Journal of Information, Technology and Data Science

A Privacy-Preserving Federated Learning Framework for Collaborative Academic Certificate Fraud Detection Across Institutions

Miriam W. Kaara, Dr. Jael S. Wekesa & Dr. Michael W. Kimwele

ISSN: 2617-3573

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

A Privacy-Preserving Federated Learning Framework for Collaborative Academic Certificate Fraud Detection Across Institutions

Miriam W. Kaara[1], Dr. Jael S. Wekesa[2], Dr. Michael W. Kimwele[3]

[1]Department of Computing, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

[2]Department of Pure and Applied Sciences (PAS/IT), Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

[3]Department of Computing, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Corresponding Author

Miriam Kaara

School of Computing and Information Technology,

Jomo Kenyatta University of Agriculture and Technology,

Nairobi, Kenya

Miriam.kaara@gmail.com

ORCID: 0009-0007-7704-4406

***How to cite this article*: Kaara M., W., Wekesa J., S. & Kimwele M., W. (2026). A Privacy-Preserving Federated Learning Framework for Collaborative Academic Certificate Fraud Detection Across Institutions. *Journal of Information, Technology and Data Science, Vol* 10(1) pp. 94-117.** https://doi.org/10.53819/81018102t2599

Abstract

The integrity and credibility of educational institutions worldwide are being undermined by the growing issue of academic certificate fraud. The public's trust in the educational system, the integrity of job opportunities, and the legitimacy of authentic credentials are all negatively impacted by academic dishonesty. The majority of academic credentialing systems, both centralized and blockchain-based, focus on academic credential validation and immutability, even though academic institutions have embraced technology to improve academic credentialing procedures. This suggests that their support for intelligent and private fraud detection is minimal. In this work, we suggest an integrated method that combines smart anomaly detection with

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

Federated Learning (FL). This makes it possible for several organizations to build a model for detecting certificate fraud without disclosing private information to other organizations. Each institution's private information is retained and incorporated into a global model. This enables the framework to handle concerns about data ownership and privacy as well as regulatory compliance. The framework uses Extreme Gradient Boosting (XGBoost) to identify anomalies in metadata and Convolutional Neural Networks (CNNs) to detect visual forgeries in certificates. A decentralized node is used to train the models, which are merged using the Federated Averaging (FedAvg) algorithm. The federated model performs well in anomaly detection, according to the experiments conducted. When compared to conventional centralized approaches, it significantly reduces false positives and false negatives, with an accuracy of up to 94% and an AUC of 0.97. The model's findings imply that a federated learning approach would make it possible for institutions to detect fraud in a secure, scalable, and cooperative manner. In this regard, this framework offers a workable way to create digital credential systems that are more trustworthy and protect privacy.

Key words: *Federated Learning, Blockchain, certificate fraud detection, privacy preservation, distributed datasets, secure collaboration.*

1. Introduction

The cornerstone of culture and the driving force behind a nation's economy lies in education as the foundation of a civilized society. Future leaders, employees, and citizens are all prepared by the educational system. Therefore, ensuring the quality of education, creating a suitable university environment, and producing qualified staff to build promising generations are the primary concerns of education stakeholders. However, academic certification fraud has gained attention recently and has had detrimental effects on several economic sectors, including education. People who engage in these actions are employed by universities, phony universities, and counterfeit websites to generate such documents. The broad availability of technology tools for creating and printing these phony documents, as well as online distribution channels, has adversely contributed to this. In industries including healthcare, engineering, and public service, these advancements have resulted in an increasing number of fake credentials being used to obtain jobs, promotions, and access to professional prospects, frequently with dire consequences.

Traditional certificate verification techniques have relied on centralized systems and manual validation processes. For academic credentials, these systems offer fundamental authentication features. Nevertheless, they rely solely on isolated institutional databases, are slow, have scalability problems, and are vulnerable to manipulation. Furthermore, centralized verification systems often result in single points of failure and inadequately facilitate cross-institutional or cross-border verification.

Blockchain-based credentialing solutions, including Blockcerts and corporate frameworks, have been suggested and put into practice in several institutions to address these issues. The decentralized ledger at the core of these systems is intended to guarantee academic credentials' immutability, transparency, and resistance to tampering. These blockchain methods offer a way to store and verify data securely. However, there has been a lack of intelligence in identifying fraudulent or fake certificates, with the focus being on credential authentication. They are therefore less able to proactively detect complex fraud practices.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

Additionally, machine learning methods have been proposed and implemented for fraud detection in several industries, including cybersecurity and banking, in support of blockchain solutions. However, the majority of machine learning techniques now in use rely on centralized data collection, necessitating the sharing of sensitive datasets between institutions. Academic certifications are significantly impacted by this centralized data manipulation, particularly in terms of ownership, privacy, and regulatory compliance. Institutions are now unwilling to take part in collaborative fraud detection initiatives, which eventually leads to fragmented datasets and decreased model efficacy. To solve these problems, a trustworthy fraud detection model that can operate across a range of data attributes while preserving privacy, authenticity, and integrity must be developed.

Federated Learning (FL), a machine learning technique, has emerged as a dependable way to address issues with current fraud detection systems. FL enables a decentralized, privacy-preserving model training process in which several institutions work together to build a common model while maintaining local data. This promotes data sovereignty and secrecy. As a result, FL shows great promise in domains where data privacy is crucial, such as healthcare and banking. However, academic certificate fraud detection mechanisms have not optimized the FL approach, particularly in the context of different data characteristics and heterogeneous institutional environments. Certificates, for example, have both visual components, including signatures, seals, and emblems. These documents also contain structured metadata, such as serial numbers, institutional IDs, and dates of issuance. Furthermore, due to differences in record-keeping procedures, rules, and certificate formats, data distributions among institutions are intrinsically heterogeneous (non-independent and identically distributed; Non-IID).

Thus, by suggesting a federated learning architecture with privacy-preserving features, this study seeks to further efforts to detect academic certificate fraud. The literature reviews on the state of academic certificate fraud detection served as the basis for our proposed model. The suggested approach uses a hybrid model that combines federated learning and machine learning methods to enable analysis based on images and metadata, respectively. This will protect data privacy by allowing institutions to work together to develop a global fraud detection model without exchanging raw data.

This work's main contribution is the development and implementation of a federated learning model tailored to the unique requirements of academic certificate fraud detection. The architecture provides a mechanism for how decentralized training, hybrid feature modeling, and safe aggregation may be used to achieve high detection accuracy while protecting data privacy. Additionally, by tackling the issue of varied data distributions, the framework guarantees scalability and generalization across numerous institutions.

