I.J. Modern Education and Computer Science, **2025, 2, 87-110**
Published Online on April 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2025.02.04

Unlocking Educational Excellence: Leveraging Federated Learning for Enhanced Instructor Evaluation and Student Success

Ariful Islam
American International University-Bangladesh, Dhaka, Bangladesh
E-mail: ariful@aiub.edu
ORCID iD: https://orcid.org/0009-0002-5537-9102

Debajyoti Karmaker*
American International University-Bangladesh, Dhaka, Bangladesh
E-mail: d.karmaker@aiub.edu
ORCID iD: https://orcid.org/0000-0003-0020-5391
*Corresponding Author

Abhijit Bhowmik
University Malaysia Pahang Al-Sultan Abdullah, Pekan, Pahang, Malaysia
E-mail: ovi775@gmail.com
ORCID iD: https://orcid.org/0000-0002-9166-347X

Md Masum Billah
American International University-Bangladesh, Dhaka, Bangladesh
E-mail: billah.masumcu@aiub.edu
ORCID iD: https://orcid.org/0000-0002-4400-6610

Md Iftekharul Mobin
American International University-Bangladesh, Dhaka, Bangladesh
E-mail: iftekhar.mobin@aiub.edu
ORCID iD: https://orcid.org/0000-0002-3065-2486

Noorhuzaimi Mohd Noor
University Malaysia Pahang Al-Sultan Abdullah Pekan, Pahang, Malaysia
E-mail: nhuzaimi@ump.edu.my
ORCID iD: https://orcid.org/0000-0002-2462-6265

Received: 16 December, 2023; Revised: 07 March, 2024; Accepted: 07 October, 2024; Published: 08 April, 2025

Abstract: Federated Learning (FL) is an emerging machine learning approach with promising applications. In this paper, FL is comprehensively examined in relation to teacher performance evaluation. Through FL, teachers can be evaluated based on data-driven metrics while preserving data privacy. There are several benefits, including data privacy preservation, collaborative learning, scalability, and privacy-preserving insights. Additionally, it faces problems related to communication efficiency, system heterogeneity, and statistical heterogeneity. To address these issues, we propose a novel clustering-based technique in federated learning. The technique aims to overcome the challenges of system heterogeneity and improve communication efficiency. We provide a comprehensive review of existing research on clustering techniques in the context of federated learning, offering insights into the state of the art in this field. In addition, we emphasize the need for advanced compression methods, enhanced privacy-preserving mechanisms, and robust aggregation algorithms for future federated learning research. To address these challenges, we present a clustering-based approach to address the merits and challenges of federated learning The clustering-based approach we propose in this research demonstrates promising results in terms of reducing communication overhead and improving model convergence in federated learning. These findings suggest that incorporating clustering techniques can significantly enhance the efficiency and effectiveness of federated learning algorithms, paving the way for more scalable and privacy-preserving distributed machine learning systems. The findings of this study suggest that clustering techniques can improve the efficiency and scalability of federated learning.

Index Terms: Feature Extraction, Fraud Review, Machine Learning, Teacher Performance, Instructor Performance

1. Introduction

Federated learning is a distributed machine learning paradigm where multiple devices collaborate to train a shared model while keeping their local data private. Where the devices only share the model parameters. The process involves a central server distributing a pre-trained model to clients, and the clients update the model based on their local data, sending the updated model back to the server for aggregation and improvement. The iterative process continues until the model reaches the desired accuracy. Federated learning can be helpful for training models on private or sensitive data because it keeps the data on the devices and lowers the possibility of privacy violations or data breaches [1]. Federated learning has the potential to greatly enhance educational outcomes, particularly in the context of teacher performance evaluation. By leveraging the power of federated learning, educational institutions can effectively evaluate and improve the effectiveness of their faculty members. By implementing federated learning techniques, educational institutions can address the challenges of data privacy and security, while still benefiting from collaborative optimization of machine learning models. Moreover, this decentralized approach has emerged as a promising solution for addressing privacy concerns in data-driven applications, particularly in domains where sensitive information is collected and analyzed. Our primary motivation for this paper is to find effective algorithms to overcome the challenges of Federated learning and how Federated learning can affect on educational sector to evaluate teacher performance. Data privacy and security are preserved by federated learning, according to several studies. This method reduces the risk of unauthorized access and data breaches by allowing training of models without sharing raw data. Additionally, federated learning has shown promising results in various fields, such as healthcare and finance, where privacy is of utmost importance.

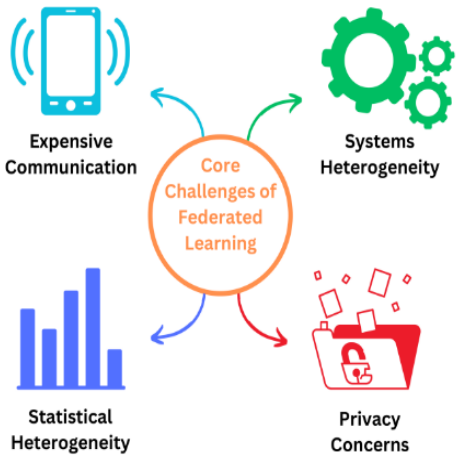

Fig. 1. Core Challenges of Federated Learning

Compared to traditional centralized machine learning approaches, FL offers several advantages:

- **Data Privacy Preservation**: Federated Learning does not require centralized data storage, as it distributes data across multiple devices or servers. As a result, there is a reduced risk of data breaches and unauthorized access, ensuring that the owners of sensitive information remain in control.
- **Collaborative Learning**: FL allows multiple parties to collaborate on learning, providing insights, and improving the model's performance without compromising the confidentiality of the data. In domains where data is siloed or geographically dispersed, this collaborative approach can be particularly beneficial.
- **Reduced Communication Overhead**: FL minimizes communication overhead and only shares model updates with the central server. The use of this technique is particularly advantageous in environments with limited resources or when dealing with large datasets.

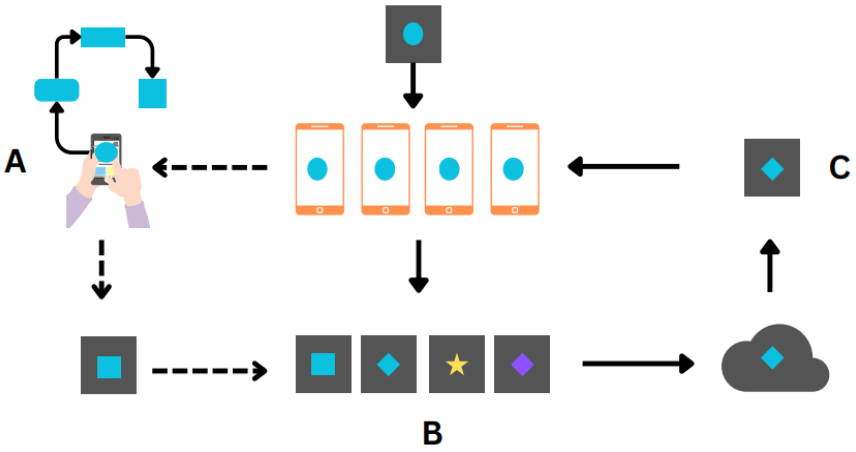

Fig. 2. Federated Learning: Collaborative machine learning without centralized training data.

Federated Learning applications spread over a number of industries, including:

- **Healthcare:** Machine learning models can be trained with FL for developing personalized treatment plans, predicting diseases, and discovering drugs in the healthcare field. FL allows healthcare organizations to share models and collaborate across multiple institutions, which can help them increase their efficiency and improve their patient care.
- **Education:** Teacher performance evaluation can be improved through the use of FL, which is more objective and data driven. FL models identify patterns and trends that may be missed by traditional evaluation methods, through the analysis of teacher-student interactions, observations of classroom activities, and student outcomes. By identifying at-risk students early on and providing targeted interventions, FL can improve student outcomes. FL models identify students who need help based on their performance and attendance patterns.
- **Finance:** Using FL, you can develop fraud detection models, risk assessment tools, and personalized financial recommendations while protecting sensitive financial information.
- **Internet of Things:** Using FL, machine learning models can be trained collaboratively on IoT devices, improving smart homes, connected vehicles, and industrial automation systems.
- **Mobile Applications:** FL can be used to train personalized models for mobile applications, such as predictive keyboards, voice recognition, and recommendation systems, without compromising user privacy.

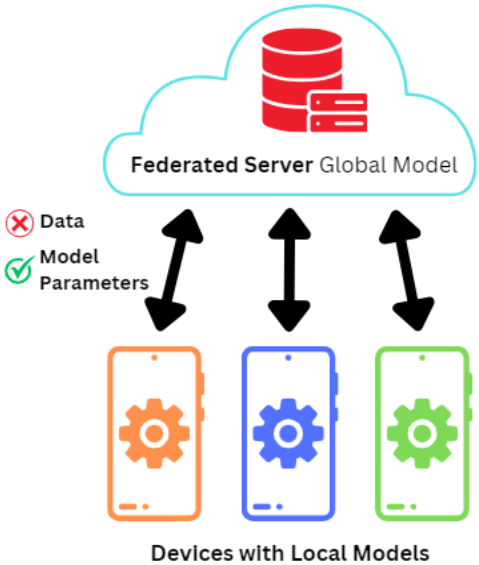

Fig. 3. Federated Learning diagram.

Teacher performance evaluation (TPE) is the systematic process of assessing a teacher's effectiveness in the classroom. Besides improving student outcomes, it also improves teacher quality. Federated learning can be an effective method for evaluating the performance of teachers. Multiple sources of data can be collected and analyzed while individual teacher data is kept private and secure. Personalized teacher evaluations can be based on federated learning, taking into account student performance, classroom dynamics, and teaching methods. By including a variety of factors

that influence teachers' performance, this approach permits a more comprehensive and accurate evaluation. To address the challenges of data privacy and heterogeneity of models, federated learning offers a promising solution. By allowing each participant to maintain a personalized model, federated learning ensures that individual teachers' unique characteristics and teaching styles are considered in the evaluation process. Moreover, personalized federated learning methods can also address the issue of non-IID (non-identically distributed) data, which is often encountered in educational settings where each teacher may have different student populations and teaching contexts.

Purposes of Teacher Performance Evaluation:

- To improve teacher quality: TPE can help to identify teachers who need additional support and to provide them with the resources they need to improve their teaching.
- To improve student outcomes: TPE can help to ensure that all students are receiving high-quality instruction.
- To promote accountability: TPE can help to hold teachers accountable for their performance.

Effective TPE systems are based on the following principles:

- Clear and consistent standards: The standards for teacher performance should be clear, consistent, and communicated to all teachers.
- Multiple measures: TPE should use multiple measures to assess teacher performance, including classroom observations, student data, teacher self-reflection, and peer feedback.
- Fair and unbiased: TPE should be fair and unbiased, and all teachers should be evaluated using the same standards.
- Focused on improvement: TPE should be focused on helping teachers to improve their performance, not on punishment.

Benefits of Teacher Performance Evaluation:

- Improved teacher quality: TPE can help to identify and support teachers who need additional support, leading to improved teacher quality overall.
- Increased student achievement: TPE can help to ensure that all students are receiving high-quality instruction, which can lead to increased student achievement.
- Greater teacher satisfaction: TPE can help teachers to become more reflective and self-aware, which can lead to greater teacher satisfaction.
- Enhanced school culture: TPE can help to create a culture of continuous improvement within a school, which can benefit all stakeholders.

