sustainability

MDPI

Article

Gradient Boosting for Health IoT Federated Learning

Sobia Wassan [1,*], Beenish Suhail [2], Riaqa Mubeen [3], Bhavana Raj [4], Ujjwal Agarwal [5], Eti Khatri [6], Sujith Gopinathan [7] and Gaurav Dhiman [8,9,10]

[1] School of Equipment Engineering, Jiangsu Urban and Rural Construction Vocational College, Changzhou 213000, China
[2] School of Economics, Shanghai University, Shanghai 201900, China
[3] School of Management, Harbin Institute of Technology (HIT), Harbin 150001, China
[4] School of Management, Institute of Public Enterprise, Hyderabad 500101, India
[5] School of Information Technology, University of Technology and Applied Sciences, Salalah 215, Oman
[6] School of Management, Nitte Meenakshi Institute of Technology, Bengaluru 560064, India
[7] School of Finance, AMU/AIMA, New Delhi 110003, India
[8] Department of Electrical and Computer Engineering, Lebanese American University, Byblos P.O. Box 13-5053, Lebanon
[9] Department of Computer Science and Engineering, University Centre for Research and Development, Chandigarh University, Mohali 140413, India
[10] Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India
[*] Correspondence: sobiaali614@gmail.com

Abstract: Federated learning preserves the privacy of user data through Machine Learning (ML). It enables the training of an ML model during this process. The Healthcare Internet of Things (HIoT) can be used for intelligent technology, remote detection, remote medical care, and remote monitoring. The databases of many medical institutes include a vast quantity of medical information. Nonetheless, based on its specific nature of health information, susceptibilities to private information, and since it cannot be pooled related to data islands, Federated Learning (FL) offers a solution as a shared collaborative artificial intelligence technology. However, FL addresses a series of security and privacy issues. An adaptive Differential Security Federated Learning Healthcare IoT (DPFL-HIoT) model is proposed in this study. We propose differential privacy federated learning with an adaptive GBTM model algorithm for local updates, which helps adapt the model's parameters based on the data characteristics and gradients. By training and applying a Gradient Boosted Trees model, the GBTM model identifies medical fraud based on patient information. This model is validated to check performance. Real-world experiments show that our proposed algorithm effectively protects data privacy.

Keywords: data privacy; federated learning; medical fraud; Internet of Things; GBTM

check for updates

Citation: Wassan, S.; Suhail, B.; Mubeen, R.; Raj, B.; Agarwal, U.; Khatri, E.; Gopinathan, S.; Dhiman, G. Gradient Boosting for Health IoT Federated Learning. *Sustainability* **2022**, *14*, 16842. https://doi.org/10.3390/su142416842

Academic Editor: Andreas Ihle

Received: 19 November 2022
Accepted: 9 December 2022
Published: 15 December 2022

1. Introduction

Federated learning helps mobile phones develop a standard prediction model collaboratively while storing all training data locally, splitting machine learning from storing data on the cloud. This extends the usage of local models that generate a prediction on mobile devices (similar to the Smartphone Vision API or On-Device Smarter Reply) by transferring model training to the device [1].

Many algorithmic and technological obstacles must be overcome before implementing FL. In a traditional machine learning system, an optimizing method such as Stochastic Gradient Descent (SGD) works on an extensive dataset split equally between cloud servers. These highly recurrent algorithms must train data interfaces with low latency and high speed. In contrast, data is distributed between millions of devices in the FL environment very heterogeneously. Furthermore, these devices have a significantly higher frequency and lower speed connections and are only rarely accessible for training. FL has recently been a hot topic of study in both business and academia. Using several machine learning techniques, communication methods, privacy-preserving techniques, and data partitioning schemes have been investigated using federated environments. Federated learning allows several companies to develop a robust machine learning technique without sharing data, allowing issues such as privacy, security, access rights, and connectivity to heterogeneous data to be addressed. Security, communications, IoT, and pharmaceutical industries are just a few areas where it is used [2]. As medical data volumes and types increase, effective mining models to evaluate these data are now needed to aid in illness diagnoses. They will also provide medical remedies and provide better patient care. Image identification, natural language processing, and healthcare are areas where machine learning has been applied. Comparing several FL algorithms provides us with a concept of how algorithmic modifications can affect the performance of the final model.

On the other hand, machine learning models can only achieve high accuracy with a large number of training samples, which is exceptionally crucial in health care since it may sometimes determine whether or not a patient's life can be saved. In this research, the authors present a customized federation learning model for intelligent IoT systems applications in a cloud system [3]. The author designed a federated learning (FL) model that employs a reputation mechanism to enable home appliance manufacturers in developing smart house systems using machine learning systems based on user data [4]. The author used federated learning to analyze cardiac activity data acquired using smart bands to track stress levels throughout various situations. The author of this research study offered a novel strategy that focuses on sleep loss in many ways [5]. By protecting the privacy of the data, the author obtained encouraging findings for employing federated learning in IoT system functionality monitoring systems [6].

This study focuses on a broad class of machine learning approaches taught using gradient-descent techniques while accounting for the practical limits of non-uniformly distributed datasets between users [7]. In this study, the authors propose a system for detecting malware threats mostly on the Industrial Internet of Things (MT-IIOT). A method based on color picture visualization and a deep learning neural network is presented for the in-depth detection of malware [8]. The author of this research study offered a novel strategy that focuses on sentimental elements of the item's qualities [9]. The author used machine learning and deep learning algorithms to forecast the risk of securing e-banking and e-commerce transactions by analyzing datasets from e-commerce and e-banking platforms [10]. In this study, the author introduces IMCFN, a new classifier that uses CNN-based deep learning architecture to recognize variations of malware families and enhance malware detection [11]. A Gradient Boosting Decision model (GBDM) classifier-based fall detection approach is developed that uses big data fusing of postural sensor and human video skeleton to improve detection accuracy [12].

This research detects the most effective and secure path for exchanging health data using the multi-objective bio-inspired heuristic cuckoo search node optimization technique [13]. The author emphasizes the roots of FL's risks, significant attacks on it, and responses, as well as their particular challenges, and also talks about interesting future research routes for more reliable FL [14]. Federated learning is vulnerable to feeding attacks, where attackers upload fraudulent model modifications to affect the whole model. Researchers propose an attack prevention method that uses synthetic stochastic networks to provide auditing data throughout the training process. It eliminates attackers by evaluating their model accuracy to identify and prevent feeding attempts in federated learning [15]. This article aims to investigate the factors of the sustainability supply chain (SHCSC) performance management through extensive literature research and the perspectives of industry experts [16].

2. Current FL Digital Health Activities

Since FL is a broad learning paradigm that reduces the need for data pooling for AI model construction, its applicability covers all aspects of AI for health care. FL may enable revolutionary advances in the future by collecting more data variation and studying patients from different demographic groups. It is already in use now.

2.1. Clinicians

Clinicians often see the same group of people because of their location and demographics, which can lead to inappropriate conclusions about how highly probably specified diseases are or how they affect each other. When they use ML-based systems, for example, as the second reader, they may be able to add accurate and reliable information from those other institutions to their experience, allowing them to make diagnoses that are more coherent than they are now. Even though this is true for ML-based systems, devices trained in a federated way can come to even less biased findings and be more sensitive to rare cases because they have access to a wider range of data. To ensure that the information is sent to collaborators in a way that is easy for everyone to understand, some preliminary work is needed, such as ensuring that the data structure, annotations, and report protocol follow the rules.

2.2. Patients

Generally, patients received primary care. Implementing FL on an international market could ensure higher clinical judgments despite the treatment site. Specifically, people seeking medical care in rural places might benefit from the same high-quality ML-assisted diagnosis provided in hospitals with a high workload. The same stands true for rare or regionally uncommon diseases, the implications of which are likely to be less severe if diagnoses can be established more quickly and precisely. FL may also reduce the barrier to being a data provider since patients may rest certain that their data remains within their hospital and that access to data can be revoked.

2.3. Hospitals and Medical Procedures

With comprehensive visibility of data access, hospitals and practices may retain the full authority of patient data, reducing the risk of third-party cyberbullying by third parties. To train and evaluate machine learning models without interruption, however, businesses must invest in physical on-premises computing resources or private cloud service delivery and adhere to standardized, open data formats. Whether a location is just engaged in assessment and testing activities or participates in training initiatives will determine the number of computing capabilities required. Even very modest institutions may join and profit from the communal models developed.

