Federated Learning for Cross-Carrier Insurance Fraud Detection: Secure Multi-Institutional Collaboration

1st Keerthi Amistapuram
Lead Software Developer
ORCID ID : 0009-0009-6408-1958

***Abstract*—Insurance fraud affects all players—individual customers and insurers alike. Yet, insurance companies are limited in their capabilities to detect fraud. Pooling data across multiple insurance carriers may enable federated machine-learning models that improve fraud-detection performance without compromising data confidentiality and privacy. A proposed system is based on horizontally or vertically partitioned federated-learning architectures, information-protection techniques such as differential privacy, and privacy-preserving mechanisms including secure multiparty computation and homomorphic encryption. Relevant data governance controls—especially aligning with data minimization and so-called data-access principles—address the key privacy concerns in a multi-institutional collaboration environment. A more effective fraud-detection model derived from data fed by multiple carriers may significantly enhance each institution's detection capability and is important in optimally allocating law-enforcement and investigation resources. Other factors that may influence the fraud-detection capability of insurers include the absence of necessary data, too little data, data noise, mislabeling, and mislabeled samples. noch Ivory-Segatta and others proposed methods to mitigate these concerns by creating a model collaboration framework for a voice-recognition task. Their framework provided an easy-to-use, secure, and reliable environment for cross-collaboration learning between multiple parties to boost label-accuracy sampling for voice-recognition applications.**

***Index Terms*—Insurance Fraud Detection, Federated Learning, Multi-Carrier Collaboration, Data Privacy, Differential Privacy, Secure Multiparty Computation, Homomorphic Encryption, Data Governance, Privacy-Preserving Mechanisms, Fraud Analytics, Horizontal and Vertical Partitioning, Data Confidentiality, Model Collaboration Framework, Label Accuracy, Cross-Institutional Learning, Fraud Prevention, Privacy-by-Design, Data Minimization, Secure Data Sharing, Law-Enforcement Optimization.**

I. Introduction and Problem Statement

Insurance fraud persists as a significant global challenge, manifesting as values and volumes increase. Centralized machine learning models offer greater detection accuracy but struggle to overcome regulatory and institutional barriers. To bridge this gap, federated learning has recently garnered considerable attention for cross-carrier fraud detection tasks. Although the federated setting enables data minimization, specific governance and compliance measures remain essential for participating institutions, especially under strict regulations such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA). Cross-institutional collaboration also requires a governance framework that carefully delineates participation parameters, data-exchange boundaries, and synchronization constraints. Insurance fraud encompasses intentional misrepresentation or deception for monetary gain, with direct consequences for both insurers and customers. Effective fraud detections models rely on substantial, high-quality labelled datasets; however, many institutions cannot devote the necessary resources. Conditioned on the relatively small number of such models and institutional appetite for privacy preservation, a cross-carrier collaboration capability can close this data gap. Effective defence relies on reliable signals, whether historical fraud occurrences or behavioural patterns. . . but carriers often cannot share this with others. Horizontally federated learning models allow for sharing broad detection strategies while maintaining institutional privacy. Specific support from differential privacy, secure multi-party computation, homomorphic encryption, and governance frameworks creates a blueprint for cross-carrier collaboration.

A. Background on Insurance Fraud

Insurance fraud is a known challenge facing all insurance carriers. Policyholders might provide false information to obtain insurance or might exaggerate claims. In recent years, cross-carrier insurance fraud has been rising. For example, a car insurance policyholder may collude with a friend who is the policyholder of a health insurance policy and whose identity has been stolen to claim additional benefits on a fictitious accident. Due to the nature of cross-carrier insurance fraud, a centralized model for all insurance companies may not be possible without cross-carrier data-sharing agreement because such insurance companies usually do not want to give other insurance companies access to their raw data for any analysis. Nevertheless, a cross-carrier insurance fraud-detection analysis

Fig. 1. Cross-Carrier Insurance Fraud Detection: Federated Encryption and Advanced Algorithms

could reveal many useful insights. A federated-signature algorithm can be applied to encrypt data such that medical records and health details of an individual cannot be decrypted by any insurance company alone. Thus, a cross-carrier insurance-fraud-detection model from individual carriers while preserving the raw data privacy and confidentiality data can be implemented. A few state-of-the-art algorithms can prevent cross-carrier fraud detection, which demand knowledge on how to augment samples so that models can obtain enhanced performance. Samples for one insurance company may not be enough to build cross-carrier fraud models, as insurance companies are located all over the world, leading to a non-synchronous distribution among the insurance companies. Therefore, techniques such as transfer learning, domain adaptation, or label projection can be used to detect if these tools can provide extra information for fraud detection. With the appropriate ethical scrutiny and safeguards in place, the insights derived from the information asymmetry could result in more accurate detection of cross-border frauds—highlighting a rich area for future research.

B. Limitations of Traditional Centralized Models

Insurance fraud detection has long been reliant on individual carriers using centralized detection models to combat fraud. Models can be hosted at one or several carriers, with all collaborating for federated training, but without sharing data. However, many models rely on data that is often not available for the policyholder of the suspected fraudulent claim (the victim) when it comes to susceptibility detection, including on-device and historical behavioral data, transaction data, etc. Insurance is not a closed ecosystem, with claims involving actors across different carriers. Therefore, detection models might yield better results if the data contributing to the solution from other carriers could be leveraged. Despite federated learning's ability to combine training data from different institutions without the need to share the real data, thus addressing privacy, regulatory, and security concerns, the framework was not well suited for this use case. The primary question was one of model training and collaboration between institutions, since the data signal for the task is often not available for the victim—the identity of which changes by claim and may not be served by any carriers at the time of fraud detection. Participating in a truly diffused collaborative digression via FL without an actor-centric structure is however also a challenging task with a detection model.

