National College of Ireland

Self-Adaptive Federated Learning System for Financial Fraud Detection

MSc Research Project
Cloud Computing

Didheemose Pananchickal Sebastian
Student ID: x23176245

School of Computing
National College of Ireland

Supervisor: Sudarshan Deshmukh

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Didheemose Pananchickal Sebastian
Student ID:	x23176245
Programme:	Cloud Computing
Year:	2024
Module:	MSc Research Project
Supervisor:	Sudarshan Deshmukh
Submission Due Date:	12/12/2024
Project Title:	Self-Adaptive Federated Learning System for Financial Fraud Detection
Word Count:	8477
Page Count:	23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	DIDHEEMOSE
Date:	11th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Self-Adaptive Federated Learning System for Financial Fraud Detection

Didheemose Pananchickal Sebastian
x23176245

Abstract

Financial fraud is a major challenge for organisations looking forward to securing key information with strict adherence to privacy legislation. Traditional fraud detection approaches rely on a centralised mechanism that raises various privacy issues and always suffers from scalability and efficient resource utilization. Herein, this paper proposes a self-adaptive FL system for detecting financial fraud, allowing team-based training of models without actually sharing sensitive information. It adopts some intelligent methods, like model pruning and quantisation, for managing heterogeneous resources across clients, hence scalable. On the Kaggle Credit Card Fraud Detection dataset, the framework first prepares the data to fix the class imbalance and scale the features, hence strengthening the training. After the FL framework combines models with FedAvg, clustering methods are used to further improve model performance using clustering labels. Experimental results show increases in accuracy, precision, recall, and F1 score with reduced computational overhead in fraud detection. Current research points out opportunities of using FL in enhancing model performance on private datasets that are common in banks, health, and IoT devices using scalable and resource-efficient solutions.

Keywords: Financial fraud, Self-adaptive Federated Learning (FL), Privacy issues, Scalability, Resource efficient solutions

1 Introduction

The increase in financial fraud significantly challenges the effort of institutions to securitize sensitive data without violating privacy conditions (Adnan and Kumar; 2024). Traditional fraud detection systems, which are mainly centralized, compromise data security and call for new approaches in the form of federated learning, FL Ahmed and Alabi; 2024. The research topic is introduced in this chapter, along with objectives and also the promise of a self-adaptive federated learning system in order to thwart the computation inefficiencies and scaling issues. while preserving data privacy in detecting fraud in financial transactions.

1.1 Background of this Research

Financial fraud is a worldwide problem, affecting individuals, companies, and financial institutions, while the methods of fraud are getting increasingly sophisticated, from payment fraud, identity theft, to insider selling (Zheng et al.; 2022). With the development

of technology, digital transfer became a hot spot for fraud. In 2021, the total losses of payment card fraud reached $32 billion (Nilson Report; 2022). Cashless payment and mobile wallet develop new approaches for fraudsters. Phishing and smishing are increasingly used.

Fraud has now become more sophisticated as AI and ML have emerged. These attackers use adversarial AI to manipulate the dataset and the prediction model to weaken the traditional fraud detection methods (Waqas et al.; 2022). Traditional systems detect known fraud patterns but are weak while dealing with new and evolving tactics (Yang et al.; 2022). These challenges are further enhanced by traditional centralized machine learning due to privacy concerns or even scalability and adaptability issues.

These challenges highlight the urgent need for more effective fraud detection solutions. Federated learning provides a decentralized model training process whereby organizations can collaborate in model training without necessarily sharing sensitive information, thus reducing several privacy risks and enhancing scalability. Probably the most promising alternative to traditional fraud detection, FL enables large-scale applications while retaining privacy. This paper presents an approach for self-adaptive FL that could enhance fraud detection, scalability, and protection of data privacy.

1.2 Research Problem

This research will focus primarily on the areas of inefficiency and scalability constraints of FL in financial fraud detection. Federated learning develops a privacy-preserving framework where multiple participants can jointly train machine learning models without data being exposed by sensitive clients (Nguyen et al.; 2021). However, the scalability of FL is limited by other significant constraints. Clients with limited computational resources frequently encounter difficulties in processing intricate fraud detection models, leading to substantial computational requirements that impede their participation (Khurana et al.; 2020). Furthermore, FL systems devoid of model compression or adaptive methods experience significant communication overhead, resulting in sluggish training convergence and heightened resource utilisation. Although federated learning provides intrinsic advantages for privacy concerns, efficient aggregation and safe transmission of model changes are the keys to avoiding potential data leakage. In light of this, this work addresses these challenges by developing a self-adaptive federated learning framework that dynamically leverages optimisation techniques such as pruning and quantization. These adjustments allow the model to change its complexity given the client capabilities, further enhancing computational efficiency, scalability, and model performance while keeping data private.

1.3 Research Question

How to develop a self-adaptive federated learning system for optimising the financial fraud detection without compromising resource consumptions with scalability and model accuracy and privacy?

1.4 Research Objectives

- Propose a suitable FL framework where the clients train the local models on their private data independently and periodically share their updates with a centre server for global aggregation.

- Introduce adaptive compression models such as pruning and quantisation, allowing the model's complexity to adaptively scale according to the computational resources of each client.
- To enhance model aggregation and synchronisation by utilising Federated Averageing (FedAvg) in order to sum up the updates of different clients with regard to sample size and performance.
- Ensuring scalability and privacy by building a system that will be able to support clients of heterogeneous hardware capability while maintaining privacy and reducing computational overhead.

1.5 Scope of this Research

The purpose of this paper is to design a self-adaptive federated learning system for financial fraud detection, which deals with key challenges such as resource efficiency, model accuracy, and data privacy. In this project, federated learning will be implemented where clients build local models on the basis of private data and upload model updates to a central server for aggregation. It does so by investigating adaptive approaches, such as pruning and quantization, for improving model complexity to better align with customers' computational capabilities. The study will be based on publicly available datasets, one of which is the Kaggle Credit Card Fraud Detection dataset, and preprocessing techniques will be employed, including handling missing values, class imbalance issues, and feature scaling. It investigates the performance of the system with regard to accuracy, efficiency, scalability, and privacy, relying on simulated clients with heterogeneous hardware specifications running atop cloud platforms like AWS EC2. The focus on financial fraud detection might also be relevant to other privacy-sensitive domains such as healthcare and IoT.

2 Literature Survey

Financial fraud has been one of the persistent issues globally in personal, business, and financial organisational aspects. As technology advances, fraudulent schemes move ahead with sophistication, including payment fraud, identity theft, and insider selling (Zheng et al.; 2022). Each of these kinds needs a different detection and preventive approach. As observed, digital transfers have completely changed the perspective of financial fraud. In this respect, the losses due to payment card fraud have mounted to US$ 32 billion in the year 2021 (Nilson Report; 2022). With the rise of cashless payments and mobile wallets, fraudsters leverage these touchpoints to commit fraud via phishing and smishing to steal personal data.

