International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Exploring Federated Learning for Fraud Detection in Finance: Evaluating Feasibility and Data Sharing Models for Privacy-Preserving Solutions

Sreenivasulu Gajula
Principal Full-Stack Engineer

Abstract

The study investigates the use of Federated Learning (FL) in fraud detection among financial institutions with respect to privacy safeguarding and performance of the models. Compared to FL and conventional centralized models, the study shows that FL has a higher recall and F1-score, hence it is more effective at detecting fraudulent transactions. Differential Privacy minimizes the risk of breach of privacy, which facilitates cooperation between institutions. It also addresses the obstacles to data heterogeneity and communication overhead in FL systems studied in the study. The results indicate that FL provides a good remedy for decentralized fraud detection and responds to privacy and scale issues in financial services.

Keywords: Federated Learning, Fraud Detection, Privacy Preservation, Differential Privacy, Decentralized Model, Financial Institutions, Centralized Model, Secure Data Collaboration, Machine Learning, Scalable Solutions.

I. INTRODUCTION

Fraud detection in the dynamic environment of financial services is becoming more demanding, considering the intricate nature of the fraudulent operations and the growing number of financial transactions. In most cases, the conventional centralized approaches to detecting fraud are not always effective in regard to privacy and regulation, particularly when a financial institution deals with sensitive information. Federated Learning (FL), a recent machine learning paradigm, is a potential solution to these issues since it allows model training on decentralized data sources without the necessity to exchange sensitive information [1].

The current study will evaluate whether Federated Learning can be implemented on the privacy-aware fraud detection framework within the financial industry. As privacy rules such as GDPR and CCPA establish, financial institutions need to make sure that their sensitive data of customers remains secure yet can be identified with ease [2]. FL allows shared data workload among several actors and preserves the privacy of the data with the help of local data processing and aggregation of models. There are some challenges in the use of Federated Learning in this area, such as technological barriers, misgivings between collaborating parties, and the harmonization of incentives. The study discusses these challenges and considers other sections of data collaboration models that have the potential to maximize utilization of FL in fraud detection without compromising privacy and regulatory adherence.

Problem Statement:

Although Federated Learning has potential, its use in the financial sector as an attempt to detect fraud has not been studied extensively. The main obstacles that should be considered are to guarantee the stability of models without sacrificing the privacy of data, building trustful partnerships, and challenging the regulations and technological barriers [3]. The proposed research will address these gaps by analysing the viability of FL in financial fraud detection and presenting relevant and efficient cooperation models that address privacy, security, and performance-related concerns.

Aims and Objectives

Aim: The aim of this research is to evaluate the feasibility of implementing Federated Learning for privacy-preserving fraud detection in the financial sector.

Objectives:

- To explore data collaboration models that enable secure, decentralized training of fraud detection models.
- To assess the impact of Federated Learning on regulatory compliance and privacy concerns in financial institutions.
- To compare the performance of Federated Learning-based fraud detection models with traditional centralized models in terms of accuracy and efficiency.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

