International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

A Robust Framework for Ensuring Data Confidentiality and Security in Modern Healthcare Networks

Muzammil Hussain[1], Nadeem Akhtar[2], Raza Hasan[3]

Submitted: 25/01/2024 **Revised**: 03/03/2024 **Accepted**: 11/03/2024

Abstract: The potential implementation of extensive data sharing in dispersed network systems is likely to give rise to concerns surrounding privacy, secrecy, and authentication inside the realm of cyberspace. The main objective of this work is to safeguard the privacy and confidentiality of data inside an unsecured environment during the exchange of audiovisual content between two Internet of Things (IoT) nodes. To effectively counteract an adversary and guarantee the confidentiality of data, we suggest the implementation of a resilient multi-level security strategy that relies on the principles of information hiding and chaos theory. While certain block-based resilient data concealing strategies based on the transform domain have demonstrated favorable outcomes, their suboptimal block and coefficient selection processes lead to inadequate performance against prevalent cyber-attacks. Therefore, we propose a Robust Framework for Ensuring Data Confidentiality and Security in EHR-based networks using federated-learning with homomorphic-encryption. A differential-privacy technique is used here to increase privacy, which entails adding noise to the aggregated model update. The suggested model outperforms the most recent techniques for data security and secrecy in networks based on electronic health records.

Keywords: *Data security in healthcare, homomorphic-encryption, differential-privacy, federated-learning*

1. Introduction

The current surge in internet and multimedia technologies is yielding a wide range of advantages across various domains. The utilization of advanced processors and sensors in several sectors of existence is demonstrating advantages owing to their multifaceted capabilities in executing diverse tasks. The widespread use of digital platforms for sharing digital information in various areas, such as healthcare, social networking, media, defense, satellite communication, security, and law enforcement, has resulted in a significant shift in the prevailing paradigm [1]. The integration of smart devices and cloud platforms has given rise to a comprehensive and manageable digital landscape known as the Internet of Things (IoT) [2]. This paradigm is anticipated to bring about a technological revolution in contemporary civilization. The IoT refers to a resilient network including interconnected physical things, which collectively establish a Smart Cyberspace [3]. The primary objective of this network is to facilitate extensive connectivity, optimize computational processes, and enable real-time data analysis. A comprehensively designed IoT system presents several application benefits, including reduced energy usage [4], heightened security measures [5], automated monitoring capabilities [6-8], and user-friendly services [9]. The aforementioned benefits have sparked significant attention within the corporate and scientific communities in recent years. Consequently, numerous prototypes and models have been offered at both the industrial and government levels. In the field of e-healthcare, the potential for significant advancements in the efficiency of service delivery is anticipated through the utilization of IoT technology in conjunction with cloud-based services [10].

The data transmitted within a distributed IoT system encompasses a range of sensitive information, such as medical records, financial data, and confidential papers [11]. These data are typically in the form of photographs, text files, and videos, which are gathered by diverse sensors and cameras and subsequently processed and stored by the system. It is imperative that the technologies used exhibit reliability, security, and computational efficiency to achieve a more advanced and efficient system that comprehensively encompasses control and intelligence inside an application. The computerized nature of controlling these systems gives rise to various dangers arising from the flow of information over insecure channels.

The process of digitizing healthcare data has been found to enhance operational efficiency within the healthcare sector [12]. However, this advancement also raises significant concerns regarding the protection of patient privacy and the confidentiality of their personal information [13]. The conventional methods for collaborative model training sometimes entail the consolidation of data, which presents potential vulnerabilities in terms of unwanted access and data breaches. Deep learning effectively extracts

[1] *Global College of Engineering and Technology,Muscat – 112, Oman
ORCID ID : 0000-0002-7353-0614*
[2] *Global College of Engineering and Technology,Muscat – 112, Oman
ORCID ID : 0009-0009-2866-0741*
[3] *Science and Eng , Solent University, Southhampton – SO140YN,UK
ORCID ID : 0000-0002-8089-837X*
** Corresponding Author Email: muzammil.h@gcet.edu.om*

meaningful information from intricate and multi-dimensional data, such as medical pictures, electronic health records, and genomic sequences [14]. Applying deep learning to healthcare data presents substantial obstacles in safeguarding the privacy and security of sensitive personal information. In this research, we present a framework to guarantee the protection of privacy in deep learning inside contemporary healthcare networks. This framework allows many parties to collectively train and assess deep learning models without jeopardizing the confidentiality of their data. The system we employ utilizes advanced techniques, includes homomorphic-encryption, secured multi-party computation, differential-privacy, and federated-learning to facilitate efficient and secure computation on encrypted data. The study offers a comprehensive examination of the trade-offs between privacy, accuracy, and efficiency within our framework, utilizing both theoretical and empirical analysis. Federated-learning presents a decentralized solution that, when integrated with homomorphic-encryption, gives an augmented level of security to protect patient's privacy.

2. Healthcare Data Privacy and Deep Learning

2.1. Healthcare Data Privacy

Healthcare data is intrinsically sensitive, encompassing a plethora of personal information such as medical records, diagnostic imaging, and treatment strategies. The process of transforming the health information into digital format and incorporating IoT devices into healthcare networks has increased the demand for strong privacy protection measures. Conventional methods for protecting data privacy are frequently inadequate when dealing with the complex algorithms employed in deep learning. This calls for creative solutions to ensure the confidentiality of patient information. These networks are increasingly utilizing data-driven technology to enhance the quality and efficiency of their services.