The increased utilisation of AI and ML has made fraud become sophisticated. Fraudsters use adversarial AI to manipulate the dataset and manipulate the prediction model, making the traditional fraud detection methods redundant (Waqas et al.; 2022). Traditional systems are good at detecting known fraud patterns but weak regarding new strategies (Yang et al.; 2022). These challenges are further exacerbated by centralised machine learning methods, which raise a number of challenges in terms of privacy, scalability, and adaptability.

These challenges highlight the urgent need for improved fraud detection systems that balance efficiency, scalability, and privacy compliance. Federated learning (FL) and decentralised approaches have prominent improvements in these aspects (Awosika et al.; 2024). Thus, FL enables different institutions to jointly train machine learning models while preserving the privacy of data by avoiding its centralization. FL presents a very attractive alternative for traditional fraud detection methods since the new approach involves lower privacy risks and can be easily scalable for large-scale applications. This paper is concerned with the proposal of a self-adaptive FL approach to boost fraud detection, scalability, and data privacy protection.

2.1 Privacy and Regulatory Challenges in Fraud Detection

Fraud affects individuals, enterprises, and financial institutions in every corner of the world, and technology increases the complexity of fraud incidents. The incidents include identity theft, payment fraud, and insider trading, which all require special methods of detection and prevention according to (PK et al.; 2024). As digital payments became a big target of financial fraud, the losses in payment card fraud reached $32 billion worldwide in 2021 (Nilson Report; 2022). The growing ecosystem of cashless payments and mobile wallets opens up newer avenues for committing fraud. There are phishing and smishing scams that have become guileful in extracting personal information.

The AI/ML-based developments have turned out to be equally sophisticated fraud. Fraudsters manipulate the datasets and models through adversarial AI that undermines the conventionally used fraud detection techniques (Yang et al.; 2022). The dynamic nature of financial fraud requires advanced detection. Though rule-based models might perform well on known patterns of fraud, they mostly perform poorly in emergent tactics. Centralized machine learning further aggravates these issues related to privacy, scalability, and adaptiveness regarding the aggregation of massive volumes of data coming from a wide variety of sources.

In centralized systems, data poses a number of privacy risks since accumulation in one place promotes the possibility of hacks. Following data protection laws like GDPR and CCPA will be more complex. Tariq et al., 2023. Data breaches are not cheap at all. The 2023 IBM Cost of a Data Breach Report pegs the global average cost of a breach to have grown to $4.45 million, thanks to a 42% increase in just the past three years to find and fix the problem.

To address these issues, novel solutions must consider efficiency, scalability, and privacy in the battle against fraud (Trompeter et al.; 2013). The decentralized approaches are pretty efficient for the same. Federated Learning (FL) is a technique that enables different parties to jointly train a machine learning model by keeping data private without collecting it in one place (Li et al.; 2020). FL serves as a scalable alternative to standard fraud detection methods with greater privacy (Neelakrishnan; 2024). This work proposes a self-adaptive FL approach to enhance fraud detection, improve scalability, and ensure data privacy.

2.2 Federated Learning: An Emerging Paradigm

Federated learning (FL) changes the game in machine learning, focusing on the privacy and the decentralised nature of the technology. In the case of use in FL of federated learning, enable model development across the globe and that too without encapsulating

raw information to a single central system. This method of working on different systems on a federated basis improves the security and protection of the information while at the same time enriching the models through federated learning.

At first, FL had been specifically developed to deal with the privacy concern of mobile applications (Yang et al.; 2022), but now it is ideal in the likes of healthcare, IoT, and the best is in financial fraud detection. From the optimisation of the FedAvg, all client updates will be folded into one, and that enables seamless, quick aggregation at the central server. On the other hand, the widespread adoption of FL comes with its own set of challenges, such as frequent client contacts, various models, and fairness between clients who have different computing power (Majeed and Hwang; 2024).

One of their targets is FL, with privacy of data being central. But then the publication of model updates might cause a leak of data points. Differential privacy, which is a noise augmentation technique, and secure multiparty computation, which is an encrypted computation, are becoming more commonplace among the FL to mitigate these issues (Nguyen et al.; 2021). These methods require more computing power, and the risk is that it can lower the accuracy of the model.

FL envisions elasticity, especially in fraud detection, which the centralised paradigm finds difficult to complement due to the growing data and number of players. In fact, FL shares computing resources among clients, which is organic growth. But still there are scalability challenges owing to the client hardware limitations, which require strategies such as model compression. Techniques like model pruning and quantisation reduce the amount of operation sizes to engage low-level clients without affecting the overall system (Wen et al.; 2023).

Though it has certain merits, its adoption remains only close to 3% because it draws models that raise stubborn concerns related to strong, secure models, issues of communications, and managing multiple environments of clients. The use of a central server for FL models has raised concerns about single points of failure (Ma et al.; 2022); hence investigations into decentralised alternatives have been conducted. Based on blockchain systems, they are one of the potential ways to increase the endurance and the resiliency of the FL in applications like those with low-tolerated privacy, such as financial fraud detection.

2.3 Adaptive Techniques in Federated Learning

Adaptive mechanisms are essential for addressing the obstacles encountered in federated learning (FL) and seamless usage of the system on different client devices (Wu et al.; 2023). Methods such as model compression, pruning, and quantisation allow federated learning systems to modify the complexity of the tasks to the current client's hardware resources. These methods enhance computing power and scalability, especially in low-resource environments.

Pruning is one of the popular adaptive methods where the model is trained by removing a number of its parameters that do not have significant contributions to maintaining accuracy and consume processing power and memory resources (Vadera and Ameen; 2022). It allows low-cost devices to be trained because pruned models only involve the essential connections in the neural network. On the other hand, quantisation reduces the accuracy of the model by lowering the bit representation for floating point calculations; thus, notakeniteda suppresses its importance score (Weng; 2021). Numerous studies suggest that these techniques could enhance the performance of federated learning systems

by improving the processing speed and decreasing the amount of memory required for low-tech devices (Banabilah et al.; 2022).

Yazici (2023) puts forth Federated Averaging (FedAvg) as the key expansion in FL. Equitable effort is maintained by making sure that model updates are sent to the clients according to the volume of their datasets and the speed of the client. This in turn enhances the convergence rate and scalability of the global model (Zhou et al.; 2021). One of the most important features in client models is their ability to be flexible. It is this property that allows effective communication with other client models while requiring minimal computation.