C. Overview of Federated Learning for Cross-Carrier Collaboration

Federated Learning (FL) enables data-holding entities to jointly learn a shared model while keeping the training data local. FL aligns with the horizontally and vertically distributed nature of datasets across insurance carriers, as well as the need for regulatory compliance in a multi-entity and potentially cross-border environment. However, the collaborative protocols require careful design to address the privacy of participating carriers, as well as fairness, incentives, and risks associated with malicious behaviors. Existing technical and business frameworks in banking and telecommunications can thus help shape the design of a federated setup for insurance carriers. Insurance fraud has a significant impact on the entire industry, and its prevention requires a cross-carrier collaborative approach. The feasibility of federated learning (FL) in a cross-carrier collaboration scenario has already been demonstrated, along with an analysis of FL's privacy-preserving techniques and their impact on the design of fraud-detection models. Building upon these foundations, the focus now turns to implementing a secure, multi-institutional collaboration framework. The organization of this work follows the necessary cross-carrier collaboration protocols, starting with a description of how participating institutions interact when collaborating, followed by a discussion of the procedures for aggregating and updating a federated model.

II. Federated Learning Fundamentals for Industry Data

Federated Learning (FL) enables multiple data owners to collaboratively train a shared model while keeping their data localized. Defined by the assembly of a federated model server on the cloud and a number of federated model clients distributed close to the data sources, FL can facilitate three common deployment topologies—Horizontal, Vertical, and Hybrid FL. Horizontal FL is applicable when multiple institutions possess similar feature sets across diverse clients, and is the primary focus of the present work. Vertical FL can be employed when data owned by different institutions cover disjoint attribute spaces but share overlapping entities.

Hybrid FL is a combination of Horizontal and Vertical FL settings. FL integrates advanced cryptography and security technology for privacy protection. Institutions within the federated network can exchange model parameters without direct access to raw data, and multiple privacy-preserving techniques such as Differential Privacy (DP), Homomorphic Encryption (HE), and Secure Multiparty Computation (SMPC) can be leveraged to address various privacy challenges. By eliminating data exchange and risk of leakage for individual data, FL ensures that the sensitive historical data of individuals or companies remain secured. Although model training and issuing data are distributed among different institutions, a centralized device is still required for hardware and software settings and environment maintenance. FL can effectively deal with system heterogeneity to a certain extent, and the communication cost of model parameter transmission is usually smaller than the huge volume of original data transmission required for centralized model training.

Equation 1 — Federated aggregation (FedAvg with sample weighting)

Setup. Carriers/clients $i = 1, \ldots, N$ each hold local data D_i of size n_i (total $n = \sum_i n_i$). Let w_t be the global model at round t.

Local update (one or more SGD epochs):

$$\text{For each } i: \; w_{i,t+1} = w_t - \eta_t \sum_{(x,y)\in \mathsf{D}_i} \nabla_w \ell(w_t; x, y)/|\mathsf{B}_i| \quad (1)$$

where η_t is the learning rate and B_i indicates mini-batching.

Server aggregation (sample-weighted average):

$$w^{t+1} = \sum_{i=1}^{N} \frac{n_i}{n} w^{i,t+1} \quad (2)$$

Equivalently in **update** form with local gradients $g_{it} = w_t - w_{i,t+1}$:

$$w_{t+1} = w_t - \sum_{i=1}^{N} \frac{n_i}{n} g_{it} . \quad (3)$$

Symbol Reference (Federated Insurance Fraud Detection)

Symbol	Meaning
N	Number of participating carriers / clients
D_i	Local dataset at carrier i
n_i	Number of samples in D_i
w_t	Global model parameters at round t
w_i^{t+1}	Local model parameters after client i trains at round t
g_i^t	Local gradient/weight update computed by client i at round t
\tilde{g}_i^t	Noised (DP) update from client i at round t
\mathcal{A}	Secure aggregation operator (e.g., HE/SMPC)
\eta_t	Learning rate at round t
\sigma	Standard deviation of DP noise
\epsilon, \delta	Differential privacy parameters
\ell(\cdot,\cdot)	Per-example loss (e.g., logistic loss)
p(y=1|x)	Fraud probability for features x
\lambda	Regularization strength
\alpha_i	Participation weight for carrier i (fairness/incentive term)

A. Federated Learning Architectures (Horizontal, Vertical, and Hybrid)

Data distribution in federated learning can be typically Horizontal, Vertical, or Hybrid in nature. However, in the discussed multi-carrier insurance fraud detection scenario, the parties operate on a Horizontal Federated Learning (HFL) architecture. Horizontal FL is adopted when heterogeneous institutions, such as different insurance carriers from various countries, face the same problem and are targeting fraud detection with a small number of labelers. These entities typically have access to different users for the same period. Users move across different carriers, which can create negative behavioral patterns. These patterns are signals indicating economic damage for at least one of the carriers. The possibly negligible users who have been highlighted as abusive by one institution trigger interesting transactions in other institutions. While user identity is not exchanged in any form, intermediate features that are sufficient to allow checks and balances remain. In this FL architecture designed for a multi-institution fraud detection project, the participating institutions might be hostile or untrustworthy, but GMV produces a model without a visible single point of failure. Insurance regulation does not permit an institution to have no control over sensitive patient information. There is always the possibility of compromise or abuse. Therefore, the carrier's information policy must be carefully configured, creating a federated environment that not only complies with the anti-money-laundering regulation but also provides a consequent collusion detection mechanism. The organization remains control of the information, receiving insight derived from data location, and the ability to carry out more targeted interaction and audits. The transfer of sensitive data is limited, and the insurance database as a whole is never exposed.

B. Security and Privacy Mechanisms (DP, MPC, HE)

Data privacy and security issues affect the establishment of a cross-carrier collaboration framework for fraud detection during online insurance claims. Secure multi-party computation, differential privacy, and homomorphic encryption are suitable techniques for addressing data privacy and security and can be integrated into a federated learning model and protocol. Although introducing these techniques incurs additional costs in accuracy, latency, and computational overhead, how they optimize concurrent

multi-carrier collaboration toward online insurance fraud detection supersedes their isolated usage. Insurance carriers should adopt data minimization and access control policies to reduce potential harm to customers and third parties that may be derived from data misuse, unauthorized access, or disclosure. At the same time, the set of data shared must be adapted to the expectations of regulators and customers, as well as to sensitivity-aware requirements originated in the field of health care. For example, insurance companies in Europe need to manage users' sensitive data to comply with local regulations such as GDPR. In particular, whenever special categories of personal data are processed, explicit consent is required unless it falls within some exceptions. GDPR is not the only regulation that has a direct impact on the sharing of customers' data since products in the insurance and health sectors are often cross-border. Thus, organizations should consider evaluating other data privacy laws and regulations in the countries where they operate.there are also principles and methodologies in other fields (e.g., health care) that can guide the design of cross-carrier data governance and compliance processes.

