Jordan Medical Journal, Supplement 1, 2024

ORIGINAL ARTICLE

Enhancing Security and Privacy in Healthcare with Generative Artificial Intelligence-Based Detection and Mitigation of Data Poisoning Attacks Software

Yasmin Makki Mohialden [*1], Saba Abdulbaqi Salman[2], Maad M. Mijwil[3,4], Nadia Mahmood Hussien[5], Mohammad Aljanabi[6], Mostafa Abotaleb[7], Klodian Dhoska[8], Pradeep Mishra[9]

[1] Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, Iraq

[2,5] Department of Computer Science, College of Education, Al-Iraqia University, Baghdad, Iraq

[3] College of Administration and Economics, Al-Iraqia University, Baghdad, Iraq

[4] Computer Techniques Engineering Department, Baghdad College of Economic Sciences University, Baghdad, Iraq

[6] Imam Ja'afar Al-Sadiq University, Baghdad, Iraq

[7] Department of System Programming, South Ural State University, Chelyabinsk, Russia

[8] Department of Production and Management, Polytechnic University of Tirana, Albania

[9] College of Agriculture, Rewa, JNKVV, (M.P.) India

*Corresponding author:
ymmiraq2009@uomustansiriyah.edu.iq

Received: May 29, 2024

Accepted: October, 13, 2024

DOI:
https://doi.org/10.35516/jmj.v58i3.2712

Abstract

This study investigated an advanced approach to enhancing security and privacy in healthcare by incorporating artificial intelligence (AI)-based strategies to detect and mitigate data poisoning attacks. The proposed method combined unified learning, homomorphic encryption, and autoencoder-based anomaly detection. It ensured that models were trained in diverse places, protected data, and improved model security. Anomaly identification and mitigation and data poisoning resistance were investigated using simulated medical data. Main results. This approach visualized and assessed model performance. This study offers a complete solution to securing medical data and models against new threats.

Keywords: Homomorphic encryption, Federated learning, Data poisoning attacks, Simulation healthcare security, Anomaly detection.

1. INTRODUCTION

Artificial intelligence (AI) techniques are currently among the most important applications in the field of data security and healthcare services. However, beyond AI techniques such as machine learning, deep learning, and expert systems, emerging possibilities with generative AI may also advance tasks by generating applications that improve services and accomplish tasks [1][2]. AI-powered chatbots are significant in growing healthcare services as they are designed to generate texts, images and music from trained parameters. Applications such as ChatGPT, Google's Bard, Microsoft Bing, Chatsonic, and Github Copilot are large language models (LLMs) and multiple neural networks that are constantly trained to emulate the human mind and make decisions with high speed and accuracy. LLMs are one of the most critical tools that rely on artificial intelligence in training large amounts of data and are used in many sectors. Generative AI models trained on medical data and patient records are expected to predict and detect changes in a dataset [3-5]. Therefore, these models will be able to interpret and analyze all data with high efficiency. Digital healthcare systems increase data breaches, theft, and other security risks. Data security and privacy increase with generative AI, federated learning, and homomorphic encryption. This research presents a data poisoning detection and mitigation strategy to strengthen healthcare AI models. The main contributions applied in this work are as follows: (1) Creating a secure model training method using federated learning and homomorphic encryption; (2) Identification and mitigation of data poisoning threats using autoencoder-based anomaly detection; (3) Reviewing the proposed strategy using simulated medical data [6-20].

Generative AI can identify anomalies or outliers that may indicate a data poisoning attack, regenerate clean versions of corrupted data, and simulate various types of data poisoning attacks to understand potential vulnerabilities [21][22]. This paper advances AI security and privacy research.

2. METHODS

Federal learning, homomorphic encryption, and autoencoder-based anomaly detection are proposed. Federated learning keeps data decentralized, minimizing exposure (risk). Data computations using homomorphic encryption improve security. Unusual data points are detected by the autoencoder, especially a variational autoencoder (VAE).