The successful deployment of such adaptive techniques would, however, require a keen focus on data protection measures alongside encryption mechanisms. FL is likely to mitigate data centralisation, although model parameters are still in some ways regarded as privacy disclosure threats because they inevitably allow sensitive information concerning the local databases to be transmitted (Casella and Fonio; 2023). Of significance is the fact that these developments have contributed to the reduction of the risk factors involved, such as the risks associated with data collaboration. Techniques such as homomorphic encryption have made such risks almost void.

To this end, the paper presents federated self-adaptive learning for financial fraud detection by applying model pruning, quantisation, and aggregation of the different models to increase the performance of the model. The federated self-adaptive learning system encourages more clients to cooperate while maintaining better privacy performance, which is a new benchmark for privacy-orientated machine learning.

2.4 Applications of Federated Learning in Financial Fraud Detection

Federated learning is increasingly applied for financial fraud detection, whereby various scholars explore its application in different contexts. (Awosika et al.; 2024) demonstrated the effectiveness of FL in detecting payment card fraud problems, where financial institutions could jointly train models without necessarily sharing sensitive information. The authors underline that FL fuses data from different sources in its attempt to pick up unusual transaction behaviour and respects legislation such as the GDPR.

Similarly, (Gandhi et al.; 2024) focused on anti-money laundering applications of FL and showed how FL aggregates transaction data across entities to detect hidden networks and sophisticated laundering schemes while maintaining data privacy. (Nguyen et al.; 2021) positions FL in fighting against identity fraud, wherein fake identities are used to exploit banking systems. They have demonstrated how FL identifies the trend indicating synthetic identities in a decentralised dataset and enables institutions to adjust to the changing landscape of fraud tactics. The current study underlined how FL might also resist in such dynamic fraud scenarios.

(Sun et al.; 2023) furthered FL by introducing Federated Averaging (FedAvg), which aggregates model updates with consideration of dataset size and performance metrics. This further solidifies collaboration without sacrificing any scalability or privacy, hence being a staple in most FL-based fraud detection systems. All these studies put together show the disruptive role of FL in financial crime detection, ranging from payment card fraud and AML to identity fraud. FL's flexibility and its preservation of privacy meet diverse challenges effectively. The present study extends this further by advocating a self-adaptive FL framework to be used against scalability and resource limitations.

2.5 Gaps in Existing Research

Federated learning (FL) research in financial fraud detection is expanding, but this area is hindered because of several gaps that impede practical applications. One primary gap is that very little adaptive strategy research has been put into improving FL systems in the resource-constrained environment within the context of fraud detection. Studies including (Wu et al.; 2023) have investigated model compression techniques such as pruning and quantisation, but more research is needed for their dynamic, real-time application, particularly for small financial institutions or mobile clients with varied hardware.

Whereas complexity adjustment by resources is central to growth, it has received relatively little research. FL has not been proven to be able to operate successfully in vast financial ecosystems. Most of the research is either performed on small-scale simulations or in a controlled environment; indeed, often with datasets like the Kaggle Credit Card Fraud Detection dataset, which may not mirror the intricacy found in global financial systems. The financial institutions are dealing with large amounts of diversified data and are subject to problems in data imbalance, real-time processing, and legal constraints.

Further research is needed in applying FL in these diverse contexts effectively, ensuring accuracy and privacy (Khurana et al.; 2020). In particular, we need to understand the trade-offs between privacy and speed, particularly for real-time scam detection (Yang et al.; 2022). It is still unclear how to implement privacy mechanisms in FL frameworks without compromising the speed of the system. Combining FL with technologies such as blockchain or edge computing remains largely conceptual. Blockchain could provide the decentralised, immutable transaction record, which may complement FL by offering more robust scam detection over what FL provides with privacy protection. (Yang et al.; 2022).

Edge computing may also save on federated learning communications through local processing of data. However, there is still a lack of research on the synergy between FL and these technologies. Much more research effort is called for to figure out how such methods can beef up fraud detection in financial transactions. Finally, very few studies are devoted to the legal and ethical impact of FL on the solution of the financial fraud detection problem. FL helps to protect privacy and compliance with data regulations; however, which would be the optimal configuration of such systems under different legal regimes is not evident (Waqas et al.; 2022).

The setup of FL systems is tricky for any global financial institution owing to the prevailing difference in legislation acts that concern privacy among nations. Most of the legal issues, such as data sovereignty and cross-border flow of data, need to be sorted out with legal oversight before effective usage in fraud detection globally.

2.6 Summary

This review of the literature reviewed the existing status of financial fraud detection, underlining the challenges involving centralised systems and privacy regulations. Federated learning has indeed turned out to be a promising paradigm, which furnishes the ground for the much-needed decentralised framework wherein the privacy of data can be preserved along with collaboration on analysis. Challenges about computational efficiency, client heterogeneity, and communication overhead thus remain the key reasons for adaptively developing FL frameworks. This paper will fill these gaps by proposing an optimised self-adaptive FL for financial fraud detection and contribute toward the improvement of knowledge academically and in practical applications.

3 Research Methodology

This research uses an experimental and computational method for financial fraud detection by developing an optimized federated learning (FL) system. It makes use of different adaptive techniques in research, such as clustering and pruning, in making those FL models even more efficient and scalable. By decentralized learning, it can be guaranteed that data privacy is preserved, with collaborative training among multiple clients to enlarge the capability of fraud detection.

3.1 Research Procedure

The rigorous approach was used to implement and test a privacy-preserving FL for detecting financial fraud. Data preparation was performed by using the Kaggle Credit Card Fraud Detection dataset, as this dataset has been the standard for financial fraud research (Dal Pozzolo et al.; 2017). This dataset contains transaction amounts, locations, timestamps, and consumer demographics. The dataset was partitioned across four virtual clients, each with a unique data subset, to imitate realistic scattered scenarios and ensure data localisation for GDPR compliance (Voigt and Von dem Bussche; 2017). The dataset was partitioned across four virtual clients, each with a unique data subset, to imitate realistic scattered scenarios and ensure data localisation for GDPR compliance (Voigt and Von dem Bussche; 2017). FedAvg weights client models by sample size to ensure equitable contributions and data privacy.

The FL framework addressed customer resource heterogeneity with adaptive techniques. Customers were categorised by dataset properties and computational capability using KMeans clustering. The Elbow Method and KneeLocator determined k=4 clusters as optimal (Satopaa et al.; 2011). Data representation during training was improved using cluster labelling. Model pruning removed insignificant parameters from neural networks, simplifying training on resource-limited devices (Han et al.; 2011). The system was tested in two scenarios: one without clustering, where all clients contributed equally, and another with clustering, where models were aggregated before global aggregation. These scenarios tested clustering's effect on model efficacy and system efficiency. The final global model was changed and provided to customers for more training iterations, enabling collaborative learning in secrecy.