3. The Advantages of Federated Learning

Federated learning is a key area in machine learning that already provides significant advantages over standard, centralized machine learning techniques. These are the advantages of federated learning:

Health: federated learning may improve the health system insurance industries since it protects private information. To diagnose uncommon diseases, federated learning models may give more data diversity by collecting data from many sites (e.g., hospitals and electronic health records databases). According to recent research titled "The future of digital health with federated learning," federated learning may assist in resolving data privacy and governance issues by enabling machine learning models to be developed from distant data [17].

Data security: the model does not require a data pool if the training datasets are maintained on the devices.

Data diversity: challenges besides data security, such as network instability in edge devices, may prevent businesses from collecting information from various sources. Federated learning allows access to massive datasets even when data sources can only communicate at certain times.

Real-time learning: using client data, models are continually improved without the need to collect data for continuous learning. However, compared to federated learning approaches, which do not require any centralized server for data processing, this system requires more sophisticated technology.

4. Challenges of Federated Learning

Investment requirements: federated learning methods may need frequent node-to-node communication. This implies that a system's storage capacity and bandwidth efficiency are two characteristics.

Data Protection: in federated learning, data is not gathered on a single entity/server; instead, multiple devices are used for data collection and analysis. This may increase the sensitivity to attack.

Even if only models, not raw data, are transmitted to the central database, it is possible to reverse engineer models to identify customer data. Federated learning may use privacy-enhancing technologies such as differential privacy, safe multiparty computing, and homomorphic encryption to boost its data privacy capabilities.

Limitations in feature: data heterogeneity: federated learning combines models from several devices to generate a superior model. Some device-specific features may restrict the applicability of models created from some devices, hence diminishing the precision of the future iteration of the model.

Indirect information leakage: researchers have studied cases in which one federation member may intentionally attack other members by installing a remote trojan into the joint global model. Federated learning is a very new method of machine learning. It requires new research and studies to boost its performance. When a central model uses the data of other devices to create a new model in federated learning, there is still a level of centralization. Researchers suggest adopting blockchain federated learning (BlockFL) or other methods to develop federated learning models with zero trust.

Alternatives for Federated Learning: the same issue of training data privacy was suggested to be solved through gossip learning. This method is entirely decentralized, as there are no servers for integrating outputs from various places. Local nodes immediately exchange models and combine them. Additionally, less technology and centralization are required for gossip learning than for federated learning, which is a benefit. Gossip learning is a unique technique, and further work is necessary to enhance its effectiveness and stability.

Section one is about the introduction, section two is about related works, section three is the methodology, the fourth section will discuss the results, and the last section will contain the conclusion, implications, limitations, and future studies.

5. Related Work

Traditional centralized training techniques typically include gathering large amounts of data from robust cloud computing, which might result in serious user privacy breaches, particularly in the medical field. Many countries have passed regulations limiting the collection of data aspects of user privacy, including the General Data Protection of the European Union (GDPR) [18].

The Health Internet of Things (HIoT) is transforming conventional industries, including healthcare, medical treatment, health policy, and community care, wherein large numbers of HIoT sensors, e.g., wearable sensors, are deployed at the network's edge to gather patient data. FL stands for federated learning [19]. The protection of personal information is a primary issue. It is essential to follow current privacy standards to protect patients' identities, especially in sectors such as medicine. On the other hand, data is essential for research and training machine learning techniques that might help identify intricate connections or individualized treatments that may go undetected. These models typically scale with the quantities of information accessible, but the current circumstance prevents massive databases from being built over several sites.

Hence, integrating comparable or related data from several sites worldwide would be advantageous as long as privacy is maintained. Since it uses machine learning models rather than raw data, federated learning is now being presented as a viable solution. That implies that personal information is never shared outside of the website or device where it was obtained. Federated learning is now a developing field of study, and several fields have been recognized for its use. This literature review examines the notion of federated learning, research into it, and its relevance to sensitive healthcare datasets [20]. The primary idea is that various CNNs produce multiple semantic representations of the picture according to their deeper architectures [21]. Using AI apps to diagnose illnesses has become standard practice as artificial intelligence (AI) has advanced, increasing illness diagnoses and decreasing patient waiting times [22]. This study's objective is to provide a summary of federated learning methods with a focus on the biomedical industry [23].

The author's secure federated-learning architecture contains vertical federated learning, horizontal federated learning, and federation learning models. The authors define, construct, and use the federated-learning framework and present a complete overview of prior publications on the subject [24]. With federated learning, a plurality of parties can develop deep learning methods using a pool of data without revealing their data sources. During training, though, a substantial amount of communication is sacrificed as a result of this kind of private information collaborative learning. Numerous compression approaches have been developed in distributed training literature to address this issue, which can significantly reduce the number of communications needed by more than three orders of magnitude [25].

The author introduces existing works on federated learning through five perspectives: database partition, security method, machine learning method, communication architecture, and systems heterogeneity to give a complete overview and encourage prospective study in this area [26]. By addressing communications, computing, and consensus delays, the authors examine an end-to-end BlockFL latency model and describe the ideal block production rate [27]. The author of this article suggests Overlap-Fed Avg. This creative structure loosens the chain-like restrictions of federated learning by simulating the methodological approach with the experimental communication phase (for example, uploading brand new designs and downloading files through the optimization method), allowing the last step to be entirely covered by the former. Overlap-Fed Avg was enhanced using a hierarchical computing method, a dataset compensating method, or a Nesterov Accelerated Gradients (NAG) method over standard Fed Avg [28].

Federated learning, a novel distributed interactive AI paradigm, holds particular promise for intelligent healthcare since it enables several users (such as hospitals) to engage in AI training while ensuring data privacy. As a result, the researchers conducted comprehensive research on using FL in intelligent healthcare [29]. The development of artificial intelligence or the proliferation of infectious diseases have accelerated the use of novel healthcare. Still, they have also raised questions about data privacy, unauthorized access, and service quality. The (MIoT) has emerged as a workable remedy to these issues, especially when paired with federated learning and blockchain technologies. To minimize a single point of failure, the blockchain is maintained by edge nodes. MIoT devices use federated learning to efficiently utilize scattered clinical data, according to a study on a blockchain-based federation learning method for intelligent healthcare [30].

This study emphasizes the integration of these two attractive innovations for use in real-time and life-critical scenarios, as well as management efficiency in innovative city-based systems. Researchers carefully investigate the different smart city-based applications of FL algorithms in DTs. The findings propose some key obstacles and prospective methods for enhanced FL-DT combined in future applications [31]. This research examines several FL approaches before proposing a real-time distributed networking system predicated on the (MQTT) protocol. This chapter focuses on SDN security issues and anomalies, including packets being lost due to an attacker's malicious behavior [32]. The author explicitly creates various machine learning network systems based on federated learning tools that rely on a Parameters Server (PS) and completely autonomous paradigms controlled by consensus procedures [33].

A shared global deep learning model and a centralized aggregating server are used in federated learning to address the above problems. At the same time, patient data remains in the hands of the appropriate entity, preserving data security and confidentiality. First, he provides a thorough, up-to-date description of works using federated learning in clinical applications in this paper. Then, from a data-centric perspective, he examines at a number of current federated learning challenges, including benchmark datasets, data distributions, and data privacy measures. He concludes by highlighting several prospective challenges and future research activities in healthcare applications [34]. This article introduces federated learning (FL) to provide remote IoT users with privacy-preserving collaboration model training at the network's edge. However, individuals in the FL system may have varying levels of Willingness To Participate (WTP), which the model owner is unsure of [35].

Every industry is now applying novel innovations such as innovative management and digitizing because of cutting-edge technology such as artificial intelligence. This development drives systematic running procedures, lowers management overhead, and increases output rate. However, it gave rise to many attacks and privacy vulnerabilities at the data store and process levels. A lack of privacy and confidence in system predictions constrains the current status of such AI-enabled smart systems' real-world use. A popular technology called blockchain can help to lessen the security concerns associated with AI applications. Since blockchain can reduce AI vulnerability and AI can improve blockchain performance, these two technologies complement one another. The use of blockchain systems to protect intelligent systems in various crucial industries, such as healthcare, finance, energy, government, and the military, is now the subject of extensive research. However, there is not a thorough review of the field's present research activities that can show how blockchain technology is being used to protect AI-based systems and increase their resilience.