II. LITERATURE REVIEW

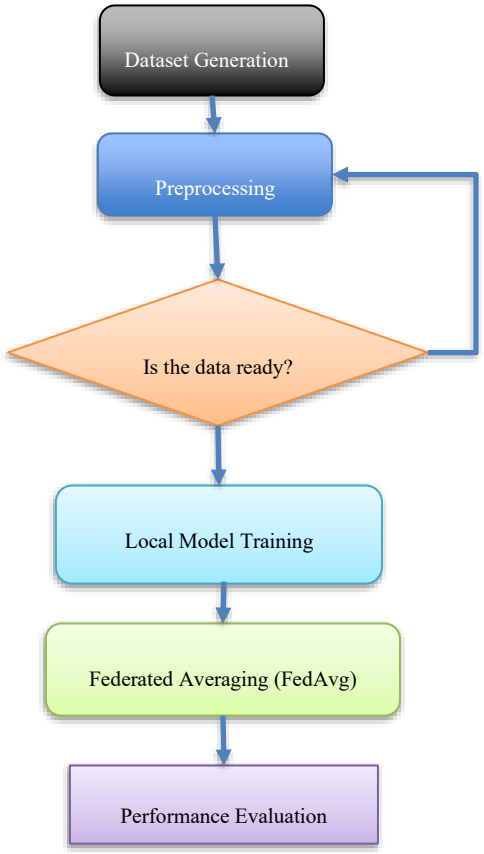

Fig 1: Flow of the research

A. The Goal of the Review:

The review's aim is to discuss the viability of Federated Learning in financial fraud detection systems and to overcome the privacy issue. Its studies data collaboration models, regulatory issues, and compares the performance of the Federated Learning and old-fashioned methods and provides some insights on the privacy-saving and effective fraud detection approaches using various journal resources

B. Study of Previous Literature

1. Federated learning in financial services

Federated Learning (FL) is a decentralized machine learning method where two or more institutions, such as banks or insurance companies, can compete to jointly train a provisional model without exchanging their confidential information [4]. In conventional machine learning methods, core models are trained on aggregated information of all subjects. This, also prompts questions of data privacy and security and in particular the financial sector, where regulatory authorities like GDPR and CCPA impose high data protection requirements. FL solves this problem because it makes sure that raw data are localized to institutions. Rather, merely updates of the model are exchanged, which considerably reduces the chances of information leakage.

Fig 2: Flow Chart of Federated learning in financial services

This flowchart illustrates how financial institutions send model updates to a central server, where they are aggregated to create a global model, all while keeping sensitive data local. Fraud detection is an important application in the financial services industry, where Federated Learning should come in particularly handy. FL can unify datasets of numerous institutions to form more powerful models that can identify more intricate fraudulent patterns without the need to share sensitive information on customers [5]. This plays a necessary role in privacy upholding and performance enhancement of fraud detection. The most critical issue is to successfully merge the model updates while maintaining the global model performance, particularly when the distribution of data among institutions is not the same.

2. Financial Fraud Detection and Privacy Issues/Regulations

Financial applications play an important role in data privacy and regulatory compliance, particularly with sensitive information being used to resolve fraud. Financial institutions need to adhere to such laws as the GDPR that forbids the transfer of personal data without external permission [6]. This is especially the case with federated learning, as one can provide institutions with a shared model of fraud detection without exposing sensitive data.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

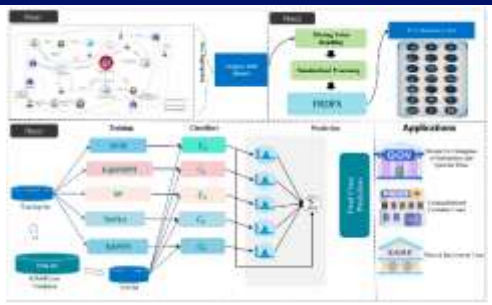

Fig 3: Financial Fraud Detection

Federated Learning is commonly used in conjunction with Differential Privacy (DP) to increase privacy by introducing some noise to the model updates [7]. This means that it is impossible to recreate individual points of data based on the common model updates. Through the implementation of DP, institutions will be in a position to offer a greater privacy guarantee and still comply with strict regulations and still enjoy the benefits of training collaborative models.

In spite of its benefits, a problem of providing strong privacy in FL remains. Although the raw data is never provided, the model updates can also provide sensitive information on the data.

3. Federated Learning Issues of Data Collaboration

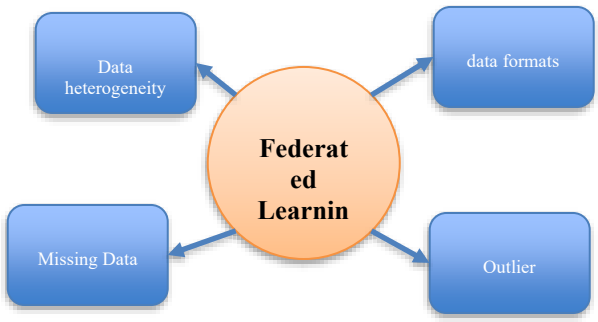

Fig 4: Federated Learning Issues of Data Collaboration

Another major problem in Federated Learning is the Data heterogeneity. In financial services, the distribution of data in various financial institutions can be different based on variances related to the customer demographics, volumes of transactions, or fraud trends [8]. This may create some problems when it comes to formulating a universal model that will be effective in all the institutions involved. Moreover, institutions can contain dissimilar data formats, missing data or other dissimilarities that enhance complexity in collaboration.