Quantification of performance was done by model accuracy, precision, recall, and F1 score (Powers; 2020). The confusion matrices of model prediction presented the trade-off that the model is making between false positives and false negatives. For the optimisation of training to support various scenarios, CPU time, memory consumption, and training time were considered for categorising clients into high, medium, or low capacity classes. Improvement in model performance and scalability with clustering and pruning justifies quite well its application under federated learning.

This research followed a strict scientific foundation and updated previous methodologies to address federated learning and financial fraud detection. Clustering, pruning, and privacy-preserving aggregation help build scalable and efficient federated learning systems for real-world applications.

3.2 Data Collection and Preprocessing

Data collection and preprocessing were crucial to this study since they prepared the dataset for model training and analysis. This study utilises the free Kaggle Credit Card Fraud Detection dataset (Dal Pozzolo et al.; 2017). This dataset is used for financial fraud research. Only 492 of 284,807 transactional records are fraudulent. The dataset includes transaction amounts, locations, timings, and consumer personal information, making it useful for fraud detection testing.

3.2.1 Data Collection

The dataset was shared across four simulated clients to approximate federated learning. Individual clients received data representing real-life activity in their region. This section prevented clients from sharing raw data in accordance with the GDPR (Voigt and Von dem Bussche; 2017). The FL framework emphasises local model training over central data gathering, making privacy-protecting strategies easier to deploy in the simulated world.

3.2.2 Data Preprocessing Steps

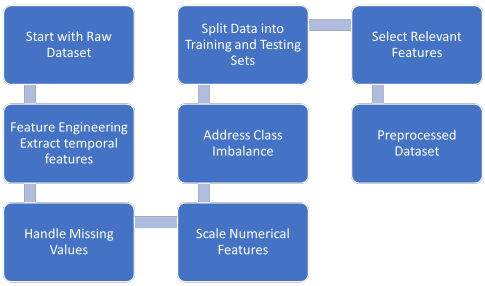

Figure 1: Data Preprocessing Steps

This study used structured data preparation methods to prepare the raw dataset for machine learning model training. First, the dataset was imported, examined, and important attributes altered to improve prediction. Time information, such as transaction day, month, year, and hour, was obtained from trans_date_trans_time. This helped academics comprehend fraud evolution. The dob column also determined clients' buying ages, adding demographics. To convert text-based data into machine learning numbers, merchant, category, gender, city, and job labels were added. To fill missing values, medians and modes of numeric and categorical categories were employed. This protected critical data.

With only a few fraudulent trades in the sample, class imbalance was high. This was fixed with SMOTE. It developed minority sample fakes to level out the dataset. Then, StandardScaler was utilised to regularise numerical characteristics and prevent broader ranges from dominating learning. To make sure the review was strong, the processed data was split into two groups: 80% for training and 20% for testing. Number elements, including transaction amounts, geographic positions, and encoded category data, were carefully selected. This systematic preparation cleaned, balanced, and standardised the dataset. This laid the groundwork for a reliable fraud detection methodology.

3.3 Federated Learning Framework

The federated learning framework was implemented to facilitate privacy-preserving collaborative training. Each client locally trains a model on its private dataset and shares only the model weights with a central server for aggregation.

Figure 2: Federated Learning Framework

This paper introduces the FL framework, which provides model sharing for collaborative training among distributed clients while guaranteeing data privacy and handling resource heterogeneity. In this paper, the dataset was pre-processed and divided into four simulated clients, each holding a portion of transactional data. FL enabled these customers to train local machine learning models on their data independently, without the need for the clients to share raw information with the central server. Instead, the clients of the system would transmit their models' parameters (weights) to a common server that aggregates the transmissions to produce a model optimised globally. It pursued the decentralisation of data protection, such as GDPR, by not allowing sensitive data to leave the local premises.

The core of the FL framework was the Federated Averaging (FedAvg) algorithm proposed (McMahan et al.; 2024). After local training, each client sent its model weights and sample size to the server. The server aggregated these weights proportionally, based on the sample sizes, to generate an updated global model. This model was then redistributed to the clients for more local training in a series of iterative rounds until convergence. Techniques such as clustering and model pruning have been applied within this framework in order to make the process more efficient, scalable, and adaptable. First of all, the clients will be clustered by KMeans according to the computation ability and data characteristics, targeting each group for optimisation accordingly. Model pruning was further considered to reduce neural network complexity so that resource-constrained clients can contribute effectively without model performance degradation. Overall, this adaptive FL framework is privacy-preserving, scalable, efficient, and thus suitable for real-world applications, including financial fraud detection.

3.4 Model Evaluation Metrics

A number of model evaluation metrics in this study were used to check how well the proposed federated learning system could identify financial fraud. Accuracy was then

used as the method of deducing the general correctness of the overall model by counting correctly predicted fraudulent and nonfraudulent transactions out of all the predictions. Precision, recall, and F1 score were also used due to the nature of the dataset not being balanced across classes and its few fraudulent deals. Precision was said to be that ability of a model to pick out fraudulent activities from all the predicted fraud cases, having as few false positives as possible. Remember, though, that we tested the model's ability to spot all real fraudulent deals, which cut down on false negatives. The F1 score, which is a harmonic mean of accuracy and recall, gave a fair evaluation of performance, which is very important when datasets aren't balanced. A confusion matrix was also used to see and study the spread of true positives, true negatives, false positives, and false negatives. This gave more information about the model's mistakes in making predictions.

In addition to performance measures, resource utilisation was another important part of the federated learning system when the resources at the clients are limited. Some metrics can be used to find out the efficiency of the usage of computation power of the model, such as CPU time, memory usage, or training time. These measures show how strong the shared platform is to hold up to local training and aggregations over different client hardware. In addition, it factored in the system's scalability regarding the number of clients the system could handle, along with additional overhead for sending model weights across different clients and the server. These review criteria ensured that the federated learning system is good to go, not only on detecting fraud but also quick, scalable, and deployable into real life, where the concerns of privacy and computing power are more imminent.

3.5 Benchmarking and Client Classification

This research used benchmarking to assess client hardware and further tune the federated learning system for a wide range of computing environments. Benchmarking measured CPU, memory, and training time. CPU time assessed how long each client took to train its local model, revealing computational efficiency. Memory utilisation assessed system memory consumption during training to ensure it could accommodate customers with different hardware resources. Training time across numerous federated learning rounds determined the system's convergent efficiency. Benchmarking also measured communication overhead, such as client-server model weight transmission time. These measurements helped detect bottlenecks and ensure federated learning framework scalability and efficiency.

