batch online learning approaches [34]. A prominent approach for deep learning falling into this category is given in [3]. An alternative view is to train local models in a multi-task setting which are subsequently combined as shown in [35]. The various local model parameters can also be safely transferred in an aggregated form by blockchain techniques as shown in [26]. One may also directly provide privacy-preserving data representations as shown for kernels in [36].

Homomorphic encryption allows to apply calculations on encrypted data without the need of decoding. One approach following this idea is additive homomorphism [37].

Differential privacy as detailed in [38] is a technique to limit information disclosure during learning. Thereby the training procedure is designed such that small modifications of the training database have no substantial impact on the

model outcome. The attacker can not obtain accurate individual information, but only a controlled piece of information, which still obeys privacy constraints. The particular strategies to implement this concept can be very simple by adding some noise contributions to the output during training or by more complex compression techniques as shown in [39].

In [36] a privacy-preserving method to compute dot-product kernels in VFL is proposed. The technique uses multi-party computation to provide theoretical guarantees on security and privacy. For a more detailed analysis, we refer to [10].

5 Applications

FL can train a united model on data from distributed sources while preserving data privacy and security and thus can play an important role in many industrial sectors, like sales, health, insurance, and others. In general, in every sector, where data cannot be directly aggregated due to privacy protection, data security, or even property rights [9].

FL has been successfully used to improve the quality of keyboard search suggestions on the Google Gboard [40]. Gboard is a virtual keyboard for mobile devices and has several features, like auto-completion and next-word prediction. The application needs to protect the privacy of users and make latency-free predictions. To avoid high data usage and battery consumption due to the FL optimization, the authors had to design mechanics to only send data to a centralized server, if the device is inside a wireless network and actively charging. The server provides every client with a training task as soon as enough clients are connected. In a related context, FL has also been used to perform mobile keyboard predictions [41].

Recently, a project has been started, which uses FL to deploy an auction for intralogistic autonomous drone transportation[4]. The goal of the project is, to create a system where drone owners can bid on transport jobs and execute them in case of winning. The bidding model is based on FL, which enables the knowledge incorporation of every drone in the global model of the system. The distribution of the global model also leads to the fact, that drones with fewer executed jobs will have the same chance to win an auction as more experienced drones.

In [42], a hybrid continual learning strategy is used to address the real-world constraints like computational and memory limits in a real-time on-device personalization task, running on a native Android application.

Other applications of FL include ranking browser history suggestions based on user-interactions [43], visual object detection [44], patient clustering to predict hospital stay time as well as mortality [45], drug discovery [46, 47] and brain tumor segmentation [48, 49, 50]. The application fields of FL are quickly increasing and are a promising research direction of ML. For a more comprehensive review, we refer to [7].

[4]mFUND program of the BMVI, project FlowPro, grant number 19F2128B

6 Conclusions

In this tutorial, we briefly discussed the evolving field of federated learning and outlined recent achievements and approaches. A detailed analysis of recent trends and problems in FL is also provided in [51, 52, 53, 54]. Due to additional privacy constraints in ML and a variety of distributed user groups of ML methods, it can be expected that FL will become even more important in the future

Acknowledgment

MH and CR are thankful for support in the FuE program Informations- und Kommunikationstechnik of the StMWi, project OBerA, grant number IUK-1709-0011// IUK530/010 and for support in the mFUND program of the BMVI, project FlowPro, grant number 19F2128B.

References

[1] Jakub Konečnỳ, H Brendan McMahan, Felix X Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016.

[2] Jakub Konečnỳ, H Brendan McMahan, Daniel Ramage, and Peter Richtárik. Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527*, 2016.

[3] H. B. McMahan, Eider Moore, D. Ramage, and B. A. Y. Arcas. Federated learning of deep networks using model averaging. *ArXiv*, abs/1602.05629, 2016.

[4] Christian Schneebeli, Saikishore Kalloori, and Severin Klingler. A practical federated learning framework for small number of stakeholders. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, WSDM 2021, pages 910–913, New York, NY, USA, 2021. Association for Computing Machinery.

[5] Daniel J. Beutel, Taner Topal, Akhil Mathur, Xinchi Qiu, Titouan Parcollet, Pedro P. B. de Gusmão, and Nicholas D. Lane. Flower: A friendly federated learning research framework, 2021. `https://flower.dev/`.

[6] Theodora S. Brisimi, Ruidi Chen, Theofanie Mela, Alex Olshevsky, Ioannis Ch. Paschalidis, and Wei Shi. Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics*, 112:59–67, 2018.

[7] Li Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin. A review of applications in federated learning. *Computers and Industrial Engineering*, 149:106854, 2020.

[8] Abbas Yazdinejad, Reza M. Parizi, Ali Dehghantanha, and Hadis Karimipour. Federated learning for drone authentication. *Ad Hoc Networks*, 120:102574, 2021.

[9] Q. Yang, Y. Liu, T. Chen, and Y. Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 2019.