Autoencoder Architecture

1. The VAE has an encoder and decoder. Encoders compress incoming data into latent spaces, which decoders recover. VAE important parameters are:
   - Number of layers: 3 (input, hidden, and output)
   - Layer types: dense and dropout
   - Size of the encoding dimension: 64
2. Federated Learning Parameters: federated learning uses a client-server. Clients train and update local models on their data. Updates are combined by the server to enhance the global model. Some key parameters are:
   - Number of clients: 10
   - Communication rounds: 50
3. Homomorphic Encryption: this technique allows computations on encrypted data without decryption. It ensures data privacy during the training process. We use a lattice-based encryption scheme for its efficiency and security [23-26].

This work consists of four major steps

that are implemented:

1. Data Generation and Poisoning Simulation: Synthetic healthcare data is generated to simulate typical datasets. This includes attributes like age, gender, BMI, blood pressure, cholesterol levels, and health labels. A portion of the data is perturbed with Gaussian noise to simulate data poisoning attacks [27-31].

2. Anomaly Detection: Anomalies are detected by comparing the mean squared error (MSE) between the original and reconstructed data. A threshold is determined to identify significant deviations, indicating potential anomalies [32].

3. Train-Test Split: Data is standardized using the StandardScaler and split into training and testing sets using the train_test_split function from scikit-learn. The proportion of data split is 80% for training and 20% for testing.

4. Data Visualization and Evaluation: Visualization techniques are used to compare original and reconstructed data. Performance metrics such as accuracy, precision, recall, and F1 score are calculated to evaluate the model's effectiveness.

This section will review the most important studies that include AI models. Khan et al. [33], suggests using lightweight RFID methods to secure healthcare systems. Additionally, Olsen [8], considers using generative AI models to secure healthcare data. This research stresses the need to use powerful AI to protect healthcare data. A study by Seh et al. [34], provides a fresh and supplementary strategy to govern data-sharing and improve medical data security. This article evaluates research papers on healthcare data breach stories to highlight current privacy and confidentiality concerns of sensitive healthcare data. According to research, healthcare businesses require proactive security measures to detect anomalous user behavior while accessing healthcare data. Studies also imply that ML approaches will protect healthcare data well. Moving forward, this study presents a conceptual framework to proactively protect healthcare data privacy and confidentiality. ML approaches identify deviating user access against Electronic Health Records in the proposed framework. In addition, the fuzzy-based Analytical Network Process (ANP), a multi-criteria decision-making method, evaluates the supervised and unsupervised ML techniques for dynamic digital healthcare data security. A study by Marulli et al. [35], examined the sensitivity of AI techniques to corrupted data to assess their dependability and resilience. This study aimed to assess the durability and dependability of artificial intelligence techniques by thoroughly analyzing how sensitive they are to tainted data. These systems must be able to recognize what is wrong, work out a solution to the ensuing issues, and then apply what they have learnt to conquer those obstacles and strengthen their resilience. By contrasting multiple models requested with reliable and tainted data, the primary research objective was to assess the susceptibility and receptivity of machine learning systems for poisoning signals. An example of the medical field was offered to help with the studies that were being conducted. The study campaign's outcomes were also assessed using ROC, F1-score, precision, specificity, and sensitivity.

Aceto et al. [36], provided a generative AI model that synthesizes anonymized traffic traces from genuine ones, addressing privacy, abundance, and representativeness. A conditional variational autoencoder (CVAE) and a preprocessing approach for traffic trace

creation underpin the idea. This method is validated by thorough empirical research using three current and publicly available datasets of benign and malicious traffic. We compare the classification performance of a robust NIDS and the quality of synthetic data to actual data for validation. In this paper, CVAE is compared to two cutting-edge AI-based traffic data generators. It is shown that NIDS trained with traces sent by our generative model have a smaller F1-score loss than real data. Competing models struggle or fail to generate traces as effective for NIDS training and statistically similar to the original. To enable reproducibility and stimulate generative AI for networking research, the author provides synthetic datasets in PCAP and tabular formats. Khalid et al. [37], this literature has focused on privacy-preserving methods and resolving clinical AI adoption barriers. This paper highlights current privacy methods for AI-based healthcare applications. Federated Learning and Hybrid Techniques are discussed together with privacy threats, security issues, and future directions. Healthcare privacy-preserving methods are thoroughly reviewed in this work. And created a privacy threat taxonomy and showed how to safeguard healthcare datasets and AI models. Finally, this paper examined privacy-preserving machine learning (PPML) approaches' limitations and hazards and identified several unresolved research topics.