The system can adapt to the different clients' capabilities through classification of these clients into high, medium, or low capacity based on benchmarking results. Large, complex models and larger datasets can be dealt with by high-capacity clients that have a minimum CPU time and enough memory. The medium-capacity clients had medium-level computing and memory capabilities and, hence, could have participated but required standard model complexity. Low-capacity clients had high CPU time with low memory and hence were resource-restricted. Generally, adaptive methods reduce neural networks to derive lowered computational costs without reductions in the model performance of the overall global model. By using this categorisation method, a fair and effective federated learning approach allowed all clients to take part in contributing to the optimisation for scalability of the system and resources.

3.6 Implementation

The federated learning architecture used numerous interrelated components to create a strong and scalable financial fraud detection system. Preprocessing and feature engineering transformed the raw dataset. Transaction day, month, and hour were calculated, and category data were numerically encoded. SMOTE balanced unbalanced data for model training. The dataset was clean and organised after these changes for federated learning. The Python-based federated learning system used TensorFlow and Keras for neural network modelling, Scikit-learn for clustering and SMOTE, and Matplotlib and Seaborn for visualisations. Local models were trained on each client's dataset and aggregated using FedAvg. KMeans clustering grouped customers based on their computing resources, while model pruning optimised resource utilisation for low-capacity clients. Implementation provided the aggregated global model, modified the dataset after pre-processing, benchmarked client classification, and calculated accuracy, precision, recall, and F1 score. These results were sufficient evidence from the paper to prove this framework regarding fraud detection, privacy, scalability, and inclusiveness across the computing settings.

3.7 Ethical Considerations

This research used sensitive information of the financial data of the customers, and ethical consideration was at a high level to make sure protection of privacy, security, and legal matters. The original dataset was acquired from a public platform called Kaggle Credit Card Fraud Detection. It had been assembled with care to protect sensitivity. Anonymised, with no PII in it, though ethical considerations went even deeper. To avoid data mistreatment, feature extraction, encoding, and scaling were done in a secure environment. This research only used the dataset for scholarly reasons, preventing unethical or unauthorised usage.

Maintaining client data security with a federated learning system improved ethical safety. This decentralised solution kept raw transactional data local to each client and never sent to the central server, preventing data breaches. We supplied just model parameters (weights) to comply with data privacy laws like the GDPR and CCPA. Ethically constructed adaptive approaches like clustering and model pruning serve clients with different computational capacities, assuring inclusiveness and justice. By balancing fraud and non-fraud instances with SMOTE, the research avoided dataset biases and misleading model predictions. These protections and an open, privacy-conscious strategy assured ethical and secure dataset management throughout the study procedure.

3.8 Limitations

This work has presented a design and evaluation of a financial fraud detection federated learning framework that has several limitations, suggesting some possible ways for improvement. First and foremost, the investigation in this study was based on the Kaggle Credit Card Fraud Detection dataset, which, because of its open access, can never really be representative of real-world financial transaction data. Most practical scenarios require heterogeneous data from a large number of sources, say financial institutions or geographies, which was beyond the scope of this study. Moreover, the data itself was pre-anonymised, further constraining the study of fraud-detection variables such as user behavioural patterns or geographical insights.

Another weak point of this work is its resource heterogeneity breadth. While the adaptive techniques of model pruning and clustering were used in the experiments to fit clients of different computational capacity, the framework did not really investigate the dynamic nature of the clients, including such very common ones like frequent dropouts or unstable network conditions. This work investigates the performance of the proposed framework using only descriptive measures: accuracy, precision, recall, and resource utilisation. This work does not consider real-time system performance or model variations in a live financial environment. As a matter of fact, these form a basis for future studies so as to make this framework robust at large-scale deployment.

4 Design Specification

The design specification describes the techniques, architecture, framework, and its derived requirements that form the basis for implementing the federated learning system. The system's core architecture applies a concept called Federated Learning, where multiple clients can train their local model independently, whereas the model updates are aggregated on the central server. It does aggregation across all clients by using the FedAvg algorithm. This allows a decentralised setup wherein sensitive data remains local to the clients, hence addressing the privacy concern. The self-adaptive technique helps to dynamically adjust the server model based on the computing resource capabilities of each client. By implementing these features, any devices with low computing resources can contribute to model building.

4.1 Federated Learning System Architecture

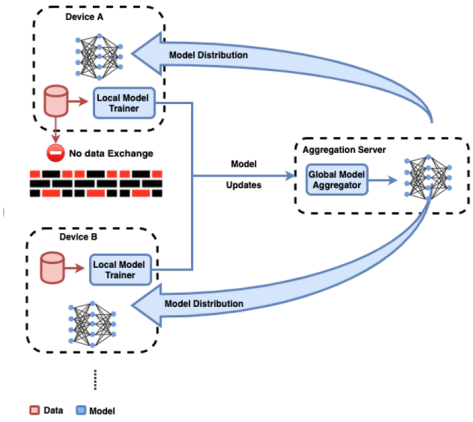

Figure 3: Architecture Diagram

- Server: The server aggregates the weights received from all clients and updates the global model. This updated model is then shared with the clients for further training.
- Clients: Each client trains a local model on its private dataset and then communicates with the server to share model weights.

4.2 Techniques

- FedAvg is a popular federated learning technique for aggregating weights from multiple client models to form a global model. In this way, model updates from each client are weighted appropriately by the amount of data they processed.
- Pruning is a technique used to optimise the model by removing less significant weights from the model to reduce its size and computational cost without effecting the model's accuracy.
- Quantisation is a technique used to reduce the precision of the model weights by converting from floating point to integer format, which will help to reduce the memory usage and speed up computation, which is beneficial for low-resource devices.
- Adaptive model technique is able to dynamically adjust to the client computing resource availability, and the model is categorised into full pruned or quantised, ensuring that clients with fewer resources are not overloaded with large models.
- Benchmarking To evaluate the system performance and ensure that the client meets the expected requirements by gathering CPU time, memory usage, and latency.

4.3 Frameworks

- TensorFlow is an open-source library for machine learning, which was developed by Google. TensorFlow builds and trains deep learning models, because it provides an easy way to create computational graphs and also executes it efficiently on various hardware platforms.
- TensorFlow Model Optimisation Toolkit TF-MOT is used for optimising machine learning models so that they can be used for deployement and execution. It supports various techniques that are used to reduce latency and inference cost for cloud and edge devices like mobiles and IoT.
- Flask is a web framework used to create server-side APIs that enable communication between the server and client in the federated learning setup.