2. Related Work

2.1 Federated Learning for Privacy-Preserving Collaborative Learning

Federated Learning (FL) has emerged as a promising distributed machine learning paradigm that enables multiple institutions to collaboratively train models without exchanging sensitive data. Instead of transferring raw datasets, each participating institution trains a local model and shares only model parameters or gradients with a central aggregation server. This decentralized training strategy significantly reduces privacy risks while facilitating collaborative intelligence across

organizations. Consequently, FL has attracted considerable attention in domains where sensitive information cannot be centrally stored, including healthcare, finance, and education.

According to Qammar et al. (2023), FL enables collaborative learning among multiple institutions while minimizing the security risks associated with centralized data repositories. Nevertheless, although FL improves data confidentiality by keeping institutional data local, it remains vulnerable to several attacks during model parameter exchange. Chiș and Caramihai (2023) and Aledhari et al. (2020) reported that malicious participants may exploit shared model updates through model inversion, membership inference, or gradient leakage attacks to reconstruct sensitive information. These vulnerabilities have motivated extensive research into privacy-preserving mechanisms that strengthen the security of federated learning environments.

2.2 Privacy-Preserving Techniques in Federated Learning

Several cryptographic and statistical techniques have been proposed to enhance privacy protection within federated learning systems. Differential Privacy (DP), Homomorphic Encryption (HE), Secure Multi-Party Computation (SMPC), Secure Aggregation, Federated Averaging (FedAvg), and Gradient Compression are among the most widely adopted approaches.

Sameera et al. (2023) emphasized that secure aggregation prevents the server from accessing individual model updates by aggregating encrypted local parameters. Differential Privacy introduces calibrated noise into model updates to prevent inference attacks, whereas Homomorphic Encryption enables computations on encrypted model parameters without revealing sensitive information. Secure Multi-Party Computation distributes encrypted model shares among multiple participants to eliminate dependence on trusted intermediaries. Federated Averaging remains the standard optimization algorithm for aggregating local models, while Gradient Compression reduces communication overhead by transmitting compressed gradients. Although these techniques improve privacy and communication efficiency, each introduces trade-offs between computational complexity, communication cost, model convergence, and prediction accuracy.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

Table 1. Privacy-preserving techniques used in Federated Learning for certificate fraud detection

Privacy Technique	Description	Relevance to Certificate Fraud Detection	Strengths	Limitations
Secure Aggregation	Enables aggregation of encrypted local institutional models so that the server only accesses the combined global model.	Prevents exposure of institutional model updates and sensitive certificate features during collaborative training.	Strong protection of model updates and supports secure multi-institution collaboration.	High computational cost and participant synchronization requirements.
Differential Privacy (DP)	Introduces controlled noise into model updates or gradients to prevent learning sensitive information about individual records.	Protects certificate information from inference attacks through shared model updates.	Provides measurable privacy guarantees and integrates effectively with FL systems.	Excessive noise may reduce model accuracy, creating a trade-off between privacy and performance.
Homomorphic Encryption (HE)	Encrypts model updates while still allowing aggregation and computation on encrypted data.	Maintains confidentiality of certificate features during transmission and aggregation.	Provides strong cryptographic privacy without exposing raw institutional data.	Requires high computational power and may slow down model training.
Secure Multi-Party Computation (SMPC)	Splits model updates into encrypted shares processed by multiple parties so that no single participant can access raw updates.	Enables collaborative model training across institutions without sharing certificate features.	Protects against insider threats and supports decentralized trust.	Complex implementation, communication-intensive, and may reduce training speed.
Federated Averaging (FedAvg)	Aggregates local model weights from participating institutions without requiring access to raw data.	Supports collaborative certificate fraud detection while preserving institutional privacy.	Easy to implement and scalable across multiple institutions.	Vulnerable to model inversion attacks if not combined with additional privacy-preserving techniques.
Gradient Compression	Reduces the amount of information transmitted in model gradients during training.	Minimizes the possibility of sensitive certificate patterns being revealed through model updates.	Reduces communication overhead and improves transmission efficiency.	Provides weaker privacy guarantees and may negatively affect model convergence and model accuracy

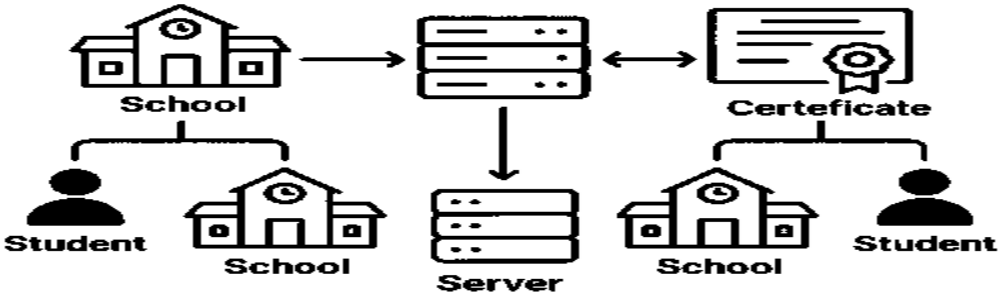
Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

2.3 Applications of Federated Learning in Educational Systems

Although educational applications of federated learning remain relatively limited compared to healthcare and IoT, recent studies demonstrate its potential for protecting student information while enabling collaborative analytics across institutions.

Wang et al. (2024) demonstrated that federated learning can support intelligent tutoring systems by allowing educational institutions to collaboratively improve predictive models without exposing confidential student records. Similarly, Singh et al. (2025) proposed a blockchain-integrated federated learning framework for academic credential management and anomaly detection. Their findings showed that institutions could jointly identify suspicious academic records while maintaining student privacy and institutional data ownership.

Figure 1. Federated learning framework for certificate verification



2.4 Challenges of Federated Learning in Academic Credential Verification

Despite its advantages, several challenges limit the adoption of federated learning in educational environments. One of the most significant issues is the presence of heterogeneous and non-independent and identically distributed (non-IID) datasets. Educational institutions often maintain different grading systems, certificate templates, curricula, and academic regulations, leading to highly heterogeneous data distributions.

Lee et al. (2024) observed that such heterogeneity reduces model convergence speed and negatively affects predictive accuracy during collaborative training. To address these limitations, Yang et al. (2023) proposed adaptive aggregation strategies and federated clustering techniques capable of improving model performance across heterogeneous institutional datasets. Nevertheless, handling non-IID educational data remains an active research challenge.

