Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 1, January 2025, pp. 288~299
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i1.pp288-299

A conceptual approach of optimization in federated learning

Farhanna Mar'i[1], Ahmad Afif Supianto[2,3]
[1]Department of Informatics, Faculty of Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia
[2]Department of ICT and Natural Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[3]Research Center for Artificial Intelligence and Cyber Security, National Research and Innovation Agency (BRIN), Jakarta, Indonesia

Article Info

Article history:

Received May 23, 2024
Revised Sep 3, 2024
Accepted Sep 7, 2024

Keywords:

Conceptual framework
Decentralized data
Federated learning
Machine learning
Optimization

ABSTRACT

Federated learning (FL) is an emerging approach to distributed learning from decentralized data, designed with privacy concerns in mind. FL has been successfully applied in several fields, such as the internet of things (IoT), human activity recognition (HAR), and natural language processing (NLP), showing remarkable results. However, the development of FL in real-world applications still faces several challenges. Recent optimizations of FL have been made to address these issues and enhance the FL settings. In this paper, we categorize the optimization of FL into five main challenges: communication efficiency, heterogeneity, privacy and security, scalability, and convergence rate. We provide an overview of various optimization frameworks for FL proposed in previous research, illustrated with concrete examples and applications based on these five optimization goals. Additionally, we propose two optional integrated conceptual frameworks (CFs) for optimizing FL by combining several optimization methods to achieve the best implementation of FL that addresses the five challenges.

Corresponding Author:

Farhanna Mar'i
Department of Informatics, Faculty of Engineering, Universitas Negeri Surabaya
Surabaya 61256, East Java, Indonesia
Email: farhannamari@unesa.ac.id

1. INTRODUCTION

Machine learning (ML) is revolutionizing various domains, driven by its ability to analyze vast amounts of data and derive actionable insights. As a key area within artificial intelligence (AI), ML requires substantial data inputs to effectively train models and make predictions. In the pratical applications of federated learning (FL), particularly within the retail industry, ML leverages consumer data to understand preferences, behaviors, and needs [1]–[5]. Traditionally, data has been collected from targeted devices and processed centrally on cloud-based servers or data centers. This data is then used to develop knowledge and inference models, serving as training for ML algorithms. Nevertheless, those approaches were no longer effective because data is certainly sensitive for the owner, and several policies had restricted access to consumer data without consent, especially if it is shared with third parties. One example of the restriction is the publication of data privacy and security policy initiated by The European Union in the general data protection regulation (GDPR) in May 2018 [6]. This policy sets boundaries for ML process to get data input which is an important point in its technical performance. In addition, centralized data collection in the data center will also be very burdensome for the learning process due to limited center resources. It would be better if consumer devices could run their learning process, because, in today's modern era, the devices were equipped with resources for high computing, advanced sensors, and excellent communication capabilities [7]. Much research has been done to overcome those problems and turn them into opportunities for a new era of ML development using a distributed computing approach called FL [8]–[10].

FL is an emerging and recent approach in ML which uses a setting where many clients collaboratively train a model using decentralization data [11]. These approaches also support information privacy and security regarding its policy to train customers' data within their mobile devices, not stored on the server [12]–[15]. Moreover, FL provides the ability to protect device privacy while ensuring high learning performance and also plays an important role in 5G mobile applications that are sensitive to computing privacy, such as edge and catching computing, networking, and spectrum management [16]. Based on all the advantages of this new approach, FL has gained widespread attention from practitioners and researchers in vast industries. In contrast, FL has been applied to several tools, such as Google Keyboard (GBoard) [17], Hey Google [18], and Hey Siri [19]. In addition, FL has also been applied to institutions that hold extra sensitive data, such as medical imaging [20], Finance Space by WeBank for money laundering detection [21], also financial fraud detection by intel and consilient [22].

FL conducts the model training process locally and then communicates with the server to perform ML through model aggregation. Because, what is sent to the server is only the aggregation result, not the raw data from the user. Despite its goals, these settings had several challenges, and the main challenge was the cost of communication round between a single device and the server to conduct the model's aggregates. Practically, the communication round cost is arguably more time-consuming than an iteration of the algorithm itself [23]. In addition, another challenge is that heterogeneous devices cause imbalanced and unstable learning and communication processes. The system's complexity is also one of the important challenges FL faces. FL was first introduced to support large-scale distributed training and security and privacy concerns. The first algorithm that has been used for FL is federated averaging (FedAvg), which combines stochastic gradient descent (SGD) on each customer with a server that executes model averaging [24]. Those research has succeeded in reducing the round of communication needed to train ML on distributed data and deal with unbalanced and heterogeneous data and devices. FL differs from common distributed learning because the distribution of the data is not uniform and independent (non-IID). The use of standard FL methods such as FedAVG was often difficult to tune and demonstrated its inability to achieve convergence [25], [26]. So it is necessary to optimize the standard method to improve the performance of FL.

However, in the practice of implementing FL in many cases, there were several things that could have been improved. Some of these problems have been tried to be solved by performing optimization techniques using several methods based on the goals or objectives to be achieved. In such a way, several studies have carried out this optimization technique with the intent that FL has shown better performance. Several types of research have been conducted to optimize FL in actual cases. One of those is the use of reinforcement learning to overcome the heterogeneity of the data sample that is not independent and identically distributed (IID) along with a framework, namely FAVOR. These frameworks could intelligently choose the customer's device to participate in each stage of the distribution process and have been proven to speed up convergence [27]. In addition, there is an adaptive optimizer version called ADAGRAD, ADAM, and YOGI to overcome the weakness of FedAVG. Adaptive optimizers have succeeded in solving problems in non-federated settings, but they also could be used in FL [25]. Another work of optimization in FL is FedNL, the Federated Newton Learn the method that also proved to reduce communication cost and complexity compared to the critical baselines [28].

The optimization of FL itself is the principal research to support the development and advancement of FL techniques and models. Based on the new approach introduced for optimization in FL, many frameworks have been proposed to overcome several problems, such as Communication Efficiency [27]–[31], Heterogeneity [32], Scalability [24], [28], [29], Privacy and Security [30], [32], [33], and Convergences rate [25], [31], [34], mix problems [35]–[37], so that it needs to be specified and categorized. In previous research, there is no discussion of the various optimization methods in FL, instead, the focus is only on one or two problems. Our research will focus on those gaps, aimed to classify the framework, and propose an integrated conceptual framework based on all of the proposed frameworks of the optimization in FL to get the best implementation of FL. In summary, we present the following contributions to this research, (1) we define the concept and introduce a typology of FL, (2) we classified the proposed framework of optimization in FL, and (3) we proposed an integrated conceptual framework for optimization in FL. This research is structured into four sections. Section 1 is about the background and the related research. Section 2 illustrates the methods, section 3 defines the results and discussion, we classifies the framework of optimization in FL, and we proposed an integrated conceptual framework for Optimization in FL. Finally, the main findings are outlined in section 4.