In the context of teacher performance evaluation, FL can be used to train ML models that predict teacher effectiveness based on various factors, such as student outcomes, classroom observations, and teacher self-assessments. The data for these models can be distributed across multiple schools or districts, ensuring that individual teacher data remains private here’s how FL algorithms work for teacher performance evaluation:

1. **Data Distribution**: Teacher data is distributed across participating schools or districts, ensuring data privacy and avoiding the need for data centralization.
2. **Local Model Training**: Each participating institution trains a local model using its own teacher data. In this way, sensitive data will remain within the institutions involved.
3. **Model Updates Aggregation**: The local model updates are aggregated at a central server. The central server does not receive or store any raw teacher data, only the model updates.
4. **Global Model Update**: The central server aggregates the model updates and updates the global model parameters. This global model represents the collective knowledge from all participating institutions.
5. **Model Sharing**: The updated global model parameters are shared with participating institutions, allowing them to improve their local models without compromising data privacy.
6. **Iterative Improvement**: The process of local model training, model update aggregation, global model update, and model sharing is repeated iteratively, leading to a continuously improving global model for teacher performance evaluation.

Analysis:

Our search identified 67 relevant publications after eliminating duplicates and applying inclusion/exclusion criteria. Using words, the results were summarized. Each publication was evaluated and discussed by the research team on a regular basis in order to assess potential biases.

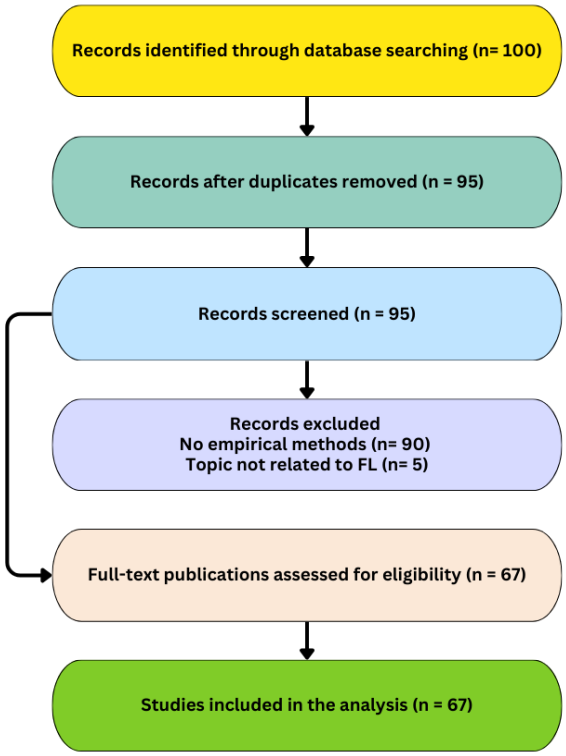

Fig. 4. The review process.

2. Survey of Related and Existing Work

Federated Learning offers a framework where teachers can collaborate, leverage localized optimizations, personalize training, and implement efficient techniques like knowledge distillation to enhance their performance while respecting data privacy and individual teaching styles. In this part, we examine and analyze the plurality of recent survey reports. The objective of federated learning algorithms is to minimize training losses across all clients, a collective model needs to be developed. A description of this objective can be given:

$$^{min}_{w} F(\omega) = \sum_{c=1}^{C} \alpha_c F_c(\omega) \quad (1)$$

Where the local loss of device c is $F_c(\omega)$, and the weight parameter is α_c with $\sum_{c=1}^{C} \alpha_c = 1$. Federated Averaging is an early FL algorithm [51]. The weighted average approach combines the local training losses from each node in the network, giving more weight to the nodes with higher accuracy. This helps to ensure that the model is learning from all of the nodes and not just the ones with the highest accuracy with $\alpha_c = \frac{n_c}{n}$, where client c consists of n_c training samples, with n being the number of training samples partitioned across all C clients.

FedAVG is an effective algorithm for FL was proposed in [51] and compares its performance to other approaches. FedAvg enables model training without centralizing sensitive data, ensuring the privacy of individual teacher or student information. FL system with one aggregation node and K users, where each user has a distinct local dataset. In this algorithm, there are K users who are indexed by k. These users have their own local datasets and models. The local mini-batch size is denoted by B, which refers to the number of training examples used in each iteration of the local training process. The number of local epochs is denoted by E, which refers to the number of times the local model is trained on the local dataset. The learning rate is denoted by η, which is a hyper-parameter that controls the step size at each iteration of the optimization process. FedAVG algorithm is used to aggregate the local models of the K users to create a global model that can be used for predictions, offering a privacy-preserving solution for evaluating teacher performance in the educational sector.

Algorithm 1: Federated Averaging Algorithm
Aggregation Node:
initialize w_G^1 // Global Model Parameters
FOR round t = 1, 2, . . . **DO**
m ← max($C \times K$, 1)
S_t ← (random set of m users)
FOR each user k ∈ S_t **DO IN PARALLEL**
w_k^{t*} ← UserUpdate(k, w_G^t) // Local Model Parameters Update
$w_G^{t+1} \leftarrow \frac{1}{m} \sum_{k=1}^{m} w_k^{t*}$
UserUpdate (k, w_G^t) // Run on user k
β ← (select batches of size)
FOR local epochs **i** from 1 to E **DO**
FOR batch b ∈ β **DO**
$w \leftarrow w - \eta \nabla l\ (w; b)$
upload w to Aggregation Node

This FedAVG method diverges due to the heterogeneity in the network. In terms of Systems Heterogeneity, there are many Stragglers in the network. In the original design of FedAveraging simply drop the slow devices. Dropping the data update from those devices can implicitly increase the bias from the training data. And In terms of Statistical Heterogeneity local updates come from different devices and they are optimized using local optimizer on local devices. Utilizing FedAvg in the education sector enhances teacher performance evaluation by leveraging collaborative model training across distributed devices while preserving data privacy and accommodating data heterogeneity. In general, FedAvg is a heuristic method and does not guarantee convergence. During data aggregation, data quality and privacy constraints are maintained by examining data distribution across participants, identifying statistical characteristics, such as mean and variance. In federated learning settings, it involves assessing data heterogeneity and bias across clients in order to facilitate informed decisions regarding aggregation strategies and model updates. Additionally, sensitive information may be protected during model aggregation through techniques such as differential privacy.

To address these previous FedAvg issues come to FedProx which is the framework for federated optimization that introduces a proximal term that penalizes large changes in weights, and this also helps convergence on highly heterogeneous data. By introducing regularization and adjustments to encourage closer alignment between local and global models, FedProx aims to enhance teacher performance evaluations within the Federated Learning paradigm. To limit the impact of variable local updates, add a proximal term to the local subproblem. Specifically, instead of minimizing just the local function, $F_k(w)$, device k by using its local solver of choice, aims to minimize the following objective h_k as closely as possible:

$$\min_w h_k(w; w^t) = F_k(w) + \frac{\mu}{2} ||w - w^t||^2 \quad (2)$$

This $||w - w^t||^2$ proximal term penalizes the model from changing too much on one single device and controls the amount that is penalized by this high parameter $\mu/2$. There are two advantages of the proximal term: It minimizes the statistical heterogeneity by limiting the number of local updates to match the initial (global) model without the need for manual adjustments. Secondly, it allows for the safe incorporation of variable amounts of local work as a result of system heterogeneity. In terms of Systems Heterogeneity FedProx allows for variable amounts of work and safely incorporates them. In terms of Statistical Heterogeneity encourage more well-behaved updates in a principled way. In general, FedProx guarantees convergence and more robust empirical performance for federated learning in heterogeneous networks [60] . A data analysis is used in this paper to study the performance of federated optimization algorithms across heterogeneous network settings. During the analysis, convergence rates, communication overhead, and model performance were evaluated under varying network conditions, such as different communication bandwidths and computational capacities. Furthermore, the study may analyze different network topologies and communication protocols to determine how they impact the efficiency and scalability of federated learning. FedProx in Federated Learning optimizes teacher performance evaluation in the education sector by incorporating proximity regularization to enhance model convergence and accuracy.

In recent years, there has been a growing body of research focused on the challenges associated with federated learning. The first challenge is that federated learning makes it difficult to train accurate models. Depending on the data distribution between the devices, it may not be possible to aggregate model parameters effectively. It can also be computationally expensive to implement federated learning, as the devices must train their models locally.

Using the **aggregation algorithm**, a global model is developed by combining the local model updates from all the train participants. Here are the most popular aggregation algorithms:

FedAvg was proposed in [51] which is a great example of how SGD can be used to develop powerful models within the framework of distributed machine learning. This algorithm consists of iteratively averaging the model parameters of neural networks across a set of clients. During each iteration, the clients train their models on local data.

After updating the model parameters, they send them to the server. Each client's model parameters are averaged by the server and returned to the clients as the averaged parameters. After averaging the model parameters, the clients train the model again on their data using the averaged model parameters. Until the model converges, this process is repeated. Federated averaging has several weaknesses. Client dropout is one of its weaknesses. Dropout occurs when a client fails to attend a training session. In some cases, this occurs because a client's battery runs low, or their network connection is interrupted. Stragglers are another weakness of the federated averaging algorithm. The Stragglers are clients who take longer to train their models than other clients. In addition to less powerful hardware, a slower network connection can also contribute to this issue. FedAvg has indeed achieved great success, but to improve its convergence properties researchers have proposed several algorithms that modify the original FedAvg algorithm such as client drifting which was discussed in [61], and a lack of adaptive learning that was discussed in [63].

SCAFFOLD (Stochastic Controlled Averaging for Federated Learning) employs a control Using SCAFFOLD was proposed in [61] where local gradients are controlled through the addition of a control variate. It can enhance teacher performance in the education sector by dynamically adjusting model aggregation, improving convergence, and preserving data privacy during collaborative model training. Control variates are functions of global model parameters that cancel out client drift. It helps the federated learning algorithm to achieve better convergence by keeping local models close to one another. SCAFFOLD has the following advantages in addition to handling client drift errors:

- Data heterogeneity and client sampling do not affect the performance of the algorithm.
- To further reduce communication costs, it can take advantage of similarity in the client's data.

They implemented their algorithm on both the MNIST and CIFAR-10 datasets in their SCAFFOLD paper. CIFAR-10 contains images of objects, and MNIST contains handwritten digits. SCAFFOLD performed similarly or better on both datasets than other federated learning algorithms. However, SCAFFOLD also has some limitations. One limitation is that it requires more computation than FedAvg. This is because SCAFFOLD needs to compute the control variates for each local update. Another weakness is that in SCAFFOLD hyper-parameter choices can be sensitive. The hyper-parameter values can affect SCAFFOLD's performance. Overall, SCAFFOLD can enhance teacher performance evaluation in the education sector by optimizing model aggregation methods for improved convergence and robustness in handling distributed data.

FedBoost is a new method that was proposed in [62] to reduce communication costs for federated learning. In federated learning, ensemble methods offer several advantages, including computational speedups, convergence guarantees, privacy, and optimality of the solution for density estimation, for which language modeling is a special case. The FedBoost algorithm trains a group of pre-trained base predictors using a communication-efficient algorithm for federated learning. The algorithm starts by randomly initializing an ensemble of base predictors. The server sends clients a weighted ensemble of the base predictors every round. Once the clients have trained their local models, they send updates to the server about their weights. After the client's weights have been updated, the server updates the ensemble's weights. The algorithm is repeated until convergence is achieved. FedBoost was implemented and evaluated on three datasets MNIST, CIFAR-10, and EMNIST datasets. FedBoost is limited to problems where the base predictors can be trained independently of each other. There are some problems that require the base predictors to be trained jointly, which is not always possible. Additionally, FedBoost can be computationally expensive since multiple base predictors must be trained. Using FedBoost in the education sector enhances teacher performance evaluation by aggregating decentralized model updates from multiple sources in Federated Learning to optimize model performance and achieve improved assessment accuracy. One potential downside of using FedBoost in the education sector is the privacy implications. If the data is centralized, there is a risk of the data being hacked or misused. Additionally, if the data is decentralized, there is a risk of the data not being consistent across different sources.