4.4 System Requirements

- Hardware: The system utilizes AWS EC2 instances, wherein the clients and server are hosted in separate instances. In this regard, each EC2 instance is allocated with at least 8 GB RAM and 2 vCPUs to make sure that the training processes are executed smoothly in the model.
- Networking: It provides HTTP requests to exchange a model and model weights between the server and clients after every round of training.
- Scalability: The system is designed to scale by adding more clients into the federated learning setup, which may lead to improved generalization and model performances.

5 Implementation/Solution Development Specification

The goal of this implementation is to develop a core function of a self-adaptive federated learning system to detect fraud in financial transactions. The final part of the implementation covers the server-side infrastructure that handles model updates from clients, performs aggregation of model weights, optimises through pruning and quantisation, and sends the updated models back to clients.

5.1 Transformed Data

In the final stage of implementation, the dataset of financial fraud detection gathered from Kaggle was preprocessed to make it suitable for federated learning. The preprocessing steps included:

- Convert dates to datetime format: The columns 'trans_date', 'trans_time', and 'dob' are converted to datetime objects to facilitate feature extraction and calculations in the financial fraud detection dataset, and also extracted features like day, month, year, and hour from the transaction date.
- Calculate Age: The age will be determined by using the difference between the transaction year and the year of birth.
- Dropping of columns that aren't necessary: The dataset columns like trans date, trans time, dob, first, last, street, cc num, and trans num are dropped in order to keep the dataset simple.
- Handling of missing values: The mean of the respective columns was used to impute missing data points.
- Encode categorical features: Label encoding is applied to columns like merchant, category, gender, city, state, and job to convert categories into numerical values.
- Balancing the data: Imbalance in the fraud labels was addressed using oversampling techniques to ensure a fair representation of fraud and non-fraud cases.
- Feature scaling: All features were scaled to a range between 0 and 1 to ensure that the model training would not be biased towards variables with larger numerical values.

5.2 Code Written

TFF is used for the implementation of the Federated Learning model. The server uses the FedAvg algorithm, which performs an average of model weights for all the clients after every round of training. The server further performs tasks like delivering models to the clients and various model optimisation techniques, including pruning and quantization. Another important role of the server is handling the communication and maintaining the models synchronised across clients.

- To simulate clients, AWS EC2 instances were used, training individual models on their own local data. At the end of each round of local training, each client sends its updated weights back to the server.
- The server code was responsible for aggregating the received weights and updating the global model.
- All communication between the server and clients was done through HTTP, using Python's HTTP server functionality.

5.3 Models Developed

The final output of the project was a self-adaptive federated learning system for fraud detection model. This model can be able to dynamically adjust to the client resource aviablity by applying Pruning and Quantization techniques also this model was trained across multiple federated learning rounds, where each client contributed to updating the model. The final model was evaluated on accuracy, percision score, recall score, and f1 score, with performance improvements observed after each federated learning round.

5.4 Tools and Languages

- Python is a programming language used for implementing systems and frameworks, and it also supports a wide range of libraries like TensorFlow, Flask, and NumPy.
- NumPy is a library, and it acts as a fundamental tool used for efficient numerical computational operations such as aggregation of model weights.
- AWS EC2: The cloud infrastructure used for hosting server and client instances. The EC2 instances simulate a federated learning environment by performing client-side computations and communicating with the server.
- Scikit-learn is used for model evaluation, such as calculating model performance metrics like accuracy and confusion matrices.
- SMOTE (Synthetic Minority Over-sampling Technique) is an efficient technique that can be used to handle class imbalance in datasets. It relies on generating synthetic samples for the minority class to balance the dataset.
- KneeLocator library makes it easy to find the elbow point from graphs, such as when determining the optimal number of clusters for k-means clustering. The term elbow point refers to the point at which the rate of change in a curve significantly decreases.
- Jupyter Notebook Used for writing and testing the federated learning code in an AWS EC2 environment.

6 Evaluation

This section evaluates the results of the federated learning framework by directly addressing the research objectives. Each objective is analysed in the context of the findings from the experiments, highlighting the extent to which it was achieved.

6.1 Experiment 1: Developing a Federated Learning Framework

The federated learning system attained competitive performance among all clients while safeguarding data privacy. In the absence of clustering, the mean accuracy among customers was 90.235%, with precision at 95.3225% and recall at 84.775%, yielding an F1 score of 89.7175%. The implementation of clustering resulted in a marginal enhancement of the framework's average accuracy to 90.3925%, while precision sustained a robust level at 94.265%. Recall experienced a rise to 86.21%, and the F1 score attained 90.0125%.

Table 1: Model with clustering

Client	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score**
Client 1	90.25	93.79	86.39	89.94
CLient 2	90.56	95.18	85.58	90.12
Client 3	90.63	95.48	85.40	90.16
Client 4	89.50	96.84	81.73	88.65
Average	90.23	95.32	84.77	89.72

Table 2: Model without clustering

Client	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score**
Client 1	89.97	94.90	84.69	89.50
CLient 2	90.80	93.75	87.55	90.54
Client 3	89.89	96.19	83.18	89.21
Client 4	80.91	92.22	89.42	90.80
Average	90.39	94.26	86.21	90.01

- Accuracy: The average accuracy across clients improved slightly with clustering (from 90.235% to 90.3925%). This demonstrates the effectiveness of clustering in balancing client contributions.
- Precision: Precision remained consistently high across all clients, indicating the model's strong ability to correctly identify true positives.
- Recall: Recall improved from 84.775% to 86.21% with clustering, reflecting better detection of fraudulent transactions.
- F1 Score: The F1 score increased from 89.7175% to 90.0125%, showcasing a balanced improvement in both precision and recall with clustering.

The results validate the federated learning framework's ability to achieve high performance while preserving privacy. The slight improvements observed with clustering highlight its utility in enhancing client participation and system efficiency. These findings align with the research objective of developing a privacy-preserving framework capable of robust fraud detection.

6.2 Experiment 2: Implementing Adaptive Model Compression Techniques

This experiment focused on the implementation of pruning as an adaptive model compression technique in the federated learning framework. The objective was to assess the effect of pruning on the model's performance, particularly focusing on metrics such as accuracy, precision, recall, and F1 score. The results highlight the effect of pruning on resource efficiency while ensuring the model's performance remains intact or improves.