[10] Viraaji Mothukuri, Reza M. Parizi, Seyedamin Pouriyeh, Yan Huang, Ali Dehghantanha, and Gautam Srivastava. A survey on security and privacy of federated learning. *Future Generation Computer Systems*, 115:619–640, 2021.

[11] J. Shi, J. Bian, J. Richter, K. Chen, J. Rahnenführer, H. Xiong, and J. Chen. Modes: model-based optimization on distributed embedded systems. *Machine Learning*, 110(6):1527–1547, 2021.

[12] Hangyu Zhu and Yaochu Jin. Multi-objective evolutionary federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 31(4):1310–1322, 2020.

[13] Enyue Yang, Yunfeng Huang, Feng Liang, Weike Pan, and Zhong Ming. Fcmf: Federated collective matrix factorization for heterogeneous collaborative filtering. *Knowledge-Based Systems*, 220:106946, 2021.

[14] Shaoqi Chen, Dongyu Xue, Guohui Chuai, Qiang Yang, and Qi Liu. Fl-qsar: a federated learning based qsar prototype for collaborative drug discovery. *bioRxiv*, 2020.

[15] Haokun Fang and Quan Qian. Privacy preserving machine learning with homomorphic encryption and federated learning. *Future Internet*, 13(4), 2021.

[16] Nuria Rodríguez-Barroso and Goran Stipcich et al. Federated learning and differential privacy: Software tools analysis, the sherpa.ai fl framework and methodological guidelines for preserving data privacy. *Information Fusion*, 64:270–292, 2020.

[17] Yang Liu, Yan Kang, Chaoping Xing, Tianjian Chen, and Qiang Yang. A Secure Federated Transfer Learning Framework. *IEEE Intelligent Systems*, 1672(c):1–1, 2020.

[18] Ce Ju, Dashan Gao, Ravikiran Mane, Ben Tan, Yang Liu, and Cuntai Guan. Federated transfer learning for eeg signal classification. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pages 3040–3045, 2020.

[19] Xingchao Peng, Zijun Huang, Yizhe Zhu, and Kate Saenko. Federated adversarial domain adaptation. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020.

[20] Dashan Gao, Yang Liu, Anbu Huang, Ce Ju, Han Yu, and Qiang Yang. Privacy-preserving heterogeneous federated transfer learning. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 2552–2559, 2019.

[21] Felix Sattler, Simon Wiedemann, Klaus-Robert Müller, and Wojciech Samek. Robust and communication-efficient federated learning from non-i.i.d. data. *IEEE Transactions on Neural Networks and Learning Systems*, 31(9):3400–3413, 2020.

[22] Yang Zhao, Jun Zhao, Mengmeng Yang, Teng Wang, Ning Wang, Lingjuan Lyu, Dusit Niyato, and Kwok-Yan Lam. Local differential privacy-based federated learning for internet of things. *IEEE Internet of Things Journal*, 8(11):8836–8853, 2021.

[23] Mingzhe Chen, Nir Shlezinger, H. Vincent Poor, Yonina C. Eldar, and Shuguang Cui. Communication-efficient federated learning. *Proceedings of the National Academy of Sciences*, 118(17), 2021.

[24] Nicola Tonellotto, Alberto Gotta, Franco Maria Nardini, Daniele Gadler, and Fabrizio Silvestri. Neural network quantization in federated learning at the edge. *Information Sciences*, 575:417–436, 2021.

[25] Xueyu Wu, Xin Yao, and Cho-Li Wang. Fedscr: Structure-based communication reduction for federated learning. *IEEE Transactions on Parallel and Distributed Systems*, 32(7):1565–1577, 2021.

[26] Hyesung Kim, Jihong Park, Mehdi Bennis, and Seong-Lyun Kim. Blockchained on-device federated learning, 2019.

[27] K. Cheng, T. Fan, Y. Jin, Y. Liu, T. Chen, D. Papadopoulos, and Q. Yang. Secureboost: A lossless federated learning framework. *IEEE Intelligent Systems*, pages 1–1, may 5555.

[28] Yang Liu, Yan Kang, Chaoping Xing, Tianjian Chen, and Qiang Yang. A secure federated transfer learning framework. *IEEE Intelligent Systems*, 35(4):70–82, 2020.

[29] Md Palash Uddin, Yong Xiang, Xuequan Lu, John Yearwood, and Longxiang Gao. Mutual information driven federated learning. *IEEE Transactions on Parallel and Distributed Systems*, 32(7):1526–1538, 2021.

[30] Hongtao Wang, Ang Li, Bolin Shen, Yuyan Sun, and Hongmei Wang. Federated multi-view spectral clustering. *IEEE Access*, 8:202249–202259, 2020.

[31] Wei Liu, Li Chen, Yunfei Chen, and Wenyi Zhang. Accelerating federated learning via momentum gradient descent. *IEEE Transactions on Parallel and Distributed Systems*, 31(8):1754–1766, 2020.