However, this also presents substantial obstacles to data privacy, as confidential health information of patients and providers may be disclosed to unauthorized entities. Data privacy is not solely a legal and ethical responsibility but also a pivotal element in establishing trust and confidence among stakeholders. As a result, healthcare networks must include suitable methods to safeguard the data they gather, retain, handle, and distribute. These protections may include encryption, anonymization, access control, and audit mechanisms.

In addition, they must comply with relevant laws and regulations that set forth the legal rights and obligations of individuals and organizations with regard to data. Healthcare networks may enhance their image, reduce liability risks, and foster a culture of accountability and respect by putting strong data privacy safeguards in place.

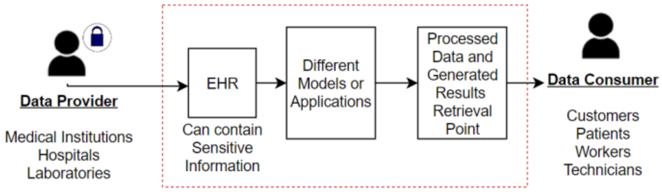

Fig 1. The flow of clinical data through EHR systems and their confidential vulnerability due to potential security breaches.

Still, this data is susceptible to misuse, theft, or exposure by malevolent entities, such as hacker groups, counterfeiters, or unapproved people. Healthcare data privacy has numerous obstacles in the digital era. Some of them are discussed here.

- These data are frequently stored and transmitted using electronic means, rendering them susceptible to cyberattacks. Hackers have the ability to take advantage of weaknesses in the systems or networks responsible for managing healthcare data. They can then get unauthorized access to the data and make changes to it.

- Healthcare data is frequently shared among various groups, including medical professionals, scientists, insurance companies, authorities, and third-party vendors. Data sharing can enhance patient outcomes and societal well-being by facilitating improved care coordination, fostering innovation, and promoting public health initiatives. Data exchange, if not properly protected or consented to, can jeopardize patient privacy.

- Healthcare data is governed by a multitude of laws and regulations designed to safeguard the confidentiality and liberties of both patients and providers. Nevertheless, these rules and regulations can vary among different governments and may lack uniformity or compatibility. This can provide difficulties in ensuring that EHR data systems and processes adhere to regulations and can work together effectively.

- Healthcare data privacy in the digital age encounters several significant challenges. In order to tackle these difficulties, it is necessary to adopt a comprehensive and cooperative approach that engages several stakeholders, including legislators, regulators, providers, researchers, patients, and technology developers. Furthermore, additional research and innovation are required to provide efficient and morally sound solutions that can effectively manage the advantages and drawbacks of healthcare data, all the while upholding the privacy and rights of patients.

2.2.EHR and IIoT

The Industrial Internet of Things (IIoT) system typically consists of numerous IoT devices that are distributed throughout the whole industrial system. The concept encompasses the interconnectivity of medical devices and sensors, which facilitates the generation of real-time data to optimize patient monitoring and treatment. The sector encompasses numerous applications, including but not limited to remote surveillance, telemedicine, intelligent watches, and implanted technologies. The objective of incorporating these technologies is to establish a cohesive and all-encompassing healthcare system, thereby enhancing the accuracy of diagnoses, the efficacy of treatments, and the overall well-being of patients. These devices have the capability to produce and send substantial quantities of sensitive health information, which can be incorporated into EHR to deliver a more comprehensive and individualized form of healthcare. Nevertheless, the integration of EHR into healthcare systems also brings up novel dangers and weaknesses pertaining to the security of these records. An instance of unauthorized individuals might exploit vulnerabilities in equipment or communication channels to gain unauthorized access, manipulate, or remove health data. In addition, they possess the capability to initiate denial-of-service assaults, causing disruption to the accessibility and operational effectiveness of the devices or EHR systems. Furthermore, individuals have the potential to utilize the data for illicit activities such as identity theft, fraudulent schemes, blackmail, or other malevolent intentions.

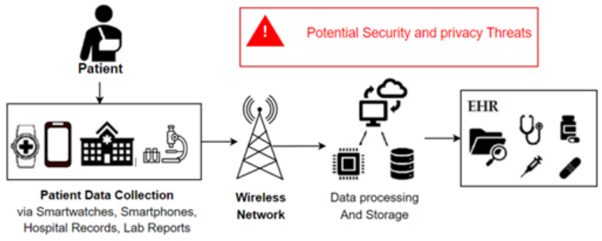
Patient
Potential Security and privacy Threats
Patient Data Collection
via Smartwatches, Smartphones, Hospital Records, Lab Reports
Wireless Network
Data processing And Storage
EHR

Fig 2: EHR and IIoT: Potential of Security and Privacy Threats.

2.3. Deep Learning in the EHR Field

Electronic health record (EHR) systems were originally designed to handle hospital basic administrative functions, allowing for the use of controlled terminology and labelling conventions [15–17]. These labeling systems lead to the creation of standardized datasets for different domains. The quantity of EHR data is progressively growing over time due to advancements in the EHR system. Consequently, there has been a proliferation of research studies exploring the potential secondary utilization of these data. Nevertheless, EHRs have notable security vulnerabilities due to their inclusion of confidential personal and financial information, which can be illicitly accessed, pilfered, or tampered with by unauthorized entities. Hence, safeguarding the security of EHR data and guaranteeing its private nature, credibility, and accessibility is of utmost importance [18-20].