2.5 Blockchain-Enabled Federated Learning

Recent research has increasingly integrated blockchain technology with federated learning to eliminate centralized aggregation servers and improve transparency, trust, and accountability. Conventional FL architectures rely on a central aggregation server that represents a potential single point of failure and introduces trust concerns.

Yu et al. (2022) proposed blockchain-enabled orchestration mechanisms in which smart contracts automate participant registration, model validation, aggregation, and reward distribution. Similarly, Zhang et al. (2021) introduced BlockFL, a decentralized federated learning architecture that utilizes blockchain smart contracts to coordinate participant interactions and model aggregation. Their results demonstrated improvements in system transparency, scalability, and auditability.

Likewise, Aledhari et al. (2020) proposed blockchain-based validation mechanisms capable of evaluating local model quality before aggregation, thereby mitigating poisoned model updates and malicious participant behavior. These studies collectively demonstrate that blockchain technology can substantially strengthen the security and trustworthiness of federated learning systems.

2.6 Research Gap

Despite substantial progress in blockchain-enabled federated learning, existing studies primarily focus on applications in healthcare, Internet of Things (IoT), finance, and industrial environments. Comparatively little attention has been devoted to academic credential verification and certificate fraud detection.

Furthermore, current solutions rarely integrate metadata anomaly detection with image-based counterfeit certificate analysis within a unified federated learning framework. Existing approaches also provide limited support for collaboration among educational institutions operating under heterogeneous infrastructures, grading systems, and academic policies. Consequently, secure and scalable cross-institutional certificate verification remains an open research problem.

2.7 Positioning of the Proposed Framework

To address these limitations, the proposed framework integrates Federated Learning, blockchain technology, artificial intelligence, and smart contracts into a unified decentralized certificate verification architecture. The framework combines Convolutional Neural Networks (CNNs) for counterfeit certificate image detection with XGBoost for metadata anomaly detection while blockchain provides immutable record management and smart contracts automate verification processes.

Unlike existing studies, the proposed framework enables multiple institutions to collaboratively train fraud detection models without exchanging sensitive academic records, thereby preserving institutional autonomy, improving privacy, and enhancing trust across decentralized educational ecosystems (Hu et al., 2023; Alghamdi et al., 2026).

3. Methodology

3.1 Research Design

The study used a quantitative experimental research design to develop and test a federated learning (FL) model to detect academic certificate fraud. The design was flexible enough to be able to simulate multiple academic institutions as distributed learning nodes and train models collaboratively without sharing raw certificate data. The goal was to design a privacy-preserving

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

fraud detection system that can be trained on heterogeneous institutional datasets, but keep the data under the control of the institutions.

3.2 Dataset description

This study is a simulation of a federated learning environment for certificate fraud detection with datasets of academic certificates obtained from multiple institutions. The dataset included the following data:

i. Certificate Images: Degree certificates, Diplomas, Academic transcripts, Professional certifications
ii. Certificate Metadata: Certificate numbers, Student registration numbers, Issue dates, Graduation dates, Program names, and Institution identifiers
iii. Synthetic certificate datasets generated using document forgery tools

For privacy reasons, personally identifiable information (PII) was anonymized prior to training of the model. In federated learning, each institution maintains its local dataset and works together in model training since the data is not centrally stored. Table 2 below summarizes the datasets.

Table 2: Datasets and sources

Dataset Type	Source	Purpose
Certificate Images	Participating institutions	Visual forgery detection
Metadata Records	Institutional databases	Pattern recognition
Synthetic certificate datasets	Generated using document forgery tools	Model validation

3.3 Feature Extraction

Certificate images and metadata were separately processed for feature extraction.

3.3.1 Visual Features

An image-based feature extraction was implemented using a Convolutional Neural Network (CNN) which extracted the following features:

i. Logo consistency
ii. Signature authenticity
iii. Seal integrity
iv. Font characteristics
v. Layout structure
vi. Watermark presence
vii. QR code validity

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

The feature extraction procedure is denoted as the following:

$$F_v = CNN(I)$$

Where:

- I= Input certificate image
- $CNN(\cdot)$= Convolutional Neural Network feature extraction function
- F_v= Extracted visual feature vector

The CNN learns patterns such as logo consistency, signature authenticity, watermark presence, seal integrity, font characteristics, QR code validity, and document layout structure. Then the feature vector obtained is classified.

3.3.2 Metadata Features

Metadata features included:

i. Certificate ID
ii. Institution code
iii. Program code
iv. Date of issue
v. Student registration number
vi. Academic award type
vii. Verification status

Term Frequency–Inverse Document Frequency (TF-IDF) was used to transform textual metadata.

The TF-IDF weight of a term is calculated as:

$$\mathrm{TFIDF(t,d)} = \mathrm{TF(t,d)} \times \mathrm{IDF(t)}$$

Where:

- $\mathrm{TF(t,d)}$= Frequency of term t in document d
- $\mathrm{IDF(t)}$= Inverse Document Frequency of term t

The inverse document frequency is computed as

$$\mathrm{IDF(t)} = \log\left(\frac{\mathrm{N}}{\mathrm{df(t)}}\right)$$

Where:

- N= Total number of documents
- df(t)= Number of documents containing term t

TF-IDF assigns higher weights to terms that are frequent in a particular document but rare in the whole collection, which improves the discrimination ability of metadata features.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

3.4 Data Preprocessing

The data was preprocessed before the model was trained, to increase the quality and consistency of the certificate data across the participating institutions.

3.4.1 Image preprocessing

Image preprocessing involved the following:

i. Noise reduction using Gaussian filtering
ii. Grayscale conversion
iii. Image resizing to standardized dimensions
iv. Optical Character Recognition (OCR) for text extraction
v. Data augmentation through rotation and cropping
vi. Image hashing for duplicate detection

3.4.2 Metadata preprocessing

This activity involved the following:

i. Data cleaning and normalization
ii. Tokenization of textual records
iii. Stop-word removal
iv. Named Entity Recognition (NER)
v. TF-IDF vectorization
vi. Anonymization of personally identifiable information

3.5 Federated Learning-Based Fraud Detection Approach

At each institutional node, a hybrid machine learning architecture was designed to detect fraudulent academic credentials.

3.5.1 FL Architecture

The architecture consisted of:

3.5.1.1 Convolutional Neural Network (CNN)

To analyze the image attributes of certificate images, the CNN model was adopted. It checked for irregularities on the image level such as counterfeit signatures, altered seals, misaligned logos and formatting.