C. Communication-Efficiency and System Heterogeneity

Cross-carrier insurance fraud detection pursues a horizontally federated setup since participating carriers provide freely available labels for legitimate claim requests. Nevertheless, features that signal fraudulent claims should be engineered along the lines of vertically federated settings, as cross-carrier feature composition enables a more accurate utility-privacy trade-off. Communication cost pragmatism motivates investing in communication-efficient protocols that compress transmitted data size; nevertheless, this requirement also requires testing the robustness and fault tolerance of the chosen communication protocols. In particular, the cross-carrier FL participant model should be compatible with stratified sampling strategies during the model-update aggregation stage to avoid high-model-latency settings in the setup of smaller carriers participating in error-free data-collection periods. System heterogeneity—e.g. in data size, feature representativeness, and party capabilities—parasitically affects cross-carrier FL performance and thus necessitates careful monitoring during the collaboration, as poorly designed aggregation strategy exacerbate neural constituent imbalance.

III. Data Governance and Compliance Across Carriers

Security measures protect the customer data shared among the participants without serious overheads, fulfilling five criteria required for cross-carrier collaboration. First, all parties should avoid sharing additional sensitive data. Second, the data that any party divulges must not be illegitimately exploited or fall into the wrong hands. Third, distribution of the data should be regulated to prevent it from being used inappropriately. Fourth, a system for consent across multiple companies is essential for using the data legally. Fifth, an auditing system that keeps everyone informed and manages the risk of legal action must be established. Minimizing the amount of data collected helps meet the first requirement. In the data-sharing process, participants must strive to divulge as little sensitive information as possible. Naturally, sensitive variables included in personal insurance contracts, such as gender, profession, health status, and claim amounts, will not be distributed. Furthermore, when collaborating with carriers in different countries, the rules of the General Data Protection Regulation (GDPR), as well as, for example, Health Insurance Portability and Accountability Act (HIPAA)-like norms, should also be considered. These regulations distinguish privacy-sensitive features from nonprivacy-sensitive features. Consent is needed to exchange sensitive features such as age, occupation, gender, and health history. A lack of consent will not prevent the sharing of physical location, behavior history, device characteristics, and other relative information because population behavior has no obvious privacy concerns. Given the previous discussions, sensitive features should have the consent module added.

Equation 2 — Secure gradient/update with clipping, DP noise, and secure aggregation

(a) Per-client clipping (bound sensitivity):

$$g^{it} = \text{clip}(g_{it}, C) = g_{it} \cdot \min\left(1, \frac{C}{\|g_{it}\|_2}\right), \quad (4)$$

with clipping norm C.

(b) Add DP noise (Gaussian mechanism):

$$g_{it} = g^{it} + \mathcal{N}\left(0, \sigma^2 C^2 I\right), \quad (5)$$

where σ is chosen to meet (ϵ, δ)-DP under your accounting method (e.g., RDP/moments accountant).

(c) Encrypt / secret-share and transmit: client sends $E(\tilde{g}_{it})$.

(d) Secure aggregation (HE/SMPC operator ΣA):

$$A(E(g_{it})_{i \in St}) = E\left(\sum_{i \in St} n_{St}\, g_{it}\right), \quad (6)$$

where St is the sampled set of participants and $n_{St} = \sum_{i \in St} n_i$.

$$w_{t+1} = w_t - \eta^t \sum_{i \in St} \frac{n_i}{n_{St}} \tilde{g}_{it} \quad (7)$$

A. Data Minimization and Access Controls

Fraud is a high-risk problem for insurance companies. Currently, there is no efficient and effective solution to prevent fraud across different insurance enterprises or different insurance types. A multi-institutional federated

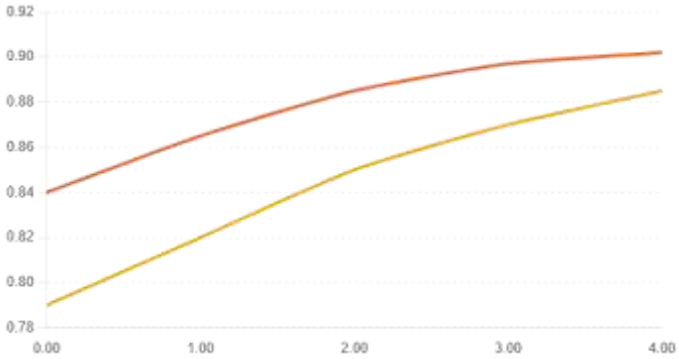

Fig. 2. Privacy–Utility Trade-off (Illustrative)

learning framework for insurance fraud prevention is proposed that enables communication and collaboration between institutions without violating data privacy. Insurance fraud detection can benefit from device linkage and cross-carrier behavior analysis. However, federated training across carriers is more challenging than training within a single carrier. Users' labels, involvement in the model and contribution to data remain crucial considerations, and secure aggregation, label-signal interaction and selection fairness must be carefully designed. Communication and model update are particularly important because of participation disparities and label delays. Insurance fraud is defined as the act of deceiving an insurance company for personal benefit, and its prevention is of vital importance in the insurance field. There are many types of insurance fraud, involving different institutions and carriers—accident fraud, commercial fraud, claims fraud, and so on. In traditional methods, fraudulent behavior is recognized without the help of other carriers, which makes it difficult to cover all fraud types and detect device-coupled fraud behavior. Fraud detection and prevention can benefit from device linkage and device-level fraud detection across carriers.

B. Regulatory Considerations (GDPR, HIPAA analogues)

Consideration of regulatory requirements is particularly relevant for an insurance fraud detection system that involves fraud prediction driven by sensitive personal data of policyholders. The General Data Protection Regulation (GDPR) principles of data minimization and access controls are particularly relevant here. GDPR requires all organizations to limit the collection of personal information to only that which is necessary for a specific legitimate purpose and involves access controls to restrict access to personal information by unauthorized users. However, multi-institution insurance fraud detection systems inherently involve the processing of personal information by a number of organizations. For such systems, data minimization ensures that each organization participating in multi-institution insurance fraud detection systems only has access to the bare minimum. At the regulatory level,

Fig. 3. Regulatory Compliance in Insurance Fraud Detection: GDPR, HIPAA, and Data Minimization

the Health Insurance Portability and Accountability Act (HIPAA) is probably the closest analogue for insurance data. HIPAA is a United States legislation that creates national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge. HIPAA outlines and mandates secure processes, protocols, and procedures for health care organizations. The same requirements can be set up for a multi-institution insurance fraud detection system, laying down rules regarding data sharing among participating institutions, the sensitive nature of the data, and the need for a legal basis for processing according to the items in GDPR. With GDPR widely adopted across Europe and parts of Asia and the HIPAA like requirements possible for insurance companies in other jurisdictions, the flow of sensitive personal data for machine learning can be made legally compliant.