2. METHOD

The development of conceptual frameworks (CFs) in Computer Science represents a set of concepts for articulating requirements and examining the system architectures [38]. The CFS developed in this research is an optimization for FL based on models development. Hence, a multi-case study approach is used to evaluate the optimization of FL. The research methodology consists of several key steps to address and optimize FL by systematically reviewing and analyzing existing literature and frameworks. The approach can be detailed as follows and visualized in Figure 1.

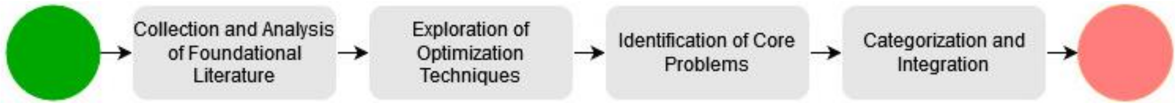

Figure 1. Research methods

A. Collection and analysis of foundational literature

Article collection: Gather a comprehensive set of articles that define FL and outline its associated weaknesses. This involves selecting research papers, reviews, and surveys that specifically focus on the core concepts and challenges of FL.

B. Exploration of optimization techniques

Optimization review: Conduct an extensive review of publications discussing various optimization techniques aimed at addressing the weaknesses identified in FL. This involves studying different optimization strategies and methods proposed in the literature to solve FL's specific issues.

C. Identification of core problems

Problem identification: Analyze the gathered research to identify and categorize the primary problems associated with FL.

D. Categorization and integration

- Categorization of optimization techniques: Classify the various optimization methods according to the five identified weaknesses. This step involves mapping each optimization technique to the specific problems it addresses.
- Framework integration: Develop a comprehensive framework by combining the identified optimization methods. The goal is to create a unified approach that addresses all five problems within a single framework, thereby enhancing FL's overall performance.

By following these steps, the research aims to develop an integrated framework that effectively resolves the core issues of FL through a systematic combination of existing optimization techniques.

3. RESULTS AND DISCUSSION

The research begins with a domain analysis to build a concept based on literature reviews in several digital publications, such as Science Direct, Springer, Arxiv.org, ACM, and IEEE. Previously, the sustainability of research in FL will be shown based on the number of articles combination on reviews/survey, original research, book chapters, and proceedings that were published using keyword FL starting in the year of FL was introduced from 2016 to 2023 to show an increasing interest in the topic of FL in Figure 2 starting from August 2, 2024. Based on the Figure 2, a significant increase occurs every year, which indicates that FL is a popular and highly developed topic. In the domain analysis stage, we first searched for articles that included FL in the title and keywords. Then, we selected papers that thoroughly defined FL and concluded the definition of FL based on its explanation. The definition study showed FL's main perspectives and concepts, especially in the internet of things (IoT), wireless communication, and future generation computer systems. Based on this analysis, we also have to build a typology of optimization in FL.

3.1. What is federated learning?

The term FL, first introduced by McMahan *et al.* [39], is one of the ML settings in which several participated devices or clients solve the model learning task and are coordinated by a server. Client's devices process their own local model with task script instruction from the central server and then send it back. Next, the central server does the aggregation process using the local training model. Finally, the central server will send the global model information to its client. In the practical application, there are two settings: cross-device FL, which targets ML for mobile devices, and cross-silo FL, which targets ML for several organizations [14]. An overview of the FL framework is presented in Figure 3. This set was made as a follow-up application of several policies related to consumer data security. Thus, the data training process on ML will be carried out locally on the client's mobile device rather than sending it to the data center for training [40]. The use of local data in reducing the risk of attacks on the server and the use of focused and temporary data minimization must be done in today's era cause technological developments go hand in hand with increasingly sophisticated levels of cyber crime [41], updates and initial aggregation follows the principle of data minimization [14], which was proposed by the 2012 White House report on the privacy of consumer data [42]. Consumers trust companies to share essential data related to public or private matters so that companies can use the data to customize consumers' preferences. Thus, the level of consumer satisfaction is met. Hence, companies must ensure that customer data remains safe and avoid these attacks. FL is needed to fulfill this aspect in line with the popularity of implementing the IoT, wireless communication, and future generation computer systems. One of the future areas of FL implementation is the IoT [43]. IoT requires an incredible amount of devices to complement each other's intended use between one device and. Likewise, the data generated from these devices will also be overloaded if deposited on a centralized server. A recent forecast estimates that 41.6 billion connected IoT devices will generate 79.4 zettabytes (ZB) of data in 2025 [44]. As a technique to overcome the phenomena mentioned above, FL is the best way to overcome them. In the first stage, the user’s device independently processes local computation iteratively using their datasets. The result of its process, that is, the local learning model updates, would be sent to the aggregation server for the global strategy. In this setting, aggregation can be done both at the edge or cloud [45].