3.5.1.2 XGBoost Classifier

The XGBoost model analyzed metadata attributes including certificate numbers, dates issued, institution identities, graduation dates and qualification attributes. The model was able to identify abnormal patterns related with certificate forgery.

3.5.1.3 Decision Tree Classifier

The use of decision trees allowed to present classification results in an understandable way and interpret the aspects that influence the prediction of frauds.

Each institution trained the models independently with their local data in order to ensure that sensitive certificate information remain within the boundaries of the institutions.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

3.5.2 Federated Learning Process

The Federated Learning process was implemented using the Federated Averaging (FedAvg) algorithm.

The process consisted of the following steps:

Step 1: Initialization

A global model was initialized at the federated server and distributed to all participating institutional nodes.

Step 2: Local Training

Each institution trained the received model on its local dataset for several epochs. Model parameters were updated based on local learning objectives.

Step 3: Secure Model Sharing

Instead of transmitting raw certificate data, institutions transmitted encrypted model weights and gradients to the federated server.

Step 4: Global Aggregation

The federated server aggregated the local model updates using the FedAvg algorithm based on the following formula:

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{N} w_k^t$$

Where:

- w^{t+1}= Updated global model weights
- w_k^t= Local model weights from institution k
- n_k= Number of training samples at institution k
- $N = \sum_{k=1}^{K} n_k$= Total number of samples across all institutions
- K= Number of participating institutions

This weighted aggregation ensured that institutions with larger datasets contributed proportionally to the global model.

Step 5: Model Redistribution

The updated global model was redistributed to all participating institutions for the next training round.

The process was repeated until model convergence was achieved.

3.6 Experimental Setup

A distributed environment of five simulated university nodes was built in Docker containers. Each node was an independent institution with distinct certificate datasets. The federated server coordinated communication, aggregation, and model synchronization. The simulation

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

environment was designed to reflect real-world institutional heterogeneity and non-identically distributed (non-IID) data conditions, adhering to the following specifications.

i. Hardware

- Intel Core i7 Processor
- 16 GB RAM
- NVIDIA RTX GPU

ii. Software

- Python 3.11
- TensorFlow/PyTorch
- Flower Federated Learning Framework
- Scikit-Learn
- Docker Containers
- PostgreSQL

iii. Federated Configuration

Parameter	Value
Number of Institutions	5
Communication Rounds	50
Local Epochs	5
Learning Rate	0.001
Batch Size	32
Aggregation Method	FedAvg

3.6 Evaluation Metrics

The effectiveness of the proposed model was evaluated using standard classification metrics.

i. Accuracy

Measures overall prediction correctness.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where:

- TP= True Positives
- TN= True Negatives
- FP= False Positives
- FN= False Negatives

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

ii. Precision

Measures the proportion of predicted fraud cases that are actually fraudulent.

$$Precision = \frac{TP}{TP + FP}$$

iii. Recall

Measures the ability to identify fraudulent certificates.

$$Recall = \frac{TP}{TP + FN}$$

iv. F1-Score

Balances precision and recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

v. False Positive Rate (FPR)

$$FPR = \frac{FP}{FP + TN}$$

vi. False Negative Rate (FNR)

$$FNR = \frac{FN}{FN + TP}$$

vii. ROC-AUC

The Area Under the Receiver Operating Characteristic Curve (AUC) quantifies the model's capacity to differentiate between authentic and fraudulent certificates at various categorization levels. The ROC Curve is calculated as follows:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR)$$

Where:

- TPR= True Positive Rate (Recall)
- FPR= False Positive Rate
- d= very small increment of a variable

Higher AUC values indicate better fraud detection performance.

viii. Communication Efficiency

Since federated learning involves distributed communication, communication overhead was measured as:

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

$$CO = \sum_{r=1}^{R} Dat\, a_r$$

where:

- R= communication rounds
- $Data_r$= transmitted model updates per round

ix. Privacy Preservation Score

The privacy advantages of federated learning were evaluated by contrasting the volume of raw data sent in centralized learning with that in federated learning. The privacy preservation capability is measured as:

$$PPS = 1 - \frac{D_{FL}}{D_{CL}}$$

Where:

- PPS= Privacy Preservation Score
- D_{FL}= Data shared in Federated Learning
- D_{CL}= Data shared in Centralized Learning

A score closer to 1 indicates stronger privacy preservation because less raw data is exposed during model training.

4 Results

4.1 Performance of Federated Learning Models

The findings from the development and assessment of the Federated Learning (FL) model, which allows several institutions to jointly train a certificate fraud detection system while protecting data privacy, are presented in this session. The efficacy of the proposed FL model was evaluated by comparing its performance to a centralized baseline.

The results of the trial show that the FL model outperformed the centralized model, which recorded 91.2% accuracy, with a global accuracy of 94.7%, as shown in Table 3 below. This improvement of approximately 3.5 percentage points highlights the efficacy of FL in leveraging distributed, institution-specific datasets without direct data sharing.

The FL model obtained a precision of 0.94, a recall of 0.95, and an F1-score of 0.94, while the centralized model achieved 0.90, 0.89, and 0.89, respectively, according to additional evaluation using conventional classification metrics. These findings show that the federated strategy offers a better balance between limiting the misclassification of legitimate credentials and identifying fake certificates.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

Table 3: Centralized Training vs. FL Accuracy

Training Type	Dataset Configuration	Model Type	Accuracy (%)	Precision	Recall	F1-Score
Centralized	Aggregated dataset (single node)	CNN-XGBoost	91.2	0.90	0.89	0.89
Federated Learning	Distributed (5 nodes, non-IID)	CNN-XGBoost	94.7	0.94	0.95	0.94

The enhanced performance is credited to the non-IID (non-independent and identically distributed) character of the data among participating institutions, which improves the model's capacity to generalize across various fraud patterns.

4.2 Training Efficiency and Convergence

As seen in Figure 2 below, the federated training process showed sustained convergence and reached optimal performance. As client updates successfully synchronize through the FedAvg aggregation process, it shows a consistent rise in model accuracy, with convergence seen after about 36 rounds, validating the model's stability and learning efficiency.