The paper is organized in the following sections: Section 2 discusses the proposed method with the main steps. Section 3 discusses the results obtained. Finally, section 4 concludes the study with the most vital points achieved.

3. PROPOSED METHOD

The suggested approach mixes federated learning, homomorphic encryption, and autoencoder-based anomaly identification without changing the original data in any way. The approach constitutes of:

- Federated Learning (FL) revolutionizes actual machine learning by bringing together protecting privacy and information decentralized utilization. This strategy keeps data on personal devices, which contributes to collaborative learning while addressing growing concerns about confidentiality and safety of data [38].
- Homomorphic encryption techniques enable computational tasks on data that is encrypted. This is a very helpful attribute for a variety of scenarios [39].
- autoencoder-based anomaly identification: - They are taught with typical data to understand how it depicts the usual condition. If an input deviates sufficiently from the learnt from throughout reasoning, the Autoencoder may certainly rebuild it incorrectly [40].

This makes it safer for both data privacy and model security. This all-around method is a strong way to stop data poisoning attempts from working. We added review measures and visualizations to give you useful information about how well the model works, especially when it comes to finding oddities in a tainted dataset. It's mostly about showing a way to use federated learning and homomorphic encryption together to make machine learning models better at finding problems and protecting against data poisoning attacks. The method carefully blends federated learning and homomorphic encryption to protect data privacy and model security at a high level. Within this method, we create fake healthcare data and save it in a CSV file. Later, we made an HTML table by putting the data from the CSV file together with the data that we made up. Example data fields in the table are "ID,"

"Age," "Gender," "BMI" (Body Mass Index), "Blood Pressure," "Cholesterol," and "Label." For creating and changing data, the implementation uses NumPy, which has methods for making HTML tables straight from both the generated data and a CSV file. This method incorporates many different areas, which not only makes the model safer but also shows what it can do by creating fake healthcare data and showing it in an HTML table.

Data Generation and Poisoning Simulation

Table 1 shows how the synthetic data function creates a dataset with a specified number of samples and attributes. The simulated data poisoning function perturbs a portion of the dataset with Gaussian noise to assault it. This process mimics training data adversaries. The general steps of the algorithm are:

Algorithm 1: Steps of the implemented algorithm

Step1: Define a function to create an HTML table

Step2: Define a function to create an HTML table from a CSV file

Step3: Accept user inputs for data type, file names, and output

Step4: Generate or read synthetic healthcare data based on user input

Step5: Save data to CSV file

Step6: Create an HTML table directly from the dataset

Step7: Create an HTML table from the CSV file

Table 1. The Generated Dataset

ID	Age	Gender	BMI	Blood Pressure	Cholesterol	Label
1	74	Male	18.571757	130/9	Normal	0
2	34	Female	36.255242	140/9	Normal	0
3	43	Male	32.581212	120/8	High	1
4	59	Male	38.209573	130/9	Normal	0
5	34	Female	19.909998	130/9	Normal	1
6	42	Female	25.237894	120/8	Normal	0
7	67	Female	24.545044	120/8	High	0
8	47	Male	18.877169	130/9	Normal	0
9	43	Female	34.449439	130/9	Normal	0
10	39	Female	37.450613	140/9	Normal	1

Data Standardization and Splitting

For feature scale consistency, StandardScaler standardizes the dataset. Using scikit-learn's train test split function, we split the data into training and testing sets.

Autoencoder Model Training

The proposed method employs an autoencoder neural network for feature learning and anomaly detection. The architecture consists of an encoder and a decoder, with the model trained using the mean squared error loss.

Anomaly Detection

After training, anomalies are detected by comparing the mean squared error (MSE) between the original and reconstructed data. A threshold is determined to identify samples with a significantly higher MSE, indicating potential anomalies.

Data Visualization and Evaluation

Visualization shows discrepancies between original and rebuilt data to explain anomaly detection. Accuracy, precision, recall, and F1 score are also calculated.

Final results—original, poisoned, and cleaned—are saved to CSV files. User's Browser: The user interacts with the web application through their browser, uploading a CSV file.