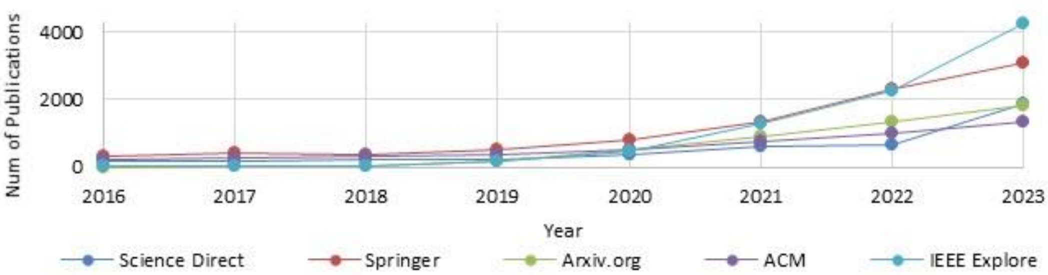

Figure 2. Number of articles published in 2016-2023

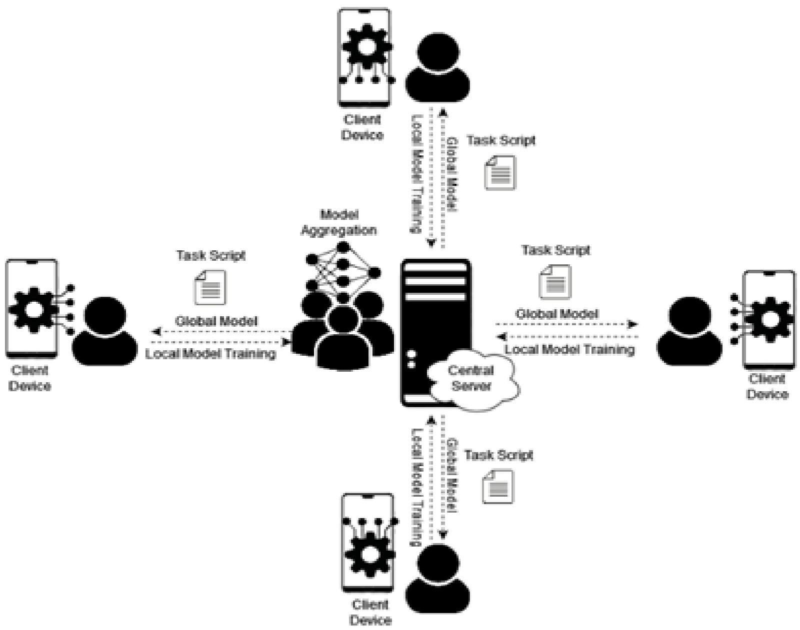

Figure 3. An overview of FL

3.2. Optimization in federated learning

FL problem can be denoted as in (1). The first algorithm used to solve the FL problem is inspired by SGD, which is the favored method for optimization.

$$\min_{x \in R^d}$$

$$\left\{f(x) \coloneqq \frac{1}{n}\sum_{i=1}^{n} f_i(x)\right\} \quad (1)$$

Information:
d = dimension of the model $x \in R^d$
n = total number of devices
$f_i(x)$ = the loss
$i \in [n]$ = data stored in machine
$f(x)$ = the empirical loss

SGD is applied to the federated optimization problem with an adaptation that is the selection of a C-fraction of clients on each round, then computing the gradient of the loss over all the data owned by these clients, so C controls the global batch size, with C=1 suitable to full-batch gradient descent, this called FedSGD [46]. However, optimization of FL is needed to get the best FL model that can be applied according to the main purpose of forming the FL model. The main objectives of the various optimizations that exist vary according to the needs of each application. Here we categorize several types of optimizations based on the main needs to be achieved.

3.2.1. Communication efficiency

FL requires communication costs between local clients and servers to send and update the aggregation model in the learning process. So that minimizing communication costs is an essential requirement to do considering FL depends on the communication relationship. Several approaches have been taken to minimize communication costs, including using a method of uploading the gradient located in a particular interval clipped by some threshold values [47], scaling down the used parameters by minimizing the complexity of the neural network models [29], also structured and sketched updates [40]. McMahan *et al.* [39] proposed FedAVG that adds more computation to each client by iterating the local update multiple times before the averaging step. In his work, FedAVG compared with SGD and FedSGD used in a large-scale next-word prediction and shows that FedAVG success trains high-quality models using relatively fewer rounds of communication than other algorithms. Many variants of FedAVG have been used in several types of research, including [48]–[52].

Zhu and Jin [53] proposed modified sparse evolutionary training (SET) algorithm to reduce the communication cost without reducing global learning accuracy. These approach optimizes deep neural network models by modifying the operator by removing operations at the last training epoch. The result shows that the modified SET Algorithm succeeded in maximizing the learning performance and minimizing the communication cost by encoding only two hyperparameters. Zhu and Jin [54] proposed FedNet2Net, a novel scheme based on a model growing with a modified training scheme. This approach used two transformations, which are called Net2Widernet and Net2DeeperNet. Then, each step, the model was developed after enhancing utilizing a functional conserving transformation on it. The result shows that FedNet2Net had an excellent performance than other algorithms like FedAVG, Federated Dropout [55], and HeteroFL [56]. In addition, communication complexity also could be reduced by constructing the training phase of FL models using personalization.

Hanzel, P. and Richtarik, P. proposed several efficient variants of SGD with and without partial participation and variance reduction [57]. Those approaches use a new optimization formulation to learn a mixture of the global model, as shown in (2).

$$\min_{x_1,\ldots,x_n \in \mathbb{R}^d} \{F(x) \coloneqq f(x) + \lambda\, \psi(x)\} \quad (2)$$

$$f(x) \coloneqq \frac{1}{n}\sum_{i=1}^{n} f_i(x_i),\ \psi(x) \coloneqq \frac{1}{2n}\sum_{i=1}^{n} \|x_i - \bar{x}\|^2$$

Information :
$\lambda \geq 0$ = *penalty parameter*
$x_1,\ldots, x_n \in \mathbb{R}^d$ = *local models*
$x \coloneqq (x_1, x_2, \ldots, x_n)$, $\bar{x} \coloneqq \frac{1}{n}\sum_{i=1}^{n} x_i$ = *the average of the local model*

3.2.2. Heterogeneity

One of the proponents of the emergence of FL is that client privacy must be maintained on the data used as modeling material in ML. However, FL standards are still vulnerable to privacy and confidentiality leakage when handling heterogeneous data from multiple sources [58]. Gao *et al.* [58] proposed heterogeneous federated transfer learning (HFTL), an approach to eliminate covariate shifts of homogeneous feature spaces and pass over different data owners' heterogeneous feature spaces using end-to-end secure multi-party learning protocols. The result shows that HFTL outperforms the local model and homogeneous FL under challenging condition settings. However, heterogeneity data between unique clients also affect model performance like accuracy reduction [11], [51], [59]. For instance, multiple branches of a bank would like to collaboratively train a customer's behavior used FL. Each bank collects customer data independently, resulting in unbalanced and heterogeneous datasets (i.e., non Identically Independently Distributed (Non-IID). One bank might not have data samples regarding a specific behavior categories. When each bank trains a model locally, the local goal may differ from the global goal. As a result, the global average model is not equal to the global objective, and this case is called model drift which causes poor model performance [60]–[63]. In addition, data heterogeneity also causes accuracy parity which is the case when a model biases regarding the unique condition in each local subject, for example, geographical and race discrimination [64]. Zhou *et al.* [64] proposed a new design setting for FL using global-local Knowledge Fusion (FedKF) to overcome both challenges in heterogeneity data, such as model drift and accuracy parity. FedKF's concept is to assign the server the task of returning global knowledge to guide the local client training, so that the local model is regularized by the global optima, which degrades the client model drift case in each training round. T1 is designed as an active-inactive model aggregation for the global model on the server side, and T2 is designed to support knowledge fusion on the client side. In the FedKF setting, the central server sets K different cache slots for saving the latest local models. In every training round, the selected clients had to upload their local models to the central server.