This research provides a bibliometric and literary evaluation of how blockchain might act as a security blanket for AI-based systems. For this analytical investigation and review, two well-known study databases (Scopus and Web of Science) were evaluated. The study found that certain journal articles and conference idea proposals had a significant influence. However, a lot more study must be conducted before implementation [36]. Our main contribution is a Gradient-Boosted Model (GBTM). We present differential privacy federated learning for local updates with the adaptive GBTM model method, which helps adapt model parameters depending on data properties and gradients. The GBTM model may detect medical fraud based on patient information by training and implementing a Gradient Boosted Trees model. To ensure performance, this model has been validated. In real-world testing, our suggested method successfully protects data privacy.

6. Methodology

Gradient boosting progressively adds weak learners so that every learner accommodates the residuals from earlier phases, thus boosting the model. The final model pulls together the findings from each phase to create a strong learner. Decision trees are used as weak learners in the gradients boosted decision trees algorithm. With the use of 19 distinct types of healthcare dataset patient information, a Gradient Boosting Tree (GBT) algorithm for the real-time monitoring of medical frauds on the patients' data is examined in this study. They were using a Gradient Boosted Trees model that has been trained and applied. This model is validated to check performance. The research methodology for detecting medical fraud based on patient information is presented in Figure 1. Figure 2 displays the screenshot of the sample dataset.

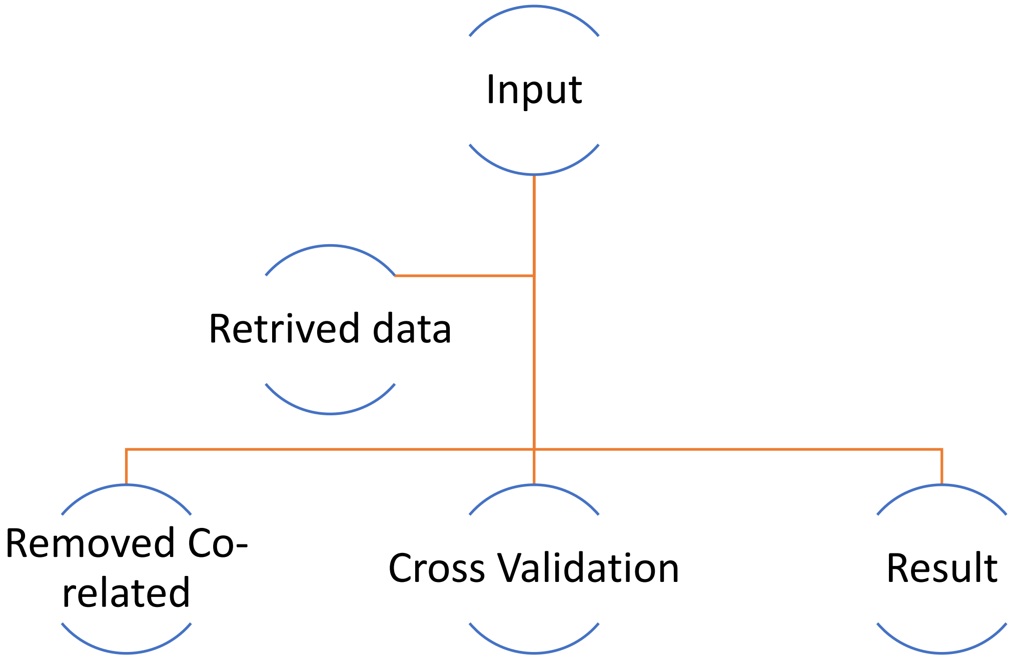
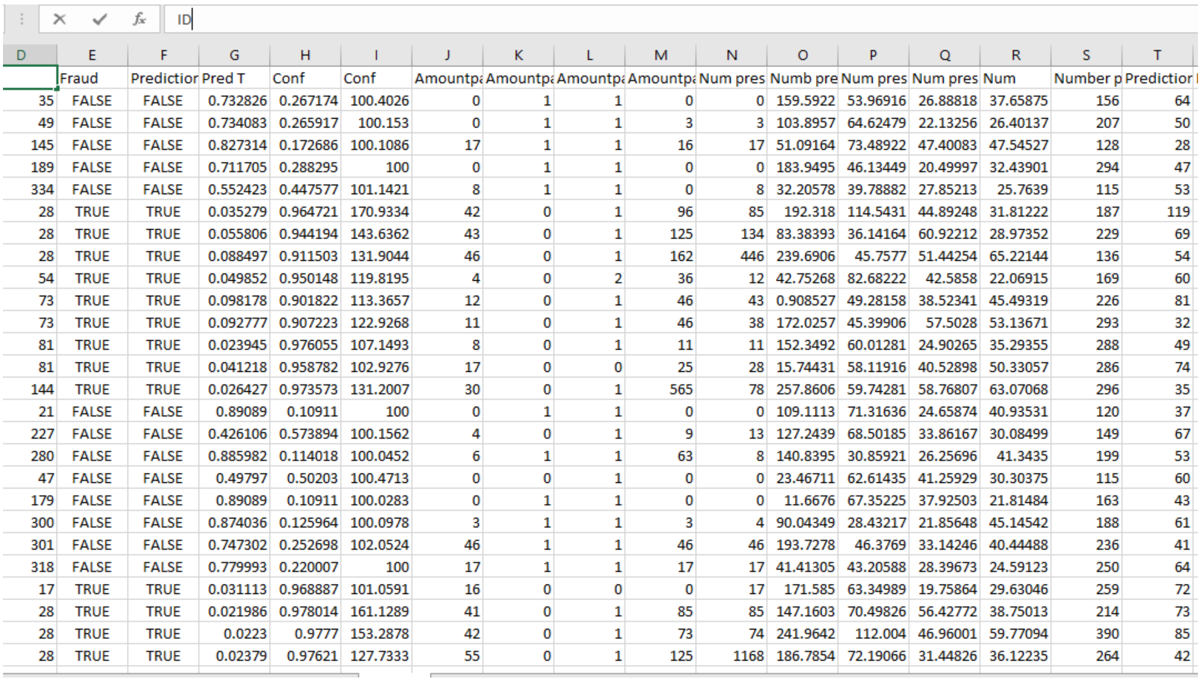


Figure 1. Display the methodology process.