1. Flask App: The Flask web application receives the CSV file upload request. It processes the data to simulate a data poisoning attack and trains an autoencoder machine learning model.
2. Data Processing: Simulates a data poisoning attack on the uploaded data.
3. Machine Learning Model: Trains an autoencoder model using the processed data.
4. Anomaly Detection: Uses the trained autoencoder to detect anomalies in the data.
5. Results and Metrics: Evaluate the performance of the anomaly detection, calculate metrics, and store the results.
6. Visualization: visualizes the original, reconstructed, and difference plots.
7. Display Results: Displays the confusion matrix, performance metrics, and data visualization on the web page.

A Confusion Matrix

A confusion matrix, sometimes referred to as an error matrix, is a table that's frequently used to explain how well the classification model performs when applied to an array of test results whose true values have been determined. It makes it possible to visualize how well an algorithm performs [2341].

Within a matrix of confusion:

- The real classes are represented by rows.
- The anticipated classes are displayed in columns.

The count (or percentage) of instances that belong to the actual class and were anticipated to belong to the predicted class is represented by each cell in the matrix. In the matrix, examples that were correctly classified are represented by diagonal elements, while cases that have been classified incorrectly are represented by off-diagonal entries. Table 2 illustrates of a confusion matrix for an issue related to binary classification may be found here:

Table 2. Confusion matrix for an issue related to binary classification.

Actual	Predicted Negative	Predicted Positive
Negative	TN	FP
Positive	FN	TP

a) The number of cases that are truly negative but were accurately anticipated to be negative is known as TN (True Negative).

b) False Positives, or FPs, are the number of cases that are truly negative but were mispredicted as positive.

c) False Negative (FN): The number of cases that are positive in reality but were miscalculated to be negative.

d) True Positives, or TPs, are the proportion of cases that are truly positive and were accurately anticipated to be

Confusion matrices are extremely helpful for assessing how well classifiers are performing, especially if there is an asymmetry in the groups or when some mistakes are priced higher than others. They may direct model changes and offer insights into the kinds of mistakes a machine learning algorithm is making positive.

4. RESULTS AND DISCUSSION

The proposed method effectively detects and mitigates data poisoning attacks. For instance, in a dataset with 50 samples and 7 features, the VAE successfully identified anomalies introduced by Gaussian noise. Visualization of the original, reconstructed, and anomaly data provided clear insights into the model's performance. The method's generalizability was evaluated across different datasets, demonstrating robust performance. However, homomorphic encryption introduces computational complexity, which may impact real-time applications. The selection of the MSE threshold is critical and may affect the detection accuracy.

Example 1

In this example, it should be taken into consideration the input data number of samples = 10 and number of features = 7 as shown in Table 3. Figure 1 displays the generated results.

Table 3: The generated dataset for example 1

ID	Age	Gender	BMI	Blood Pressure	Cholesterol	Label
1	46	Male	28.14756	120/8	Normal	1
2	43	Male	21.23251	120/8	High	0
3	35	Female	36.74323	140/9	Normal	1
4	18	Male	32.79949	120/8	High	1
5	39	Female	31.33941	120/8	High	0
6	27	Male	38.93679	140/9	High	0
7	66	Female	33.50282	130/9	Normal	0
8	49	Male	23.97282	140/9	Normal	0
9	46	Male	39.24726	130/9	High	1
10	63	Male	33.07402	120/8	High	0