In order to guarantee the security of healthcare systems, deep learning is crucial [21]. EHRs comprising both structured and unstructured data, including prescriptions, lab test results, diagnosis information, and clinical notes, can be evaluated by deep learning algorithms [22]. These models are able to do this analysis at remarkably fast speeds without sacrificing accuracy. Deep learning networks are transforming healthcare delivery and are essential to the integration of health systems in clinical settings. Researchers have looked into how EHRs could automate the prescription process, giving doctors more information to consider when choosing and writing prescriptions. EHRs provide a full overview of patients' medical histories, encompassing many elements such as prior prescriptions, assessments, tests conducted by laboratories, medication regimens, and diagnostic imaging scans [23]. These entities serve as the primary conduits for individualized medical research data [24]. Furthermore, the recent advancements in the quality of EHRs have garnered the attention of researchers, primarily due to their prospective uses in the fields of medical diagnosis and recommendation [25].

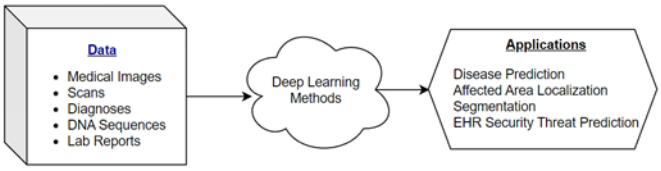

Fig 3: Application of Deep Learning in the EHR Field.

These algorithms can be employed to detect fraudulent insurance claims and forecast forthcoming risks [26]. Thanks to the emergence of telemedicine, wearables, and remote patient monitoring, the models can now be utilized for real-time patient monitoring and risk prediction [27]. Furthermore, the models can be utilized to identify and proactively mitigate cyber assaults on healthcare systems [28]. They possess the ability to discern recurring behavioural patterns that suggest a possible danger and promptly notify security personnel to respond accordingly [29]

3. Literature Review

Deep learning has the ability to drastically change the medical and healthcare sectors. Unfortunately, a number of inference attack models have demonstrated that deep learning may jeopardise critical patient data. The deep neural networks' large capacity is the main reason behind the loss of privacy. To be more precise, a deep network may unintentionally memorise patient data that is included in the training set. Comparing health data transportation and storage to traditional public blockchain solutions reveals a distinct set of difficulties. Healthcare data has unique privacy and security requirements compared to other forms of data [30]. Zhang et al. introduced an approach for safeguarding privacy when training deep neural networks. This method incorporates the use of declining Gaussian noise applied to the gradients [31]. Alzubi et al. proposed a novel strategy to protect the privacy of electronic health records that blends deep learning and blockchain technologies. The system is able to identify and remove anomalous individuals from the database and guarantees restricted access to health records by combining blockchain technology with an encryption-based federated-learning module [32]. The Collective Learning protocol was created by Paul et al. to safely exchange classified time-series data across entities for the purpose of partially training a binary classifier network's variables [33]. Using differential-privacy, Beaulieu-Jones et al. have created a unique distributed training technique with cycle weight transfer to meet the need for formal privacy assurances [34]PriMIA is a complementary and easily-accessible software framework developed by Kaissis et al. to perform encryption-based and differentially private, securely integrated federated-learning inference on clinical imaging datasets [35]. An intriguing first step in resolving privacy concerns with EHR data is to rely on synthetic data. Yoon et al. proposed EHR-Safe, a generative modelling technique to generate realistically synthetic EHR data while maintaining patient privacy. Sequential encoder-decoder networks and generative adversarial networks form the core of this paradigm.

4. Methodology

4.1. Basic Architecture

In the first stage, every healthcare facility conducts training for its local model using its own EHR data. Next, the local model undergoes encoding through the utilization of homomorphic-encryption. The approach is predicated on the premise of a synchronous federated-learning framework, in which a central server has the responsibility of orchestrating the entire procedure. The incorporation of the differential-privacy noise term is intended to provide protection against potential attempts to reconstruct the individual's identity. The utilization of suitable encrypted protocols, including encrypted communication channels and cryptography libraries, is of utmost importance to guarantee a seamless implementation.

4.2. Federated-learning

Federated-learning refers to a machine learning approach that eliminates the requirement for data exchange between multiple distributed workstations or servers that each own local data specimens and allows the training of a model on each one. Each contributing machine uses its local data for carrying out model changes in a typical Federated-learning instance. Subsequently, these modifications are consolidated to enhance the global model. If w is the global-model here, then is the local model update from the device i. Equation 1 represents the global-model update obtained through the aggregation of local changes.

$$w_{global} = \sum_i w_i \qquad (1)$$

The method operates under the assumption of a synchronous federated-learning framework, wherein a central-server is accountable to coordinate the entire process. Properly setting the learning rate η and other hyperparameters is necessary for convergence. The technique serves as the fundamental framework of federated-learning , wherein multiple devices work together to enhance a global-model with no need to exchange the raw data.