The algorithm of FedKF mention as follows:

1. **phase t-1 :** The central server aggregates active clients by uploading local models to obtain active client aggregates (ACA), whereas FedKF aggregates all clients cached models to obtain overall client aggregates (OCA).
2. **phase t :** A portion of the general clientele is chosen to be active clients $\{a_1, \dots, a_m\}$. τ is the rate of selection that follows that $m = \tau K$. All active clients receive both the ACA and OCA models from the central server.
3. Each active client determines the ACA models as w_s (student model) and OCA models as w_T (teacher model)
4. Uses data-free Knowledge Distillation to transfer knowledge from the teacher to the student models. Meanwhile, the student model is trained using the local dataset of each active client. All global knowledge is combined and transmitted to the student model.
5. All active clients upload their local student model as the updated local model to the central server.
6. The weight in the cache slots is updated by the central server. If the model has been fully trained, either exit or proceed to step 2.

Based on the study, The result is FedKF had suceed to be a better solution in heterogeneous agnostic FL with high model performance, privacy-preserving, and fairness [64].

3.2.3. Scalability

The scalability of FL could be calculated using several vital factors, like the number of overlapping samples, also the dimension of hidden representation, and a total of the features [65], the System's ability to handle the number of client devices [66], the amount of learning on large volumes of data [67], accuracy, computational, storage, and network overhead. Han and Han [35] proposed DeFL: Decentralized Weight Aggregation for Cross-Silo FL with two key ideas that are reducing local updates in the server by collecting all the local updates itself, then weight that's only in the current training are maintained and synchronized so the network and the storage overhead could be reduced [35]. In the DeFL approach abstracts, each participating node has different roles: client and replica. A client aggregates correct weight from actual nodes with a Multi-Krum filter [68]. The result shows that DeFL had successfully defended the standard threat models with the best performance and produced efficient storage and network usage.

3.2.4. Privacy and security

Privacy and Security are the advantages highlighted in FL because the training data is not sent to the server but remains on each client device and runs the training process independently. However, in reality, privacy still cannot be guaranteed since training models from users will still be sent to servers for global aggregation [69]. The trained models shared on the server could lead to privacy attacks that are model inversion [70]. To overcome this problem, an optimization framework on FL was developed using secure aggregation (SA) so that the server only learn the global model update not the individual model update [69].

SA is a part of secure multi-party computation algorithms that hold each private value and collaborate to compute an aggregate value without knowing other information about its value but only what is learnable from the aggregate value itself [71]. Bonawitz *et al.* [71] proposed a SA protocol consisting of four rounds: Share Keys, masked input collection, consistency check, and unmasking in a synchronous network. A variant of SA called secure indexing also had been proposed to guarantee privacy and security [72].

In addition, since the trained model was sent to the server, the system could get poisoning model attacks by malicious clients. It was challenging to identify the malicious clients because no local client data was available on the server [36]. Ma *et al.* [36] proposed a model update aggregation (MUB) to defend against the threat of byzantine attacks, which are additive noise and sign-flipping attacks. This research proved could enhance both privacy and security using a combination of MUB and initial client model initialization (ICMI).

3.2.5. Convergences rate

Training processes in FL need to be fast and accurate because there are several types of training batches on-device that should be finished to get the best result. Most of the research had focused on designing the best aggregation strategies for improving convergence rates, but Qiu *et al.* [73] had different views and proposed ZeroFL, which is the framework that is devoted to accelerating on-device training by using highly sparse operations. There are three strategies used in ZeroFL for doing local sparsification to produce both performance improvisation and reduce communication costs. This approach believes that not all of the weights need to be transferred to the central server for the aggregation and uses local sparsification instead. Three local sparsifications are used: Top-K-Weights, Diff on Top-K-Weight, and Top-KWeights Diff. The result shows that ZeroFL enhances the accuracy performance by 1.5% while reducing 1.9 x uplink communication [73]. ZeroFL also shows the best performance in Non-IID Data, even though these types of data cause slow model training and enforce additional communication rounds for FL to converge [74]. Wu, Hong, Wang, and Ping proposed federated adaptive weighting (FedAdp) that also accelerates model convergences under the pressure of non-IID data [74]. The main idea of FedAdp is to assign a different weight for different models at each communication round based on smoothed angle adaptively. There are two steps used in this setting: A non-linear mapping function using a variant of the Gompertz function [75] and Weighting using the Softmax function [74]. FedAdp succeeded in getting better performance compared with FedAvg by reducing the communication rounds by up to 54.1% on the MNIST dataset and up to 45.4% on the FashionMNIST dataset.

3.2.6. A conceptual framework of optimization in federated learning

Based on the optimization carried out in previous studies, it is proven to be able to improve FL performance compared to pre-existing standard algorithms. So if the optimization methods based on these different objectives are put together, it can be guaranteed that FL's performance will be even better and more perfect. In order to prove that optimization in FL can also be carried out on more than one approach, several other studies have also been conducted such as optimization in federated principal components analysis (PCA) based on Grassmann to cover privacy concern and the scalability of the limited device for anomaly detection in IoT Networks [71]. Those research yields drastically reducing the analysis time of the system. Not only that, privacy concern also be optimized with convergences rate when the clients had been clustered into a pervasive social connection between users. Phong *et al.* [76] proposed a novel social-aware clustered federated learning (SCFL) to achieve the best performance of FL without sacrificing the model because the noise given for privacy concern in differential privacy. Efficient FL could be achievable when the heterogeneity and the convergence rate had been provided with the best performances. One of the characteristics of non-IID data is statistical heterogeneity that could be solved with SGD in traditional centralized learning [77], But in the FL setting statistical heterogeneity had become a severe problem that SGD could not solve and yields degradation of model performances [78]. So it cannot be denied that solving heterogeneity problems will simultaneously increase the performance model as the convergence rate.

FedCG [79] is also one of the frameworks that prove Efficient FL could be developed with more than one goal of optimizations. FedCG using adaptive client selection and gradient compression. This approach selects a representative client regarding statistical heterogeneity, and after training locally, compressed model updates would be uploaded with matching their capabilities to the parameter server for the aggregation phase. In a mobile wireless device, a spectrum allocation optimization was done to enhance FL, minimize time consumption, and ensure fast convergences by implementing a robust device selection [80]. FedHP is also proposed to overcome heterogeneity as the critical challenge on FL, which integrates an adaptive control of local updating frequency and the network topology [81]. FedHP successfully reduced the time competition by about 51% and increased model accuracy up to 5% in heterogeneous conditions. Furthermore, the previous research, which requires re-weights local updates, causes other challenges that as poorer optima of the client model when it is in more heterogeneous conditions [82]. FedSkip is the proposed method by Fan *et al.* [82] which could improve the client optima by systematically skipping FedAvg and spreading local models to the cross devices. In addition to solving statistical heterogeneity problems and improving model performance, FedSkip has also reduced communication rounds in cases with many clients, for example, on CIFAR-100 data. Thus, to find an effective and efficient combination of optimizations in FL, thorough testing of all types of optimizations found in previous studies is required, so that optimization in FL will be an up-and-coming research.