Federated Matched Averaging (FedMA) is a federated matrix averaging algorithm developed for convolutional neural networks (CNNs) and long short-term memory networks (LSTMs) in [64]. This method starts by building a shared global model layer by layer. The hidden elements (e.g., convolution channels, LSTM hidden states, and neurons for fully connected layers) are matched and averaged according to their feature extraction signatures. Throughout the model, this process is repeated. FedMA was implemented and evaluated on three datasets MNIST, CIFAR-10, and ImageNet datasets. Compared to FedAvg and FedProx, the FedMA aggregation algorithm offers several advantages. First, it reduces communication overhead by exchanging only the weights of matched hidden elements. Secondly, it can improve global model performance by averaging weights of matched hidden elements from different clients. The third advantage is that it can handle imbalanced datasets better. FedMA has one main limitation, which is that clustering the hidden elements of local models can be computationally expensive. This is especially true for deep neural networks with a large number of hidden components. FedMA appears to be a promising federated learning algorithm. By reducing communication overhead, it achieves comparable accuracy to other federated learning algorithms. Utilizing Federated Matched Averaging improves Teacher Performance Evaluation in the education sector by enabling collaborative model updates while addressing data heterogeneity across distributed sources. This is because FedMA only requires a single communication round to update the model, unlike traditional federated learning algorithms which require multiple rounds. This reduces the communication overhead and enables efficient model updates. Additionally, Federated Matched Averaging allows for the aggregation of data from multiple distributed sources, which reduces the bias and variance of the predictions.

To address the limitations of existing theoretical analyses of local Stochastic Gradient Descent (local SGD) was proposed in [59], a widely used optimization algorithm for machine learning. The algorithm works by first dividing the data into M subsets, each stored on a different device. On each subset of data, each device then runs local SGD for H iterations. As a result of H iterations, the devices communicate with each other and average their parameters. Until the algorithm converges, this process is repeated. Client drift error is a phenomenon in federated learning where the local models of different clients drift away from each other due to the differences in their data distributions. The "bounded gradient dissimilarity" is a new notion of variance. The bounded gradient dissimilarity can help prevent client drift by penalizing parameters and gradients that are too different from the global model. This ensures that the local model is not too far from the global model and helps to keep the predictions accurate and the model performance highly measures the difference in gradients between clients based on the loss function. It has been shown the convergence rate of local SGD can be improved by using a step size that is inversely proportional to the bounded gradient dissimilarity.

$$\min_{t\in\{0,\dots,T-1\}} E\left[\| \nabla F\left(\mathrm{x}^{(t)}\right) \|^2\right] \leq O\left(\frac{F\left(\mathrm{x}^{(0)}\right)-F^*}{\eta\tau T}\right) + O\left(\frac{\eta L\sigma^2}{m} + \eta^2 L^2(\tau-1)\sigma^2\right) + O\left(\eta^2 L^2 \tau(\tau-1)\sigma_q^2\right) \tag{3}$$

Assuming this kind of bounded dissimilarity the error which is shown on the left side of the equation can be bounded as given by these 3 terms of the right-hand side of the equation, first 2 terms mainly come from performing mini batches SGD and the third term is the client drift error, we can see if the increases with the amount of data heterogeneity which is denoted by σ_q^2 and it also increases the tau which is the number of local steps. We do more and more local steps there is more heterogeneity in the system this error bound becomes worse However, it has the weakness of requiring careful tuning of the noise term. There is a possibility that the algorithm will not converge if the noise term is too large. The algorithm may not be able to mitigate the effects of heterogeneous data if the noise term is too small. Using the Stochastic Gradient Descent algorithm in teacher performance evaluation within the education sector can enhance efficacy through iterative weight updates based on small batches of data, optimizing model parameters to improve teaching strategies.

q-Fair Federated Learning (q-FFL) algorithm aimed at achieving fair resource allocation between clients that was proposed in [65]. The algorithm assigns a weight to each client based on its fairness score. It is a measure of how fair the data is to the client. A higher weight is assigned to clients who have higher fairness scores. According to the algorithm, resources are allocated to clients based on their weights. Clients with a higher weight receive more resources. A fairness-accuracy trade-off hyper-parameter determines the number of resources allocated to each client. This paper acknowledges some weaknesses of q-FFL. There is a weakness in the fact that it is not always possible to allocate resources fairly. As a result, a client's fairness score may not accurately reflect the fairness of its data. Another weakness of q-FFL is its computational complexity. The reason for this is that the algorithm needs to calculate each client's fairness score before allocating resources according to their weight. Using the q-Fair Federated Learning (q-FFL) algorithm enhances teacher performance evaluation in the education sector by ensuring fairness and equitable model updates among participating teachers, promoting balanced improvements across diverse datasets.

AdaComm is an adaptive communication strategy that was proposed in [66] for federated learning that provides local updates. The model starts with a high communication frequency and then gradually reduces it as it converges. It is done in order to achieve a good trade-off between the error rate and the algorithm's runtime. As a first step, the algorithm calculates the error of the global model after each round of communication. In the case of a small error, the frequency of communication decreases. In the case of a large error, the communication frequency will be increased. It is designed to be robust in the face of heterogeneous data and stragglers. All MNIST, CIFAR-10, and Fashion-MINIST datasets, also demonstrated better accuracy and faster convergence than other communication strategies. This model has the disadvantage of being sensitive to the choice of hyperparameters. Moreover, calculating the error of the global model after each round of communication can be computationally expensive. Enhancing teacher performance evaluation in the education sector via adaptive communication strategies facilitates personalized feedback and support tailored to individual teaching methodologies and student needs.

FedVARP is a federated learning algorithm that was proposed in [67] to address variance caused by partial client participation. Each non-participating client is provided with a surrogate model on the server. Based on the updates received from participating clients, the surrogate models are updated. To estimate the gradients of the loss function at non-participating clients, surrogate models are used. To achieve this, backpropagation is used. In backpropagation, the gradients of a loss function are calculated for the parameters of the model. Using the estimated gradients, the global model is updated. By doing so, we can reduce the variance due to partial client participation. FedVARP was implemented on the MNIST and CIFAR-10 datasets in the paper. Handwritten digits are in the MNIST dataset, and natural images are in the CIFAR-10 dataset. FedVARP has the weakness of requiring the server to maintain surrogate models for each non-participating client. Clients who do not participate in large numbers, can harm the server. It is also possible for FedVARP to be computationally expensive to calculate the estimated gradients. The reason for this is that backpropagation is a computationally intensive technique. Enhancing teacher performance evaluation in the education sector using FedVARP involves leveraging variable selection and regularization techniques within the Federated Learning framework to optimize model performance and assess teaching effectiveness.

Table 1. Represents the different comparison of related works in Federated Learning

Ref.	Methods	Contributions	Limitations
[51]	FedAVG	• A Distributed learning method	• Client dropout
[67]	FedVARP	• Loss functions are calculated of the model	• Maintain the server for each non-participating client. • Calculate the estimated gradients
[66]	AdaComm	• Trade-off between the error rate and runtime	• Choice of hyperparameters
[65]	q-FFL	• Achieving fair resource allocation between clients • A measure of data is to the client.	• Fairness score may not accurately reflect the fairness of its data. • Computational complexity
[64]	FedMA	• Matched hidden elements from different clients.	• Clustering model can be computationally expensive
[62]	FedBoost	• Reduces communication costs.	• Require the base predictors to be trained

2.1 Communication efficiency

Federation learning presents a number of challenges, one of them being communication between clients and the central server. Compression techniques, reducing the number of communication rounds, and asynchronous communication are some methods to improve communication efficiency.

2.1.1 Local Updating

Local Updating is one of the methodologies that address the challenge of expensive communications in federated learning. By allowing each client to update its model using only local data, the need for frequent communication with the central server is reduced. The process of computing updates locally on devices participating in Federated Learning (FL) and it is not very efficient to send those updates to the central server because it requires a singular communication round, which means that all devices must send their updates to the central server at the same time. However, when networks have poor connectivity, device dropouts or synchronization delays can cause delays in this process, which can create bottlenecks in communication [11]. To reduce communication rounds within the FL protocol, a method is used in [2] that uses hierarchical clustering to group clients together based on the similarity of their local updates, once the clients are grouped, they are trained independently on specialized models. There could be a communication bottleneck when so many connected devices update their parameters with the central server. Each cluster of similar clients c in FL+HC is trained with a special model $f_c(w)$. To achieve different objectives,

$$\forall_c \in C, E_{D_k}[F_k(w)] = f_c(w) \; where \; k \in c \tag{4}$$

The server-side operation is usually much faster than the client-side since it can be processed on a larger machine, and the updates are all merged into the same operation using this special model, thereby reducing the amount of communication between the server and the clients.

FedPAQ's local updating process using a technique called gradient compression, improves communication efficiency through local updates. Gradient compression reduces the size of the gradients that need to be uploaded to the server by quantizing them. As a result, the communication cost can be further reduced. In scenarios where the devices have limited communication bandwidth or where the data is sensitive and cannot be shared, this can be especially useful [6]. FedPAQ, an extension of FedAvg leveraging personalized model updates, enhances teacher performance evaluation in the education sector by addressing individualized instructional needs through personalized federated learning approaches.

2.1.2 Compression techniques

Another solution to address the challenges of communication in federated learning is by implementing compression techniques. These compression techniques aim to reduce the amount of data that needs to be communicated between the central server and the client devices. Several methods can be used to perform this, such as sparsity, quantization, and subsampling. These techniques aim to reduce the communication overheads and computation costs during the training process. However, Model compression approaches face novel challenges in federated environments due to low device participation, local data that is not identically distributed, and local updating schemes. For instance, local data compression techniques can be used to detect outliers and compensate for the incompleteness of the data, while global compression techniques can be used to detect and reduce model redundancy.

Meanwhile, two compression techniques in the paper for decentralized training: extrapolation compression and difference compression. Using extrapolation compression, gradient updates are compressed by extrapolating from previous updates. In difference compression, only the difference between the current and previous gradient updates is sent. Additionally, they demonstrate that both compression techniques can reduce communication costs while maintaining good accuracy on various datasets and tasks. There are three main compression techniques for federated learning discussed [20]. First of all, the quantization technique reduces the number of bits used to represent model parameters by rounding them. Secondly, in sketching, model updates are projected onto a lower-dimensional space to

reduce their size. Finally dithering reduces the sensitivity of model updates to quantization errors by adding noise. However, one of their limitations is that they can all reduce model accuracy. It depends on the technique and compression level used to determine the degree of accuracy loss [13].

2.2 System heterogeneity

There may be differences in the hardware and software specifications among clients in a federated learning system, it is difficult to train a model that will work for all clients. To tackle the challenge of system heterogeneity in federated learning, various approaches aimed at handling asynchronous communication, fault tolerance, and active device sampling strategies.