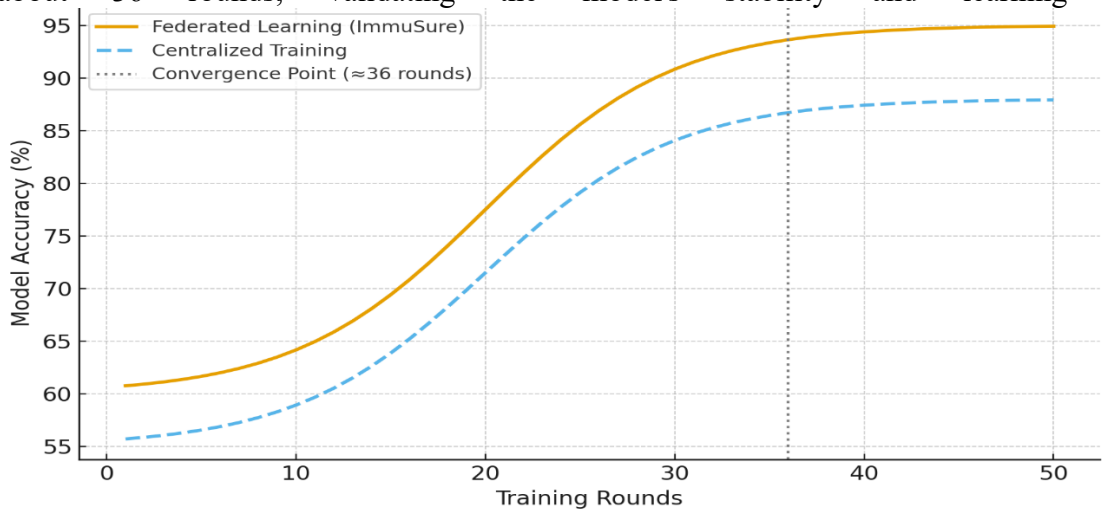


Figure 2: Accuracy trend across federated learning training rounds

The steady increase in model correctness over training rounds confirms the efficacy of the Federated Averaging (FedAvg) algorithm in aggregating dispersed model updates. This convergence pattern suggests that high model performance can be attained through collaborative learning among dispersed nodes in a manageable number of iterations, rendering the method computationally viable for practical implementation.

4.3 ROC Curve and Discriminative Performance

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) were used to assess the FL model's discriminative power. As seen in Table 4 below, the FL model

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

performed exceptionally well in classification with an AUC of 0.972, which was much higher than the centralized model's 0.931.

Table 4: Comparative Performance

Model Type	True Positive Rate (TPR)	False Positive Rate (FPR)	AUC	Interpretation
Centralized	0.91	0.08	0.931	Good discrimination, higher misclassification risk
Federated Model	0.96	0.04	0.972	Excellent discrimination, minimal false positives

The ROC analysis also revealed that the FL model maintained a True Positive Rate (TPR) over 0.95 while keeping the False Positive Rate (FPR) below 0.05, exhibiting strong sensitivity and specificity as shown in Figure 3 below.

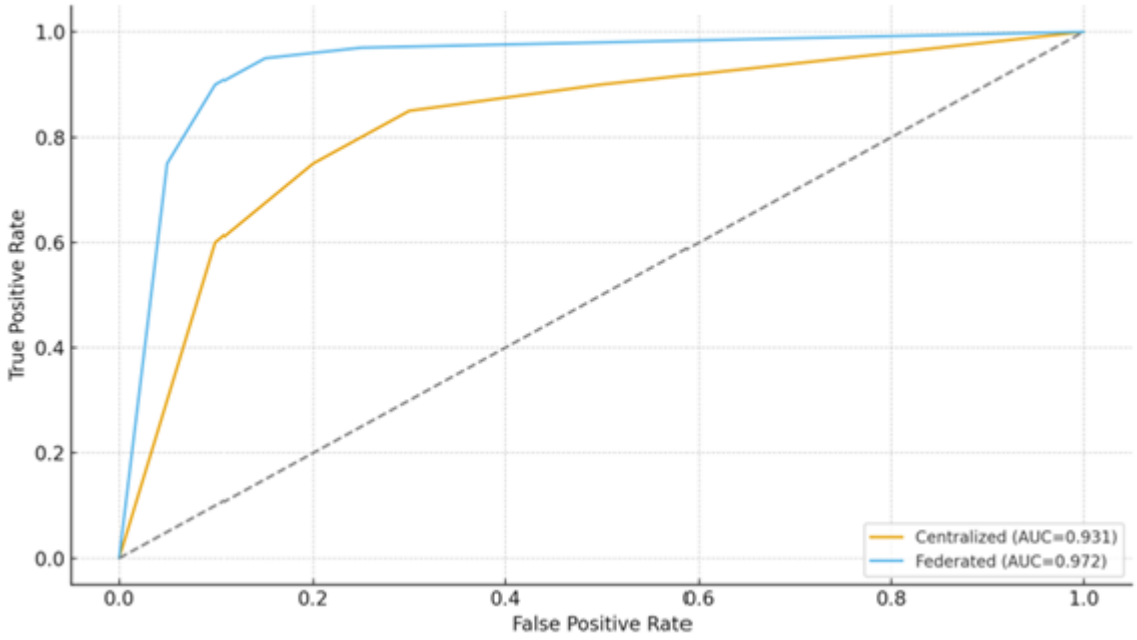


Figure 3: Comparison of ROC Curves

These results demonstrate the great efficacy of the federated model in differentiating between authentic and fraudulent certificates according to different categorization criteria.

4.4 Confusion Matrix Analysis

For both models, confusion matrices were created to offer a more in-depth understanding of categorization behavior. Table 5 below provides a summary of the findings.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information, Technology and Data Science
Volume 10||Issue 1||Page 94-117 ||August||2026|
Email: info@stratfordjournals.org ISSN: 2617-3573

Table 5: Confusion Matrix

Model Type	True Positives (TP)	False Positives (FP)	True Negatives (TN)	False Negatives (FN)
Centralized Model	452	38	426	54
Federated Model	478	22	445	25

The classification results were greatly enhanced by the FL model by:

i. A 42% decrease in false positives (38 → 22)
ii. A 53% reduction in false negatives (54 → 25)
iii. A rise in both true positives and true negatives

These enhancements show that the FL model lowers the possibility of rejecting valid certificates, which is crucial in real-world verification systems, in addition to improving fraud detection accuracy.

4.5 Privacy Preservation and Collaborative Learning

Enabling collaborative model training without providing raw institutional data was one of the FL model's main goals. The experimental configuration verified that:

i. Every participating node learned models locally
ii. Only model parameters that were encrypted were exchanged
iii. There was no exchange of raw certificate data.

This indicates that the suggested FL framework is appropriate for sensitive domains like academic credential verification since it effectively preserves data privacy while maintaining excellent model performance.