D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	Fraud	Predictior	Pred T	Conf	Conf	Amountpa	Amountpa	Amountpa	Amountpa	Num pres	Numb pre	Num pres	Num pres	Num	Number p	Predictior
35	FALSE	FALSE	0.732826	0.267174	100.4026	0	1	1	0	0	159.5922	53.96916	26.88818	37.65875	156	64
49	FALSE	FALSE	0.734083	0.265917	100.153	0	1	1	3	3	103.8957	64.62479	22.13256	26.40137	207	50
145	FALSE	FALSE	0.827314	0.172686	100.1086	17	1	1	16	17	51.09164	73.48922	47.40083	47.54527	128	28
189	FALSE	FALSE	0.711705	0.288295	100	0	1	1	0	0	183.9495	46.13449	20.49997	32.43901	294	47
334	FALSE	FALSE	0.552423	0.447577	101.1421	8	1	1	0	8	32.20578	39.78882	27.85213	25.7639	115	53
28	TRUE	TRUE	0.035279	0.964721	170.9334	42	0	1	96	85	192.318	114.5431	44.89248	31.81222	187	119
28	TRUE	TRUE	0.055806	0.944194	143.6362	43	0	1	125	134	83.38393	36.14164	60.92212	28.97352	229	69
28	TRUE	TRUE	0.088497	0.911503	131.9044	46	0	1	162	446	239.6906	45.7577	51.44254	65.22144	136	54
54	TRUE	TRUE	0.049852	0.950148	119.8195	4	0	2	36	12	42.75268	82.68222	42.5858	22.06915	169	60
73	TRUE	TRUE	0.098178	0.901822	113.3657	12	0	1	46	43	0.908527	49.28158	38.52341	45.49319	226	81
73	TRUE	TRUE	0.092777	0.907223	122.9268	11	0	1	46	38	172.0257	45.39906	57.5028	53.13671	293	32
81	TRUE	TRUE	0.023945	0.976055	107.1493	8	0	1	11	11	152.3492	60.01281	24.90265	35.29355	288	49
81	TRUE	TRUE	0.041218	0.958782	102.9276	17	0	0	25	28	15.74431	58.11916	40.52898	50.33057	286	74
144	TRUE	TRUE	0.026427	0.973573	131.2007	30	0	1	565	78	257.8606	59.74281	58.76807	63.07068	296	35
21	FALSE	FALSE	0.89089	0.10911	100	0	1	1	0	0	109.1113	71.31636	24.65874	40.93531	120	37
227	FALSE	FALSE	0.426106	0.573894	100.1562	4	0	1	9	13	127.2439	68.50185	33.86167	30.08499	149	67
280	FALSE	FALSE	0.885982	0.114018	100.0452	6	1	1	63	8	140.8395	30.85921	26.25696	41.3435	199	53
47	FALSE	FALSE	0.49797	0.50203	100.4713	0	0	1	0	0	23.46711	62.61435	41.25929	30.30375	115	60
179	FALSE	FALSE	0.89089	0.10911	100.0283	0	1	1	0	0	11.6676	67.35225	37.92503	21.81484	163	43
300	FALSE	FALSE	0.874036	0.125964	100.0978	3	1	1	3	4	90.04349	28.43217	21.85648	45.14542	188	61
301	FALSE	FALSE	0.747302	0.252698	102.0524	46	1	1	46	46	193.7278	46.3769	33.14246	40.44488	236	41
318	FALSE	FALSE	0.779993	0.220007	100	17	1	1	17	17	41.41305	43.20588	28.39673	24.59123	250	64
17	TRUE	TRUE	0.031113	0.968887	101.0591	16	0	0	0	17	171.585	63.34989	19.75864	29.63046	259	72
28	TRUE	TRUE	0.021986	0.978014	161.1289	41	0	1	85	85	147.1603	70.49826	56.42772	38.75013	214	73
28	TRUE	TRUE	0.0223	0.9777	153.2878	42	0	1	73	74	241.9642	112.004	46.96001	59.77094	390	85
28	TRUE	TRUE	0.02379	0.97621	127.7333	55	0	1	125	1168	186.7854	72.19066	31.44826	36.12235	264	42

Figure 2. Display a sample of a medical patient information data set.

Step 1: in Step 1, we input the dataset in the retrieved system.

Step 2: in Step 2, we obtain medical information from patients and previous information about possible fraudulent activities. To feed the GBT algorithm, the data is turned into integers.

Step 3: we have numerous attributes, but only a few are connected (e.g., totals and partial counts). We automatically delete variables with a correlation more significant than 95%.

Step 4: to detect fraudulent conduct, the GBT method is applied. To ensure performance and eliminate statistical bias, the model is verified. Data is balanced on the train side of validation to support the model in detecting odd fraudulent situations.

Step 5: finally, in step 5, we received the results. This model may be used to forecast fraud—the original data and the model's output score. The accuracy, confusion matrix, AUC, and other parameters are included in the performance results. This output port sends the Gradient Boosted classification and regression problems model. This classification and regression problems model may be used to predict the label attribute on unknown data sets.

Metric's Evaluations

In Figure 3, we presented the 200-sample data set with 19 attributes. The data set distributed with equal amounts of 100 samples is true class, and 100 is false class. Two colors differentiate the positive and negative classes. The green color describes a positive class with a true label, while the blue color describes the negative class with a false label.

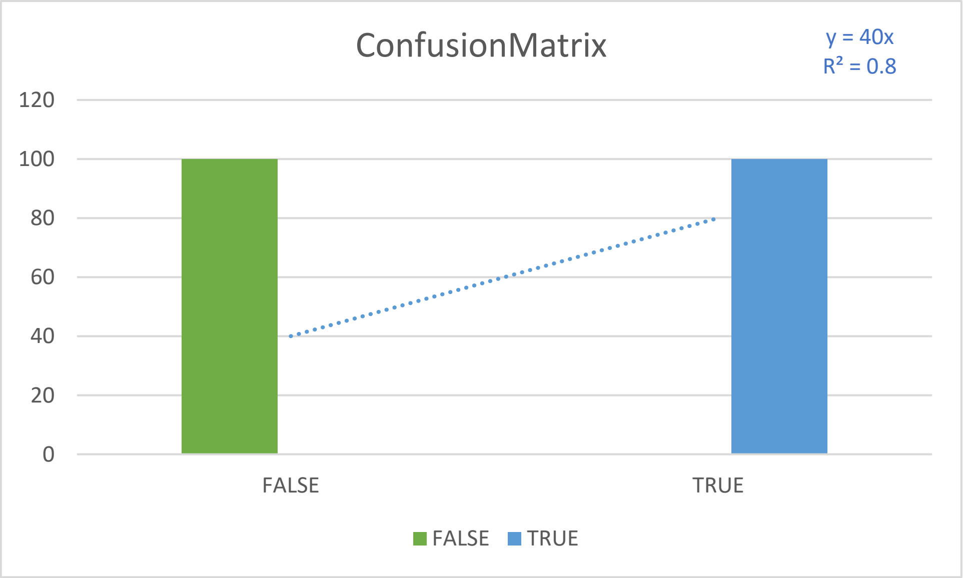


Figure 3. Presented the classes sample true vs. false. (The green color describes a positive class with a true label, while the blue color describes the negative class with a false label).

7. Result and Discussion

7.1. Gradient Boosted Model

A gradient-boosted model is a combination of regression or classification tree algorithms integrated into one. Both of these forward-learning ensemble techniques provide predictions by iteratively improving initial hypotheses. A flexible nonlinear regression method for boosting tree accuracy is called "boosting". An ensemble of weak predictive methods is created by applying weak classification approaches to modify data gradually, including a set of decision trees. While adding more trees increases accuracy, it also complicates systems and makes it harder for people to interpret them. The gradient boosting technique expands on tree boosting to solve these issues.

The operator creates a local H2O cluster with one node and executes the algorithms. Although it only requires one node, the operation is in parallel. The gradient boosted decision trees technique use decision trees as weak learners. To detect residuals, a loss function is utilized. In a regression investigation, mean squared error (MSE) may be used, whereas logarithmic loss (log loss) might be employed in a categorization study. Another essential aspect of the RMSE is that because the errors are squared, more incredible mistakes are given a considerably higher weight. As a result, a tenth-of-a-percentage-point error is

100 times worse than a one-percentage-point error. The inaccuracy scales linearly when employing the MAE. The RMSE model is shown in Figure 4.

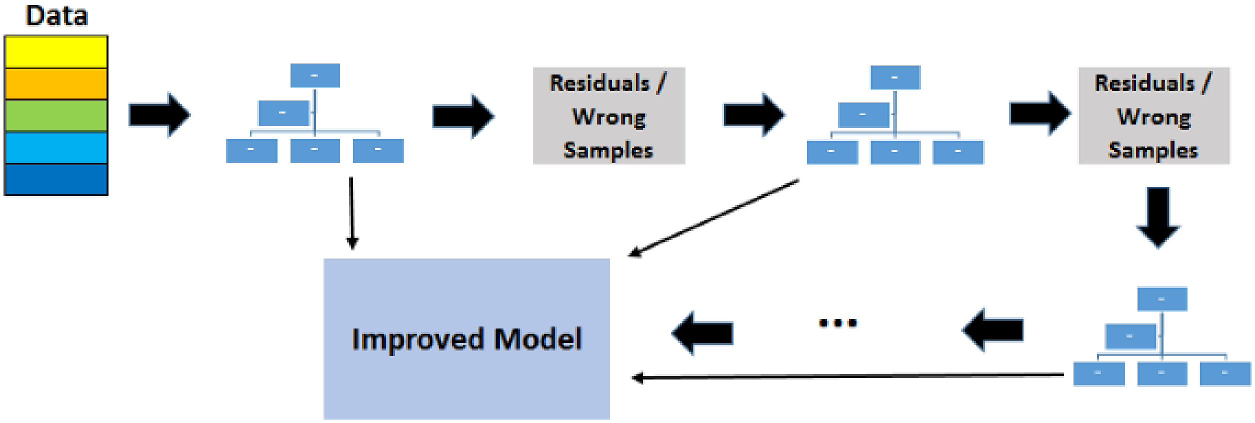


Figure 4. Represent the RMSE model.