Based on the literature studies that have been carried out on several optimization approaches to FL, we have categorized into several sub-categories based on the optimization objectives achieved, which are Communication efficiency, heterogeneity, scalability, privacy and security, also convergences, rate. In this section, we propose a conceptual framework based on those optimizations in the previous studies. The proposed conceptual framework is based on the implementation of optimization in two objects, that is, the server and client sides. In detail, the proposed conceptual framework of optimization FL in this study is depicted in Figure 4. The new form in optimization for FL to optimize performance using several approaches from the five aspects of FL optimization needs. Besides that, the framework also divides the optimization into two sides: clients and servers, based on the implementation of optimization methods. In practice, this framework is possible to run, considering that each optimization approach has its algorithm to improve different optimization aspects so that if all of these optimization strategies are implemented, it will produce a perfect FL setting. There are two optional frameworks of optimization in FL, as depicted in Figure 4. In communication efficiency, the SET algorithm is executed for the first time, which is a setting for the deep neural network used in the learning process, then FedNet2Net optimization is also executed in the client area as well as in the learning stage by providing a condition for the preserving transformation function. After that, to overcome the problem of heterogeneity, there are two options that we could choose between HFTL using an end-to-end secure multi-party learning protocol and FedKF on two sides, namely servers and clients using knowledge fusion.

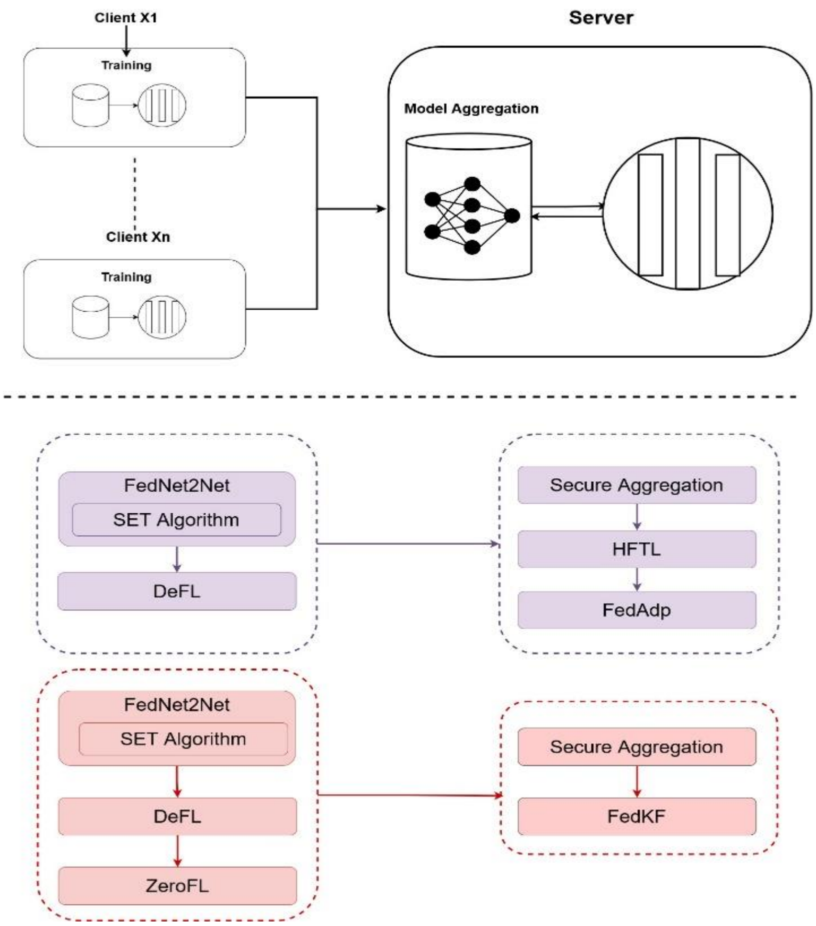

Figure 4. Conceptual framework of optimization in FL

FL was designed to address the privacy and security of data used during the ML process. However, in practice, FL still faces security threats because, after all, the local training results must still be sent to the server for further global aggregation, in handling FL as well can be optimized to handle privacy and security using a secure aggregation protocols model. In addition, the fundamental problem for FL is scalability which can also be optimized to get the best performance using decentralized weight aggregation (DeFL). Then, no less important is the issue of convergence rate because FL carries out the training process separately. In general, the resulting convergence results are definitely lower, so this is also an important challenge for FL, one of the optimization strategies that can be implemented is by implementing ZeroFL using a highly sparse operation to accelerate on-device training. Apart from that, in order to optimize the convergence level better, FedAdp operations can also be added using Adaptive Weighting by assigning different weights to each different model. Based on the proposed framework of optimization in FL, we can conduct experiments to prove how the performance of various optimization methods is carried out simultaneously. So hopefully, we can find the best form of optimization from FL.

4. CONCLUSION

In many practical applications of FL, various optimization techniques have been employed to address these issues, utilizing different methods depending on the specific goals or objectives. Optimizing FL is a key area of research that drives the advancement and enhancement of FL techniques and models. This research has analyzed how optimization can be made FL performs better. Our research indicates that, according to current studies, the optimization of FL is divided into five functions of optimization, that are reducing communication cost between customer and data center, solving heterogeneity, privacy and security concerns, overcoming scalability, and increasing convergences rate. This research has proposed a conceptual framework for designing and implementing optimization in FL based on the successful optimization used in previous studies with the hope that it could be used to optimize FL while addressing five main problems. There are two optional frameworks for optimization in FL. In terms of communication efficiency, the SET algorithm is first executed, which sets up the deep neural network used in the learning process. Then, FedNet2Net optimization is performed both on the client side and during the learning stage, with a condition for maintaining the transformation function. To address the issue of heterogeneity, there are two options to choose from: HFTL using an end-to-end secure multi-party learning protocol, or FedKF, which involves Knowledge Fusion on both servers and clients. For future works, it is necessary to prove whether the combination of optimizations proposed in this research can improve the performance of FL.

REFERENCES

[1] P. Chou, H. H. C. Chuang, Y. C. Chou, and T. P. Liang, “Predictive analytics for customer repurchase: Interdisciplinary integration of buy till you die modeling and machine learning,” *European Journal of Operational Research*, vol. 296, no. 2, pp. 635–651, 2022, doi: 10.1016/j.ejor.2021.04.021.

[2] E. Samunderu and M. Farrugia, “Machine learning with applications predicting customer purpose of travel in a low-cost travel environment — a machine learning approach,” *Machine Learning with Applications*, vol. 9, no. April 2021, p. 100379, 2022, doi: 10.1016/j.mlwa.2022.100379.