C. Consent, Auditing, and Legal Risk Management

The consent processes of most jurisdictions adopt the principle of data ownership by the data subject and require clear explanations of data usage. In multi-institution scenarios involving multiple data owners, data access for fraud signal generation within a certain scope, and even features necessary for monolithic detection, is minimized through carefully designed consent and access control. When system owners grant external partners access to their sensitive data, the protection of data subjects has to be highlighted. A good example is the Health Insurance Portability and Accountability Act (HIPAA) in the USA,

which regulates the privacy of sensitive information by disclosure permission, enables data transferability, and reduces the restriction on pre-existing medical conditions when changing health insurance. Similar restrictions exist in the insurance domain of Europe, e.g., the General Data Protection Regulation (GDPR), which defines how personal data can be collected and processed and enables individuals to gain access to their data. With the quick development of fraud detection, mitigation and control, and the rapid revolution of AI technology, a fraud AIFC pipeline emerges, a risk-in-control approach can mitigate the auditing and legal risk. A risk audit committee can periodically audit the system and rectify identified issues of generalised data as well as usage policy. These discussions indicate the pressing need for any multi-party business to balance normal practices of consent, usage, and control with privacy of all participating parties, especially in cross-border insurance activities.

IV. Fraud Detection Modeling and Features

Although multiple insurers have suggested using similar approaches and analyzing the same type of features, their data are typically kept separate, making supervised training difficult. Thus, the most relevant fraud-related Signals may be considered for semi-supervised learning to provide label information for these models. Given the various fraud discovery tasks across different insurers (finding user behavior patterns, detecting high-risk users, and identifying blacklisted accounts), multi-task and transfer learning—along with data and label sharing across partners—may enhance the models' performance. The feature Signal that can be shared to support cross-carrier fraud detection relates to customer behavior patterns. In particular, behavioral Signals are more actionable than temporal Signals. While the distribution differences of active Signal patterns among carriers are quite distinct, the concluded detection models are still deniable learning-acceptable for different carriers at user-device–indexing levels—making reliable extrapolation and Tracking abilities available to avoid Service attack patterns. Similar fundamental patterns can also be detected for other devices, given that different devices mainly used for Payment transactions may exhibit different Labeling patterns. Detecting potential fraud relying on sparse and tailored fraud patterns provides less label information for training. The sparse fraud targets converging at high-global-risk areas such as Listing, Buying, and Selling Signals allow for introducing a very simple Logistic Regression detection model. Logistic Regression provides a far higher visual explanation degree than model types such as DNN and GCN. In addition, the temporal relationship between devices, behaviors, and Service Coverage also can form Label Information for Device-Wise Anti-Fraud Detection. Multi-Feature Graph Convolutional Networks built without a Fraud coverage Label might find some anomalous behavior patterns and

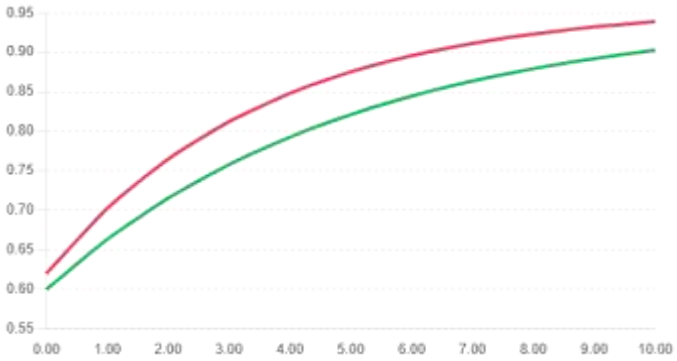

Fig. 4. Convergence Under Heterogeneous Carriers (Illustrative)

hotspots, even Edge devices. Nevertheless, these edges are usually ignored by the SPs for further action, including Labeling; advanced Fraud coverage Labels would help to improve the detection, at least for available Label area regions. Moreover, when the applied Applications or Cycles are identified, Advanced features based on the same device for authentication and Protection should be added to reliability reinforcement.

Equation 3 — Cross-carrier fraud probability (logistic model)

Let $x \in \mathbb{R}^d$ be engineered features (behavioral/device/support signals), and $y \in \{0, 1\}$ the fraud label.

Probability model:

$$p_\theta(y = 1 \mid x) = \sigma(\theta^\top x) = \frac{1}{1 + e^{-\theta^\top x}} \tag{8}$$

Regularized empirical risk (local):

$$L_i(\theta) = n_{i1}(x, y) \in D_i \quad [-y \log p_\theta - (1-y) \log(1-p_\theta)] + 2\lambda \|\theta\|^2$$

Global target: minimize $L(\theta) = \sum_{i=1}^{n} L_i(\theta)$ via the secure DP-FL update in Eq. 2. (9)

A. *Feature Engineering for Fraud Signals*

Insurance carriers possess valuable data on the normal behavioral patterns of their users, including information from insurance applications, claims, claims qualification, and support records. Such records help determine whether fraud is occurring in the insurance context. Information pertinent to identifying potential frauds that can be shared by carriers, across a federated learning (FL) setting, typically includes user behavioral patterns over mobile and Internet devices, device signals (traces), and support center records. In a federated learning setting, a user's device signatures and behavioral patterns can be utilized to detect fraudulent activities through their traversal features (anomalous non-temporal sequences). These features can serve as relevant signals in determining fraud from horizontally-distributed sources. Behaviors,

including negative patterns over a short period (even sequence anomalies over a set period), can be indications of insurance fraud. The support team's interaction with users represents another signal. Records of legitimate behavior patterns across front-ending systems can be used to model and encode the distribution signature of benign users toward superior support teams. In addition, claims data from the insured devices can be examined over time to analyze whether these users are making abnormal claims; making claims regularly/infrequently; requesting claims/items or records from the support team intentionally, unintentionally, or suspiciously; or cyclical claims over time through different devices.