The fact that the mistakes are squared means that more significant errors are given a significantly higher weight in the RMSE. As a result, a tenth-of-a-percentage-point mistake is 100 times as bad as a single-percentage-point error scales linearly when utilizing the MAE. The MSE is 0.03001235 in Table 1, but the RMSE is 0.17324072. The RM-h2o-model-gradient boosted trees are shown in Table 1.

Table 1. Represented the rm-h2o-model-gradient boosted trees.

MSE	0.03001235
RMSE	0.17324072
R^2	0.7839111
AUC	0.9925
pr_auc:	0.99860364
Log loss	0.12644695
mean_per_class_error	0.01
default threshold	0.723754823

A binary classification is used to produce predictions with two alternative results: positive and negative. Furthermore, each example's prediction may be correct or incorrect, resulting in a 2 × 2 confusion matrix with four entries:

- TP—the amount of "true positives" or positive instances detected accurately.
- FP—the amount of "false positives" or mistakenly detected negative examples.
- FN—the amount of "false negatives" or positive examples that were misidentified.
- TN—the amount of "true negatives" or incorrectly detected negative examples.

In this paper, the sample data is 200 with 19 attributes. We labeled with True and False. The true samples are 125, which is represented with green color, while the false ratio is 75, denoted with blue in Figure 5.

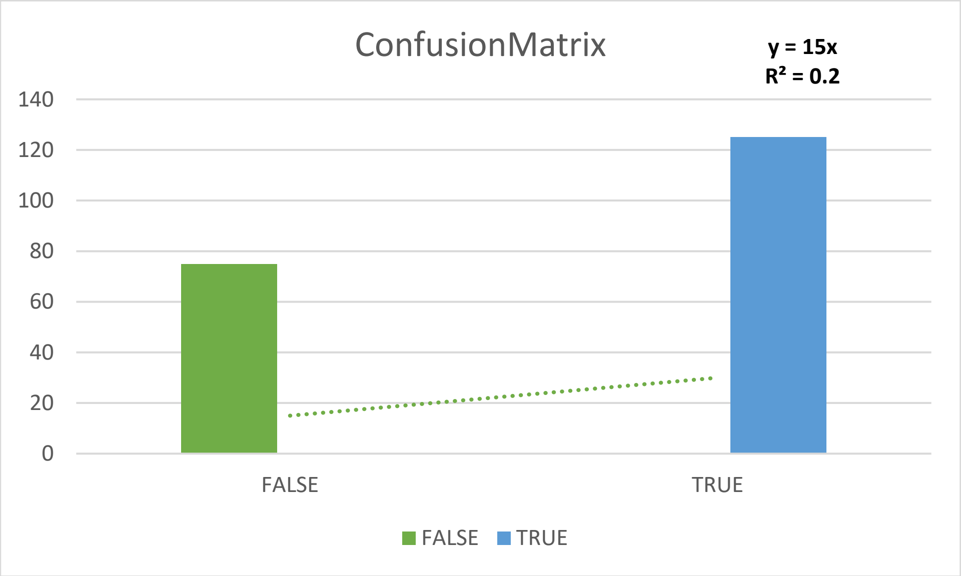


Figure 5. Presented the classes true vs. false. (The green color describes a positive class with a true label, while the blue color describes the negative class with a false label).

In Table 2, we presented the class's error rate. There are two classes labeled false and true. The total number of samples is 22. The false class error is 00.0000, whereas the true class error is 980.0200.

Table 2. Table 2 presents the class error rate.

Label	Sample	Error	Rate
False	20	00.0000	0/20
True	2	980.0200	2/100
Total	22	980.0167	2/120

Boosting combines learning algorithms to create a strong learner from a group of weak learners. In the gradients enhanced decision trees approach, decision trees serve as the weak learners. Each tree seeks to minimize the errors made by prior ones. Boosting trees are weak learners; however, by stacking them in a sequence and focusing on the shortcomings of the previous one, boosting transforms into a very efficient and accurate model. Bootstrap sampling is not required for boosting, unlike bagging. As illustrated in Table 3, a new tree is generated and put in a modified manner of the original dataset.

Table 3. Presented the GBTM classification error.

Number of Trees	Training RMSE	Training Log Loss	Training AUC	Training pr_Auc	Training Lift	Classification Error
0	0.37268	0.45056	0.5	0.83333	1	0.16667
1	0.35129	0.39956	0.95125	0.9906	1.2	0.9167
20	0.17324	0.12645	0.9925	0.9986	1.2	0.01667

Because trees are created in a certain order, learning boosting algorithms take a long time. When a new tree is created, it is placed on a modified method of the initial dataset, as shown in Table 3. Boosting algorithms take a long time to learn because trees are added sequentially. Models that train gradually perform better in statistical learning.

Gradient boosting is an approach that helps a machine learning ensemble reduce variance and bias. By mixing N number of learners, the method aids in the turning of weak learners into strong learners. In Table 4, we presented the GBT model tree description.

Table 4. Display model summary of the gradient boost model.

Model Summary	
Number of Trees	20
Number of Internal Trees	20
Model Size in Bytes Min.	2907
Depth Min	3
Depth Max	5
Depth Mean Depth Min	4.6
Depth Min	4
Leaves Max	10
Mean Leaves	7.05

Table 5 presents the performance vector where the recall obtained the best result with 95.00% compared to accuracy, precision, etc., where the classification error is 17.50% with micro average +/− 9.20%. The specificity is 70%, where the AUC is 0.906.

Table 5. Display performance vector.

Performance Vector		
Accuracy	82.50%	+/−9.20%
micro average	82.50%	
classification_error	17.50%	+/−9.20%
micro average	17.50%	
AUC	0.906	+/−0.074
micro average	0.906	
Precision	77.11%	+/−11.32%
micro average	77.11%	
Recall	95.00%	+/−7.07%
micro average	95.00%	
Specificity	70%	/−15.63%

Table 6 presented the confusion metrics of the performance vector: accuracy, precision, recall, AUC, specificity classification error micro average.

Table 6. Table 6 presented the confusion matrices of the performed vector.

Confusion Matrix		
CM	**FALSE**	**TRUE**
FALSE	70	3
TRUE	30	95

7.2. Based Classification

To forecast the medical fraud detection attribute of the patients' information dataset, the H2O GBT operator is utilized. Because the labeling is nominal, it will be classified. The GBT settings have been modified somewhat. To avoid overfitting, the number of trees is started with one when the end point is 1.2, and to prevent lifting; for similar reasons, the learning rate has been raised to 0.3. The generated model is integrated into an Apply Method operator, which runs the GBT model on the sample data for medical fraud detection. The Accuracy measure is calculated using the labeled dataset and a Performance (Binominal Classifier) operator. Tables 5 and 6 show the Performance Vector and the Gradient Boosted Trees Model for the process output. Figure 6 shows the trees of the Gradient Boosted model.