[3] Z. Xu, G. Zhu, N. Metawa, and Q. Zhou, “Machine learning based customer meta-combination brand equity analysis for marketing behavior evaluation,” *Information Processing and Management*, vol. 59, no. 1, p. 102800, 2022, doi: 10.1016/j.ipm.2021.102800.

[4] G. Pangestu, F. Mar’I, F. A. Bachtiar, and F. I. Maulana, “The heron formula approach for head movements detection for student focus detection during pandemic online class,” *Proceedings of 2021 International Conference on Information Management and Technology, ICIMTech 2021*, no. August, pp. 678–683, 2021, doi: 10.1109/ICIMTech53080.2021.9535065.

[5] F. Mar’i and A. A. Supianto, “Clustering credit card holders based on bill payments using improved K-means with particle swarm optimization,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 6, pp. 737–744, 2018, doi: 10.25126/jtiik.20185858.

[6] GDPR, “General data protection regulation (GDPR) – official legal text,” *General Data Protection Regulation*. pp. 1–99, 2018.

[7] W. Y. B. Lim *et al.*, “Federated learning in mobile edge networks: a comprehensive survey,” *IEEE Communications Surveys and Tutorials*, vol. 22, no. 3, pp. 2031–2063, 2020, doi: 10.1109/COMST.2020.2986024.

[8] D. Upreti, H. Kim, E. Yang, and C. Seo, “Defending against label-flipping attacks in federated learning systems using uniform manifold approximation and projection,” *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 459–466, 2024, doi: 10.11591/ijai.v13.i1.pp459-466.

[9] M. Chen and D. K. Halim, “Federated learning for scam classification in small Indonesian language dataset: an initial study,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 1, pp. 325–331, 2023, doi: 10.11591/ijeecs.v30.i1.pp325-331.

[10] M. N. Meqdad, A. H. Hussein, S. O. Husain, and A. M. Jawad, “Classification of electrocardiogram signals based on federated learning and a gaussian multivariate aggregation module,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 936–943, 2023, doi: 10.11591/ijeecs.v30.i2.pp936-943.

[11] P. Kairouz *et al.*, “Advances and open problems in federated learning,” *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021, doi: 10.1561/2200000083.

[12] D. Li, Z. Luo, and B. Cao, “Blockchain-based federated learning methodologies in smart environments,” *Cluster Computing*, vol. 0123456789, 2021, doi: 10.1007/s10586-021-03424-y.

[13] R. C. Geyer, T. Klein, and M. Nabi, “Differentially private federated learning: a client level perspective,” no. Nips, pp. 1–7, 2017, doi: 10.48550/arXiv.1712.07557.

[14] J. Wang *et al.*, “A field guide to federated optimization,” *arXiv preprint arXiv:2107.06917,* pp. 1–88, 2021, doi: 10.48550/arXiv.2107.06917.

[15] Y. M. Saputra and G. Alfian, "Privacy aware-based federated learning framework for data sharing protection of internet of things devices," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 2, pp. 979–985, 2023, doi: 10.11591/ijeecs.v31.i2.pp979-985.
[16] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Blockchain for 5G and beyond networks: A state of the art survey," *Journal of Network and Computer Applications*, vol. 166, no. April, p. 102693, 2020, doi: 10.1016/j.jnca.2020.102693.
[17] A. Hard *et al.*, "Federated learning for mobile keyboard prediction," 2018, doi: 10.48550/arXiv.1811.03604.
[18] Google, "Your voice & audio data stays private while Google Assistant improves - Google Assistant Help," 2021.
[19] Apple, "Designing for Privacy - WWDC 2019 - Videos - Apple Developer," 2019.
[20] E. Darzidehkalani, M. Ghasemi-rad, and P. m. a. van Ooijen, "Federated learning in medical imaging: Part II: methods, challenges, and considerations," *Journal of the American College of Radiology*, 2022, doi: 10.1016/j.jacr.2022.03.016.
[21] WeBank, "Utilization of FATE in anti money laundering through multiple banks," p. 2020, 2020 [online]: https://www.fedai.org/cases/utilization-of-fate-in-anti-money-laundering-through-multiple-banks/
[22] Intel, "Intel and consilient join forces to fight financial fraud with AI," [online]: https://www.intel.com/content/www/us/en/newsroom/news/fight-financial-fraud-ai.html#gs.f4po72
[23] J. Konečný, B. McMahan, and D. Ramage, "Federated optimization: distributed optimization beyond the datacenter," *arXiv preprint arXiv:1511.03575*, no. 1, pp. 1–5, 2015, doi: 10.48550/arXiv.1511.03575.
[24] E. Jeong, S. Oh, H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Communication-efficient on-device machine learning: federated distillation and augmentation under non-IID private data," no. Nips, 2018, doi: 10.48550/arXiv.1811.11479.
[25] S. Reddi *et al.*, "Adaptive federated optimization," *arXiv preprint arXiv:2003.00295,* no. 2, pp. 1–38, 2020, doi: 10.48550/arXiv.2003.00295.
[26] F. Mar'i, A. A. Supianto, and F. A. Bachtiar, "Comparison of federated and centralized learning for image classification," *Computer Science Research, Embedded Systems and Logic*, vol. 11, no. 225, pp. 393–400, 2023, doi: 10.33558/piksel.v11i2.7367.
[27] H. Wang, Z. Kaplan, D. Niu, and B. Li, "Optimizing federated learning on Non-IID data with reinforcement learning," *Proceedings - IEEE INFOCOM*, vol. 2020-July, pp. 1698–1707, 2020, doi: 10.1109/INFOCOM41043.2020.9155494.
[28] M. Safaryan, R. Islamov, X. Qian, and P. Richtárik, "FedNL: making newton-type methods applicable to federated learning," *arXiv preprint arXiv:2106.02969,* 2021, doi: 10.48550/arXiv.2106.02969.
[29] H. Zhu and Y. Jin, "Multi-objective evolutionary federated learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 4, pp. 1310–1322, 2020, doi: 10.1109/TNNLS.2019.2919699.
[30] R. Kanagavelu *et al.*, "CE-Fed: communication efficient multi-party computation enabled federated learning," *Array*, vol. 15, no. June, p. 100207, 2022, doi: 10.1016/j.array.2022.100207.
[31] Z. Lin, G. S. Member, and H. Liu, "CFLIT : coexisting federated learning and information transfer," *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 8436–8453, 2023, doi: 10.1109/TWC.2023.3263148.
[32] X. Dong, S. Q. Zhang, A. Li, and H. T. Kung, "SphereFed: hyperspherical federated learning," *In European Conference on Computer Vision,* no. i, pp. 1–30, 2022, doi: 10.48550/arXiv.2207.09413.
[33] D. Jhunjhunwala, "FedVARP : tackling the variance due to partial client participation," *In Uncertainty in Artificial Intelligence,* no. Uai, 2022, doi: 10.48550/arXiv.2207.14130.
[34] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: distributed machine learning for on-device intelligence," *arXiv preprint arXiv:1610.02527,* pp. 1–38, 2016, doi: 10.48550/arXiv.1610.02527.
[35] J. Han and Y. Han, "DeFL : decentralized weight aggregation for cross-silo federated learning," *arXiv preprint arXiv:2208.00848,* vol. August, 2022, doi: 10.48550/arXiv.2208.00848.
[36] X. Ma, H. Sun, R. Q. Hu, and Y. Qian, "A new implementation of federated learning for privacy and security enhancement," In *GLOBECOM 2022-2022 IEEE Global Communications Conference,* doi: 10.48550/arXiv.2208.01826.
[37] C. Xie, P.-Y. Chen, C. Zhang, and B. Li, "Improving privacy-preserving vertical federated learning by efficient communication with ADMM," *In 2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML),* pp. 1–22, 2022, doi: 10.48550/arXiv.2207.10226.
[38] P. Antunes, N. H. Thuan, D. Johnstone, and N. Hoang, "Nature and purpose of conceptual frameworks in design science," *Scandinavian Journal of Information Systems*, vol. 33, no. 2, pp. 59–96, 2021.
[39] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep network from decentralized data," in *20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017, vol. 54, p. 10, doi: 10.48550/arXiv.1602.05629.
[40] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492,* pp. 1–10, 2016, doi: 10.48550/arXiv.1610.05492.
[41] M. Lubis and D. O. D. Handayani, "The relationship of personal data protection towards internet addiction: Cyber crimes, pornography and reduced physical activity," *Procedia Computer Science*, vol. 197, no. 2021, pp. 151–161, 2021, doi: 10.1016/j.procs.2021.12.129.
[42] White House, "Consumer data privacy in a networked world: a framework for protecting privacy and promoting innovation in the global digital economy white house report * foreword," *Journal of Privacy and Confidentiality*, vol. 4, no. 2, pp. 95–142, 2012.
[43] K. Bogacka, K. W. Michiniewska, M. Paprzycki, M. G. Anatasyia Danilenka, L. Tassakos, and E. Garro, "Introducing federated learning into internet of things ecosystems – preliminary considerations," *arXiv*, no. i, p. 7, 2022, doi: 10.1109/WF-IoT54382.2022.10152142.
[44] Globaldots, "41.6 Billion IoT devices will be generating 79.4 Zettabytes of data in 2025," 2019, [online]: https://www.globaldots.com/resources/blog/41-6-billion-iot-devices-will-be-generating-79-4-zettabytes-of-data-in-2025/
[45] L. U. Khan, W. Saad, Z. Han, E. Hossain, C. Seon Hong, and S. Member, "Federated learning for internet of things: recent advances, taxonomy, and open challenges," *IEEE Communications Surveys and Tutorials,* pp. 1–41, doi: 10.48550/arXiv.2009.13012.
[46] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017*, vol. 54, 2017, doi: 10.48550/arXiv.1602.05629.
[47] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," *2015 53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015*, pp. 909–910, 2016, doi: 10.1109/ALLERTON.2015.7447103.
[48] H. Zhao, Z. Li, and P. Richtárik, "FedPAGE: A fast local stochastic gradient method for communication-efficient federated learning," *arXiv preprint arXiv:2108.04755,* 2021, doi: 10.48550/arXiv.2108.04755.
[49] H. Yuan and T. Ma, "Federated accelerated stochastic gradient descent," *Advances in Neural Information Processing Systems*, vol. 2020-Decem, no. NeurIPS, pp. 1–13, 2020, doi: 10.48550/arXiv.2006.08950.