B. Model Architectures Suitable for FL (Logistic Regression, DNNs, Graph Models)

Insurance fraud detection using FL typically relies on FL architectures like generalized logistic regression or distributed DNNs. Stand-alone and ensemble models such as XGBoost can also be employed in a non-FL format. However, the setup is inherently cross-carrier-aware. Constant factors in the model should ideally remain similar in scale and distribution across carriers to ensure better convergence with fewer rounds. Significant differences in feature distributions may lead to divergence, but stratified sampling during training can help mitigate this issue. While the label distribution might not reveal the same signalling features, some features remain crucial for discrimination. For instance, the claim amount acts as a signal for both prediction pairs. Multiple carriers may be required to train a model on rare labels. Labeling delays further complicate the scenario, which can be addressed by a dedicated model for fraud detection under risk, typically with a deluge of negative samples, allowing the model to learn the characteristics of clean samples.

C. Labeling Strategies and Ground Truth Challenges

The complexity of insurance fraud detection often stems from the difficulty of creating an effective signal for labeling. However, when multiple carriers form a protocol for secure model training, a decentralized yet efficient labeling strategy can be established, serving as a basis for fraud signal training. A signal may require a set of rules designed by a legal or forensic expert and validated by law enforcement in the jurisdiction covered by the participating carriers. Such a ruleset would also be essential when monitoring model performance. Recent developments in clear-text data detection models may also help. However, extra care is needed due to possible information leakage from one signal across the participation space. Another way of generating training and validation labels is a semi-automatic method, which collects and computes possible fraud scores for each claim. These scores need to be postponed several months or years after the claim approval and payment, thus slowing down the data exchange in a FL cross-carrier protocol or loop. Moreover, cross-institutional samples with fraud flags should also be noted. Nonetheless, the general cost of labeling should be lower than in a centralized solution using different ML techniques.

V. Secure Multi-Institutional Collaboration

Cross-carrier collaboration relies on clearly defined protocols across three dimensions: 1) participation criteria for the cross-carrier consortium, 2) boundaries on information exchange among participants beyond the model parameters themselves, and 3) synchronization protocols on the frequency of parameter information exchange among participants 9.2. Aggregation and model update protocols govern how the weight updates received from participants are aggregated to obtain a new global model, as well as the mechanism for determining the contributing parties, the method used to sample participants for the aggregation and the fault tolerance of the process. These protocols are tightly intertwined with the privacy-preserving techniques deployed in federated learning and their concrete instantiations have an impact on the choice of such techniques. 9.3. Developing a federated learning solution for cross-carrier fraud detection goes beyond the technologies involved and requires considering trust, dynamism and participation incentives during the collaboration. Most importantly, fairness concerns must be addressed to avoid limiting the scope of the cross-carrier collaboration. Participation fairness guarantees, at all times, that an insurer cannot be excluded from model training based on its decision to join or leave the collaboration. Data contribution incentives are designed to encourage participation in the cross-carrier training process without relying on the availability of an external authority, and are based on the understanding that the training data of one insurer can be used by the other participants in the model updates. Finally, anti-collusion measures are incorporated to mitigate the risk of information leakage associated with the fact that the parties can exchange the labels in the training data freely.

A. Cross-Carrier Collaboration Protocols

To facilitate cross-carrier cooperation, several protocols outline participation conditions, define data-sharing boundaries, and establish synchronization requirements. First, participants must share data sources and be willing to run machine-learning models on their devices, while being open to using the final optimized model. Second, partnering insurance companies should identify exactly which features can be shared with partners, and set conditions on when this information can be disclosed. The list of features and categories must be final. If key features of the model are disclosed, the model has limited protection against loss. Third, stratified sampling techniques guarantee enough samples from each category for optimal model performance; if any category has too few samples, then the model collapses. The Secure aggregator should control aggregating timing and

Carrier	Samples	Fairness-Adjusted Weight
A	50000	0.49
B	30000	0.294
C	10000	0.098
D	7000	0.069
E	3000	0.049

monitor each institution's contribution. Institutions that contribute less than expected are likely attackers, and better fault-tolerant sampling methods can recover model quality. Ensure that StratifiedDPM sampling considers sampled DPM ratio P INPUT, to maintain quality when critical data are very small.

Equation 4 — Collaborative convergence (heterogeneity-aware sketch)

Under standard smoothness assumptions and unbiased gradient noise, a common FedAvg-style bound is

$$E[F(w_{t+1}) - F(w^*)] \leq (1 - \eta_t \mu) E[F(w_t) - F(w^*)] + \eta_t^2 \quad (10)$$

(stochastic noise σ_g^2 + client drift Γ^2), (11)

where F is the global risk, μ is strong convexity, σ_g^2 models SGD variance, and Γ captures **data heterogeneity / client drift** that stratified sampling seeks to reduce. With decreasing η_t the error term vanishes (convex case), motivating the paper's emphasis on stratified participation.

Carrier Participation & Fairness Weights (Illustrative)

B. Aggregation and Model Update Protocols

Secure aggregation schemes ensure that only the aggregated output is revealed, not the individual updates. A prominent approach involves workers encrypting their updates with a additive homomorphic encryption scheme in such a way that the secret key is shared among the server and the remaining workers. During aggregation, the server combines the received partial results with its own encrypted output, enabling it to compute the aggregate without learning any individual updates. In contrast to standard homomorphic encryption schemes, the encryption function must not leak information that allows a worker to forge encrypted values using the shared secret key. This condition is naturally satisfied for multiplicative homomorphic encryption but needs careful treatment for additive homomorphic schemes. In addition to security, the above aggregation process introduces significant communication overhead, potentially slowing down training. Stratified sampling is often adopted to preserve communication efficiency. Normally, workers are assigned different subgroups of the overall dataset for training, and only the participants of a certain subgroup can perform a weight update in the same round; workers outside the subgroup receive the latest model and perform no computations. Latency, scalability, and fault tolerance are core issues in secure FL systems, particularly in practical settings with many parties. The protocols above are designed for multiple rounds of computation and take into account potential server and worker failures.