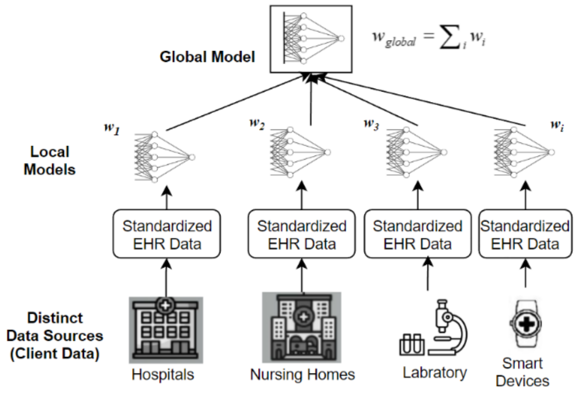

Fig 4: Basic Framework of the Federated-learning Model.

An increased client count can enhance the diversity of datasets, hence potentially enhancing the resilience and generalizability of the global model. As the client count grows, the communication overhead proportionally increases. Every individual client is required to transmit its model updates to the central server, and the server is responsible for consolidating these updates. It is necessary to make trade-off a balance which minimizes the expenses associated with communication while maximizing the advantages of having a wide range of data. The computing capacities of each client are a determining factor in the number of clients. Insufficient computational resources across clients can result in longer training times or higher latency during the federated-learning process when there are numerous clients. The number of clients can impact the stability of the federated-learning system. The presence of a large number of clients might lead to increased unpredictability in model updates, posing a challenge to achieving convergence.

4.3 Homomorphic-encryption with Differential-privacy

Homomorphic-encryption enables to carrying out of computations on the encrypted data without the need for decryption. The aforementioned concept holds significant implications within the realm of EHR, wherein the safeguarding of sensitive patient data from unauthorized access is paramount. However, it is equally imperative to ensure that said information may be effectively utilized for a multitude of objectives, including but not limited to billing, analytics, research, and diagnosis. The utilization of Homomorphic-encryption allows for the delegation of data processing tasks to external entities in the context of EHR systems while maintaining the integrity of privacy and security measures. As an illustration, a healthcare facility has the capability to employ encryption techniques to secure its EHR data. After that, the cloud service provider can receive the encrypted data and use it to do any necessary actions. The provider is able to do these operations without accessing the original unencrypted data, thereby ensuring the confidentiality of the plaintext information. Subsequently, the institution possesses the capability to decipher the obtained data and subsequently employ it for its designated objective.

In order to execute a secure aggregate while preserving the confidentiality of each of the model revisions, we make use of homomorphic-encryption. Let us take $E(x)$ is the Homomorphic-encryption of input EHR data (x). If *p* and *q* are plaintext input data, then the encrypted data can be represented as follows.

$$Encrypted\_Information = E(p+q) = E(p).E(q) \quad (2)$$

Here $E(w_i)$ is the Homomorphic-encryption update of a local model w_i .

$$E(w_{global}) = \sum_i E(w_i) \quad (3)$$

Here, $E(w_{global})$ represents the encrypted global-model update.

This enables the central server to acquire the aggregated model update without being granted accessibility to any specific variables in the model.

In order to augment the level of privacy, the implementation of differential-privacy is proposed, which involves the incorporation of noise into the aggregated model update. Differential-privacy is a method employed to safeguard the privacy of people inside a dataset through the introduction of random-noise to any of the data or the queries. Homomorphic-encryption is a cryptographic method that enables the execution of calculations on encrypted data without the need for decryption. The integration of differential-privacy and Homomorphic-encryption can facilitate the achievement of safe and confidential data analysis in situations where the entities responsible for data ownership and data analysis are distinct. As an illustration, a medical facility has the capability to employ encryption techniques to safeguard the confidential medical records of its patients. These encrypted records can then be transmitted to an external researcher, who can conduct statistical analysis of the encrypted data by utilizing Homomorphic-encryption methods. The researcher has the option to introduce noise into the obtained results in order to guarantee differential-privacy, subsequently transmitting them back to the hospital. The decryption of the results by the hospital enables the acquisition of valuable insights while maintaining the confidentiality of both patient and researcher data. This noise can be described by mathematical equations as below.

Let ϵ denote the privacy parameter, and let N represent the noise term that is randomly selected from a Laplace distribution. Reduced values of ϵ are associated with heightened levels of privacy assurances. This occurrence implies that while having access to the collective update, it becomes difficult to determine the individual device’s impact due to the introduction of additional noise.

$$E(w_{global})_{DP} = E(w_{global}) + \mathrm{N}(\varepsilon, \Delta\varepsilon) \quad (4)$$

4.4 Multi-party Decryption for Model Update

While homomorphic-encryption protects data while it is being computed, the process of decrypting it presents a risk to the system's overall security. The application of multi-party computation is used to address this problem. This protocol is a cryptographic approach that allows many parties to cooperatively compute a function while maintaining the privacy of each party's input. In the case of EHR, this approach is employed to assure the distribution of the decryption key among many entities, hence necessitating collaborative efforts to acquire the ultimate decrypted outcome. If $x_1, x_2, x_3, ...x_n$ are confidential input from distinct device parties, then the function that should be decrypted by each party without knowing the input data can be presented as follows:

$$Decryption\_function = f(x_1, x_2, x_3, ...x_n) \quad (5)$$

The approach employed in this study utilizes a secure multi-party computation protocol, enabling many entities to collaboratively decrypt the model parameters while maintaining the confidentiality of their respective secret keys. The method ensures the preservation of privacy for both the encrypted data and the model parameters. When the central server receives an encrypted global-model upgrade that is differentially confidential, it decrypts the update to get the final version.