[50] X. Liang, S. Shen, J. Liu, Z. Pan, E. Chen, and Y. Cheng, "Variance reduced local SGD with lower communication complexity," pp. 1–24, 2019, doi: 10.48550/arXiv.1912.12844.

[51] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, "SCAFFOLD: stochastic controlled averaging for federated learning," *37th International Conference on Machine Learning, ICML 2020*, vol. PartF16814, pp. 5088–5099, 2020, doi: 10.48550/arXiv.1910.06378.

[52] E. Gorbunov, F. Hanzely, and P. Richtárik, "Local SGD: unified theory and new efficient methods," *In International Conference on Artificial Intelligence and Statistics,* vol. 130, no. 2, 2020, doi: 10.48550/arXiv.2011.02828.

[53] H. Zhu and Y. Jin, "Multi-Objective Evolutionary Federated Learning," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 4, pp. 1310–1322, 2020, doi: 10.1109/TNNLS.2019.2919699.

[54] A. K. Kundu and J. Jaja, "FedNet2Net: saving communication and computations in federated learning with model growing," 2022, doi: 10.48550/arXiv.2207.09568.

[55] S. Caldas, J. Konečny, H. B. McMahan, and A. Talwalkar, "Expanding the reach of federated learning by reducing client resource requirements," *arXiv 2018. arXiv preprint arXiv:1812.07210,* 2018, doi: 10.48550/arXiv.1812.07210.

[56] E. Diao, J. Ding, and V. Tarokh, "HeteroFL: Computation and communication efficient federated learning for heterogeneous clients," *arXiv preprint arXiv:2010.01264,* pp. 1–24, 2020, doi: 10.48550/arXiv.2010.01264.

[57] F. Hanzely and P. Richtárik, "Federated learning of a mixture of global and local models," *arXiv preprint arXiv:2002.05516,* vol. 1, no. 1, pp. 1–40, 2020, doi: 10.48550/arXiv.2002.05516.

[58] D. Gao, Y. Liu, A. Huang, C. Ju, H. Yu, and Q. Yang, "Privacy-preserving heterogeneous federated transfer learning," *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019*, pp. 2552–2559, 2019, doi: 10.1109/BigData47090.2019.9005992.

[59] J. Wang, Q. Liu, H. Liang, G. Joshi, and H. Vincent Poor, "Tackling the objective inconsistency problem in heterogeneous federated optimization," *Advances in Neural Information Processing Systems,* vol. 2020-Decem, pp. 1–34, 2020, doi: 10.48550/arXiv.2007.07481.

[60] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," 2018, doi: 10.48550/arXiv.1812.06127.

[61] T. Lin, L. Kong, S. U. Stich, and M. Jaggi, "Ensemble distillation for robust model fusion in federated learning," *Advances in Neural Information Processing Systems*, vol. 2020-Decem, no. NeurIPS, 2020, doi: 10.48550/arXiv.2006.07242.