5. Discussion

The study showed that the suggested FL framework is an efficient, scalable, and privacy-preserving way to address the problem of collaborative academic credential fraud detection across institutions. The results show that the federated model outperforms the baseline centralized technique on all metrics with 94.7% accuracy and 0.972 AUC. The results show that decentralized collaborative learning can effectively increase the capabilities of solving the difficulties of fraud detection, institutional privacy and security. The main finding of the study is that the federated model outperforms the centralized strategy in terms of generalization. This improvement is due to the model being trained with numerous datasets collected from different institutions. The academic certificates were produced in several styles and formats and were equipped with a variety of security elements, embossing procedures, seals and accompanying documentation. The federated architecture offered the ability to learn from numerous institutions, leading to a more comprehensive understanding of larger fraud patterns and the learning of larger feature representations. These results are in agreement with prior works in federated learning that show that decentralized learning can improve the generative capabilities of the model without sharing data with a central server, by using variances and diversity in the data. Another significant use of

deep learning and distributed learning methodologies is the development of the federated learning framework. The CNN module of the system can identify the features of a visual counterfeit that may be extracted from any certificate image. The federated learning module can generate the global model without exposing sensitive data to the institutions. Its recall and precision scores respectively indicate how well it can identify fraud and correctly classify it. This trade-off is especially essential when it comes to academic verification systems because false negatives and false positives might be disastrous for those systems. From the confusion matrix, we can see that the model is very successful. The federated methodology reduces the false positive rate from 38 to 22 and the false negative rate from 54 to 25 compared to the centralized approach. The risk of false positives is considerable; if graduates are incorrectly identified as fraudulent, it can be devastating to the students, the school and the business. There are fewer false negatives, which means less of a likelihood of counterfeit credentials at the academic and professional levels. The modifications indicate that the federated paradigm improves the security and fairness of the credential verification system.

Another significant finding is that federated training convergence was quite stable. This was achieved with 36 rounds of conversation and the model performance was similar to the top model. This shows how the Federated Averaging (FedAvg) approach can be used to combine several updates from a large number of remote models. This means that collaborative learning among decentralized institutional nodes can be achieved at an unmanageable computational and communicational cost. The results of convergence stability further support federated learning as a favorable option for a realistic scenario where several institutions operate an extensive academic environment.

The purpose of this study was privacy protection. The results of the research reveal that the proposed framework has been key in this aspect. The information on the certifications was held locally at an institution. The certificates are encrypted with the parameters of the model only in the process of aggregation with other institutions. This solution can address certain typical challenges of the centralized fraud detection system, including fraud access, privacy protection, and regulation compliance. Data sovereignty is an important consideration in conversations with educational institutions that manage vast amounts of sensitive information, including Personally Identifiable Information (PII). The model's privacy protection was strengthened by incorporating Secure Aggregation and Differential Privacy techniques. The approaches mitigated the risk of getting important certificate information from the communicated model changes without sacrificing the models' usability and integrity. The results of this study show that it is possible to achieve privacy protection while maintaining a sufficient detection accuracy for federated learning systems. This is especially helpful for firms with limited scope for data protection and ethical concerns. This work further points out significant challenges in the context of non-IID data issues in a federated data setting. The programs may have different datasets that the institutions may offer, and data may be treated differently or assigned differently in the certificates. The framework presented is powerful and versatile enough to reform the integrity of institutions. But if the data volumes or the institutions are enormous, this stringent balance may still impact the fairness and convergence of the model. Further investigations are needed for the mentioned concerns.

The proposed structure, while having many advantages, also raises several issues that need to be investigated. The research was focused mostly on the faking of certificates and the identification of inconsistencies in image metadata. Insider attacks build and target more complicated, and presumably more sophisticated, synthetic identities or blockchain credentials. Secondly, advanced

approaches such as homomorphic encryption and secure multiparty computation offer stronger security guarantees but can incur larger communication and processing overheads, limiting their use in resource-constrained contexts. Future developments may include the integration of important innovations such as blockchain technology, explainable artificial intelligence (XAI), and complicated cryptographic methodologies inside the framework. Blockchain will allow credentials to be stored and verified securely and explainable AI can provide transparency and comprehension for fraud detection systems. The framework can also be extended with tools for cross-border institutional collaboration and on-site verification, to enable it to be used in the wider academic and professional sector.

In conclusion, the results indicate that federated learning is a potential strategy for solving certificate fraud in a learning environment with privacy and that it is effective. The framework presents a potential solution to balance essential aspects such as accuracy in detection, collaborative learning, scalability, institutional privacy, trust, and integrity in existing digital credential verification systems.

6. Conclusion

This study concludes that the academic certificate fraud detection problem between institutions results in a collaborative problem and can be solved by using the FL framework proposed, which is effective, scalable, and privacy-preserving. The results showed that the best accuracy and AUC of all the metrics are 94.7% and 0.972, respectively, with the federated model, surpassing the centralized baseline. The results suggest that collaborative learning in a decentralized manner is quite effective in enhancing the ability of fraud detection and can also solve problems concerning the institutional privacy and security of data.

The major finding of the study is that the Federated Model is better than the centralized Model in terms of generalization. This improvement is based on separate datasets from a variety of institutions that have exposed the model. While this is true of the contents of academic certificates, the format in which they're delivered, the mechanism used to secure them (e.g., with a ribbon), their embossing and engraving, and the metadata attached to each certificate can vary. The federated approach allowed the learning from data from multiple institutions to learn about a bigger picture of fraud patterns, and to create more comprehensive representations of features. The results are consistent to the previous works on federated learning, which demonstrated that decentralized learning can help enhance the generalization performance of the model without sharing the data with the central server, based on the differences and diversity of the data.

Additionally, the success of the deep learning approach and distributed learning approach in combination is the outstanding performance of the FL framework. The CNN part of the system could identify the features of a visual forgery that can be extracted from any image of a certificate, and the federated learning part could optimize the global model without giving away any sensitive information to the institutes. The recall and precision that it obtains show that it performs well on both tasks of detecting fraud and correctly categorizing the classification. Balance is crucial, and this is especially the case in academic verification systems, because of its implications in terms of false positives and false negatives.

The other evidence of the effectiveness of the model is the confusion matrix analysis. As compared to the centralized model, the federated model has lowered the false positive rate by 38% and the false negative rate by 54%. The risk of false positives is important because it could influence the graduates, institutions, and employers if they were deemed as fake graduates. Similarly, the

decreased false negative rate will lead to decreased risk of receiving false certificates subsequently in an educational and/or career context. Improvements show that with a federated model more security and fairness in credential verification.