Fig. 5. Cross-Carrier Collaboration in Insurance: Trust, Fairness, and Anti-Collusion Mechanisms

C. Trust, Incentives, and Fairness Among Participants

Insurers face a challenge related to the fair distribution of trust among parties engaged in collaborative project development. Five main aspects promote trust among parties: (1) the block-size fairness of participation encompasses various influences on reliability, applicability, desirability, and operational feasibility; (2) fairness of distributed data contribution encourages valuable sharing; (3) low risk of collusion among insurers contributes towards promoting interest in participating; (4) discharge law mechanisms mitigate collusion; (5) an equitable reputation mechanism influences the reliable and trustworthy implementations through a deposit-based approach. Pseudocode visualizes the whole cross-carrier collaboration process. The following parts indicate the procedure for collaboration among multiple carriers in the cross-carrier setting. First, the participants agree on the scope, protocol, and strategies of the collaborative model, especially the involved institutions. Second, they examine the data contained by all participants and agree on data exchange types. The third part ensures that the involved participants synchronize Gradients.

VI. Privacy-Preserving Techniques and Security Risks

Cross-carrier insurance fraud detection exemplifies a federated learning scenario in which security and privacy are paramount. Insurance data are subject to stringent regulations analogous to GDPR and often contain sensitive information (e.g., health records) governed by HIPAA-like laws. Even participation in the detection effort can be sensitive in these circumstances — any connection between entities' models may lead to the leakage of private information concerning customers or fraud rings with links to multiple carriers. Consequently, the design must treat privacy and security as first-class citizens. Differential privacy (DP) is one of the strongest defenses against privacy leakage and is one of the few approaches able to protect against so-called "inference attacks," where an adversary exploits a third party's learned model to infer details about its own data. Nonetheless, a model preserving DP requires binary no-free-lunch decision-making if signals are to be effective toward distinct fraud types with potential high correlations across GTs. Although provably secure methods from secure multiparty computation (MPC) and homomorphic encryption (HE) can be employed for the secure aggregation of local model updates on a central server, these generally suffer from excessive latency, which it is desirable to avoid in production. The mix of horizontality and verticality in the cross-carrier detection scenario creates a custom generalization of DeepSecure that incorporates a novel approach for efficiently and securely aggregating vertically-shared data. Cross-carrier detection is also vulnerable to typical data-poisoning behavior that disrupts model quality and to leaking sensitive information about the local data of non-participating entities into the final model.

Equation 5 — Privacy-preserving loss (objective perturbation view)

An equivalent way to see DP's effect is **objective perturbation:**

$$L(\theta) = L(\theta) + 2\lambda\|\theta\|^2 + \text{DP perturbation } b^{\top}\theta,\ b \sim \mathcal{N}(0, \tau^2 I) \quad (12)$$

or, in **gradient perturbation** form during training, the DP noise in Eq. 2 induces

$$E[\nabla L] = \nabla L,\ Var[\nabla L] = \sigma^2 C^2 I, \quad (13)$$

causing the classic privacy–utility trade-off you report.

A. Differential Privacy for Fraud Signals

Police expect citizens to report crimes to enable a successful investigation. However, insurance fraud detection via Centralized Machine Learning (CML) systems suffers from an inherent data communication problem—participants in a CML system may not possess the exact information needed to capture a fraud signal across

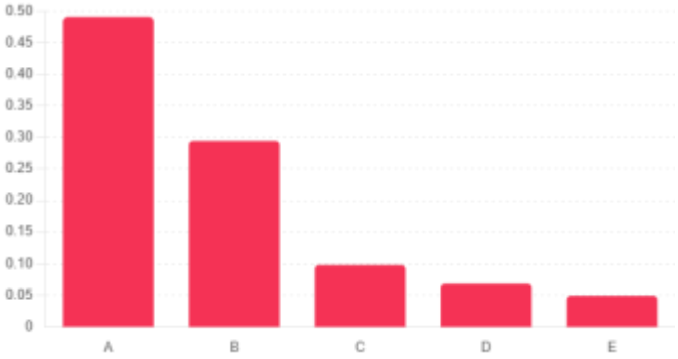

Fig. 6. Fairness-Adjusted Aggregation Weights (Illustrative)

entities. For instance, different types of insurance products may produce distinct patterns or signals in the corresponding data; one insurance type may not possess enough "data-rich" patterns, yet others with a wealth of samples may be "data-poor." Although an institution may possess a specific combination of signals that can identify fraud, it cannot do so without access to others with complementary signals. Thus, two requirements for specific confirmatory testing exist: (1) the existence of a fraud signal for the specific coverage through supervised ML, and (2) a necessary and sufficient condition satisfied by a critical combination of signals between different institutions or companies. Differentially Private (DP) designs encourage participating parties to release sensitive information without compromising individual privacy. Insurance companies must keep client information confidential yet desire the best fraud protection. A DP approach may test utility at each layer of the ML pipeline and gauge the trade-off applied to the dataset. Conducting an initial test for each feature in a specific coverage may determine whether DP is warranted. If so, a DP design should be entered into the pipeline to quantify the accuracy drop from utility loss. Each stage of the ML pipeline must validate the signal and ascertain whether it is viable for integration with DP.

B. Secure Multiparty Computation and Homomorphic Encryption Trade-offs

Federated learning relies on secure multiparty computation (SMPC) or homomorphic encryption (HE) to protect the central server from potential exposure of either model weights or gradients. These cryptographic tools complement other security measures by safeguarding the quality of aggregated information, enabling clients to exchange strategic or sensitive business information without fear of being observed, and mitigating the risk of collusion by trusted but curious participants. Yet, they also incur performance overhead, especially latency, and may constrain attention budgets. As a result, the choice between SMPC and HE in real-world applications ultimately hinges on a trade-off between privacy protection and performance. During

ε	TP	TN	Accuracy
0.5	430	1480	0.9095238095238095
1.0	445	1490	0.9170616113744076
2.0	462	1497	0.927117841930904
4.0	471	1502	0.9297832233741753
8.0	474	1504	0.931261770244821

communication, one participant is usually responsible for deploying and managing the communication server, while others send encrypted messages and receive decrypted updates. In some scenarios, these roles are defined per training global epoch, thereby enhancing fault tolerance. Stratified sampling on the client device may also be employed to ensure that only the advanced data cluster based on timing, location, and feature value is communicated for training.