[32] M. Fernandes, C. Silva, J. Arrais A. Cardoso, and B. Ribeiro. Decay momentum for improving federated learning. In Michel Verleysen, editor, *Proceedings of the 29. European Symposium on Artificial Neural Networks ESANN 2021*, page numbers to be obtained from ToC of this proceedings book, Evere, Belgium, 2021. D-Side Publications.

[33] J. Brinkrolf and B. Hammer. Federated learning vector quantization. In Michel Verleysen, editor, *Proceedings of the 29. European Symposium on Artificial Neural Networks ESANN 2021*, page numbers to be obtained from ToC of this proceedings book, Evere, Belgium, 2021. D-Side Publications.

[34] C.M. Bishop. *Neural networks for pattern recognition*. Oxford University Press, USA, 1995.

[35] Mikhail Yurochkin, Mayank Agarwal, Soumya Ghosh, Kristjan H. Greenewald, Trong Nghia Hoang, and Yasaman Khazaeni. Bayesian nonparametric federated learning of neural networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 7252–7261. PMLR, 2019.

[36] M. Polato, A. Gallinaro, and F. Aiolli. Privacy-preserving kernel computation for vertically partitioned data. In Michel Verleysen, editor, *Proceedings of the 29. European Symposium on Artificial Neural Networks ESANN 2021*, page numbers to be obtained from ToC of this proceedings book, Evere, Belgium, 2021. D-Side Publications.

[37] Stephen Hardy, Wilko Henecka, Hamish Ivey-Law, Richard Nock, Giorgio Patrini, Guillaume Smith, and Brian Thorne. Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption, 2017.

[38] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[39] Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H. Yang, Farhad Farokhi, Shi Jin, Tony Q. S. Quek, and H. Vincent Poor. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15:3454–3469, 2020.

[40] Timothy Yang, Galen Andrew, Hubert Eichner, Haicheng Sun, Wei Li, Nicholas Kong, Daniel Ramage, and Françoise Beaufays. Applied federated learning: Improving google keyboard query suggestions. *CoRR*, abs/1812.02903, 2018.

[41] David Leroy, Alice Coucke, Thibaut Lavril, Thibault Gisselbrecht, and Joseph Dureau. Federated learning for keyword spotting, 2019.

[42] L. Pellegrini, V. Lomonaco, G. Graffieti, and D. Maltoni. Continual learning at the edge: Real-time training on smartphone devices. In Michel Verleysen, editor, *Proceedings of the 29. European Symposium on Artificial Neural Networks ESANN 2021*, page numbers to be obtained from ToC of this proceedings book, Evere, Belgium, 2021. D-Side Publications.

[43] Florian Hartmann, Sunah Suh, Arkadiusz Komarzewski, Tim D. Smith, and Ilana Segall. Federated learning for ranking browser history suggestions, 2019.

[44] Yang Liu, Anbu Huang, Yun Luo, He Huang, Youzhi Liu, Yuanyuan Chen, Lican Feng, Tianjian Chen, Han Yu, and Qiang Yang. Fedvision: An online visual object detection platform powered by federated learning, 2020.

[45] Li Huang, Andrew L. Shea, Huining Qian, Aditya Masurkar, Hao Deng, and Dianbo Liu. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Journal of Biomedical Informatics*, 99:103291, 2019.

[46] Zhaoping Xiong, Ziqiang Cheng, Chi Xu, Xinyuan Lin, Xiaohong Liu, Dingyan Wang, Xiaomin Luo, Y. Zhang, Nan Qiao, M. Zheng, and Hualiang Jiang. Facing small and biased data dilemma in drug discovery with federated learning. *bioRxiv*, 2020.

[47] Shaoqi Chen, Dongyu Xue, Guohui Chuai, Qiang Yang, and Qi Liu. Fl-qsar: a federated learning based qsar prototype for collaborative drug discovery. *bioRxiv*, 2020.

[48] Suyi Li, Yong Cheng, Y. Liu, Wei Wang, and Tianjian Chen. Abnormal client behavior detection in federated learning. *ArXiv*, abs/1910.09933, 2019.

[49] Micah J. Sheller, G. Anthony Reina, Brandon Edwards, Jason Martin, and Spyridon Bakas. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. *CoRR*, abs/1810.04304, 2018.

[50] Wenqi Li, Fausto Milletarì, Daguang Xu, Nicola Rieke, Jonny Hancox, Wentao Zhu, Maximilian Baust, Yan Cheng, Sébastien Ourselin, M. Jorge Cardoso, and Andrew Feng. Privacy-preserving federated brain tumour segmentation. In Heung-Il Suk, Mingxia Liu, Pingkun Yan, and Chunfeng Lian, editors, *Machine Learning in Medical Imaging*, pages 133–141, Cham, 2019. Springer International Publishing.

[51] P. Kairouz and H. B. McMahan et al. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1-2):1–210, 2021.

[52] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao. A survey on federated learning. *Knowledge-Based Systems*, 216, 2021.

[53] Mohammed Aledhari, Rehma Razzak, Reza M. Parizi, and Fahad Saeed. Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, 8:140699–140725, 2020.

[54] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.