2.2.1 Asynchronous Communication

In federated learning, by using asynchronous communication, each client does not need to wait for other clients to send updates before sending their own. Rather than having to wait for all clients to be ready before starting to train the model, the central server can start training the model straight away. This reduces the overall waiting time because clients can update the central server as soon as they are ready, and it also helps to address the challenge of system heterogeneity by allowing devices with different capabilities to participate in the training process. Several efficient algorithms have been presented. For instance, a number of federated learning algorithms are asynchronous [12]. The following challenges are handled differently by these algorithms:

- Straggler avoidance: The purpose of this method is to prevent devices from falling behind and delaying the convergence of the global model. The use of a deadline is one method of preventing stragglers, which specifies the maximum time a device has to upload its updates. It is also possible to use a voting mechanism, which allows the server to proceed with the aggregation even if some devices have not yet uploaded their updates.
- Staleness control: In this context, stale gradients refer to gradients that are outdated and may not reflect the current state of the data. An effective method of controlling staleness is to use a decay factor, which decreases the weight of older gradients. There is also the option of using a threshold, which only considers gradients that have been updated within a certain period.
- Weighted aggregation: This refers to the methods used to assign different weights to the gradients of different devices. The amount of data on the device or the device's computational power may need to be taken into account in this process.

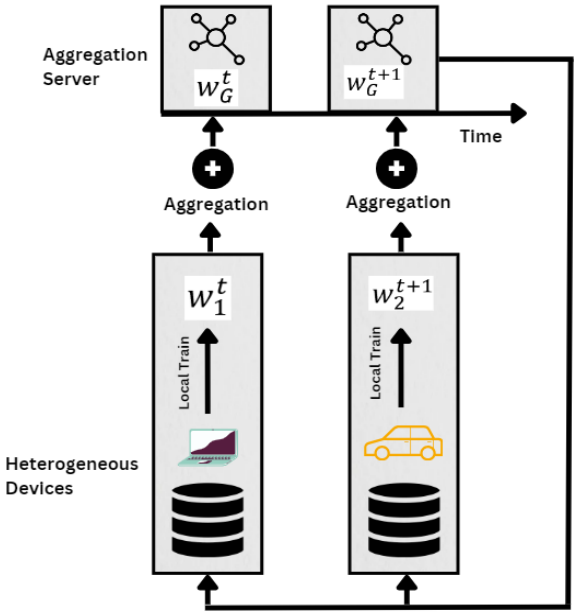

Fig. 5. Asynchronous Federated Learning.

Although staleness control can improve accuracy, it can also increase communication costs. In addition to improving convergence speed, straggler avoidance can also reduce accuracy. Choosing an asynchronous AFL method will depend on the specific application and the desired trade-offs [12]. Utilizing Asynchronous Communication in Federated Learning can improve Teacher Performance Evaluation in the education sector by accommodating varying schedules and enabling decentralized model updates without real-time synchronization, fostering broader teacher participation.

2.2.2 Active Device Sampling

The active device sampling method involves selecting a subset of clients for each training session. As a result, the central server only has to communicate with a subset of clients, resulting in improved efficiency in communication. Moreover, Presents federated learning with heterogeneous resources and the problem of client selection. When heterogeneous resources are involved, random client selection may not always lead to optimal performance.

As a solution to this problem, an active sampling algorithm FedCS was proposed in [34] that mitigates the problem of inefficient training processes when clients with limited computational resources or poor wireless channel conditions. There are two phases to the algorithm: The first step of the algorithm is to estimate the uncertainty of each client's local model. An analysis of model parameters can be conducted by calculating their variance, distance from the global model parameters, or entropy. Next, a subset of clients will participate in the next round of training. Based on the utility function that takes uncertainty and diversity into account, clients are selected. Although its advantages are numerous, there are also limitations, such as the computation costs of uncertainty estimation and sensitivity to the distribution of data. Increasing teacher performance evaluation in the education sector using FedCS Federated Learning involves optimizing collaborative model aggregation techniques to enhance accuracy and efficiency while addressing non-IID data challenges. Similarly, an incentive mechanism based on contract theory was proposed in [35] that will lead to mobile devices participating in federated learning by producing high-quality local training data. To overcome the issue of information asymmetry, the mechanism maps the contributed resources into appropriate rewards. A type-n data owner's utility function is as follows:

$$U_D(f_n) = R_n - \mu E_n^t = R_n - \mu[\frac{\psi}{\theta_n}\zeta C_n S_n f_n^2 + E_n^{com}] \quad (5)$$

Therefore, using high-quality data, the incentive mechanism simulates the use of mobile devices in federated learning. Furthermore, this approach is cost-effective and time efficient, since it requires fewer samples than a random sampling approach.

2.2.3 Fault Tolerance

A fault tolerance strategy is designed to deal with situations in which clients fail or become unavailable. To achieve this, redundancy, checkpointing, or byzantine fault tolerance can be used. Especially in large-scale deployments, fault tolerance is an important consideration for federated learning. With fault tolerance techniques, federated learning can continue even if a failure occurs. Various fault tolerance methods were discussed in [39] for federated learning, including the following:

- Checkpointing: Using this technique, the state of federated learning is periodically saved on a central server. If the process fails, it can be restarted from the checkpoint.
- Resilience to stragglers: Stragglers, or devices that are slow to participate in the federated learning process, can be reduced using techniques such as early stopping and deadline scheduling.
- Robust aggregation: In this technique, aggregation methods are used that are less sensitive to noise and outliers.

2.3 Statistical heterogeneity

It can be challenging to train a model that generalizes well to the population if the data on the clients are different. To address statistical heterogeneity, a variety of methods are available, such as data augmentation, data balancing, and federated learning with multiple rounds of communication.

2.3.1 Data balancing

Data balancing involves distributing data to all clients in an equitable manner in order to make them representative of the overall population. A variety of methods can be used to accomplish this, including randomly sampling data from each client.

The data balancing technique in federated learning prevents the global model from being biased toward the data of any particular client, thereby addressing statistical heterogeneity. Different clients' data can differ significantly, which can be helpful. This ensures that the global model can be improved by aggregating the data from different clients with diverse datasets. Furthermore, it enables collaboration between different clients in a privacy-preserving manner. This technique also ensures that the global model cannot be accessed by any particular client, protecting the clients' data privacy. Additionally, it helps to avoid overfitting the global model on a single client's data, improving the overall accuracy of the model. Adynamic weighted model aggregation algorithm was proposed in [49] where using statistical heterogeneity to overcome the constraint of non-IID data divergence on model convergence. Using the dynamic weighted model aggregation algorithm in Federated Learning enhances teacher performance evaluation in the education sector by adaptively assigning varying weights to model updates from different teachers, optimizing the collective model while considering individual teacher contributions and local dataset variations. Besides, considering local learning generality instead of proximal restriction when addressing data heterogeneity was proposed in [50]. Describe different approaches for addressing federated learning's data skew problem that can result in good machine learning

models even in environments where the data may be highly skewed [52]. There is strong evidence to suggest that data balancing methods can help address statistical heterogeneity in federated learning.

2.3.2 Model Averaging:

This involves averaging the local models from all clients before sending them to the central server. This can help to reduce the impact of statistical heterogeneity on the global model. A study found that model averaging aggregates multiple client models trained on heterogeneous data to obtain a well-performed global model [53]. In another study, federated versions of popular Domain Generalization algorithms were proposed in [54] and showed that by applying appropriate data augmentation, we can mitigate data heterogeneity in the federated setting. Overall, Statistical heterogeneity in federated learning can be mitigated by augmentation of data and dynamic weighted model aggregation.

2.4 Privacy concerns

In federated learning, one of the main goals is to protect the privacy of client data. Multiple approaches exist to address privacy concerns, such as differential privacy, homomorphic encryption, and secure multiparty computations. Here are some privacy and security techniques.

- **Homomorphic encryption** (HE): It is a type of encryption that allows computations to be performed on encrypted data. This can be used in federated learning to train a model on encrypted data without decrypting the data. A privacy-preserving federated learning (PPFL) algorithm was proposed in [57] that aggregates encrypted parameters without decryption on the centralized server. Furthermore, the algorithm uses a distributed cryptosystem to allow every node to use a different homomorphic encryption private key. A data structure that can transfer multiple parameters can alleviate the communication overhead associated with the HE schemes. Weights w_i consist of a bit representing the sign, a zero bit for preventing homomorphic additions from overflowing, and the rest of the bits for the value. There is a method for calculating the number of weights included in a ciphertext as D= $\frac{K}{L_0}$.
- **Secure Multiparty Computation** (SMC): A privacy-preserving federated learning framework was proposed in [58] based on chained SMC that achieves practical privacy preservation without impairing accuracy and convergence speed. In another study, a defense strategy was proposed [55], based on obfuscating the gradients of sensitive data while concealing data, which offers the highest level of protection while preserving FL performance. Specifically, the authors alter a few samples within a mini-batch to mimic the sensitive data at the gradient levels. By using this technique, sensitive data can be obscured without sacrificing FL performance. By using gradient projection, the proposed method ensures privacy without affecting performance. On the other hand, the use of sketching algorithms provides both privacy and performance benefits while maintaining accuracy [56]. Utilizing Secure Multiparty Computation in Federated Learning can enhance teacher performance evaluation in the education sector by preserving data privacy while enabling collaborative model training across distributed devices or teachers.

3. Discussion

Federated learning can empower teachers to easily evaluate their performance and identify areas for improvement. This can be done through data analytics that tracks student performance data, such as grades, attendance, and survey responses, and compares it to the teacher's performance data, such as the number of hours spent teaching, the number of classes taught, and the number of students taught. This will help teachers to better understand their strengths and weaknesses and make more informed decisions about their teaching. By comparing data sets, teachers can identify any areas where they are underperforming or need additional help. They can also use this data to determine the best strategies for optimizing their teaching, such as changing teaching methods or focusing on certain areas of instruction. This will enable them to provide more personalized and effective instruction to their student. Exploring the implications of the findings can help teachers make informed decisions about curriculum development, instructional resources, and targeted interventions for individual students. This data-driven approach allows teachers to address specific learning needs and ensure that each student receives the support necessary for academic success. The FL offers several advantages for teacher performance evaluation (TPE):

1. Data Privacy: FL protects sensitive teacher data by keeping it decentralized and avoiding data sharing.
2. Collaborative Learning: FL enables collaborative model training without data sharing, allowing institutions to benefit from collective knowledge.
3. Scalability: FL can accommodate large and diverse datasets from multiple institutions without compromising performance.
4. Privacy-Preserving Insights: FL provides insights into teacher performance trends and patterns without compromising data privacy.

However, there are also some challenges associated with FL:

- Communication Overhead: FL requires communication between participating institutions and the central server, which can introduce overhead and latency.
- Heterogeneous Data: The distributed nature of FL data can lead to heterogeneities in data quality and distribution, affecting model performance.
- Privacy Vulnerabilities: While FL aims to protect data privacy, there are still potential vulnerabilities that need to be addressed.

Overall, FL presents a promising approach for privacy-preserving teacher performance evaluation, enabling collaborative learning and insights without compromising data privacy. As FL technology continues to mature, it is expected to play an increasingly important role in teacher evaluation and professional development.