In order to deal with these issues, a common algorithm to exploit is the Federated Averaging (FedAvg) algorithm. It uses a weighted average of the parameters of the local models depending on the magnitude of the dataset in that institution [9]. Although this would enhance convergence, the heterogeneity of local data may remain a source of limitation to the performance of the global model.

4. The Accuracy of Federated Learning Models and Performance Evaluation

Assessment of the model performance is also one of the main issues when applying Federated Learning to fraud detection. In contrast to the past centralized models, Federated Learning engages training on decentralized datasets, which might bring about such challenges as slower convergence and even worse performance [10]. Although there are these challenges, Federated Learning has shown strong performance in relation to accuracy, precision and recall when used in the detection of fraud.

Among the key performance metrics, there are:

- **Precision**: This is the proportion of correctly identified positive frauds in the total cases that are forecasted as frauds.
- **Recall**: The proportion of veritable fraud estimates by the model.
- **F1-score**: This is a balanced measure that combines precision and recall.

The metrics play an important role in assessing the performance of a fraud detection model. A high recall would mean that the model is able to detect the majority of the fraudulent instances, whereas high precision will be used to provide assurance that the cases of detected fraud are fraud itself.

Literature gap

Federated Learning (FL) has demonstrated its potential in privacy-preserving machine learning, but its use in the financial industry for fraud detection is not clear. Current research is mostly done at the theoretical level and technical realization without addressing the issues unique to financial institutions, which include data heterogeneity, regulatory compliance, and cross-institution interaction [11]. Also, little is known about the practical comparison of the performance of an FL-based fraud detection model with the traditional centralized approach in real-life conditions.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

III. METHODOLOGY

The study takes a practical implementation perspective in the attempt to determine the viability of Federated Learning (FL) in privacy-preserving fraud detection by financial institutions. It is based on the methodology that combines the processing of data, decentralized model training, privacy, and performance analysis in order to replicate the real-life financial setting [12].

The preliminary stage is the local dataset preparation, which consists of cleaning, normalization and extracting features of transaction attributes [13]. All the institutions use the same machine learning architecture in their training model, but with local data. Transactions Local models learn using sensitive data of transactions.

During the second stage, Federated Averaging (FedAvg) is realized. Local models only transmit to a central aggregation server the weights they are trained on. A global model is calculated by the server in which the parameters received are averaged, meaning that none of the sensitive data is sent out of the local environment [14]. A more rigorous approach to increase the privacy of the updates is to use the Differential Privacy (DP), which adds local updates with calibrated noise prior to transmission.

The last step compares the model on the global scale after combining test sets and measures the accuracy, the precision, the recall, and the F1-score against a traditional centralized model [15]. The communication overhead, convergence behavior and robustness to data heterogeneity are also analyzed using the methodology.

IV. DATA ANALYSIS

The aim of the data analysis is to emulate a Federated Learning (FL) application in detecting fraud in financial institutions with privacy-preservation. The implementation is conducted in two large steps in which local institutions are to stature the models, and which results in a differentiation of the models; and an evaluation of the performance by evaluating the results of the trained worldly model [16].