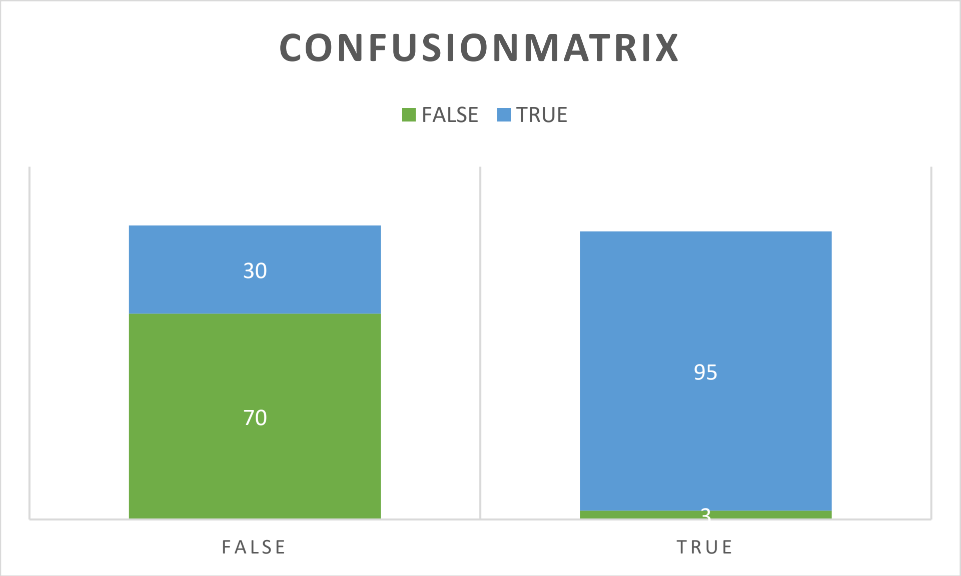


Figure 6. Presented the column views with label data of the confusion matrix. (The green color describes a positive class with a true label, while the blue color describes the negative class with a false label).

Gradient boosting, on the surface, appears to be a stage-wise additive approach for generating learners, i.e., existing trees in the model are not modified while new trees are added at each stage). The stochastic gradient descent procedure identifies factors contributing to the weak learner in the ensemble. The computed contribution from each tree is predicated on minimizing the strong learner's total error. Table 7 presents the GBTM training parameters result. The 20 leaves are also presented with different training results with different parameters. Figure 7 is presented the graphical view of Table 7 trained training parameters result.

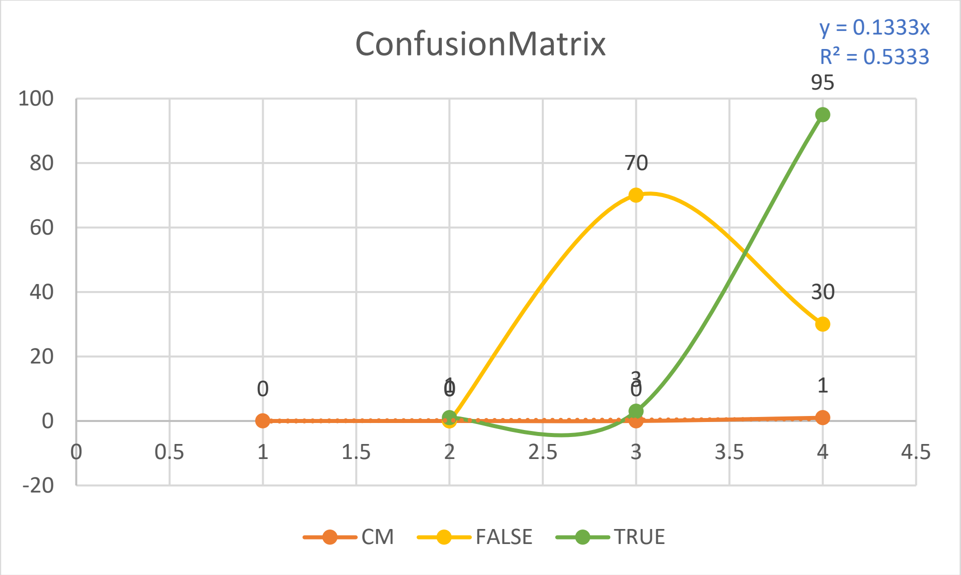


Figure 7. Presents the interpretation result of the confusion matrix.

Table 7. Presented the tree training parameters result.

Number of Trees	Training RMSE	Training Log Loss	Training AUC	Training Lift Training	Training pr_auc	Training Classification Error
0	0.37268	0.45056	0.5	1	0.83333	0.16667
1	0.35129	0.39956	0.95125	1.2	0.9906	0.09167
2	0.33415	0.36475	0.95125	1.2	0.9906	0.09167
3	0.32004	0.33835	0.95125	1.2	0.9906	0.09167
4	0.3047	0.31315	0.9645	1.2	0.99289	0.06667
5	0.29033	0.29008	0.97225	1.2	0.99449	0.06667
6	0.28032	0.27347	0.97375	1.2	0.99484	0.06667
7	0.26652	0.25329	0.978	1.2	0.99564	0.05
8	0.2556	0.23746	0.981	1.2	0.99628	0.05
9	0.24503	0.2226	0.9815	1.2	0.99636	0.04167
10	0.23617	0.20972	0.9865	1.2	0.9974	0.04167
11	0.22753	0.19751	0.987	1.2	0.99752	0.04167
12	0.2199	0.18684	0.9885	1.2	0.99783	0.04167
13	0.2127	0.17715	0.99	1.2	0.99812	0.025
14	0.20638	0.16843	0.99	1.2	0.99812	0.025
15	0.19895	0.15923	0.9905	1.2	0.99823	0.025
16	0.19418	0.15262	0.9905	1.2	0.99825	0.01667
17	0.18912	0.14567	0.9915	1.2	0.99843	0.01667
18	0.18518	0.14004	0.9915	1.2	0.99843	0.01667
19	0.1769	0.13123	0.9925	1.2	0.9986	0.01667
20	0.17324	0.12645	0.9925	1.2	0.9986	0.01667

Gradient-boosted trees use a technique known as the ensemble method. Boosting continuously combines weak learners (often decision trees with a single split, known as decision stumps), so each small tree tries to fix the errors of the former one. Figure 8 presented the GBTM gradient boosted decision tree, while the Figure 9 presented a graphic of overall results, and Figure 10 presented a linear result of trained parameters.

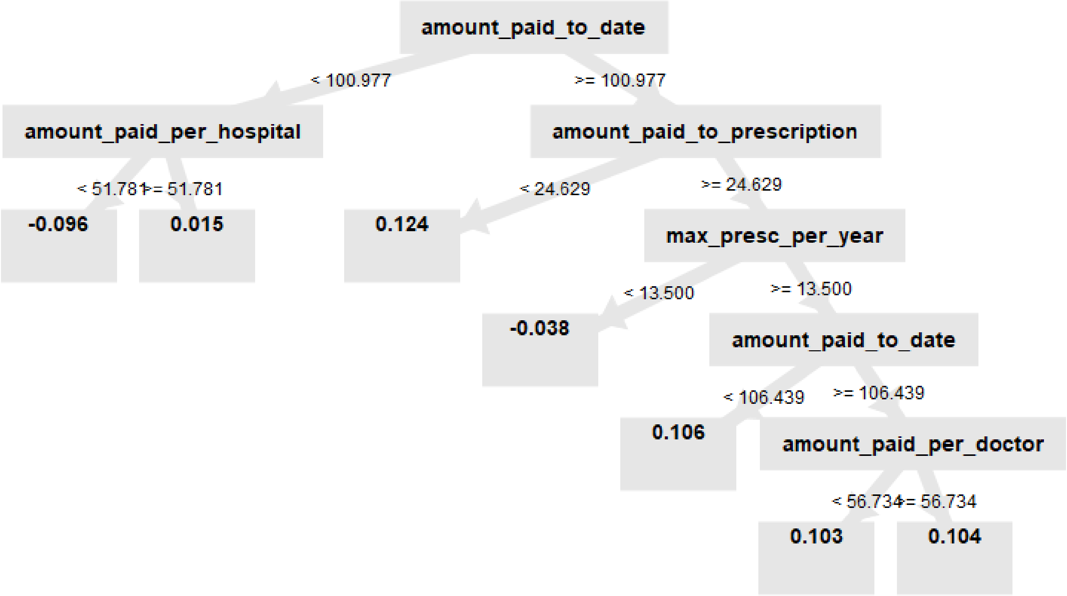


Figure 8. Presents the gradient boosted decision tree. Tree description. GBTM has 20 leaves, but here we display the final tree, which contains 20 leaves. amount_paid_to_date < 100.977; | amount_paid_per_hospital < 51.781: −0.096 {}; | amount_paid_per_hospital >= 51.781: 0.015 {}; amount_paid_to_date >= 100.977; | amount_paid_to_prescription < 24.629: 0.124 {}; | amount_paid_to_prescription >= 24.629; | | max_presc_per_year < 13.500: −0.038 {}; | | max_presc_per_year >= 13.500; | | | amount_paid_to_date < 106.439: 0.106 {}; | | | amount_paid_to_date >= 106.439; | | | | amount_paid_per_doctor < 56.734: 0.103 {}; | | | | amount_paid_per_doctor >= 56.734: 0.104 {}.