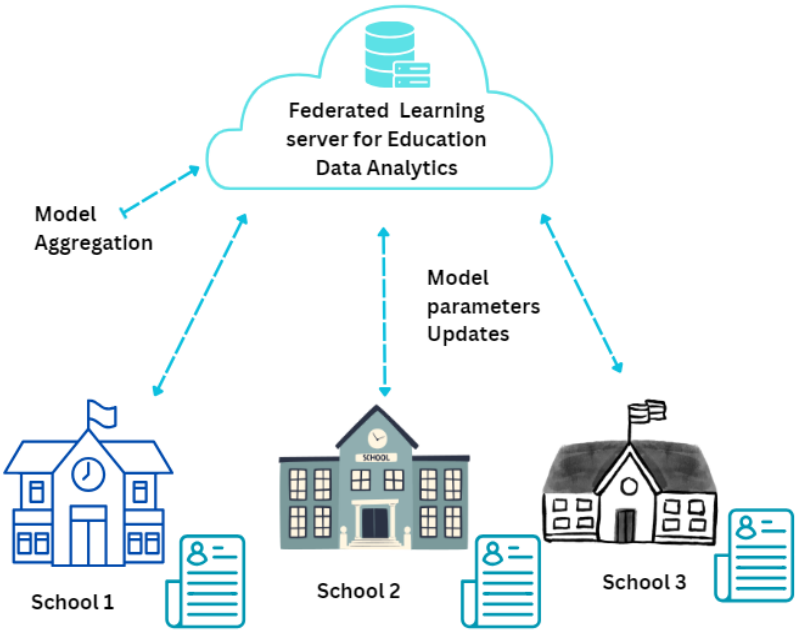

Fig. 6. To improve teacher performance evaluation using federated learning.

The different methodologies for overcoming challenges in federated learning have different strengths and weaknesses. Some methodologies may prioritize data privacy and security at the expense of communication efficiency, while others may focus on improving model performance but with potential risks to data privacy and security.

3.1 Communication Efficiency

One of the major challenges in federated learning is communication efficiency, which plays a crucial role in reducing the communication cost and time required to train a model. Communication efficiency refers to the ability to transmit and process data and model updates with minimal overhead, latency, and network bandwidth. It is especially critical in federated learning since the data is distributed across a large number of clients, which can cause significant communication overhead and latency.

In Federated Learning, the training data is distributed over a large number of clients, each with unreliable and relatively slow network connections. This presents challenges in terms of communication efficiency. Some of the challenges are:

- The typical clients in this setting are mobile phones, which have limited battery life and bandwidth.
- The communication costs can be high due to the large number of clients and the need to transfer model updates from each client to the central server.
- The network connections can be unreliable, leading to dropped or delayed updates, which can affect the quality of the trained model.

To address these challenges, Structured updates and sketched updates methods was proposed in [16] to reduce the uplink communication costs. These methods directly learn an update from a restricted space or compress a full model update before sending it to the server, respectively. Experiments show that these methods can reduce the communication cost by two orders of magnitude. The proposed methods are tested on both convolutional and recurrent networks, and the results show that they are effective in reducing communication costs while maintaining the quality of

the trained model. However, It assumes that the clients are honest and follow the protocol, but in practice, clients may be malicious and try to manipulate the model or the communication process. They do not consider the impact of the proposed methods on the convergence rate or the quality of the trained model, and it is unclear how much performance is sacrificed to achieve communication efficiency.

There are different methods to reduce communication costs in federated learning while maintaining model accuracy and minimizing the use of time and space resources [15]. A method of distributing a model was proposed in [17] with a structure different from that of the server model, distributing a model suitable for clients with different data sizes, and training a server model using the reconstructed model trained by the client. An algorithm RingFed was proposed in [18], a framework that reduces communication overhead by transmitting updated parameters between each client in turn, and only the final result is transmitted to the central server. Another algorithm FedPSO was proposed in [19], that uses particle swarm optimization to update the global model by transmitting score values rather than large weights, which reduces the amount of data used in network communication and improves the accuracy of the global model they also found to be effective in reducing communication costs compared to standard methods in experiments conducted on an image classification model. Specifically, the proposed method was able to accomplish training with a cost of 0.229 times smaller than the standard method [19]. On the other hand, FedSCA was proposed in [20], that reduces data communication by transferring score principles rather than all client models' weights and utilizing the Sine Cosine Algorithm (SCA) mechanism as a weight updating technique to improve the clients' models. All of these methods have been shown to reduce communication costs while maintaining model accuracy, with some methods showing improvements in accuracy over existing algorithms.

3.2 System Heterogeneity

Devices in federated networks may have different storage, communication, and computational capabilities, and battery limits due to variability in hardware (CPU, GPU) and network connectivity [14]. It may cause a lot of stragglers directly. This problem can also cause the accuracy reduction also. Sometimes, multiple clients are connected to the learning process who are not active at the processing time or can upload an unreliable model which can decrease the speed of the training process.

Federated learning can be used to overcome the challenge of system heterogeneity in distributed computing architectures. A scalable production system was described in [4] for federated learning in mobile devices, while the challenges and methods of federated learning in heterogeneous and potentially massive networks and they examine the specific characteristics and challenges of FL [3], as well as how it is now done and what can be done in the future. An adaptive federated learning algorithm was proposed in [21], for learning model parameters from data distributed across multiple edge nodes, while the authors in [22] focus on the statistical challenge of federated learning when local data is non-IID. Here they can collectively suggest that federated learning can enable model training on decentralized data while keeping data localized and that there are various approaches and challenges to consider when implementing federated learning. Besides, to reduce communication costs in the presence of system heterogeneity the authors in [23] propose a multistage optimization scheme that nearly matches the lower bound on communication complexity across all heterogeneity levels. The authors of [24,25] propose two-stream models with MMD constraint and feature fusion modules, respectively, to reduce required communication rounds. FedLin, a new algorithm was proposed on [26] that exploits past gradients and employs client-specific learning rates to achieve linear convergence in the presence of objective and systems heterogeneity. Heterogeneous systems are handled by using the FedSAE [36] algorithm which selects only devices having a big amount of global model that reduces communication cost. It reduces stragglers and utilizes only active devices. Another approach FedSkel was proposed in [37] for heterogeneous systems. Client sampling [38] can handle both heterogeneous and statistical heterogeneity and converges to real-time. Another future direction in federated learning research is the development of more robust aggregation algorithms.

3.3 Statistical Heterogeneity

Data may be distributed unevenly across devices, leading to non-IID (independent and identically distributed) data. Addressing the biases and challenges caused by skewed data distributions is crucial for accurate model training. The definition of statistical heterogeneity in FL is the variation in data distributions and model parameter distributions among different parties [27]. The training process can be affected by these differences, leading to issues such as slow model convergence, overfitting, and degradation of model accuracy. Addressing the underlying statistical heterogeneity is necessary to tackle these issues. Several approaches can be taken to tackle the issue of statistical heterogeneity in FL. Data transformation, model adaptation, data preprocessing, and weight regularization are among those included:

- Data Transformation: Transforming data from different parties is a method to address the statistical heterogeneity issue in FL and make it more homogeneous [28]. Training the data from various sources to have the same distribution or applying statistical transformations such as normalization or standardization can accomplish this.
- Model Adaptation: It is a method of addressing the statistical heterogeneity issue in FL by adapting the model to the data from different parties based on their heterogeneity levels [29]. This can be done by training different

models for each party, or by training a single model on the aggregated data from all parties. These techniques ensure that nodes with more representative or higher-quality data have a greater influence on the final model.

- Data Preprocessing Techniques: Another method is to apply data preprocessing techniques that can standardize or normalize the data across different nodes, making it more comparable and reducing the impact of heterogeneity. For example, techniques such as feature scaling, data augmentation, or applying dimensionality reduction methods can help address variations in data distribution across nodes.
- Weight Regularization: It is a method of addressing the statistical heterogeneity issue in FL by constraining the weights of the model to be more homogeneous across different parties. This can be done by adding a regularization term to the loss function that penalizes weights that differ significantly between parties.

These techniques typically involve adding penalties or constraints to the learning process that encourage the model to be more robust against data heterogeneity. For example, adding a penalty on the norm of learned results in the objective function can slow down the influence of heterogeneity rather than eliminate it.

3.4 Privacy Restriction:

Ensuring data privacy and the security of participants' information during the training process is essential. Developing robust privacy-preserving techniques, such as differential privacy and secure multi-party computation, helps protect sensitive data. In federated learning, privacy is a big concern area. Though data is preserved locally in the device, there is some chance of data leakage during model transmission. As the encrypted method is used during sending the model to the central server, unreliable clients can affect the learning process. We should concentrate on privacy preservation so that malicious and unreliable clients can reduce the accuracy of the result. We should find some effective algorithms to increase data privacy.

The papers suggest several ways to overcome challenges in ensuring that the privacy of data is not compromised during the training process in federated learning. A split-learning-based framework was proposed in [40] that enables FL clients to maximize available resources within their local network without compromising the benefits of an FL approach. PruneFL, a novel FL approach was proposed in [41] with adaptive and distributed parameter pruning, which adapts the model size during FL to reduce both communication and computation overhead and minimize the overall training time while maintaining a similar accuracy as the original model. The study extends the PyTorch library to address some limitations. Further improvements in efficient sparse matrix multiplication can address these limitations in the future. The dataset used includes images from FEMINIST and CelebA, with model architectures VGG-11, ResNet-18, and MobileNetV3-Small. The experiments are conducted on a real edge computing prototype and a simulated setting. The results are compared with five baselines, but comparisons may not be directly comparable due to differences in model size. PruneFL achieves good results while limiting the model size and can be combined with other compression techniques. The analysis provides proof for theorems related to the algorithm's optimality [41]. On the other hand, a federated weighted average with differential privacy (DP-FedAW) algorithm was proposed in [42] which quantifies the degree of non-IID for different user datasets and adjusts the aggregation weights of each user, effectively alleviating the model convergence problem caused by differences in Non-IID data during the training process. Finally, a method for distributed selecting relevant data was proposed in [43] where a benchmark model trained on a small benchmark dataset is used to evaluate the relevance of individual data samples at each client and select the data with sufficiently high relevance.

In a different study, Laplace distributed noise and weighted aggregation algorithm was proposed in [30] that uses with mutual information as a scoring mechanism to protect user data privacy. Privacy-preserving and communication-efficient method in an FL framework with one-shot offline knowledge distillation using unlabelled, cross-domain, non-sensitive public data [31]. Besides these in [32] the authors combine cryptographic tools with federated learning to design privacy-preserving protocols, which protect every participant parameter’s information. HybridAlpha was proposed in [33] which is an approach for privacy-preserving federated learning employing an SMC protocol based on functional encryption. The SMC network allows multiple parties to compute a function simultaneously without disclosing their individual data. The data never leaves the devices of participants, so it can be used to train machine learning models in a privacy-preserving manner.

4. Proposed Clustering Technique

Overcoming system heterogeneity and statistical heterogeneity is a big challenge in Federated Learning. A clustering approach can reduce this heterogeneity issue. If we can cluster the same type of devices in a cluster and make a cluster head that only communicates with the central server, it can reduce communication rounds and overhead as well. Multiple clusters can be used for a large number of connected devices during the training process. Clustering can also help to improve the scalability and performance of the federated learning system by reducing the number of devices that need to communicate with the central server. According to particular characteristics, such as their physical location, network connectivity, or hardware capabilities, clustering entails putting similar devices in one group. The devices can then cooperate to train the model locally, sending only updates to the central server. In federated learning environments where devices may have limited network access or processing capabilities, this might lessen the amount of

communication that is necessary between devices and the central server. By clustering devices based on their characteristics, federated learning can be expanded to more thoroughly explore the potential of collaborative training. This approach allows devices with similar capabilities to work together, optimizing the training process and reducing the burden on the central server.