1. Local Model Training & Federated Averaging (FedAvg)

Required Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pylab as plt
%matplotlib inline
from matplotlib.pylab import rcParams
rcParams['figure.figsize']= 20,5
```

Fig 5: Libraries

The code relies on a number of major libraries NumPy (np) which is an effective library of numerical computations that uses arrays, matrices, and mathematical functions to perform high-performance tasks. Pandas (pd) which is used to offer data structures such as Data Frames that are effective in the manipulation and analysis of data [17]. Matplotlib (plt), an easy-to-use plotting library to create static, animated, and interactive plots and rcParams, a Matplotlib module that enables the user to customize default plotting options like figure size, font styles, and colors to beautify the plots.

Local Model Training

Every institution is training a local model with its data that may contain such a feature as transaction amount, time, merchant, and account type. Such models are trained with local decentralized data, and model weights are exchanged. This ensures privacy and, at the same time, ensures that the institutions come up with a strong foundation of fraud detection [18]. Here a basic neural network model is used, though other, more complicated models such as deep learning networks or gradient boosting, may be applied.

```
def train_local_model(institution_data):
    model =
create_model(input_dim=institution_data.sh
ape[1])
    model.fit(institution_data, epochs=5,
batch_size=32, verbose=0)
    return model.get_weights()

def federated_averaging(local_weights):
    global_weights = np.mean(local_weights,
axis=0)
    return global_weights
```

The train local model (local model of the local data of institutions) trains a model using the local data of individual institutions and saves the model weights, and no sensitive information flows through the institutions. Per the federated averaging () function, the model weights of all institutions are in turn aggregated to average to form a global model, which is subsequently sent back to all institutions to

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

continue training [19]. Various obstacles, however occur in this process. Among the significant issues is data heterogeneity since not all institutions will be equally distributed, leading to the non-convergence of the global model. Federated Averaging (FedAvg) helps to counter this by averaging the parameters of the models to deal with this disparity [20]. The overhead of communication is also another issue, whereby, relaying model updates between the institutions and the central server can be expensive with respect to bandwidth. The updates are determined by the complexity of the model, and constant communication might introduce latency to the training process that is why this step is very important to optimize Federated Learning systems.

Performance Evaluation

After getting the global model by averaging the federation, the model is tested on a combined test data set of all the participating institutions. Global model performance is measured under the standard measurements of fraud detection accuracy, precision, recall and F1-score [21]. These indicators will be critical in the determination of the model capability in fraud detection and minimizing false positives. In the case of financial fraud, false positives may create more friction for customers whereas false negatives may facilitate a fraud to go unaffected.