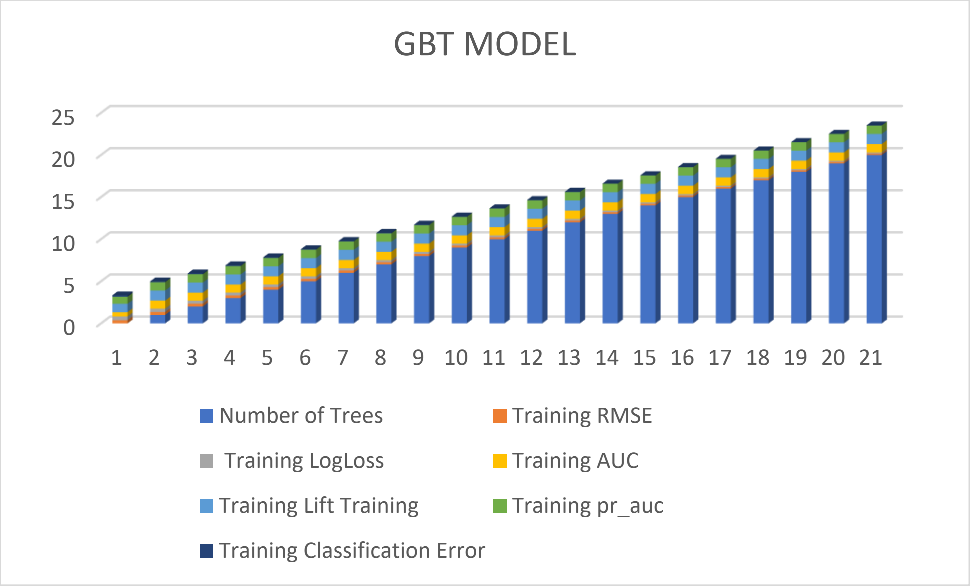


Figure 9. Presented a graphic of overall results.

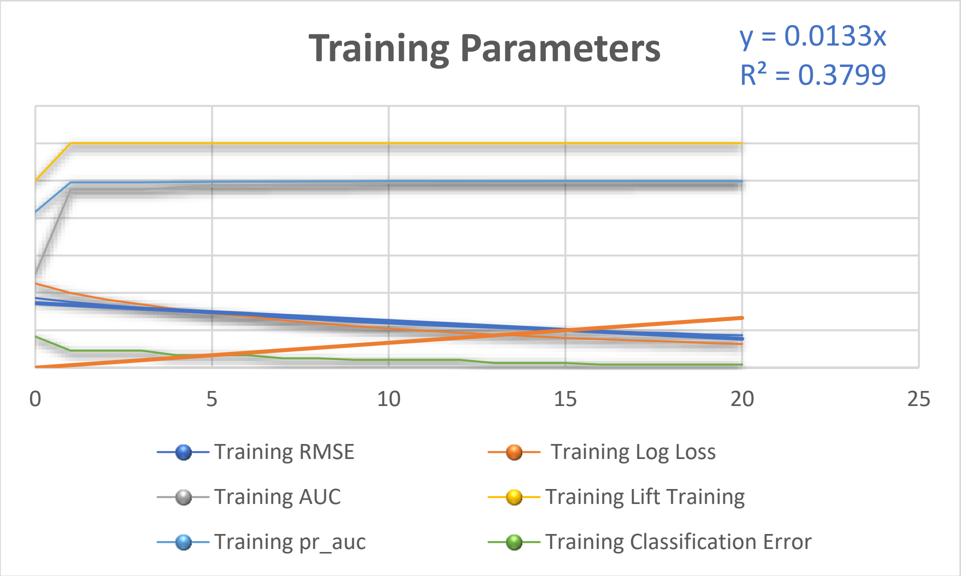


Figure 10. Presented a linear result of trained parameters.

8. Conclusions

Many innovations in the field of data healthcare have emerged from ML and DL, in particular. Since all machine learning (ML) methods significantly benefit from the ability to collect data that approximate the real global distributions, FL is a possible method for producing strong, accurate, safe, robust, and unbiased models. By enabling several parties to train collaboratively without the need to share or centralize data sets, FL effectively addresses issues related to the leakage of private medical information.

The most fascinating and trending technologies in the intelligence healthcare industry are Federated Learning (FL), Machine Learning, and Artificial Intelligence. The healthcare system has always depended on centralized employees exchanging raw data. However, with AI integration, the system would consist of several agent contributors capable of efficiently connecting with their desired host. Furthermore, FL is an important feature that operates independently; it maintains communication in the preferred system based on a mathematical model without exchanging raw data. Combining FL and AI approaches can reduce several limitations and issues in the healthcare system. Federated learning could be an effective approach for facilitating IoT information security (i.e., intrusion detection systems in the IoT context) because when preserving data privacy and preventing the high communication of centralized cloud-based methods (for example, high-frequency documents from time-series sensors). An adaptive Differential Protection Federated Learning Health IoT (DPFL-HIoT) architecture is proposed in this study.

Under federated learning, several individuals exchange their data remotely to train a single deep learning approach collaboratively and iteratively, similar to a team presentation or study. Each party downloads the model, often a pre-trained foundation model, from a cloud-based server. They train the model using their confidential data and summarize and encode its new configuration. Model modifications are sent to the cloud, encoded, aggregated, and included in the centralized system. The collaborative training repeats iteration after iteration until the model is thoroughly trained.

We present a differential privacy federated learning method based on the adaptive GBTM model technique, which may introduce noise to model parameters based on the training data's features and gradient. The GBTM model may detect medical fraud based on patient information by training and using a Gradient Boosted Trees model. This model is validated to check performance. The results of our suggested algorithm on real-world data indicate that it can effectively secure data privacy.

9. Implication, Limitations, and Future Study

The present world is fascinated with data analytics, especially in the medical field. Healthcare data has grown more significant for data analysis, including pharmacy data and supplies, patient data, medical professional information, and associated businesses' responsibility for insurance or similar financial-related operations. The data on the healthcare industry, on the other hand, is fragmented. It is not in the original format, so the data is vulnerable since it contains medical industry information. The most sensitive is data from the insurance industry, which cannot be transferred from one industry to another.

Gradient boosting is a technique used in machine learning to tackle classification and regression issues. It is a sequential ensemble learning approach in which the model's performance improves over time. The model is created in a stage-by-stage manner using this procedure. It infers the model by allowing a differentiable loss function to be optimized. However, this study is limited to minimum sample data. In the future, we can build a model with more sample data. The predictor variables are assessed more accurately with each weak learner added to the model.

Author Contributions: Conceptualization; S.W.; methodology, S.W.; software S.W.; validation, S.W. and B.S.; formal analysis, S.W. and G.D.; investigation, S.W.; resources, B.R., S.G. and R.M.; data curation, U.A. and E.K.; writing—original draft preparation, S.W.; writing—review and editing, S.W.; visualization, S.W.; supervision, S.W.; project administration, S.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Casado, F.E.; Lema, D.; Criado, M.F.; Iglesias, R.; Regueiro, C.V.; Barro, S. Concept drift detection and adaptation for federated and continual learning. *Multimed. Tools Appl.* **2022**, *81*, 3397–3419. [CrossRef]
2. Hu, L.; Yan, H.; Li, L.; Pan, Z.; Liu, X.; Zhang, Z. MHAT: An efficient model-heterogenous aggregation training scheme for federated learning. *Inf. Sci. (Ny)* **2021**, *560*, 493–503. [CrossRef]
3. Wu, Q.; He, K.; Chen, X. Personalized Federated Learning for Intelligent IoT Applications: A Cloud-Edge based Framework. *IEEE Comput. Graph. Appl.* **2020**, *1*, 35–44. [CrossRef] [PubMed]
4. Zhao, Y.; Zhao, J.; Jiang, L.; Tan, R.; Niyato, D.; Li, Z.; Lyu, L.; Liu, Y. Privacy-Preserving Blockchain-Based Federated Learning for IoT Devices. *IEEE Internet Things J.* **2021**, *8*, 1817–1829. [CrossRef]