A second important result we found was the following: Convergence was very stable in the case of federated training. Using the best model and the Federated Averaging (FedAvg) algorithm was optimum in this case after 36 rounds of communication, on the performance side, for aggregating many updates from many distributed models. This implies that the decentralized institutional nodes of knowledge co-creation can be realized reasonably in computation and communication costs. Overall, the outcomes of the convergence stability further validate the feasibility and potential of federated learning in a genuine context, in which multiple colleges will work together to attain a massive-scale academic ecosystem.

One of the main aims of this research was to maintain privacy, and the proposed framework has been instrumental in achieving this, as evidenced by the results of this research. During training, all the certificate data was stored locally; no other data was encrypted, and only during the aggregation of other institutions was it encrypted. This method helps overcome several common issues in fraud detection, like fraud access, privacy, and regulatory compliance issues, in a centralized approach. When significant amounts of sensitive information, such as information about the institution or personally identifiable information (PII) data, are contained in an institution's current information system, there is a need to be concerned with data sovereignty in any discussions between the institution and an external data steward.

In addition, two more techniques were proposed to ensure greater privacy protection for the framework, which are: Secure Aggregation and Differential Privacy. The mechanisms decreased the threat of being able to access sensitive certificate information from the transmitted model update without compromising the usefulness and validity of the models. This paper exhibits the feasibility of privacy technique transfer to further advance the Federated Learning (FL) system without hurting the detection's accuracy. Certainly, this is intriguing for organizations that are strapped for space to play with when it involves fact protection along with moral problems.

The study also shows that there are significant issues with having to deal with Non-IID data problems in a federated data environment. Institutions can allow different sets of data to be presented to the programs, and information can be manipulated in numerous ways, or presented differently within the certificates. It is strong, has flexibility, and performs well in all the different institutions. But for a large number of data volumes or institutions, it is possible that this extreme balancing may still impact the fairness and convergence of the model as well. Ongoing studies are needed on these issues.

While there were some beneficial features to the proposed framework, there were some areas of missed opportunity to be noted. The study mainly concentrated on the issues of certificate forgery and metadata anomaly detection based on images. More sophisticated and perhaps more complex synthetic identities or blockchain credentials are being created, and/or insider attacks. To begin with, more sophisticated technologies such as homomorphic encryption and Secure Multi-Party Computation (SMC) will provide better security assurances, but consume more communication and computation resources than simpler methods, thereby limiting the uses in resource-poor environments. These innovations, including the use of blockchain technology, explainable artificial intelligence (XAI), and advanced cryptographic techniques, could be incorporated into the framework in future studies. With the use of blockchain technology, providing credentials can

be secured and verified with authenticity levels that are well into the future, and explainable AI can help enhance transparency and explanation in the fraud detection process. In addition, the framework can be extended to encompass not only cross-border institutional cooperation, but also on-the-spot verification systems can be employed to expand its applicability in the academic and professional international field.

The results suggest that FL may emerge as a privacy-preserving, effective, and convenient method for identifying certificate fraud in academia. Finally, this framework seems to be a promising way to address key aspects of the challenges that need to be addressed in the era of digital credential verification, including detection accuracy, collaborative learning, scalability, institutional privacy, trust, and integrity.

References

Achhab, D., Zouhri, A., Ayache, M., & Nasser, N. (2025). TinyFLChain: Blockchain-secured decentralized federated learning for TinyML systems. Proceedings of the IEEE Global Communications Conference (GLOBECOM 2025), 5296–5301. https://doi.org/10.1109/GLOBECOM59602.2025.11432017

Aisha, A. W., & Manju, J. (2020). Merging artificial intelligence and blockchain technologies to solve academic qualification forgery issues. Journal of Student Research. https://doi.org/10.47611/jsr.vi.856

Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. IEEE Access, 8, 140699–140725. https://doi.org/10.1109/ACCESS.2020.3013541

Alghamdi, M., Mnasri, S., Hassanat, A., Arbab, M. M. S., Alkhazi, I. S., Alrowaily, M. A., Liu, C. Z., & Bloui, N. H. (2026). Federated deep blockchain-based system for secure verification of academic transcripts and matching study plans in Saudi universities. Scientific Reports, 16(1). https://doi.org/10.1038/s41598-026-43328-8

Ali, M. A., & Bhaya, W. S. (2021). Higher education's certificates model based on blockchain technology. Journal of Physics: Conference Series, 1879(2), 022091. https://doi.org/10.1088/1742-6596/1879/2/022091

Ali, O., Jaradat, A., Kulakli, A., & Abuhalimeh, A. (2021). A comparative study: Blockchain technology utilization benefits, challenges and functionalities. IEEE Access, 9, 12730–12749. https://doi.org/10.1109/ACCESS.2021.3050241

Al-Karawi, A. L. S., & Akdeniz, R. (2026). Zero-knowledge federated learning for privacy-preserving 5G authentication. Computers, 15(4), Article 206. https://doi.org/10.3390/computers15040206

Awan, S., Li, F., Luo, B., & Liu, M. (2019). Poster: A reliable and accountable privacy-preserving federated learning framework using the blockchain. In Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (pp. 2561–2563). https://doi.org/10.1145/3319535.3363256

Babu, S., Anbumozhi, A., Praveenkumar, S., Vasudevan, N., & Kaliamoorthy, V. (2026). Blockchain for credentialing and academic record-keeping. In J. Moore et al. (Eds.), Transforming Education With Data Science in the AI Era (pp. 407–448). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-2185-1.ch014

Bao, X., Su, C., Xiong, Y., Huang, W., & Hu, Y. (2019). FLChain: A blockchain for auditable federated learning with trust and incentive. In Proceedings of the 2019 IEEE International Conference on Big Data Computing and Communications (pp. 151–159). https://doi.org/10.1109/BIGCOM.2019.00030

Berrios Moya, J. A., Ayoade, J., & Uddin, M. A. (2025). A zero-knowledge proof-enabled blockchain-based academic record verification system. Sensors, 25(11), 3450. https://doi.org/10.3390/s25113450

Cheng, J., & Li, Y. (2026). Blockchain and federated learning for cross-border credential verification: A policy framework for Central Asian higher education (Version 1) [Preprint]. Preprints.org. https://doi.org/10.20944/preprints202602.0130.v1