Equation 6 — Detection accuracy & related metrics for class imbalance

Given counts TP,FP,TN,FN at threshold τ:

$$Accuracy(\tau) = TP + FP + TN + FNTP + TN, \quad (14)$$

$$Precision = TP + FPTP, \quad (15)$$

$$Recall = TP + FNTP. \quad (16)$$

For **AUC**, integrate TPR vs FPR over τ . In FL experiments, report curves over communication rounds and over privacy budgets ϵ to visualize the trade-offs emphasized in the draft.

Confusion Matrices Across Privacy Levels Illustrative

C. Threat Models and Attack Mitigations (Poisoning, Inference)

The cross-carrier setup carries several security risks related to the data-sharing system. If the data-sharing mechanism is not properly protected, malicious parties can utilize the information to attack a specific insurance company or influence the joint model training. Threats can be classified into three common attack types: betrayal, poisoning, and inference attacks. Betrayal attacks refer to participants that are dishonest during model training or testing. Such threats can be addressed by establishing a trusted third party who can protect the private data parts. However, collaboration remains a critical factor in these scenarios. Poisoning attacks refer to malicious institutions that attempt to degrade the model performance via data-sharing patterns. Secure model aggregation approaches such as EdgeDP-AD must be utilized. Secure aggregation contractors can be offered maintenance and training, and explore secure aggregation techniques in the context of data-hungry FL frameworks operating on horizontally partitioned datasets with possible label corruption and absence. The solutions achieve strong model utility through a validated positive trade-off between the privacy budget and the edge-community privacy utility. Security frameworks are supported for data-collaboration scenarios among multiple participants using the FL paradigm. Such security frameworks deal with data bribes in an insurance-related trust-based collaboration using the secure multi-party computation, and examine the problem of model updating in FL. Inference attacks entail unauthorized institutions leveraging auxiliary information to extract sensitive training data attributes of other institutions, making it possible to predict the training data issued by others. Employing differential privacy when distributing auxiliary signals is an effective safeguard against inference attacks. Given the aforementioned challenges and relevant defense mechanisms, establishing an incentivized trust reinforcement mechanism is critical for an effective and credible FL platform.

VII. Conclusion

Despite a recent surge in reported fraudulent claims—as substantial as rising economic losses—the evident systematic limitations of existing fraud detection models restrict their usefulness. To alleviate this impediment, FL has been proposed to federate models across carriers, thereby providing a wider detection horizon that is especially suited to the characteristics of insurance fraud and offers substantial potential without transgressing data privacy. Nevertheless, FL models are difficult to design, combine, train, and deploy, as these processes require comorbidity preparedness, careful selection of participating carriers, cross-institutional communications, adequate sharing of model updates, the availability of infrastructure and technology, maintenance of mutual trust, and the establishment of sufficient incentives for fraud detection. As a consequence, these processes have not been discussed in the insurance context, mitigating the possibility of real-world cross-carrier fraud detection employing FL. By systematically addressing these practical requirements, a practically useful and complete cross-carrier FL collaboration framework for insurance fraud detection has been presented. Beyond current applications and directions, the cross-carrier FL paradigm holds further potential for the development of an anti-fraud ecosystem in the insurance industry. For instance, the willingness of participating carriers—especially financial regulatory entities—to share and keep data open can strengthen the overall prediction capacity, fortify FL label/ground-truth transition through the use of relationships and representations, and avoid labeling delays. Improved endorsement mechanisms can further push the development of industry anti-fraud alliances at all levels. By creating model update sharing platforms among third-party technical service providers, multi-party donations can be stimulated, advancing an honest-threshold anti-fraud model. From a broader perspective, only with the coordinated efforts of regulators, legislators, and judicial organs—dynamically shaping reg-

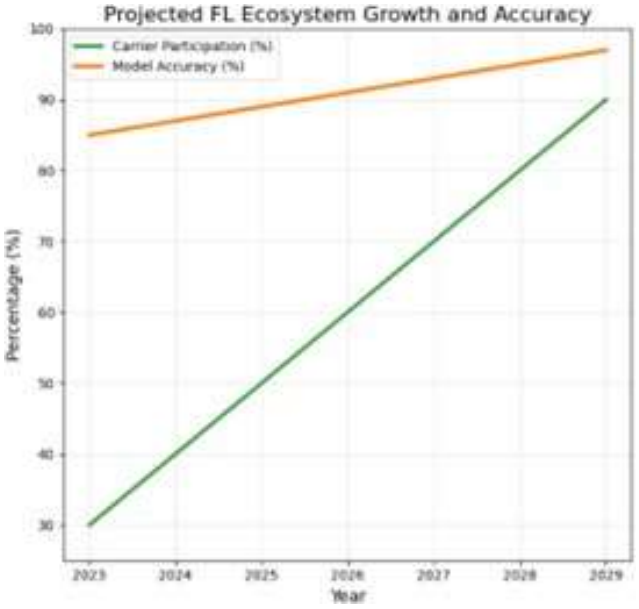

Fig. 7. Projected FL Ecosystem Growth and Accuracy

ulations, supporting confidential computing technologies, creating omni-directional cooperation mechanisms, and jointly preventing cross-border interest crimes—can the intelligent level of the FL model be improved and finally prevent various malicious activities.

A. Final Thoughts and Future Directions in Insurance Fraud Detection

Despite the challenges of a multi-institution deployment involving third-party insurers and regulatory compliance, experience in other sectors shows that establishing the rights infrastructure enables substantial benefits. Organizing insurance fraud detection as a federated learning problem enables privacy-preserving information sharing on sensitive data and combines cross-carrier knowledge for improved prediction performance. Data governance, collaboration protocols, data exchange boundaries, model aggregation methods and incentives for fairness and participation have to be designed carefully to gain the trust of all involved organizations. Detection relies on the detectability of fraud signals across carriers and transferability across jurisdictions in the case of international collaborations. Labels can be expensive and sometimes impossible to share, especially between competing entities, and proxies should be investigated for use in these situations. The systemic nature of fraud may also induce sudden detection loss due to shared detection codes. These concepts are relevant to many applications of cross-carrier language detection; in insurance fraud detection, the evolution of features and models together with the support for institutions in data minimization are thus key aspects. Enabling the cross-carrier sharing of models, data augmentation and boosted methods via communication also helps alleviate biennial claims-staging issues.