$$w_{final} = D(E(w_{global})_{DP}) \quad (6)$$

The aforementioned model update is thereafter employed to enhance the overall model without disclosing any particular information on individual contributions. Equation 7 represents the comprehensive mathematical formulation of the privacy-preserving Federated-learning update, incorporating the utilization of Homomorphic-encryption and differential-privacy techniques.

$$w_{final} = D(\sum_i E(w_i) + \mathrm{N}(\varepsilon, \Delta\varepsilon)) \qquad (7)$$

The privacy of individual EHR readings of a case is preserved by computing the average on the encrypted numbers. The decryption procedure, which entails acquiring the average of the plaintext, necessitates cooperation among the participants engaged in the multi-party computation, thereby ensuring that no individual entity can gain entire access to the decrypted outcome.

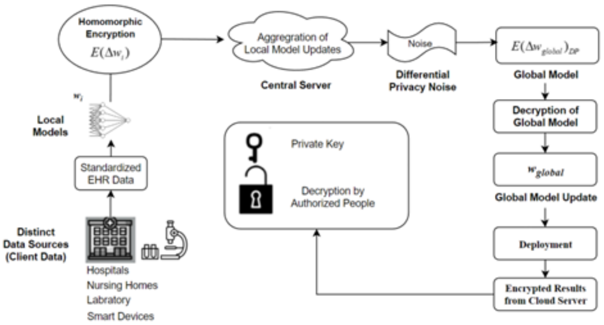
Homomorphic Encryption
$E(\Delta w_i)$
Aggregration of Local Model Updates
Central Server
Noise
Differential Privacy Noise
$E(\Delta w_{global})_{DP}$
Global Model
Decryption of Global Model
w_{global}
Global Model Update
Deployment
Encrypted Results from Cloud Server
w_i
Local Models
Standardized EHR Data
Private Key
Decryption by Authorized People
Distinct Data Sources (Client Data)
Hospitals
Nursing Homes
Labratory
Smart Devices

Fig 5: The Full Algorithm Overview Flow Diagram

4.5 Overall Algorithm

The overall algorithm is discussed here. Initially, each healthcare institution trains its local model w_i on its EHR data. Then, the local model is encoded using homomorphic-encryption. The method operates under the assumption of a synchronous federated-learning framework, wherein a central server is responsible for coordinating the overall process. The inclusion of the differential-privacy noise term N(ϵ, Δϵ) serves the purpose of safeguarding against potentially recovering its identity attempts. It is imperative to utilize appropriate encrypted protocols, such as encrypted channels of communication and cryptography libraries, in order to ensure an uninterrupted deployment.

The Federated-learning with Homomorphic-encryption Algorithm:

```
1: Initiate model:
        Select n clients (healthcare institution).
        Each healthcare institution initializes a local model $w_i$ and encrypts it: $E(w_i)$.
2: Iterative Federated-learning :
        Each healthcare institution trains its local model $w_i$ on its EHR data.
3:      Homomorphic-encryption of Local Model Update: $E(\Delta w_i)$
4:      Transmit $E(\Delta w_i)$ to the central server in a secure way.
5:      At the central server, aggregate the encrypted local updates: $E(\Delta w_{global}) = \sum_i E(\Delta w_i)$
6:      Add differential-privacy noise: $E(\Delta w_{global})_{DP} = E(\Delta w_{global}) + \mathrm{N}(\varepsilon, \Delta\varepsilon)$
7: Decrypt the Global Model:
        Deciphering the differentially confidential, aggregated model update-
            $\Delta w_{global} = D(E(\Delta w_{global})_{DP})$
8: Update Global Model: $w_{global} = w_{global} + \Delta w_{global}$
9: This process (Steps 2-8) should be repeated numerous times till the overall model converges.
10: Prediction
        Deploy the final global-model ( $w_{global}$ ) to healthcare institutions securely.
11:     If necessary, aggregate predictions from multiple institutions without exposing individual predictions.
12:     Secure Model Prediction.
```

5. Results and Discussion

5.1 Dataset

The Premier Healthcare database, widely recognised as one of the biggest clinical databases in the US, provided the EHR data used in this investigation. It is a 12-month project that includes data from millions of patients. 415 institutions from across the United States provided the data [36]. According to some, these facilities represent the general experience of hospitals in the US [36]. Discharge files, which are official records of all chargeable items such as medicine, laboratory usage, and therapeutic and diagnostic services, are provided by each hospital listed in the database. Each of these things has a connection to a particular patient's admittance [36]. The EHR data of 1,271,733 hospital admissions are initially present in the database we used for our analysis without any pre-processing. Because of the wide variety of drugs that might be prescribed, every patient has a total of 24,428 features.

Experimental setup

The architecture of our model is a fully connected neural network consisting of an input layer, 20 hidden layers, with each having 512-dimensions, as well as scalar output-layer. For binary classification, At the output layer, we make use of the logistic function and log loss. The model is optimised with a mini-batch size of 128 instances using the Adam optimizer. The default value of 0.001 was utilised as the learning rate. The weights of the model were periodically taken at 200 mini-batch iterations, and the final model was selected retrospectively from the snapshot with the best performance on the validation test. No clear regularisation was thought to be required. A comprehensive investigation of hyperparameters, such as various network depths (from 2 to 32) and the 'ReLU' activation function, was used to select the network configuration. The Python programming language was utilized in the development of the software..