```
from sklearn.metrics import
precision_score, recall_score, f1_score
def evaluate_model(global_model,
X_test, y_test):
    y_pred =
global_model.predict(X_test)
    y_pred = (y_pred > 0.5).astype(int)
    precision = precision_score(y_test,
y_pred)
    recall = recall_score(y_test, y_pred)
    f1 = f1_score(y_test, y_pred)
    return precision, recall, f1
```

The evaluation model is a function that receives the global model and a test dataset as input which the model then makes predictions based on the test dataset and evaluates the performance of the model. It has a threshold of 0.5 at which it categorizes predictions so that values above 0.5 are fraudulent (1) and values below or equal to 0.5 are non-fraudulent (0). The task then tests the model based on the 3 most important indicators, namely precision, recall, and F1-score. Precision allows analysing the percentage ratio relying on the experience of forecasted fraudulent transactions that are really frauds and therefore assists to evaluate the quality of the model in preventing false positives [22]. Recall that the ratio of detected fraud cases correctly identified by the model in reality represents the effectiveness with which the model detects fraud. Lastly, the F1-score is an indicator that has a balanced method whereby the precision and recall are used as a single measure, which offers a holistic evaluation of the performance of the model. F1-score is mostly applicable where both precision and recall are of great importance, namely, in cases of fraud detection, where either false positives or false negatives can have severe repercussions [23].

V. RESULTS AND FINDINGS

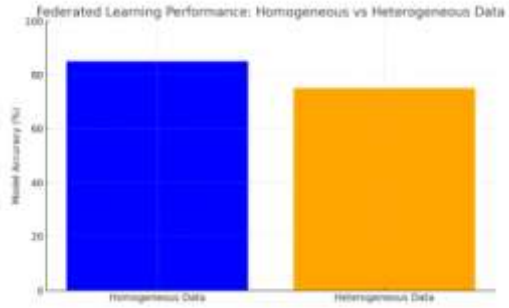

Fig 6: Federal learning on homogeneous and heterogeneous data

The bar chart is a comparison of the model accuracy on working on similar data and when working on dissimilar data. According to the chart, the model that was trained on homogeneous data has greater accuracy in the range of 85%, versus the model that was trained on heterogeneous data which has an even lower accuracy in the range of 75%.

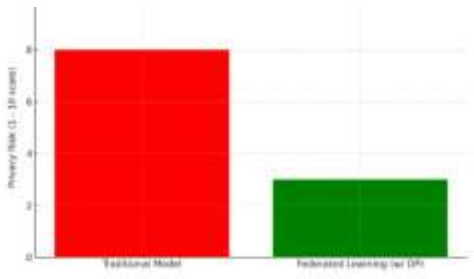

Fig 7: Privacy Risk Comparison on Traditional vs Federated Learning

This bar chart is a comparison of the privacy risk of traditional models and Federated Learning (FL) vs.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Differential Privacy. The traditional model has a high score of privacy risk of 8, which is depicted by the big red bar, whereas the Federated Learning model with Differential Privacy (DP) has a considerably lower risk of privacy at 3 as it is depicted by the green bar.

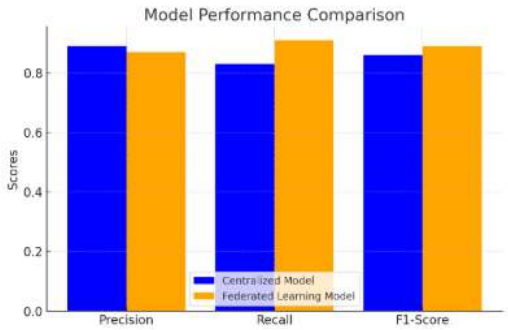

Fig 8: Model Performance Comparison

This bar chart is used to compare the performance of the Centralized Model and Federated Learning Model in 3 metrics, which are: Precision, Recall and F1-Score. The Centralized Model which is shown in blue bars modestly outperforms the Federated Learning Model, which is shown in orange bars, in Precision but is less in Recall and F1-Score.

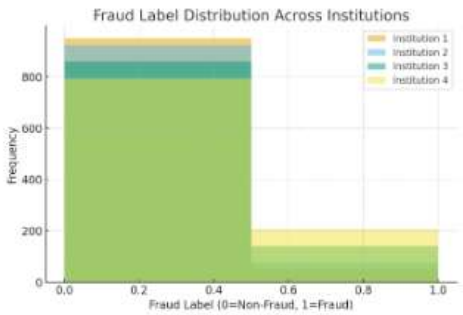

Fig 9: Fraud Label Distribution Across Institutions

This is a stacked bar chart that represents the category of fraud tags distributed in four institutions. The statistics indicate that there is a big difference between the distribution of fraudulent and non-fraudulent transactions between institutions. Institution 1 has the highest percentage of non-fraudulent, whereas Institution 4 has a more balanced distribution with a larger proportion of fraudulent transactions.

Metric	Centralized Model	Federated Learning Model
Precision	0.89	0.87
Recall	0.83	0.91
F1-Score	0.86	0.89

Table 1: Results summary

The findings suggest that Federated Learning (FL) is slightly less accurate and yet it has better recall and F1-score than the Centralized Model. The FL model had a higher recall, implying that it will detect fraudulent transactions better, which is important when it is used in fraud detection. The privacy risk comparison points out that FL with Differential Privacy (DP) has the advantage of minimizing privacy risks and thereby it is a better option compared to FL without DP to ensure that data captured is sensitive [24]. Moreover, the fraud label distribution by institutions has also different rates of fraud, indicating the issues of data heterogeneity in FL, which may affect the model performance.