REFERENCES

[1] Paleti, S., Mashetty, S., Challa, S. R., ADUSUPALLI, B., & Singireddy, J. (2024). Intelligent Technologies for Modern Financial Ecosystems: Transforming Housing Finance. Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions. Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions (July 02, 2024).

[2] Dong, P., Feng, R., Quan, Z., & Wang, T. (2024, February). Federated learning for insurance companies: Unlocking the potential of privacy-preserving data sharing. Society of Actuaries Research Institute.

[3] Innan, N., Marchisio, A., Bennai, M., & Shafique, M. (2024, April). QFNN-FFD: Quantum Federated Neural Network for Financial Fraud Detection. arXiv preprint arXiv:2404.02595.

[4] Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).

[5] Nkosi, D. (2024). Multimodal AI Interfaces for Enhanced Financial Security: Integrating Voice, Gesture, and Behavioral Analytics. Nature Financial Technology, 2(2), 156-178. https://doi.org/10.1038/s44296-024-00009-8

[6] Pati, S., Singh, A., Paul, S., & Biswas, S. (2024). Privacy preservation for federated learning in health care. Patterns, 5(7), 100974.

[7] Wang, H. (2024). Advancements of Credit Card Fraud Detection Based on Federated Learning. In Proceedings of SCI-VERSE (Vol. , Paper 135274). SCITEPRESS

[8] Challa, S. R., Challa, K., Lakkarasu, P., Sriram, H. K., & Adusupalli, B. (2024). Strategic Financial Growth: Strengthening Investment Management, Secure Transactions, and Risk Protection in the Digital Era. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 97-108.

[9] Sinaci, A. A., Gencturk, M., Alvarez-Romero, C., Laleci Erturkmen, G. B., Martinez-Garcia, A., Escalona-Cuaresma, M. J., & Parra-Calderon, C. L. (2024). Privacy-preserving federated machine learning on FAIR health data: A real-world application. Computational and Structural Biotechnology Journal, 24, 136–145. https://doi.org/10.1016/j.csbj.2024.02.014

[10] El Hallal, T., Wang, H. (2024). Advances of Credit Card Fraud Detection Based on Federated Learning. In Proceedings of SCI-VERSE (Paper 135274). SCITEPRESS.

[11] Yellanki, S. K. (2024). Leveraging Deep Learning and Neural Networks for Real-Time Crop Monitoring in Smart Agricultural Systems. American Data Science Journal for Advanced Computations (ADSJAC) ISSN: 3067-4166, 2(1).

[12] Dalmaz, O., Mirza, M. U., Elmas, G., Ö zbay, M., Dar, S. U. H., Ceyani, E., Avestimehr, S., & C¸ ukur, T. (2024). One model to unite them all: Personalized federated learning of multi-contrast MRI synthesis. Medical Image Analysis, 94, 103121.

[13] Sha, X. (2024). Research on Financial Fraud Algorithm Based on Federated Learning and Big Data Technology. arXiv preprint arXiv:2405.03992.

[14] Motamary, S. (2024). Transforming Customer Experience in Telecom: Agentic AI-Driven BSS Solutions for Hyper-Personalized Service Delivery. Available at SSRN 5240126.

[15] Fu, H., et al. (2024). Personalized federated learning for abdominal multi-organ segmentation.

[16] Abadi, A., Doyle, B., Gini, F., Guinamard, K., Murakonda, S. K., Liddell, J., Mellor, P., Murdoch, S. J., Naseri, M., Page, H., Theodorakopoulos, G., & Weller, S. (2024). Starlit: Privacy-Preserving Federated Learning to Enhance Financial Fraud Detection. arXiv preprint arXiv:2401.10765.

[17] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[18] hung, H., & Lee, J. S. (2024). Federated influencer learning for secure and efficient collaborative learning in realistic medical database environment. Scientific Reports, 14, 22729.

[19] Jillo, G. (2024). Advances and Challenges in Fraud Detection in Medical Insurance. SSRN Working Paper.
[20] Pandiri, L., & Chitta, S. (2024). Machine Learning-Powered Actuarial Science: Revolutionizing Underwriting and Policy Pricing for Enhanced Predictive Analytics in Life and Health Insurance.
[21] Yahiaoui, M. E., Khlifi, M., & Alimi, A. M. (2024). Federated learning with privacy-preserving techniques for multi-institutional brain tumor segmentation. Diagnostics, 14(24), 2891.
[22] Hong, B., et al. (2024). Health insurance fraud detection based on multi-channel heterogeneous graph structure learning. Heliyon, (e30045).
[23] Nandan, B. P. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. International Journal of Medical Toxicology and Legal Medicine, 27(5), 759-772.
[24] Khan, M. S. I., Gupta, A., Seneviratne, O.,& Patterson, S. (2024, August 3). Fed-RD: Privacy-preserving federated learning for financial crime detection. arXiv preprint arXiv:2408.01609.
[25] Du Preez, A., Bhattacharya, S., & van der Schyff, K. (2024). Fraud detection in healthcare claims using machine learning: A systematic review. Artificial Intelligence in Medicine, 155, 103061.
[26] Agentic AI in Data Pipelines: Self OptimizingSystems for Continuous Data Quality, Performance,and Governance. (2024). American Data Science Journal for Advanced Computations (ADSJAC) ISSN: 3067-4166, 2(1). https://adsjac.com/index.php/adsjac/article/view/23
[27] Methuku, V. (2024). AI-Enabled Federated Learning System for Privacy-Preserving Health Insurance Underwriting and Fraud Detection. Technical Disclosure (dpubs).
[28] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).
[29] World Health Organization. (2024). Use of machine learning for fraud detection within the claims management process in the Philippines. WHO Publications.
[30] Rashid, N. S., & Yasin, H. M. (2024). Privacy-preserving machine learning: A review of federated learning techniques and applications. International Journal of Scientific World, 11(1), 30-39.
[31] Meda, R. (2024). Predictive Maintenance of Spray Equipment Using Machine Learning in Paint Application Services. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 2(1).
[32] Nelson, J., Lawson, A., & Conlins, W. (2024). Cutting-Edge Research in Fraud Detection: Self-supervised learning, graph neural networks, and federated learning. Advances in Economics, Management and Political Sciences, 98, 43-49.