5. Wassan, S.; Xi, C.; Shen, T.; Gulati, K.; Ibraheem, K.; Rajpoot, R.M.A.L. The Impact of Online Learning System on Students Affected with Stroke Disease. *Behav. Neurol.* **2022**, *2022*, 1–14. [CrossRef] [PubMed]
6. Can, Y.S.; Ersoy, C. Privacy-preserving Federated Deep Learning for Wearable IoT-based Biomedical Monitoring. *ACM Trans. Internet Technol.* **2021**, *21*, 1–17. [CrossRef]
7. Abdellatif, A.A.; Mhaisen, N.; Mohamed, A.; Erbad, A.; Guizani, M.; Dawy, Z.; Nasreddine, W. Communication-efficient hierarchical federated learning for IoT heterogeneous systems with imbalanced data. *Futur. Gener. Comput. Syst.* **2022**, *128*, 406–419. [CrossRef]
8. Nie, L.; Ning, Z.; Wang, X.; Hu, X.; Cheng, J.; Li, Y. Data-Driven Intrusion Detection for Intelligent Internet of Vehicles: A Deep Convolutional Neural Network-Based Method. *IEEE Trans. Netw. Sci. Eng.* **2020**, *7*, 2219–2230. [CrossRef]
9. Wassan, S.; Chen, X.; Shen, T.; Waqar, M.; Jhanjhi, N.Z. Amazon Product Sentiment Analysis using Machine Learning Techniques. *Rev. Argent.* **2021**, *2021*, 695.
10. Wassan, S.; Xi, C.; Jhanjhi, N.; Raza, H. A smart comparative analysis for secure electronic websites. *Intell. Autom. Soft Comput.* **2021**, *29*, 187–199. [CrossRef]
11. Vasan, D.; Alazab, M.; Wassan, S.; Naeem, H.; Safaei, B.; Zheng, Q. IMCFN: Image-based malware classification using fine-tuned convolutional neural network architecture. *Comput. Netw.* **2020**, *171*, 107138. [CrossRef]
12. Cai, W.Y.; Guo, J.H.; Zhang, M.Y.; Ruan, Z.X.; Zheng, X.C.; Lv, S.S. GBDT-Based Fall Detection with Comprehensive Data from Posture Sensor and Human Skeleton Extraction. *J. Healthc. Eng.* **2020**, *2020*, 1–15. [CrossRef] [PubMed]
13. Mohanaprakash, K.; GunaSekar, T. Efficient and Secure Remote Health Management in Cloud in Vehicular Adhoc Network Environment. *J. Med. Imaging Heal. Inform.* **2022**, *11*, 2966–2975. [CrossRef]
14. Bouacida, N.; Mohapatra, P. Vulnerabilities in Federated Learning. *IEEE Access* **2021**, *9*, 63229–63249. [CrossRef]
15. Zhao, Y.; Chen, J.; Zhang, J.; Wu, D.; Blumenstein, M.; Yu, S. Detecting and mitigating poisoning attacks in federated learning using generative adversarial networks. *Concurr. Comput. Pr. Exp.* **2020**, *34*, e5906. [CrossRef]
16. Hossain, M.K.; Thakur, V. Drivers of sustainable healthcare supply chain performance: Multi-criteria decision-making approach under grey environment. *Int. J. Qual. Reliab. Manag.* **2022**, *39*, 859–880. [CrossRef]
17. Rieke, N.; Hancox, J.; Li, W.; Milletarì, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K.; et al. The future of digital health with federated learning. *Npj Digit. Med.* **2020**, *3*, 1–7. [CrossRef]
18. Śmietanka, M.; Pithadia, H.; Treleaven, P. Federated learning for privacy-preserving data access. *Int. J. Data Sci. Big Data Anal.* **2021**, *1*, 1–13. [CrossRef]
19. McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.Y. Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics*; PMLR: Fort Lauderdale, FL, USA, 2017.
20. Pfitzner, B.; Steckhan, N.; Arnrich, B. Federated Learning in a Medical Context: A Systematic Literature Review. *ACM Trans. Internet Technol.* **2021**, *21*, 1–31. [CrossRef]
21. Vasan, D.; Alazab, M.; Wassan, S.; Safaei, B.; Zheng, Q. Image-Based malware classification using ensemble of CNN architectures (IMCEC). *Comput. Secur.* **2020**, *92*, 101748. [CrossRef]
22. Dai, W.; Brisimi, T.S.; Adams, W.G.; Mela, T.; Saligrama, V.; Paschalidis, I.C. Prediction of hospitalization due to heart diseases by supervised learning methods. *Int. J. Med. Inform.* **2015**, *84*, 189–197. [CrossRef] [PubMed]
23. Xu, J.; Glicksberg, B.S.; Su, C.; Walker, P.; Bian, J.; Wang, F. Federated Learning for Healthcare Informatics. *J. Healthc. Informatics Res.* **2021**, *5*, 1–19. [CrossRef] [PubMed]
24. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]
25. Sattler, F.; Wiedemann, S.; Muller, K.R.; Samek, W. Robust and Communication-Efficient Federated Learning from Non-i.i.d. Data. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 3400–3413. [CrossRef] [PubMed]
26. Zhang, C.; Xie, Y.; Bai, H.; Yu, B.; Li, W.; Gao, Y. A survey on federated learning. *Knowl.-Based Syst.* **2021**, *216*, 106775. [CrossRef]
27. Kim, H.; Park, J.; Bennis, M.; Kim, S.L. Blockchained on-device federated learning. *IEEE Commun. Lett.* **2020**, *24*, 1279–1283. [CrossRef]
28. Zhou, Y.; Ye, Q.; Lv, J. Communication-Efficient Federated Learning with Compensated Overlap-FedAvg. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *33*, 192–205. [CrossRef]
29. Nguyen, D.C.; Pham, Q.-V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.; Hwang, W.-J. Federated Learning for Smart Healthcare: A Survey. *ACM Comput. Surv.* **2022**, *55*, 1–37. [CrossRef]
30. Chang, Y.; Fang, C.; Sun, W. A blockchain-based federated learning method for smart healthcare. *Comput. Intell. Neurosci.* **2021**, *2021*, 1–12. [CrossRef]
31. Ramu, S.P.; Boopalan, P.; Pham, Q.-V.; Maddikunta, P.K.R.; Huynh-The, T.; Alazab, M.; Nguyen, T.T.; Gadekallu, T.R. Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions. *Sustain. Cities Soc.* **2022**, *79*, 103663. [CrossRef]
32. Nazar, M.J.; Iqbal, S.; Altaf, S.; Qureshi, K.N.; Usmani, K.H.; Wassan, S. Software-Defined Networking (SDN) Security Concerns. In *Information Security Handbook*; CRC Press: Boca Raton, FL, USA, 2022. [CrossRef]
33. Tedeschini, B.C.; Savazzi, S.; Stoklasa, R.; Barbieri, L.; Stathopoulos, I.; Nicoli, M.; Serio, L. Decentralized Federated Learning for Healthcare Networks: A Case Study on Tumor Segmentation. *IEEE Access* **2022**, *10*, 8693–8708. [CrossRef]

34. Prayitno; Shyu, C.-R.; Putra, K.T.; Chen, H.-C.; Tsai, Y.-Y.; Hossain, K.S.M.T.; Jiang, W.; Shae, Z.-Y. A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications. *Appl. Sci.* **2021**, *11*, 11191. [CrossRef]
35. Lim, W.Y.B.; Garg, S.; Xiong, Z.; Niyato, D.; Leung, C.; Miao, C.; Guizani, M. Dynamic Contract Design for Federated Learning in Smart Healthcare Applications. *IEEE Internet Things J.* **2021**, *8*, 16853–16862. [CrossRef]
36. Shinde, R.; Patil, S.; Kotecha, K.; Ruikar, K. Blockchain for securing ai applications and open innovations. *J. Open Innov. Technol. Mark. Complex.* **2021**, *7*, 189. [CrossRef]