[62] D. Yao *et al.*, "Local-global knowledge distillation in heterogeneous federated learning with non-IID data," *arXiv preprint arXiv:2107.00051,* pp. 1–16, 2021, doi: 10.48550/arXiv.2107.00051.

[63] Z. Zhu, J. Hong, and J. Zhou, "Data-free knowledge distillation for heterogeneous federated learning," *In International Conference on Machine Learning,* 2021, doi: 10.48550/arXiv.2105.10056.

[64] X. Zhou, X. Lei, C. Yang, Y. Shi, X. Zhang, and J. Shi, "Handling data heterogeneity in federated learning via knowledge fusion," *International Conference on Machine Learning,* vol. July, pp. 1–15, 2022, doi: 10.48550/arXiv.2207.11447.

[65] Y. Liu, Y. Kang, C. Xing, T. Chen, and Q. Yang, "A secure federated transfer learning framework," *IEEE Intelligent Systems,* vol. 35, no. 4, pp. 70–82, 2020, doi: 10.1109/MIS.2020.2988525.

[66] S. K. Lo, Q. Lu, L. Zhu, H. Y. Paik, X. Xu, and C. Wang, "Architectural patterns for the design of federated learning systems," *Journal of Systems and Software*, vol. 191, p. 111357, 2022, doi: 10.1016/j.jss.2022.111357.

[67] P. Boobalan *et al.*, "Fusion of federated learning and industrial internet of things: a survey," *Computer Networks*, vol. 212, no. May, p. 109048, 2022, doi: 10.1016/j.comnet.2022.109048.

[68] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer, "Machine learning with adversaries: Byzantine tolerant gradient descent," *Advances in Neural Information Processing Systems*, vol. 2017-Decem, no. Nips, pp. 119–129, 2017, doi: https://papers.nips.cc/paper_files/paper/2017/hash/f4b9ec30ad9f68f89b29639786cb62ef-Abstract.html.

[69] A. R. Elkordy and Y. H. Ezzeldin, *How much privacy does federated learning with secure aggregation guarantee ?*, vol. 1, no. 1. Association for Computing Machinery, 2023.

[70] M. Fredrikson, S. Jha, and T. Ristenpart, "Model inversion attacks that exploit confidence information and basic countermeasures," *Proceedings of the ACM Conference on Computer and Communications Security*, vol. 2015-Octob, pp. 1322–1333, 2015, doi: 10.1145/2810103.2813677.

[71] K. Bonawitz *et al.*, "Practical secure aggregation for federated learning on user-held data," *arXiv preprint arXiv:1611.04482 13,* no. Nips, 2016, doi: 10.48550/arXiv.1611.04482.

[72] K. Prasad, S. Ghosh, and G. Cormode, "Reconciling security and communication efficiency in federated learning," *arXiv preprint arXiv:2207.12779,* 2022.

[73] X. Qiu, J. Fernandez-marques, P. P. B. Gusmao, Y. Gao, T. Parcollet, and N. D.Lane, "ZEROFL : Efficient on-device training for federated learning with local sparsity," in *ICLR,* 2022, pp. 1–16, doi: 10.48550/arXiv.2208.02507.

[74] H. Wu and P. Wang, "Fast-convergent federated learning with adaptive weighting," *IEEE International Conference on Communications*, no. June, 2021, doi: 10.1109/ICC42927.2021.9500890.

[75] M. N. Gibs and D. J. C. Mackay, "Variational Gaussian processes classifiers," p. 12, 2000.

[76] L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai, "Privacy-preserving deep learning via additively homomorphic encryption," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1333–1345, 2018, doi: 10.1109/TIFS.2017.2787987.

[77] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, "A survey on distributed machine learning," *ACM Computing Surveys*, vol. 53, no. 2, 2020, doi: 10.1145/3377454.

[78] Q. Li, Y. Diao, Q. Chen, and B. He, "Federated learning on non-IID Data Silos: an experimental study," *Proceedings - International Conference on Data Engineering,* vol. 2022-May, pp. 965–978, 2022, doi: 10.1109/ICDE53745.2022.00077.

[79] Y. Xu, Y. Liao, H. Xu, Z. Ma, L. Wang, and J. Liu, "Adaptive control of local updating and model compression for efficient federated learning," *IEEE Transactions on Mobile Computing,* 2022, doi: 10.1109/TMC.2022.3186936.

[80] T. Zhang, K.-Y. Lam, J. Zhao, F. Li, H. Han, and N. Jamil, "Enhancing federated learning with spectrum allocation optimization and device selection," *IEEE/ACM Transactions on Networking,* pp. 1–18, 2022, doi: 10.48550/arXiv.2212.13544.

[81] Y. Liao, Y. Xu, H. Xu, L. Wang, and C. Qian, "Adaptive configuration for heterogeneous participants in decentralized federated learning," *In IEEE INFOCOM 2023-IEEE Conference on Computer Communications,* 2022, doi: 10.48550/arXiv.2212.02136.

[82] Z. Fan, Y. Wang, J. Yao, L. Lyu, Y. Zhang, and Q. Tian, "FedSkip: combatting statistical heterogeneity with federated skip aggregation," *In 2022 IEEE International Conference on Data Mining (ICDM),* 2022, doi: 10.48550/arXiv.2212.07224.

BIOGRAPHIES OF AUTHORS

Farhanna Mar'i received the bachelor degree in Informatics Engineering from Institute of Technology and Business Yadika Pasuruan and the master's degree in Computer Science in Universitas Brawijaya (UB), Indonesia. Currently, serving as an Assistant Professor at the Department of Informatics, Faculty of Engineering at Universitas Negeri Surabaya (UNESA), East Java, Indonesia. Her research interest focuses on optimization methods, federated learning, and data science. She can be contacted at email: farhannamari@unesa.ac.id.

Ahmad Afif Supianto was a distinguished researcher at the Research Center for Artificial Intelligence and Cyber Security, National Research and Innovation Agency (BRIN), Indonesia. His work focused on developing advanced machine learning algorithms for data mining applications. He received his Bachelor of Science degree in Mathematics and Master of Computer degree in Informatics Engineering from the Sepuluh Nopember Institute of Technology, where he made significant contributions to the fields of clustering algorithms and pattern recognition. He furthered his academic journey at Hiroshima University, Japan, where he earned his Ph.D. in Information Engineering in 2017, focusing on leveraging machine learning techniques for educational data mining. In 2022, He joined the Department of ICT and Natural Sciences at the Norwegian University of Science and Technology (NTNU) as a postdoctoral fellow. His research during this time was centered around the innovative application of machine learning in the realm of digital twins. Ahmad Afif Supianto's legacy lives on through his contributions to the fields of artificial intelligence and data science. He will be remembered for his dedication to advancing technology and his commitment to academic excellence. He can be contacted at email: ahmad.a.supianto@ntnu.no.