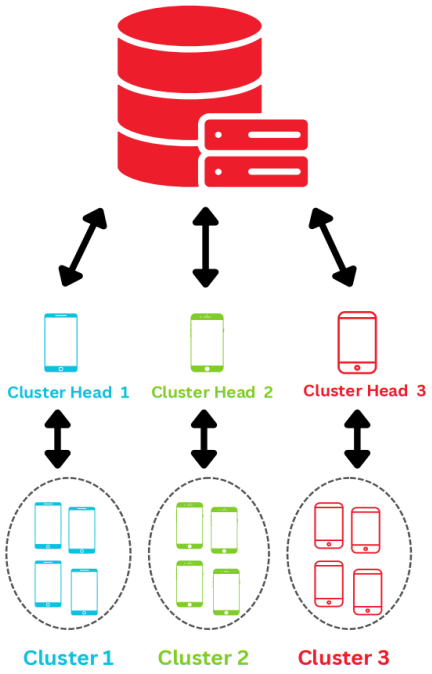

Fig. 7. Proposed cluster-based federated learning.

In federated learning, clustering can be implemented in several ways to overcome system heterogeneity and statistical heterogeneity:

1. **Implement a centralized clustering algorithm:** Devices can be clustered based on their hardware specifications, such as CPU, memory, and storage capacity. In this way, similar devices will be grouped.
2. **Decentralize clustering algorithms:** The devices can cluster themselves based on the distribution of their data. Various clustering algorithms can be used to accomplish this, including k-means clustering, hierarchical clustering, and density-based clustering.
3. **Use a hybrid clustering algorithm**: Initially, the central server clusters the devices based on their hardware specifications, and then the devices cluster themselves based on their data distributions. In this way, the best of both worlds can be achieved.

The cluster heads will be able to communicate with the central server once the devices have been clustered. As a result, only the cluster heads will need to communicate with the central server, reducing the number of communication rounds and overhead.

Clustering can also help to improve the scalability and performance of the federated learning system by reducing the number of devices that need to communicate with the central server. According to particular characteristics, such as their physical location, network connectivity, or hardware capabilities, clustering entails putting similar devices in one group. The devices can then cooperate to train the model locally, sending only updates to the central server. In federated learning environments where devices may have limited network access or processing capabilities, this might lessen the amount of communication that is necessary between devices and the central server.

Here is the most popular recent research on the use of clustering algorithms to reduce the burden on the central server in federated learning:

A hierarchical clustering algorithm was proposed in [44] for federated clustered learning, based on the similarity between clients. To perform the algorithm, the clients are first clustered based on how similar their data is. A local model is then trained on clients within each cluster. In the next step, the local models are aggregated to create a global model. MNIST and CIFAR-10 datasets were used to evaluate the authors' algorithm. They demonstrated that their algorithm is comparable to other algorithms based on federated learning. Although the framework is useful, it does have some limitations. It is computationally expensive to use the hierarchical clustering algorithm, especially if there are many clients. A second issue is that the framework requires clients to share their data with the central server, which could pose a privacy issue.

Another algorithm for dealing with heterogeneity in federated learning using resource-aware clustering was proposed in [45]. When clustering the clients, the algorithm takes into account the resources available at each client, such as the data amount and the computational capacity. According to the algorithm:

- Calculate the similarity of resources between each pair of clients first. The Euclidean distance is used for this purpose.
- In the second step, the clients are clustered according to the similarity of their resources. The process can be carried out using a clustering algorithm such as k-means or hierarchical clustering.
- The third step involves assigning each client to a cluster. Clients are assigned to the cluster with the highest similarity in resources.

Based on the MNIST and CIFAR-10 datasets, the authors evaluated the algorithm.

Stochastic clustered federated learning (StoCFL) algorithm was proposed in [46] for handling non-ID data in federated learning. As a first step, StoCFL clusters clients based on how similar their data is. After that, stochastic gradient descent (SGD) is used to train the clients in each cluster on a local model. To create a global model, the updated local models are aggregated. StoCFL was evaluated on the MNIST and CIFAR-10 datasets. Despite this, the algorithm has some limitations. As the global model is not trained based on all of the data, the algorithm may not be as accurate as other federated learning algorithms. In addition, the algorithm may not be robust to heterogeneous data, since each cluster may have a different distribution of clients' data.

To tackle the challenges of data heterogeneity and communication overhead in federated learning FedSoft was proposed in [47]. Each client belongs to multiple clusters in FedSoft, which clusters the clients into soft clusters. Using data from their clusters and clusters near them, clients train their local models. A proximal gradient method is then used to update the local models to ensure that they do not overfit the data. The authors evaluated FedSoft on two datasets: the MNIST dataset and the CIFAR-10 dataset. However, the performance of FedSoft may degrade if the number of clients is large and the theoretical analysis of FedSoft is based on several assumptions, which may not hold in practice.

An another study [48] present an algorithm where clients are grouped based on how closely their local updates match using hierarchical clustering. Each cluster's clients then exchange updates among themselves and send them to the central server after aggregating them. In this way, the global model is more accurate and communication overhead is reduced. Two non-IID datasets were used to evaluate the proposed algorithm: MNIST and CIFAR-10. The proposed algorithm was compared to other federated learning algorithms, such as FedAvg and FedProx. Compared to the other algorithms, the proposed algorithm achieved a higher level of accuracy. However, the proposed algorithm has some limitations. It requires that clients communicate with each other first. In some circumstances, this may not be feasible, such as when clients reside in different countries. Secondly, if the clusters are not well-defined, the algorithm may not achieve good accuracy [48].

In real-world scenarios, clustered federated learning algorithms face the following challenges:

- Limited communication bandwidth: Many clients and servers share a limited communication bandwidth. In federated learning algorithms, this can cause a bottleneck.
- Unreliable communication channels: There may be problems with the communication channels between the clients and the central server. It is possible for packets to be lost and delays to occur as a result.
- Privacy concerns: Clients may be concerned about the privacy of their personal information. Federated learning algorithms should be designed in a manner that protects the privacy of data.
- Heterogeneous devices: There may be different types of devices with different computational resources among the clients. It is therefore difficult to train a global model that is accurate for all clients as a result.

Despite these challenges, clustered federated learning has the potential to provide a promising method for building machine learning models that preserve privacy. Clustered federated learning is capable of solving a wide range of real-world problems by addressing the challenges mentioned above. For instance, it can be used to build large-scale machine learning models without revealing any sensitive information, such as training data, to the federated learning server. Additionally, clustered federated learning can be used to build powerful models by leveraging the distributed resources of a local network of machines. Finally, clustered federated learning can be used to improve the performance of machine learning models by taking advantage of the data-rich environment of a local network. Clustered federated learning is an important tool for organizations looking to leverage the power of machine learning while retaining the privacy of their data. It can also help to reduce the computational requirements for training machine learning models. This can help organizations to save on the cost of computing resources, while still achieving the same performance. Additionally, clustered federated learning can also be used to ensure that no single organization holds too much control over a shared dataset.

Table 2. Represents the different comparison clustering technique in Federated Learning

Ref.	Methods	Contributions	Limitations
[46]	StoCFL	• Handling non-ID data in federated learning.	• Requires all clients to participate in the FL process and necessitates information about the number of clusters, which are often difficult to obtain.
[47]	FedSOFT	• Significantly reduce client workload compared to traditional clustered federated learning.	• Limited scalability • Sensitivity to hyperparameters • Assumption of data similarity:
[48]	FL+ HC	• A modification to federated learning (FL) by introducing a hierarchical clustering step (FL+HC) to separate clusters of clients	• This method requires additional communication rounds and may not be effective for all types of non-IID data

5. Future Directions

Federated learning (FL) holds the potential to empower the educational sector, particularly in Teacher Performance Evaluation (TPE), allowing educators to assess their performance with ease and precision. Through data analytics that tracks student performance metrics such as grades, attendance, and survey responses, and compares them to the teacher's performance data, including teaching hours, class load, and the number of students taught, FL can offer valuable insights. These insights enable educators to gain a comprehensive understanding of their strengths and areas for improvement, enabling data-informed decisions in their teaching practices.

Federated learning offers several distinct advantages for Teacher Performance Evaluation (TPE):

1. **Data Privacy**: FL upholds stringent data privacy standards, ensuring that sensitive teacher-related data remains decentralized and shielded from external access. This protection is paramount, particularly in the education sector, where privacy is a fundamental concern.
2. **Collaborative Learning**: FL's collaborative model training paradigm fosters collective knowledge-building across institutions without the need for actual data sharing. Teachers, schools, and education authorities can harness the collective wisdom while preserving data privacy.
3. **Scalability**: FL's architecture exhibits remarkable scalability, accommodating diverse and extensive datasets from multiple institutions without compromising performance. This scalability is advantageous for educational institutions of varying sizes and data volumes.
4. **Privacy-Preserving Insights**: FL's unique approach enables the extraction of valuable insights into teacher performance trends and patterns without compromising individual data privacy, aligning perfectly with the ethical considerations in the education sector.

However, challenges persist within the realm of federated learning:

Challenge 1: Communication Overhead

FL mandates communication between participating institutions or nodes and a central server, which can introduce significant communication overhead and latency. In the context of the educational sector, this challenge hinders the timely dissemination of insights and recommendations regarding Teacher Performance Evaluation (TPE).

Possible Solutions:

1. **Asynchronous Communication**: Implementing asynchronous communication allows participating nodes to update the central server without waiting for others to finish, reducing communication delays. This is particularly useful when there are a large number of clients or limited communication bandwidth.
2. **Edge Processing**: Utilize edge computing capabilities to enable preliminary data processing and model updates at the local level before transmitting relevant information to the central server. This minimizes data transfer needs and reduces communication overhead.
3. **Efficient Data Compression**: Employ efficient data compression techniques to reduce the volume of data transferred between clients and the central server. This can help mitigate communication overhead without compromising data integrity.

Challenge 2: Heterogeneous Data

In the educational sector, diverse sources of data can lead to heterogeneities in data quality and distribution. Heterogeneous data can affect model performance, making it challenging to derive accurate insights for TPE.

Possible Solutions:

1. **Resource-Aware Clustering**: Implement clustering techniques to group clients with similar data types or data characteristics. By clustering nodes based on their data attributes, you can create subsets of clients with comparable data, enabling more accurate model training and performance assessment.
2. **Hybrid Models**: Develop hybrid federated learning models that accommodate heterogeneity by allowing different client groups to train distinct model variations tailored to their respective data. The central server can then combine these models intelligently.
3. **Data Transformation and Preprocessing**: Employ data transformation and preprocessing techniques to standardize or normalize data across different nodes. Techniques like feature scaling, data augmentation, and dimensionality reduction can help mitigate variations in data distribution.

Challenge 3: Privacy Vulnerabilities

Privacy remains a critical concern in FL, particularly in the educational sector, where sensitive student and teacher data must be safeguarded. Despite FL's privacy-preserving approach, potential vulnerabilities exist that could compromise data privacy.

Possible Solutions:

1. **Differential Privacy**: Implement differential privacy techniques to add noise to data before sharing, ensuring that individual data points remain indistinguishable. This method enhances data privacy and reduces the risk of information leakage.
2. **Secure Multi-Party** Computation (SMC): Utilize SMC protocols to allow multiple parties to compute a function collectively without revealing their data. SMC ensures that data remains on the participants' devices, preserving privacy during the model training process.
3. **Enhanced Encryption**: Explore advanced encryption methods and cryptographic tools to protect data at all stages of FL, from local model updates to transmission to the central server. This approach further fortifies data privacy.

Challenge 4: Scalability

The scalability of FL is vital in the educational sector, where the number of participating institutions or nodes can be substantial. Scaling FL without compromising performance is an ongoing challenge.

Possible Solutions:

1. **Hierarchical Structures**: Organize FL into hierarchical structures, where local aggregations occur at different levels before reaching the central server. This reduces the computational load on the central server and enhances scalability.
2. **Efficient Aggregation Algorithms**: Develop more efficient aggregation algorithms that can handle a large number of clients or nodes while minimizing the computational demands. These algorithms ensure that scalability is maintained as FL extends to more participants.
3. **Adaptive Communication**: Implement adaptive communication strategies that prioritize interactions with nodes that contribute the most valuable data, thus reducing the overall communication burden on the central server.