Performance Metrics of pruned model with and without clustering

Table 3: Performance Metrics of pruned model with and without clustering

Metric	**With Clustering**	**Without Clustering**
Accuracy (%)	90.235	90.39
Precision (%)	95.32	94.26
Recall (%)	84.78	86.21
F1 Score (%)	89.72	90.01

6.3 Discussion

The results from Experiment 1 emphasise the effectiveness of the federated learning system implemented in ways that ensure more effectiveness while upholding privacy. The results of the application of the clustering presented improvements that were of moderate significance. It was particularly in recall and F1 score. Mean accuracy moved relatively from 90.235% to 90.3925% with clustering, which means that the clustering collaborated to boost the model's performance in regards to the handling of data from different clients. It is important to note that recall went from 84.775% to 86.21% which meant that the framework was becoming better at detecting the fraudulent transactions. With this improvement, the importance of the fraud detection systems maximisation of detection of fraud cases that are within set thresholds comes into play. The modest increase in F1 score, from 89.7175% to 90.0125% on the other hand, is additional evidence that the clustering enhances the sensitivity and specificity of fraud detection while also ensuring that the overall detection efficiency is improved without adversely affecting any of the metrics. These results are good and support the purpose of anti-fraud systems because it is clear that there is evidential value of clustering in the federated learning framework.

The second experiment, in which pruning was considered as an adaptive model compression technique, was able to show the benefits of resource utilisation in the federated learning system. The results showed that the performance of the system was able to maintain weight performance; it was noted that pruning less important parameters was able to enable models to have reduced memory and computational requirements. This proved to be advantageous, especially for clients with lean resources. Post-pruning, the accuracy of the model reduced slightly, with the percentage dropping from 90.3925% to 90.235%, This differentiation or range indicates clearly that pruning had very little impact on the model's ability to detect fraud in threatening models. The recall to precision measurement saw a linear reduction in precision moving from 95.3225% to 94.265%. The increase was rather encouraging, climbing from 84.775% to 86.21%,, this increase narrates a different story from the precision measurement. More specifically, it tells us that the model is still good at detecting a fraudulent transaction but was not as effective as it was

before. The F1 of the score was also affected positively with an increase from 89.7175% to 90.0125%. This showed that the joint attack by pruning and clustering techniques increased the model's ability in terms of accuracy and recall on fraud detection tasks.

To summarise, both experiments underscore the potential of enhancing efficiency in federated learning frameworks while optimising available resources and together addressing the protection of personal information. Clustering was useful in increasing model performance, especially recall and F1 score, while pruning enhanced efficient usage of resources without compromising on accuracy or recall. These results confirm the potential of these strategies proposed in this paper to address the challenges posed by scaling in practice federated learning models, especially in distributed and resource-scarce settings.

7 Conclusions and Future Work

The main objective of this project was to investigate the use of a federated learning framework for financial fraud detection that guarantees both superior performance and data privacy. The study aimed to address the question: How can a federated learning architecture be designed to enhance fraud detection while reducing resource utilisation and assuring scalability and privacy? A federated learning system was developed, enabling clients to train local models on their own data and regularly transmit changes to a central server for aggregation. The aims of this investigation encompassed:

- Establishing a federated learning architecture that guarantees data privacy while attaining strong performance.
- Utilising adaptive model compression methods, including pruning, to enhance computing efficiency and model efficacy.
- Improving model aggregation using Federated Averaging (FedAvg) to tackle client heterogeneity and model convergence issues.
- Facilitating scalability and privacy by developing a system that accommodates diverse client hardware capabilities while safeguarding data privacy and minimising computational overhead.

The experimental findings indicate the efficacy of the suggested federated learning system. Experiment 1 validated that the framework effectively achieved superior performance for accuracy, precision, recall, and F1 score while maintaining data privacy. The marginal enhancements noted with the implementation of clustering (from 90.235% to 90.3925% accuracy) indicate that clustering may further refine the framework, especially for fraud detection jobs involving customers with diverse data. Experiment 2 demonstrated the advantages of employing pruning as an adaptive model compression method, which decreased the model size and computing demands without substantially impacting performance, as seen by the stability or marginal enhancement of the F1 score. Federated Averaging (FedAvg), employed for model aggregation, facilitated the equilibrium of client contributions and guaranteed the system's efficiency and scalability.

The principal conclusions derived from this research are as follows:

- The federated learning approach effectively safeguards privacy while delivering robust fraud detection performance.

- Clustering boosts performance by augmenting recollection and F1 score, indicating superior fraud detection.
- Pruning lowers model size and computing demands while maintaining essential performance metrics.
- FedAvg improves model aggregation by facilitating fair contributions from clients with differing data volumes and hardware capacities.

7.1 Consequences of the Research

This research has important implications for both academia and industry. This research advances the application of federated learning for privacy-preserving fraud detection, proving the practical aspect of its usefulness in the real world. It further expands the knowledge of pruning-based adaptive model compression techniques and their use in federated learning. This study proves that federated learning can be successfully deployed in the preventive measures against financial fraud where privacy and limited resources are always critical issues. The improvements in recall and F1 score indicate increasing prospects of more effective detection of fraud in real-time settings without compromising the privacy of the data.

This system has the potential for use in many other sectors, such as banking, insurance, and e-commerce, where privacy and fraud detection management are essential. Financial institutions could use this method to design efficient fraud detection systems that preserve critical data from clients whilst allowing the use of the models to improve through shared data use. The techniques investigated in this study could also be relevant in health care or IoT (Internet of Things) systems in which the issues of maintaining confidentiality in data sharing are very vital.

7.2 Limitations

Notwithstanding the encouraging outcomes, some limitations were identified during the investigation. The scalability of the federated learning system has not been thoroughly evaluated with a substantial number of clients. The studies undertaken utilised a limited subset of customers, which may not adequately represent the issues encountered in large-scale deployments. Moreover, although pruning demonstrated the capacity to enhance the model, its enduring effect on the model's adaptation to developing fraudulent strategies was not investigated. The research predominantly concentrated on the technological dimensions of federated learning, neglecting a thorough examination of user experience and operational problems faced by stakeholders, including financial institutions and end-users.

7.3 Future Work

Expanding the Framework: By taking part in this research one would logically want to test the federated learning framework on a wider scale with far more clients in order to see how it performed in terms of breaking scalability barriers especially in a real-world scenario. This involves checking the delay of federated updates, the efficiency of communication, and control of the non-IID problem among clients.

Long-Term Adaptation and Model Robustness: As a prototypical model of future attempts, a priori, it assumes a change of the framework in accordance to the change of the fraud pattern over a period of time; research in the future may focus on determining how this change has occurred. The technique described above involves using a fixed model; however, this may not be so ideal, as a common feature of such fake strategies is that they often change, and hence the model must be routinely replaced. The online or continuous learning embedded within the federation may be an area of interest in order to retain the model's precision and relevance to the incoming types of fraud.

Client Heterogeneity: More interestingly, clients with heterogeneity in a sense that they have different levels of computation power and how to manage such clients is further an interesting area to look into. This possibly involves having to improve model synchronisation processes in order to reduce the time required for training to converge on clients with limited resources without compromising fairness in the global model aggregate.