The testing were carried out on an Ubuntu 18.04 LTS server equipped with two 16GB NVIDIA P5000 GPU cards, an Intel(R) CPU running at 2.21GHz, 190GB of RAM, and 190GB of RAM. To run our models and tests, we use NumPy version 1.14.1 and Keras version 2.2.1 with a TensorFlow backend version 1.12.0. To make it easier to replicate our findings, we run our code within a Docker container using Python 3.6.2.

5.2 Results

The value of η was adjusted from 0.01 to 0.08 for every scheme, with an increment of 0.01. A batch was selected from all the batches in this study, where a client's data size can be different from others. Within the federated schemes, an extra parameter exists known as the number of global rounds, denoted as TCL. The federated design selects clinical facilities from a range of options, specifically [1,2,3,4,5]. We must select a size that is sufficiently large to avoid impeding the rate of convergence while also being small enough to prevent compromising privacy by increasing the likelihood of sampling.

We used balanced accuracy, precision, and area under the curve as metrics in this study. Balanced accuracy serves as a suitable statistic for evaluating the performance of a model in binary classification. The metric in question is the average of sensitivity and specificity. It is particularly useful when working with imbalanced data, where one of the target classes is significantly more prevalent than the other.

$$Balanced\_accuracy = \frac{sensitivity + specificity}{2} \quad (8)$$

Sensitivity, also known as the true positive rate or recall, quantifies the percentage of actual positive instances that are accurately predicted out of all the positive predictions that the model can make. Specificity is a statistical measure that quantifies the accuracy of identifying negative outcomes by a model. It is calculated by dividing the number of successfully detected negatives by the total number of negative predictions that might be made.

A single federated run consists of 200 cycles, and the highest and lowest scores of every performance parameter among the ten federated runs are recorded. During each round, a random selection is made to aggregate three medical institutions. The outcomes of this study are presented in Table 1.

Algorithm Trials	Privacy	AUC	Balanced Accuracy	Precision
Trial 1	ε=0	0.731	0.72	0.76
Trial 2	ε=1	0.854	0.86	0.82
Trial 3	ε=2	0.872	0.87	0.85
Trial 4	ε=3	0.885	0.85	0.86
Trial 5	ε=4	0.892	0.88	0.85

Table 1. Outcomes of Federated-learning Model in Different Trials.

When compared to local training, Federated-learning experiences a decrease in performance because of the need for network communications. This decrease is further amplified by the inclusion of Homomorphic-encryption and Multi-party decryption, resulting in a threefold rise in the timeframe required for training. Training large neural network topologies takes more time due to the need for network transfer. We have executed four distinct trials for this study depending on the degree of privacy. We used the Premier Healthcare database for this study investigation. It was observed that trial 1, with almost no privacy layer employed, reached a balanced accuracy of 72%, with a precision score of 0.76. When we started increasing the privacy score criteria, the performance of the model started to improve. When the value of ε increased to 2, the balanced accuracy reached a score of 87%. In the end, it was observed that when the privacy parameter reaches a score of 5, the model performance is far better. The balanced accuracy achieves a score of 88%. The value of the area under the precision-recall curve was 0.892 at that time.

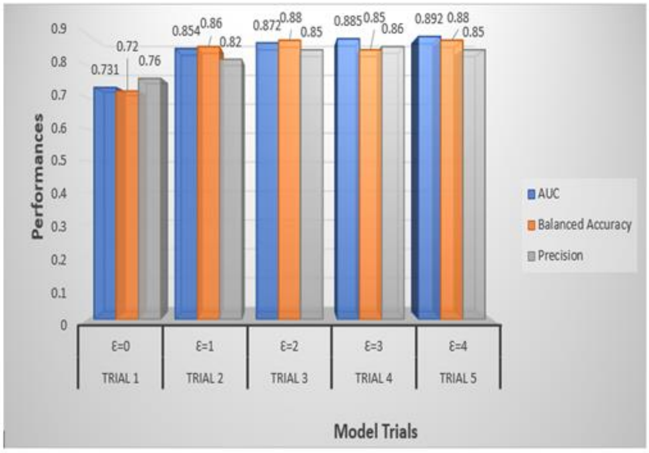

Fig 6. Model Performances at Various Conditions.

Consider conducting a more in-depth analysis of the AUC and the receiver operating characteristic (ROC) curves that were produced for the various methods and for the round that yields the most favorable outcomes in terms of AUC for the various trials. During trial 5, it was noted that the AUC reached its highest value of 0.892. This suggests that the trial requires a shorter amount of time to compute compared to the other situations.

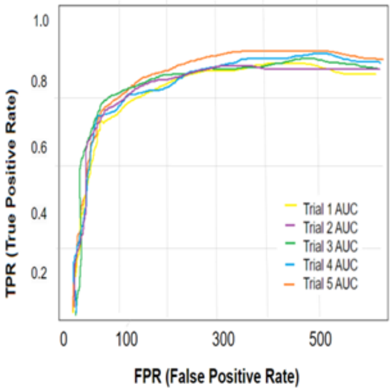

Fig 7. ROC curve for Each trial in this study.