Addressing these challenges with the proposed solutions is essential for the successful implementation of federated learning in the educational sector, particularly in the context of Teacher Performance Evaluation. By improving communication efficiency, managing data heterogeneity, strengthening privacy protections, and enhancing scalability, FL can play a pivotal role in elevating the quality of education and teacher assessments while safeguarding sensitive data.

6. Conclusion

A brief overview of federated learning has been provided in this article. In conclusion, the literature review highlights that clustering different types of nodes is a valuable approach to address challenges in federated learning. By grouping nodes with high computational power and reliable network connections together, computationally intensive tasks can be efficiently performed, while nodes with lower computational power or intermittent network connections can form separate clusters to contribute with smaller models or local updates. This clustering approach optimizes the communication process, reducing the number of rounds and overall data transmission. By addressing data heterogeneity through clustering nodes with similar data types, federated learning can mitigate the impact on model performance. Fl collaborates with multiple decentralized nodes to train the model locally without exchanging raw data to the central server for aggregation. It has been a big area of research for both industries and academics in recent years. It enables privacy and security for leakage of data. We conduct various papers related to the challenges and analyze to overcome

the issues as well. Finally, we outline a promising future direction for further research. Overall, the findings emphasize the importance of clustering as a valuable technique to overcome challenges and optimize the effectiveness of federated learning.

Acknowledgment

The work has been supported by UMPSA RDU GRANTRDU230352, titled “A fraud detection model for instructor’s evaluation based on semantic keyword extraction using ma-chine learning”.

References

[1] S. Shen, T. Zhu, D. Wu, W. Wang, and W. Zhou, “From Distributed Machine Learning To Federated Learning: In The View Of Data Privacy And Security,” Concurrency and Computation, vol. 34, no. 16, Jul. 2022, doi: 10.1002/cpe.6002.

[2] C. Briggs, Z. Fan, and P. Andras, “Federated learning with hierarchical clustering of local updates to improve training on non-iid data,” arXiv preprint arXiv:2004.11791, 2020

[3] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated Learning: Challenges, Methods, and Future Directions,” IEEE Signal Process. Mag., vol. 37, no. 3, pp. 50–60, May 2020, doi: 10.1109/MSP.2020.2975749.

[4] K. Bonawitz et al., “Towards Federated Learning at Scale: System Design.” arXiv, Mar. 22, 2019. Accessed: Mar. 20, 2023. [Online]. Available: http://arxiv.org/abs/1902.01046

[5] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, “A review of applications in federated learning,” Computers & Industrial Engineering, vol. 149, p. 106854, Nov. 2020, doi: 10.1016/j.cie.2020.106854.

[6] A. Reisizadeh, A. Mokhtari, H. Hassani, A. Jadbabaie, and R. Pedarsani, “Fedpaq: A communication-efficient federated learning method with periodic averaging and quantization,” in International Conference on Artificial Intelligence and Statistics. PMLR, 2020, pp. 2021–2031

[7] “Challenges and future directions of secure federated learning: a survey | SpringerLink.” https://link.springer.com/article/10.1007/s11704-021-0598-z (accessed Mar. 21, 2023).

[8] C. Ma, J. Li, M. Ding, K. Wei, W. Chen, and H. V. Poor, “Federated Learning with Unreliable Clients: Performance Analysis and Mechanism Design.” arXiv, Jul. 31, 2021. doi: 10.48550/arXiv.2105.06256.

[9] Latif, U., Khan., Madyan, Alsenwi., Zhu, Han., Choong, Seon, Hong. (2020). Self Organizing Federated Learning Over Wireless Networks: A Socially Aware Clustering Approach. 453-458. doi: 10.1109/ICOIN48656.2020.9016505

[10] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated Optimization in Heterogeneous Networks,” 2018, [Online]. Available: http://arxiv.org/abs/1812.06127.

[11] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated learning with non-iid data,” arXiv preprint arXiv:1806.00582, 2018

[12] C. Xu, Y. Qu, Y. Xiang, and L. Gao, “Asynchronous Federated Learning on Heterogeneous Devices: A Survey.” arXiv, Mar. 30, 2023. Accessed: Jul. 21, 2023. [Online]. Available: http://arxiv.org/abs/2109.04269

[13] Hanlin Tang, Shaoduo Gan, Ce Zhang, Tong Zhang, and Ji Liu. 2018. Communication compression for decentralized training. In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). Curran Associates Inc., Red Hook, NY, USA, 7663–7673.

[14] M. Asad, A. Moustafa, and T. Ito, “FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning,” Applied Sciences, vol. 10, no. 8, Art. no. 8, Jan. 2020, doi: 10.3390/app10082864.

[15] M. Chen, N. Shlezinger, H. V. Poor, Y. C. Eldar, and S. Cui, “Communication-efficient federated learning,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 17, p. e2024789118, 2021, doi: 10.1073/pnas.2024789118.

[16] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, “Federated Learning: Strategies for Improving Communication Efficiency.” arXiv, Oct. 30, 2017. Accessed: May 10, 2023. [Online]. Available: http://arxiv.org/abs/1610.05492

[17] Kang, D.; Ahn, C.W. Communication Cost Reduction with Partial Structure in Federated Learning. Electronics 2021, 10, 2081. https://doi.org/10.3390/ electronics10172081

[18] G. Yang, K. Mu, C. Song, Z. Yang, and T. Gong, “RingFed: Reducing Communication Costs in Federated Learning on Non-IID Data.” arXiv, Jul. 19, 2021. Accessed: May 10, 2023. [Online]. Available: http://arxiv.org/abs/2107.08873

[19] Park, S.; Suh, Y.; Lee, J. FedPSO: Federated Learning Using Particle Swarm Optimization to Reduce Communication Costs. Sensors 2021, 21, 600. https:// doi.org/10.3390/s21020600

[20] A. K. Abasi, M. Aloqaily, M. Guizani and F. Karray, "Sine Cosine Algorithm for Reducing Communication Costs of Federated Learning," 2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, 2022, pp. 55-60, doi: 10.1109/MeditCom55741.2022.9928614.

[21] S. Wang *et al.*, “Adaptive Federated Learning in Resource Constrained Edge Computing Systems.” arXiv, Feb. 16, 2019. Accessed: May 10, 2023. [Online]. Available: http://arxiv.org/abs/1804.05271

[22] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated Learning with Non-IID Data,” 2018, doi: 10.48550/arXiv.1806.00582.

[23] Hou, C., Thekumparampil, K.K., Fanti, G.C., & Oh, S. (2021). Reducing the Communication Cost of Federated Learning through Multistage Optimization. ArXiv, abs/2108.06869.

[24] X. Yao, C. Huang, and L. Sun, “Two-Stream Federated Learning: Reduce the Communication Costs,” in *2018 IEEE Visual Communications and Image Processing (VCIP)*, Taichung, Taiwan: IEEE, Dec. 2018, pp. 1–4. doi: 10.1109/VCIP.2018.8698609

[25] X. Yao, T. Huang, C. Wu, R.-X. Zhang, and L. Sun, “Federated Learning with Additional Mechanisms on Clients to Reduce Communication Costs.” arXiv, Sep. 01, 2019. Accessed: May 10, 2023. [Online]. Available: http://arxiv.org/abs/1908.05891

[26] Mitra, A., Jaafar, R.H., Pappas, G.J., & Hassani, H. (2021). Achieving Linear Convergence in Federated Learning under Objective and Systems Heterogeneity. ArXiv, abs/2102.07053.
[27] H. Zhu, J. Xu, S. Liu, and Y. Jin, "Federated Learning on Non-IID Data: A Survey." arXiv, Jun. 12, 2021. doi: 10.48550/arXiv.2106.06843.
[28] K. Bonawitz et al., "Towards Federated Learning at Scale: System Design," Proceedings of Machine Learning and Systems, vol. 1, pp. 374–388, Apr. 2019.
[29] R. Sun et al., "FedMSA: A Model Selection and Adaptation System for Federated Learning," Sensors, vol. 22, no. 19, p. 7244, Sep. 2022, doi: 10.3390/s22197244.
[30] S. Wang, Y. Li, A. Zhao, and Q. Wang, "Privacy Protection in Federated Learning Based on Differential Privacy and Mutual Information," *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture*, pp. 428–435, Oct. 2021, doi: 10.1145/3495018.3495093.
[31] X. Gong *et al.*, "Preserving Privacy in Federated Learning with Ensemble Cross-Domain Knowledge Distillation," *AAAI*, vol. 36, no. 11, pp. 11891–11899, Jun. 2022, doi: 10.1609/aaai.v36i11.21446.
[32] W. Yang, B. Liu, C. Lu, and N. Yu, "Privacy Preserving on Updated Parameters in Federated Learning," *Proceedings of the ACM Turing Celebration Conference - China*, pp. 27–31, May 2020, doi: 10.1145/3393527.3393533.
[33] R. Xu, N. Baracaldo, Y. Zhou, A. Anwar, and H. Ludwig, "HybridAlpha: An Efficient Approach for Privacy-Preserving Federated Learning," *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*, pp. 13–23, Nov. 2019, doi: 10.1145/3338501.3357371.
[34] T. Nishio and R. Yonetani, "Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge," in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), May 2019, pp. 1–7. doi: 10.1109/ICC.2019.8761315.
[35] Kang, Jiawen & Xiong, Zehui & Niyato, Dusit & Yu, Han & Liang, Ying-Chang & Kim, Dong In. (2019). Incentive Design for Efficient Federated Learning in Mobile Networks: A Contract Theory Approach. 1-5. 10.1109/VTS-APWCS.2019.8851649.
[36] L. Li, M. Duan, D. Liu, Y. Zhang, A. Ren, X. Chen, Y. Tan, C. Wang, FedSAE: A Novel Self-Adaptive Federated Learning Framework in Heterogeneous Systems, IJCNN 2021
[37] J. Luo, J. Yang, X. Ye, X. Guo, W. Zhao, FedSkel: Efficient Federated Learning on Heterogeneous Systems with Skeleton Gradients Update, In Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM '21), pp-(1–5), 2021
[38] B. Luo, W. Xiao, S. Wang, J. Huang, L. Tassiulas, Tackling System and Statistical Heterogeneity for Federated Learning with Adaptive Client Sampling, 2112.11256v1 [cs.LG] , 2021
[39] S. Liu, N. Gupta, and N. H. Vaidya, "Redundancy in cost functions for Byzantine fault-tolerant federated learning," in Proceedings of the First Workshop on Systems Challenges in Reliable and Secure Federated Learning, Virtual Event Germany: ACM, Oct. 2021, pp. 4–6. doi: 10.1145/3477114.3488761.
[40] S. Bharti and A. Mcgibney, "Privacy-Aware Resource Sharing in Cross-Device Federated Model Training for Collaborative Predictive Maintenance," *IEEE Access*, vol. 9, pp. 120367–120379, 2021, doi: 10.1109/ACCESS.2021.3108839.
[41] Y. Jiang *et al.*, "Model Pruning Enables Efficient Federated Learning on Edge Devices." arXiv, Apr. 06, 2022. Accessed: Jul. 14, 2023. [Online]. Available: http://arxiv.org/abs/1909.12326
[42] Tan, Q., Wang, B., Yu, H., Wu, S., Qian, Y., & Tao, Y. (2023). DP-FEDAW: FEDERATED LEARNING WITH DIFFERENTIAL PRIVACY IN NON-IID DATA. International Journal of Engineering Technologies and Management Research, 10(5), 34–49. https://doi.org/10.29121/ijetmr.v10.i5.2023.1328
[43] T. Tuor, S. Wang, B. J. Ko, C. Liu, and K. K. Leung, "Overcoming Noisy and Irrelevant Data in Federated Learning." arXiv, Jun. 22, 2020. Accessed: Jul. 14, 2023. [Online]. Available: http://arxiv.org/abs/2001.08300
[44] A. Ghosh, J. Chung, D. Yin, and K. Ramchandran, "An Efficient Framework for Clustered Federated Learning," in Advances in Neural Information Processing Systems, Curran Associates, Inc., 2020, pp. 19586–19597. Accessed: Jul. 10, 2023. [Online].Available:https://proceedings.neurips.cc/paper_files/paper/2020/hash/e32cc80bf07915058ce90722ee17bb71-Abstract.html4
[45] R. Mishra, H. P. Gupta, and G. Banga, "Resource Aware Clustering for Tackling the Heterogeneity of Participants in Federated Learning." arXiv, Jun. 07, 2023. Accessed: Jul. 10, 2023. [Online]. Available: http://arxiv.org/abs/2306.04207
[46] D. Zeng, X. Hu, S. Liu, Y. Yu, Q. Wang, and Z. Xu, "Stochastic Clustered Federated Learning." arXiv, Mar. 01, 2023. Accessed: Jul. 10, 2023. [Online]. Available: http://arxiv.org/abs/2303.00897
[47] Y. Ruan and C. Joe-Wong, "FedSoft: Soft Clustered Federated Learning with Proximal Local Updating." arXiv, Mar. 22, 2022. Accessed: Jul. 10, 2023. [Online]. Available: http://arxiv.org/abs/2112.06053
[48] C. Briggs, Z. Fan, and P. Andras, "Federated learning with hierarchical clustering of local updates to improve training on non-IID data." arXiv, May 06, 2020. Accessed: Aug. 20, 2023. [Online]. Available: http://arxiv.org/abs/2004.11791
[49] A. Chen, Y. Fu, L. Wang, and G. Duan, "DWFed: A statistical- heterogeneity-based dynamic weighted model aggregation algorithm for federated learning," *Frontiers in Neurorobotics*, vol. 16, 2022, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnbot.2022.1041553
[50] M. Mendieta, T. Yang, P. Wang, M. Lee, Z. Ding, and C. Chen, "Local Learning Matters: Rethinking Data Heterogeneity in Federated Learning." arXiv, Apr. 13, 2022. Accessed: Jul. 22, 2023. [Online]. Available: http://arxiv.org/abs/2111.14213
[51] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data." arXiv, Jan. 26, 2023. Accessed: Jul. 22, 2023. [Online]. Available: http://arxiv.org/abs/1602.05629
[52] D. Verma, G. White, S. Julier, S. Pasteris, G. Cirincione, and S. Chakraborty, "Approaches to address the Data Skew Problem in Federated Learning," May 2019, p. 50. doi: 10.1117/12.2519621.
[53] T. Zhou, Z. Lin, J. Zhang, and D. H. K. Tsang, "Understanding Model Averaging in Federated Learning on Heterogeneous Data." arXiv, May 20, 2023. doi: 10.48550/arXiv.2305.07845.
[54] "[2206.09979] Mitigating Data Heterogeneity in Federated Learning with Data Augmentation." https://arxiv.org/abs/2206.09979 (accessed Jul. 23, 2023).
[55] J. Wu, M. Hayat, M. Zhou, and M. Harandi, "Defense against Privacy Leakage in Federated Learning." arXiv, Sep. 13, 2022. doi: 10.48550/arXiv.2209.05724.