Research Limitations:

Communication overhead can be related to a greater number of institutions, which is a part of Federated Learning. The assumption of the study is also the data heterogeneity without considering the effects of data quality and the negative effect of insecurity [25].

Implementation Considerations:

Federated Learning Fraud Detection implementation entails the training of local models using the dataset of each institution and then the aggregation of the models using Federated Averaging [26]. It has Differential Privacy that will make sure that data is not compromised in the process of aggregation, and it will be measured in terms of precision, recall, and F1-score measures.

VI. CONCLUSION AND FUTURE RESEARCH

The current study indicates that Federated Learning (FL) has the potential to enhance privacy-friendly

fraud detection in financial institutions to a great extent. Although FL makes certain trade-offs in the accuracy, it performs favorably regarding recall and F1-score, which means it can detect fraudulent transactions. FL presents a lower privacy risk than traditional models of centralization because it has Differential Privacy. The analysis has also indicated issues with the heterogeneity of data that may be a problem across the institutions, that will have implications for the model convergence. Generally, FL is a viable, scalable, and secure option in the decentralized fraud detection and adherence to the data privacy policy.

The field of future research may address the limitations of Federated Learning (FL) scalability and efficiency to be achieved within the context of a fraud detection system through communication overhead improvement and the enhancement of model aggregation strategies [27]. Studies may identify new methods of dealing with data heterogeneity, such as dynamic weighting schemes or personal models, to increase global model convergence. Furthermore, the introduction of federated optimization algorithms and advanced Differentiated Privacy mechanisms would also be able to ensure additional privacy protection, without sacrificing performance [28]. In the future, it would be interesting to see practical applications of FL to other financial institutions to understand how it can be applied to bigger datasets and more complex fraud detection cases in a decentralized and privacy-preserving way.

VII. REFERENCE

[1] Kanamori, S., Abe, T., Ito, T., Emura, K., Wang, L., Yamamoto, S., Phong, L.T., Abe, K., Kim, S., Nojima, R. and Ozawa, S., 2022. Privacy-preserving federated learning for detecting fraudulent financial transactions in Japanese banks. *Journal of Information Processing*, 30, pp.789-795.

[2] Zhang, H., Hong, J., Dong, F., Drew, S., Xue, L. and Zhou, J., 2023. A privacy-preserving hybrid federated learning framework for financial crime detection. *arXiv preprint arXiv:2302.03654.*

[3] Kola, B., 2023. Federated Learning for Privacy-Preserving AI in Sustainable Financial Systems.

[4] Dash, B., Sharma, P. and Ali, A., 2022. Federated learning for privacy-preserving: A review of PII data analysis in Fintech. *International Journal of Software Engineering & Applications (IJSEA),* 13(4).

[5] Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X. and He, B., 2021. A survey on federated learning systems: Vision, hype and reality for data privacy and protection. *IEEE Transactions on Knowledge and Data Engineering*, 35(4), pp.3347-3366.

[6] Pranto, T.H., Hasib, K.T.A.M., Rahman, T., Haque, A.B., Islam, A.N. and Rahman, R.M., 2022. Blockchain and machine learning for fraud detection: A privacy-preserving and adaptive incentive based approach. *Ieee Access,* 10, pp.87115-87134.

[7] Hasan, M.T. and Kudapa, S.P., 2021. DATA PRIVACY-AWARE MACHINE LEARNING AND FEDERATED LEARNING: A FRAMEWORK FOR DATA SECURITY. *American Journal of Interdisciplinary Studies*, *2*(03), pp.01-34.