6. Conclusions

The possible integration of widespread data sharing in distributed network systems is expected to generate difficulties regarding privacy, confidentiality, and authentication within the domain of cyberspace. The main goal of this study is to protect the privacy and confidentiality of data inside an insecure environment while exchanging audiovisual content between two IoT nodes. In order to successfully combat an opponent and ensure the secrecy of information, we propose the adoption of a robust security approach that utilizes a multi-tiered system based on the concepts of concealing information and chaos theory. Although certain block-based resilient data hiding schemes that rely on the transform domain have shown promising results, their inefficient block and coefficient selection methods result in insufficient performance against common cyber-attacks. Hence, we provide a Resilient Framework for Guaranteeing Data Confidentiality and Security in EHR-based networks by employing federated-learning with homomorphic-encryption. In this case, a differential-privacy strategy is employed to enhance privacy by introducing random perturbations to the aggregated model update. The suggested model exhibits superior performance compared to the most advanced methods now available for ensuring data confidentiality and security in networks based on EHR. Our system incorporates sophisticated methods such as homomorphic-encryption, safe multi-party computation, differential-privacy, and federated-learning to enable efficient and secure computation on encrypted data. The article provides a thorough investigation of the compromises between privacy, accuracy, and efficiency within our framework, including both theoretical and empirical analysis. Federated-learning is a decentralized approach that, when combined with homomorphic-encryption, provides an enhanced level of security to safeguard the privacy of patients. Upon observation, it was noted that the model's performance significantly improves when the privacy parameter achieves a value of 5. The balanced accuracy score is 88%. The precise recall curve had an area of 0.892 at that specific time.

There can be several limits in the implementation of this model. When dealing with uneven data, the use of differential-privacy might result in the model becoming overly focused on the majority class, which can have negative effects on both fairness and effectiveness.Ensuring compliance with privacy rules, such as HIPAA, during the implementation of privacy-preserving measures necessitates meticulous deliberation. Achieving a harmonious equilibrium between safeguarding privacy and adhering to regulatory requirements might prove to be a formidable task. Federated-learning (FL) is an advantageous method that allows for cooperative learning from decentralized data sources, eliminating the need for data centralization. Nevertheless, FL still discloses certain information regarding the local data through the exchanged model parameters or gradients. In the proposed model, Homomorphic-encryption enables the encryption of data while allowing computations to be performed on the encrypted data, hence maintaining data confidentiality. Privacy. DP introduces random perturbations into the model updates to safeguard against inference assaults relying on statistical analysis.MPD facilitates the secure combination of encrypted model updates while keeping the individual contributions undisclosed. We assess the performance of our system using a standardized dataset of EHRs and demonstrate that it attains exceptional levels of accuracy and privacy while also diminishing the expenses associated with communication and computing in comparison to current approaches.

Acknowledgements

This research was supported by Global College of Engineering and Technology by providing research laboratory and funds. We thank our colleagues from Global College of Engineering and Technology. who provided insight and expertise that greatly assisted the research. We thank Nadeem Akhtar, Assistant Professor for assistance with literature review, methodology, and Raza Hasan, Senior Lecturer, Solent University,UK for comments that greatly improved the manuscript.

Author contributions

Muzammil Hussain: Conceptualization, Methodology, Software, Field study **Nadeem Akhtar:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Raza Hasan:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] R. Agrawal and S. Prabakaran, “Big data in digital healthcare: lessons learnt and recommendations for general practice,” -*Heredity,* Vol.-124, no. 4, pp. 525-534, 2020.

[2] P. Muralidhara, “IoT applications in cloud computing for smart devices,” *International Journal of Computer Science Technology,* Vol.-1, no. 1, pp. 1-41, 2017.

[3] B. Sharma and M. S. Obaidat, “Comparative analysis of IoT based products, technology and integration of IoT with cloud computing,” *IET Networks,* Vol.-9, no. 2, pp. 43-47, 2020.

[4] R. Arshad, *et al.,* “Green IoT: An investigation on energy saving practices for 2020 and beyond,”*IEEE Access,* Vol.-- 5, pp. 15667-15681, 2017.

[5] G. Lulla, *et al.,* “IoT based smart security and surveillance system,” in *2021 international conference on emerging smart computing and informatics (ESCI)*, 2021, pp. 385-390: IEEE.

[6] M. Valinejadshoubi, *et al.,* “Development of an IoT and BIM-based automated alert system for thermal comfort monitoring in buildings,”*Sustainable Cities,* Vol.- 66, p. 102602, 2021.

[7] S. S. Vedaei *et al.*, “COVID-SAFE: An IoT-based system for automated health monitoring and surveillance in post-pandemic life,”*IEEE access,* Vol.- 8, pp. 188538-188551, 2020.

[8] D. Mrňa, B. Badánik, and A. Novák, “Internet of things as an optimization tool for smart airport concept,”*European Transport-Trasporti Europei,* Vol.- 82, 2021.

[9] S. A. ElRahman and A. S. Alluhaidan, “Blockchain technology and IoT-edge framework for sharing healthcare services,”*Soft Computing,* Vol.- 25, no. 21, pp. 13753-13777, 2021.

[10] J. L. Shah, *et al.,* “Integration of Cloud and IoT for smart e-healthcare,” in *Healthcare paradigms in the internet of things ecosystem*: Elsevier, 2021, pp. 101-136.