[56] Z. Liu, T. Li, V. Smith, and V. Sekar, "Enhancing the Privacy of Federated Learning with Sketching," *ArXiv*, Nov. 2019, Accessed: Jul. 23, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Enhancing-the-Privacy-of-Federated-Learning-with-Liu-Li/cb68e7849eb9ba864926fd869570f04f6ec4edd1#citing-papers
[57] Park, Jaehyoung, and Hyuk Lim. 2022. "Privacy-Preserving Federated Learning Using Homomorphic Encryption" *Applied Sciences* 12, no. 2: 734. https://doi.org/10.3390/app12020734
[58] Y. Li, Y. Zhou, A. Jolfaei, D. Yu, G. Xu and X. Zheng, "Privacy-Preserving Federated Learning Framework Based on Chained Secure Multiparty Computing," in IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6178-6186, 15 April15, 2021, doi: 10.1109/JIOT.2020.302291
[59] A. Khaled, K. Mishchenko, and P. Richtárik, "Tighter Theory for Local SGD on Identical and Heterogeneous Data." arXiv, Apr. 14, 2022. Accessed: Aug. 27, 2023. [Online]. Available: http://arxiv.org/abs/1909.04746
[60] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks." arXiv, Apr. 21, 2020. Accessed: Aug. 25, 2023. [Online]. Available: http://arxiv.org/abs/1812.06127
[61] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, "SCAFFOLD: Stochastic Controlled Averaging for Federated Learning." arXiv, Apr. 09, 2021. Accessed: Sep. 02, 2023. [Online]. Available: http://arxiv.org/abs/1910.06378
[62] Hamer, J., Mohri, M. and Suresh, A.T., 2020, November. Fedboost: A communication-efficient algorithm for federated learning. In *International Conference on Machine Learning* (pp. 3973-3983). PMLR.
[63] S. Reddi et al., "Adaptive Federated Optimization." arXiv, Sep. 08, 2021. Accessed: Sep. 02, 2023. [Online]. Available: http://arxiv.org/abs/2003.00295
[64] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, "Federated Learning with Matched Averaging." arXiv, Feb. 15, 2020. Accessed: Sep. 02, 2023. [Online]. Available: http://arxiv.org/abs/2002.06440
[65] Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair Resource Allocation in Federated Learning." arXiv, Feb. 14, 2020. Accessed: Aug. 25, 2023. [Online]. Available: http://arxiv.org/abs/1905.10497
[66] J. Wang and G. Joshi, "Adaptive Communication Strategies to Achieve the Best Error-Runtime Trade-off in Local-Update SGD." arXiv, Mar. 07, 2019. Accessed: Aug. 25, 2023. [Online]. Available: http://arxiv.org/abs/1810.08313
[67] D. Jhunjhunwala, P. Sharma, A. Nagarkatti, and G. Joshi, "FedVARP: Tackling the Variance Due to Partial Client Participation in Federated Learning." arXiv, Jul. 28, 2022. Accessed: Aug. 27, 2023. [Online]. Available: http://arxiv.org/abs/2207.14130

Teacher Evaluation by Students

Teacher's name: **Your name:**

The number rating stands for the following: **1** = rarely **2** = once in a while **3** = sometimes **4** = most of the. If it doesn't apply, leave it blank. Circle the answer that fits with your experience of this teacher for each:

EXPLICIT CURRICULUM: How well does the teacher teach the core subject?							**If you circ** Write why Write the
1	Teacher is prepared for class.	1	2	3	4	5	
2	Teacher knows his / her subject.	1	2	3	4	5	
3	Teacher is organized and neat.	1	2	3	4	5	
4	Teacher plans class time and assignments that help students to problem solve and think critically. Teacher provides activities that make subject matter meaningful.	1	2	3	4	5	
5	Teacher is flexible in accommodating individual student needs.	1	2	3	4	5	
6	Teacher is clear in giving directions and explaining what is expected on assignments and tests.	1	2	3	4	5	
7	Teacher allows you to be active in the classroom learning environment.	1	2	3	4	5	
8	Teacher manages the time well.	1	2	3	4	5	
9	Teacher returns homework in a timely manner.	1	2	3	4	5	
10	Teacher has clear classroom procedures so students don't waste time.	1	2	3	4	5	
11	Teacher grades fairly.	1	2	3	4	5	
12	I have learned a lot from this teacher about this subject.	1	2	3	4	5	
13	The teacher gives me good feedback on homework and projects so that the teacher is creative in developing activities and lessons. can improve.	1	2	3	4	5	
14	Teacher is creative in developing activities and lessons.	1	2	3	4	5	
15	Teacher encourages students to speak up and be active in the class.	1	2	3	4	5	
16	Teacher follows through on what he/she says. You can count on the teacher's word.	1	2	3	4	5	
17	Teacher listens and understands students ' points of view; he/she may not agree, but students feel understood.	1	2	3	4	5	

Authors' Profiles

Ariful Islam has completed his Bachelor's in Computer Science and Engineering and Master's in Information & Database Management from American International University-Bangladesh (AIUB) in 2021 and 2023, respectively. He is currently working as a Lecturer of the Faculty of Science and Technology at American International University-Bangladesh (AIUB). He is interested in research areas including Federated Learning, Machine Learning, Data Analytics, Graph Theory, and a wide variety of Algorithms and Data Structures.

Debajyoti Karmaker is currently working as an Associate Professor of the Faculty of Science and Technology at American International University-Bangladesh (AIUB). His research interests include computer vision, machine learning, and deep learning. I am particularly interested in the areas of object detection, action recognition, and biologically inspired collision avoidance strategies for quadcopters.

Abhijit Bhowmik is currently working as an Associate Professor of the Faculty of Science and Technology at American International University-Bangladesh (AIUB). His research and development areas include Natural Language Processing (NLP), Sentiment Analysis, e-Learning, Software Engineering, Software QA and Testing, System Analysis, Machine Learning, Data Mining, Networking, Algorithm and computing, Wireless Sensor Network, Video on Demand, Consultancy, Project Management, Organizational Leadership & People Management.

Md Masum Billah is currently working as an Assistant Professor of the Faculty of Science and Technology at American International University-Bangladesh (AIUB). His research interests are NLP, Machine Learning, data mining, and IoT applications in various fields.

Md Iftekharul Mobin is currently working as an Assistant Professor of the Faculty of Science and Technology at American International University-Bangladesh (AIUB). His research interests are NLP, time-series prediction, health informatics, IoT, Image processing etc.

Noorhuzaimi Mohd Noor has been in the academic, research & consultancy field since 2003. She is currently a Head of Program (Entrepreneurship) at Centre of Creative Entrepreneur Development and Senior Lecturer at The Universiti Malaysia Pahang, Malaysia. She received her B.Sc. in Computer Science from the Universiti Putra Malaysia, Malaysia, in 1999, followed by a Master degree in Science from the same university, in 2003. She received her Ph.D. degree in Computer Sciences from Universiti Kebangsaan Malaysia, Malaysia, in 2016. She is the author of more than 20 research articles. Her research interests include natural language processing, expert system, and computer security. She is also a Reviewer for the Journal of Information and Communication Technology (JICT) and editor in charge for the International Journal of Software Engineering and Computer Systems (IJSECS). Dr. Noorhuzaimi is also a certified Professional Technologist from the Malaysia Board of Technologists (MBOT) and Malaysian Qualifications Agency (MQA) Assessor Panel where she is actively involved as Assessor Panel for Technology and Technical Academic Programs Accreditation.

How to cite this paper: Ariful Islam, Debajyoti Karmaker, Abhijit Bhowmik, Md Masum Billah, Md Iftekharul Mobin, Noorhuzaimi Mohd Noor, "Unlocking Educational Excellence: Leveraging Federated Learning for Enhanced Instructor Evaluation and Student Success", International Journal of Modern Education and Computer Science(IJMECS), Vol.17, No.2, pp. 87-110, 2025. DOI:10.5815/ijmecs.2025.02.04