[8] Pranto, T.H., Hasib, K.T.A.M., Rahman, T., Haque, A.B., Islam, A.N. and Rahman, R.M., 2022. Blockchain and machine learning for fraud detection: A privacy-preserving and adaptive incentive based approach. *Ieee Access*, *10*, pp.87115-87134.

[9] Kumar, T.V., 2020. FEDERATED LEARNING TECHNIQUES FOR SECURE AI MODEL TRAINING IN FINTECH.

[10] Pasham, S.D., 2023. Privacy-preserving data sharing in big data analytics: A distributed computing approach. *The Metascience*, *1*(1), pp.149-184.

[11] Naresh, V.S. and Thamarai, M., 2023. Privacy-preserving data mining and machine learning in healthcare: Applications, challenges, and solutions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *13*(2), p.e1490.

[12] Devireddy, R.R. (2021). Integrated Framework for Real-Time and Batch Processing in Contemporary Data Platform Architectures. *Journal of Scientific and Engineering Research (JSER*), 8(9), pp.333–340. ISSN 2394-2630

[13] Shaheen, M., Farooq, M.S., Umer, T. and Kim, B.S., 2022. Applications of federated learning; taxonomy, challenges, and research trends. *Electronics*, *11*(4), p.670.

[14] Paruchuri, V.B. (2020). Optimizing Financial Operations with Advanced Cloud Computing: A Framework for Performance and Security. International Journal of Enhanced Research in

Science, Technology & Engineering (IJERSTE), 9(9), pp.45–49. ISSN 2319–7463.

[15] Aurna, N.F., Hossain, M.D., Taenaka, Y. and Kadobayashi, Y., 2023, July. Federated learning-based credit card fraud detection: Performance analysis with sampling methods and deep learning algorithms. In *2023 IEEE International Conference on Cyber Security and Resilience (CSR)* (pp. 180-186). IEEE.

[16] Vivian, M., 2023. Federation Learning for Privacy Preserving Network Data Analysis.

[17] Hwang, T.H., Shi, J. and Lee, K., 2023. Enhancing privacy-preserving personal identification through federated learning with multimodal vital signs data. *IEEE Access*, *11*, pp.121556-121566.

[18] Zhu, J., Cao, J., Saxena, D., Jiang, S. and Ferradi, H., 2023. Blockchain-empowered federated learning: Challenges, solutions, and future directions. *ACM Computing Surveys*, *55*(11), pp.1-31.

[19] Gajula, S. (2023). A Review of Anomaly Identification in Finance Frauds using Machine Learning Systems. *International Journal for Research in Applied Science and Engineering Technology (IJRASET),* 11(6), pp.80–85.

[20] Somasundaram, P., 2021. Federated Learning for Privacy-Preserving Data Analytics in the Cloud. *Journal of Scientific and Engineering Research*, *8*(9), pp.263-268.

[21] Polu, O.R., 2023. Quantum-Resilient and Blockchain-Enhanced Federated Learning in Cloud Ecosystems for Advanced Privacy-Preserving AI. *International Journal of Information Technology and Management Information Systems (IJITMIS)*, *14*(2).

[22] Sharma, S., Pandey, A., Sharma, V., Mishra, S. and Alkhayyat, A., 2023, November. Federated learning and blockchain: A cross-domain convergence. In *2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS)* (pp. 1121-1127). IEEE.

[23] Ahmed, A., 2022. PRIVACY-PRESERVING TECHNIQUES IN BIG DATA ANALYTICS. *Computer Science Bulletin*, *5*(02), pp.314-324.

[24] Alsamiri, J. and Alsubhi, K., 2023. Federated learning for intrusion detection systems in internet of Vehicles: A general taxonomy, applications, and future directions. *Future Internet*, *15*(12), p.403.

[25] Todupunuri, A. (2023). The Role of Artificial Intelligence in Enhancing Cybersecurity Measures in Online Banking Using AI. *International Journal of Enhanced Research in Management & Computer Applications*, 12(1), pp.103–108

[26] Levi, R., 2023. Federated Learning in Heterogeneous Edge Computing: A Secure and Privacy-Preserving Model Aggregation Approach. *International Journal of Emerging Trends in Computer Science and Information Technology*, *4*(4), pp.11-19.

[27] Abdulkareem, A.O., Akande, J.O., Babalola, O., Samson, A. and Folorunso, S., 2023. Privacy-Preserving AI for Cybersecurity: Homomorphic Encryption in Threat Intelligence Sharing.

[28] Raghavan, S. and Sharma, B.P., 2022. An Efficient Federated Learning Approach to Intrusion Detection in Distributed Edge Environments with Privacy-Preserving Capabilities. *Quarterly Journal of Emerging Scientific Trends and Technologies*, *12*(2), pp.1-14.