[11] L. Sha, F. Xiao, W. Chen, and J. Sun, “IIoT-SIDefender: Detecting and defense against the sensitive information leakage in industry IoT,”*World Wide Web,* Vol.- 21, pp. 59-88, 2018.

[12] L. D. Serbanati, “Health digital state and Smart EHR systems,”*Informatics in Medicine Unlocked,* Vol.- 21, p. 100494, 2020.

[13] W. Bani Issa *et al.*, “Privacy, confidentiality, security and patient safety concerns about electronic health records,”*International nursing review,* Vol.- 67, no. 2, pp. 218-230, 2020.

[14] M. Mahmud, M. S. Kaiser, T. M. McGinnity, and A. J. C. c. Hussain, “Deep learning in mining biological data,” Vol.- 13, pp. 1-33, 2021.

[15] M. J. Bowie, "*Understanding Current Procedural Terminology and HCPCS Coding Systems," 2021*. Cengage Learning, 2021.

[16] J. Xu, *et al.,* “A survey of deep learning for electronic health records,”*Applied Sciences,* Vol.- 12, no. 22, p. 11709, 2022.

[17] S. K. Tayebati, *et al.,*, “Identification of World Health Organisation ship’s medicine chest contents by Anatomical Therapeutic Chemical (ATC) classification codes,”*International maritime health,* Vol.- 68, no. 1, pp. 39-45, 2017.

[18] N. N. Basil, S. *et al.,* “Health Records Database and Inherent Security Concerns: A Review of the Literature,”*Cureus,* Vol.- 14, no. 10, 2022.

[19] S. Qu, “Combining Structured and Unstructured Data in Electronic Health Record for Readmission Prediction via Deep Learning,” The Ohio State University, 2020.

[20] I. Li *et al.*, “Neural natural language processing for unstructured data in electronic health records: A review,”*Computer Science Review,* Vol.- 46, p. 100511, 2022.

[21] S. Qamar, “Healthcare data analysis by feature extraction and classification using deep learning with cloud based cyber security,”*Computers Electrical Engineering,* Vol.- 104, p. 108406, 2022.

[22] D. Zhang, *et al.,* “Combining structured and unstructured data for predictive models: a deep learning approach,”*BMC medical informatics decision making,* Vol.- 20, no. 1, pp. 1-11, 2020.

[23] Y. Zhang, *et al.,* “LEAP: learning to prescribe effective and safe treatment combinations for multimorbidity,” in *proceedings of the 23rd ACM SIGKDD international conference on knowledge Discovery and data Mining*, 2017, pp. 1315-1324.

[24] C. Su, *et al.,* “GATE: graph-attention augmented temporal neural network for medication recommendation,”*IEEE Access,* Vol.-8, pp. 125447-125458, 2020.

[25] A. Ponselvakumar *et al.,* “Advancement in precision medicine and recommendation system for clinical trials using deep learning methods,” in *IOP conference series: materials science and engineering*, 2021, Vol.- 1055, no. 1, p. 012110: IOP Publishing.

[26] G. Zhang, *et al.,* “Identifying fraud in medical insurance based on blockchain and deep learning,” *Future Generation Computer Systems,* Vol.- 130, pp. 140-154, 2022.

[27] [27]A. Meyer *et al.*, “Machine learning for real-time prediction of complications in critical care: a retrospective study,”*The Lancet Respiratory Medicine,* Vol.- 6, no. 12, pp. 905-914, 2018.

[28] S. Silvestri, *et al.,* “A Machine Learning Approach for the NLP-Based Analysis of Cyber Threats and

Vulnerabilities of the Healthcare Ecosystem," *Sensors,* Vol.- 23, no. 2, p. 651, 2023.

[29] L. Priya, *et al.,* "Deep Learning in Healthcare," *Deep Learning Edge Computing Solutions for High Performance Computing,* pp. 121-133, 2021.

[30] D. Berdik, *et al.,* "A survey on blockchain for information systems management and security," *Information Processing Management,* Vol.- 58, no. 1, p. 102397, 2021.

[31] X. Zhang, J. Ding, M. Wu, S. T. Wong, H. Van Nguyen, and M. Pan, "Adaptive privacy preserving deep learning algorithms for medical data," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 1169-1178.

[32] J. A. Alzubi, *et al.,* "Cloud-IIoT-based electronic health record privacy-preserving by CNN and blockchain-enabled federated-learning ," *IEEE Transactions on Industrial Informatics,* Vol.- 19, no. 1, pp. 1080-1087, 2022.

[33] J. Paul, *et al.,* A. Al Badawi, B. Veeravalli, and K. M. M. Aung, "Privacy-preserving collective learning with homomorphic-encryption,"*IEEE Access,* Vol.- 9, pp. 132084-132096, 2021.

[34] B. K. Beaulieu-Jones, *et al.,* "Privacy-preserving distributed deep learning for clinical data," *arXiv preprint arXiv:.01484,* 2018.

[35] G. Kaissis *et al.*, "End-to-end privacy preserving deep learning on multi-institutional medical imaging,"*Nature Machine Intelligence,* Vol.- 3, no. 6, pp. 473-484, 2021.

[36] A. Fejza, *et al.,* "Scalable and interpretable predictive models for electronic health records," in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 2018, pp. 341-350: IEEE.