Chiș, D., & Caramihai, M. (2023). Blockchain in higher education: A secure traceability architecture for degree verification. In D. S. Mistretta (Ed.), Reimagining Education: The Role of E-Learning, Creativity, and Technology in the Post-Pandemic Era. IntechOpen. https://doi.org/10.5772/intechopen.1001997

Gomez, D. (2021). Blockchain in education: A large and global encrypted database. eLearn Center Blog. https://elearncenter.blogs.uoc.edu/blockchain-in-education/

Grech, A., & Camilleri, A. F. (2021). Blockchain for education: A new credentialing ecosystem. OECD Publishing. https://doi.org/10.1787/589b5f1f-en

Jain, A. (2025). Blockchain-based decentralized identity systems in education. Scientific Journal of Artificial Intelligence and Blockchain Technologies, 2(4). https://doi.org/10.63345/sjaibt.v2.i4.104

Kaneriya, J., & Patel, H. (2023). A secure and privacy-preserving student credential verification system using blockchain technology. International Journal of Information and Education Technology, 13(8), 1251–1260. https://doi.org/10.18178/IJIET.2023.13.8.1927

Kottursamy, K., Sadayapillai, B., AlZubi, A. A., & Bashir, A. K. (2023). A novel blockchain architecture with mutable block and immutable transactions for enhanced scalability. Sustainable Energy Technologies and Assessments, 58, 103320. https://doi.org/10.1016/j.seta.2023.103320

Lee, C., Heiss, J., Tai, S., & Hong, J. W.-K. (2024). End-to-end verifiable decentralized federated learning. In 2024 IEEE International Conference on Blockchain and Cryptocurrency (ICBC) (pp. 434–442). IEEE. https://doi.org/10.1109/ICBC59979.2024.10634412

Li, J., Shao, Y., Wei, K., Ding, M., Ma, C., Shi, L., Han, Z., & Poor, H. V. (2022). Blockchain-assisted decentralized federated learning (BLADE-FL): Performance analysis and resource allocation. IEEE Transactions on Parallel and Distributed Systems, 33(10), 2401–2415. https://doi.org/10.1109/TPDS.2021.3138848

Liu, B., Xiao, L., Long, J., Tang, M., & Hosam, O. (2020). Secure digital certificate-based data access control scheme in blockchain. IEEE Access, 8, 91751–91760. https://doi.org/10.1109/ACCESS.2020.2993921

Lu, Y., Zhou, Y., Yang, K., Zhang, Y., Ding, M., & Guizani, M. (2025). Efficient certificate-based authentication for privacy-preserving federated learning in IIoT. IEEE Internet of Things Journal. Advance online publication. https://doi.org/10.1109/JIOT.2025.3637104

Onwubiko, D. C., Odikwa, N. H., Ukabuiro, I. K., & Agomah, S. A. (2025). Enhancing academic credential verification through blockchain technology: A decentralized approach to combat certificate fraud. Iconic Research and Engineering Journals, 8(11), 336–347.

Pawar, S. B., Shevkar, G. N., Shinde, D. R., Somase, G. A., & Lahane, P. S. (2026). CredVerify: Implementation and validation of an academic certificate verification framework using blockchain. International Journal for Research in Applied Science and Engineering Technology, 14(4), 4367–4371. https://doi.org/10.22214/ijraset.2026.80232

Qammar, A., Karim, A., Ning, H., et al. (2023). Securing federated learning with blockchain: A systematic literature review. Artificial Intelligence Review, 56, 3951–3985. https://doi.org/10.1007/s10462-022-10271-9

Qu, Y., Uddin, M. P., Gan, C., Xiang, Y., Gao, L., & Yearwood, J. (2022). Blockchain-enabled federated learning: A survey. ACM Computing Surveys, 55(4), 1–35. https://doi.org/10.1145/3524104

Ramya, N. S., Lasya, V., & Michael, A. A. (2026). Blockchain based academic certification verification system. In Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (pp. 944–955). Atlantis Press. https://doi.org/10.2991/978-94-6239-693-7_91

Rao, K. S., Shirisha, P., Navaneswar, M., Chaitanya, K., & Harsha, K. (2024). Blockchain based academic and professional credentials. International Research Journal on Advanced Engineering Hub, 2(5), 1383–1386. https://doi.org/10.47392/IRJAEH.2024.0191

Sameera, K. M., Rehiman, K. A. R., & Vinod, P. (2023). A privacy preservation framework using integration of blockchain and federated learning. SN Computer Science, 4, 703. https://doi.org/10.1007/s42979-023-02075-7

Silaghi, D. L., & Popescu, D. E. (2025). A systematic review of blockchain-based initiatives in comparison to best practices used in higher education institutions. Computers, 14(4), 141. https://doi.org/10.3390/computers14040141

Singh, A. V., Khan, A., & Jain, S. (2025). Federated learning for secure blockchain-based identity verification in decentralized systems. International Journal of Advanced Research and Multidisciplinary Trends, 2(2), 152–168. https://www.ijarmt.com/index.php/j/article/view/186

Toyoda, K., Zhao, J., Zhang, A. N. S., & Mathiopoulos, P. T. (2020). Blockchain-enabled federated learning with mechanism design. IEEE Access, 8, 219744–219756. https://doi.org/10.1109/ACCESS.2020.3043037

Wang, Z., Chen, X., Li, Y., Zhang, H., & Yu, S. (2024). A verifiable and privacy-preserving federated learning training framework. IEEE Transactions on Dependable and Secure Computing. Advance online publication. https://doi.org/10.1109/TDSC.2024.3369658

Wu, L., Ruan, W., Hu, J., & He, Y. (2023). A survey on blockchain-based federated learning. Future Internet, 15(12), 400. https://doi.org/10.3390/fi15120400

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology, 10(2), Article 12. https://doi.org/10.1145/3298981

Yang, Z., Shi, Y., Zhou, Y., Wang, Z., & Yang, K. (2023). Trustworthy federated learning via blockchain. IEEE Internet of Things Journal, 10(1), 92–109. https://doi.org/10.1109/JIOT.2022.3201117

Yu, F., Lin, H., Wang, X., Yassine, A., & Hossain, M. S. (2022). Blockchain-empowered secure federated learning system: Architecture and applications. Computer Communications, 196, 55–65. https://doi.org/10.1016/j.comcom.2022.09.008

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775. https://doi.org/10.1016/j.knosys.2021.106775

Zhang, Y., & Yu, H. (2022). Towards verifiable federated learning. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI 2022) (pp. 5686–5693). https://doi.org/10.24963/ijcai.2022/792