Upgrades in Privacy and Safety: Initially, this study paid attention to data protection, but future studies can focus on privacy enhancement techniques like secure aggregation, differential privacy, suitable encryption schemes to strengthen the federated learning process and reduce the chances of model inversion or violation of secrecy. The future work could include a more broad-level investigation of performance metrics of federated learning in real-life applications such as a fraud detection system. This could include metrics such as the cost-benefit factors, user feedback, and performance of fraudulently impacted systems in real time so as to improve the system for commercialisation purposes.

This research shows that federated learning, combined with clustering, pruning, and federated averaging, is a great combination for fraud detection in data-sensitive environments. Future work may improve these techniques with issues of scaling, novel fraudulent techniques, and even variability within customer populations to further increase the efficacy of the system and extend its reach.

References

Adnan, S. A. and Kumar, P. (2024). Financial crimes and fintech in india, *E-banking, Fintech, & Financial Crimes: The Current Economic and Regulatory Landscape*, Springer, pp. 97–109.

Ahmed, A. A. and Alabi, O. O. (2024). Secure and scalable blockchain-based federated learning for cryptocurrency fraud detection: A systematic review, *IEEE Access* **12**: 102219–102241.

Awosika, T., Shukla, R. M. and Pranggono, B. (2024). Transparency and privacy: The role of explainable ai and federated learning in financial fraud detection, *IEEE Access* **12**: 64551–64560.

Banabilah, S., Aloqaily, M., Alsayed, E., Malik, N. and Jararweh, Y. (2022). Federated learning review: Fundamentals, enabling technologies, and future applications, *Information processing & management* **59**(6): 103061.

Casella, B. and Fonio, S. (2023). Architecture-based fedavg for vertical federated learning, *Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing*, pp. 1–6.

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C. and Bontempi, G. (2017). Credit card fraud detection: a realistic modeling and a novel learning strategy, *IEEE transactions on neural networks and learning systems* **29**(8): 3784–3797.

Gandhi, H., Tandon, K., Gite, S., Pradhan, B. and Alamri, A. (2024). Navigating the complexity of money laundering: Anti–money laundering advancements with ai/ml insights, *International Journal on Smart Sensing and Intelligent Systems* **17**(1).

Han, H., Gu, B., Wang, T. and Li, Z. (2011). Important sensors for chiller fault detection and diagnosis (fdd) from the perspective of feature selection and machine learning, *International journal of refrigeration* **34**(2): 586–599.

Khurana, R. et al. (2020). Fraud detection in ecommerce payment systems: The role of predictive ai in real-time transaction security and risk management, *International Journal of Applied Machine Learning and Computational Intelligence* **10**(6): 1–32.

Li, L., Fan, Y., Tse, M. and Lin, K.-Y. (2020). A review of applications in federated learning, *Computers & Industrial Engineering* **149**: 106854.

Ma, C., Li, J., Shi, L., Ding, M., Wang, T., Han, Z. and Poor, H. V. (2022). When federated learning meets blockchain: A new distributed learning paradigm, *IEEE Computational Intelligence Magazine* **17**(3): 26–33.

Majeed, A. and Hwang, S. O. (2024). A multifaceted survey on federated learning: Fundamentals, paradigm shifts, practical issues, recent developments, partnerships, trade-offs, trustworthiness, and ways forward, *IEEE Access* **12**: 84643–84679.

McMahan, H., Xu, Z. and Zhang, Y. (2024). A hassle-free algorithm for strong differential privacy in federated learning systems, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pp. 842–865.

Neelakrishnan, P. (2024). Data security fundamental requirements, *Autonomous Data Security: Creating a Proactive Enterprise Protection Plan*, Springer, pp. 1–39.

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J. and Vincent Poor, H. (2021). Federated learning for internet of things: A comprehensive survey, *IEEE Communications Surveys & Tutorials* **23**(3): 1622–1658.

Nilson Report (2022). Payment card fraud losses reach $32.34 billion, *GlobeNewswire News Room* . Available at: https://www.globenewswire.com/news-release/2022/12/22/2578877/0/en/Payment-Card-Fraud-Losses-Reach-32-34-Billion.html.

PK, R., Khaparde, A., Bendre, V. and Katti, J. (2024). Fraud detection and prevention by face recognition with and without mask for banking application, *Multimedia Tools and Applications* pp. 1–24.

Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation, *arXiv preprint arXiv:2010.16061* .

Satopaa, V., Albrecht, J., Irwin, D. and Raghavan, B. (2011). Finding a" kneedle" in a haystack: Detecting knee points in system behavior, *2011 31st international conference on distributed computing systems workshops*, IEEE, pp. 166–171.

Sun, T., Li, D. and Wang, B. (2023). Decentralized federated averaging, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**(4): 4289–4301.

Trompeter, G. M., Carpenter, T. D., Desai, N., Jones, K. L. and Riley, R. A. (2013). A synthesis of fraud-related research, *Auditing: A Journal of Practice & Theory* **32**(Supplement 1): 287–321.

Vadera, S. and Ameen, S. (2022). Methods for pruning deep neural networks, *IEEE Access* **10**: 63280–63300.

Voigt, P. and Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr), *A Practical Guide, 1st Ed., Cham: Springer International Publishing* **10**(3152676): 10–5555.

Waqas, M., Tu, S., Halim, Z., Rehman, S. U., Abbas, G. and Abbas, Z. H. (2022). The role of artificial intelligence and machine learning in wireless networks security: Principle, practice and challenges, *Artificial Intelligence Review* **55**(7): 5215–5261.

Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J. and Zhang, W. (2023). A survey on federated learning: challenges and applications, *International Journal of Machine Learning and Cybernetics* **14**(2): 513–535.

Weng, O. (2021). Neural network quantization for efficient inference: A survey, *arXiv preprint arXiv:2112.06126* .

Wu, X., Huang, F., Hu, Z. and Huang, H. (2023). Faster adaptive federated learning, *Proceedings of the AAAI conference on artificial intelligence*, Vol. 37, pp. 10379–10387.

Yang, Z., Chen, M., Wong, K.-K., Poor, H. V. and Cui, S. (2022). Federated learning for 6g: Applications, challenges, and opportunities, *Engineering* **8**: 33–41.

Zheng, Y., Lai, S., Liu, Y., Yuan, X., Yi, X. and Wang, C. (2022). Aggregation service for federated learning: An efficient, secure, and more resilient realization, *IEEE Transactions on Dependable and Secure Computing* **20**(2): 988–1001.

Zhou, Y., Ye, Q. and Lv, J. (2021). Communication-efficient federated learning with compensated overlap-fedavg, *IEEE Transactions on Parallel and Distributed Systems* **33**(1): 192–205.