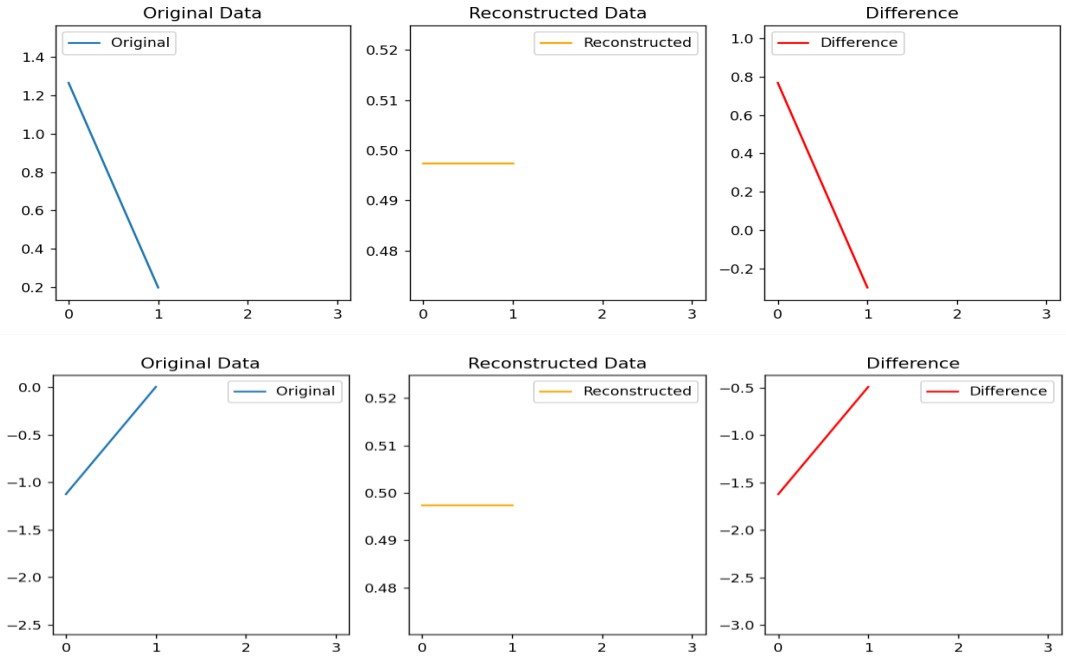

Figure 1. Anomaly Detection Data Visualization and Evaluation for example 1

Example 2

Consider that if the input sample number is 50 and the number of features is 7, Table 4 displays the generated data set for example 2. Figure 2 illustrates the generated results.

Table 4. The generated dataset for example 2

ID	Age	Gender	BMI	Blood Pressure	Cholesterol	Label
1	58	Female	24.55172	130/9	High	1
2	18	Male	20.56821	120/8	High	1
3	62	Male	31.11844	120/8	High	1
4	73	Female	29.09497	130/9	High	1
5	71	Male	30.70869	140/9	Normal	1
6	27	Male	22.42974	140/9	High	1
7	18	Male	35.28213	120/8	Normal	1
8	39	Male	20.24089	130/9	Normal	0
9	36	Female	38.4441	120/8	Normal	1
10	26	Female	35.11298	140/9	Normal	0
11	24	Female	24.02172	130/9	High	1
12	69	Female	29.9367	120/8	High	1
13	41	Male	23.72804	120/8	Normal	1
14	31	Male	28.17952	140/9	Normal	0
15	51	Female	32.72435	120/8	High	1
16	65	Male	24.84793	140/9	Normal	0
17	54	Female	20.2436	140/9	Normal	1
18	32	Female	35.92968	130/9	Normal	0
19	64	Female	23.35415	130/9	High	0
20	74	Male	34.39943	140/9	High	1
21	64	Female	34.67962	140/9	Normal	0
22	53	Female	22.28153	130/9	Normal	1
23	64	Male	19.81694	130/9	High	0
24	44	Female	22.14533	140/9	Normal	0
25	22	Female	28.60677	140/9	High	0
26	26	Female	29.83639	120/8	High	0
27	22	Male	24.80421	140/9	Normal	0
28	53	Male	32.05131	140/9	Normal	1
29	62	Female	19.55667	130/9	Normal	0
30	66	Male	36.62335	140/9	Normal	1
31	48	Male	19.53846	120/8	High	0
32	46	Male	27.46242	140/9	High	0
33	57	Female	19.31368	130/9	Normal	0
34	73	Male	33.68613	120/8	Normal	1
35	67	Female	38.72478	120/8	Normal	1
36	65	Male	30.7964	120/8	Normal	1
37	66	Female	29.29924	130/9	Normal	1
38	54	Male	25.45463	120/8	High	1
39	26	Female	30.26693	120/8	Normal	1
40	41	Female	36.60552	140/9	Normal	0
41	72	Male	33.68595	120/8	Normal	1
42	40	Female	21.80319	130/9	High	1
43	62	Female	27.03105	120/8	Normal	1
44	30	Male	32.88432	130/9	Normal	1
45	50	Male	27.33584	120/8	Normal	0
46	64	Female	35.66981	130/9	High	0
47	63	Male	34.44979	120/8	High	0
48	51	Male	20.44219	140/9	Normal	0
49	49	Female	33.50677	130/9	High	1
50	64	Male	38.0699	130/9	High	0

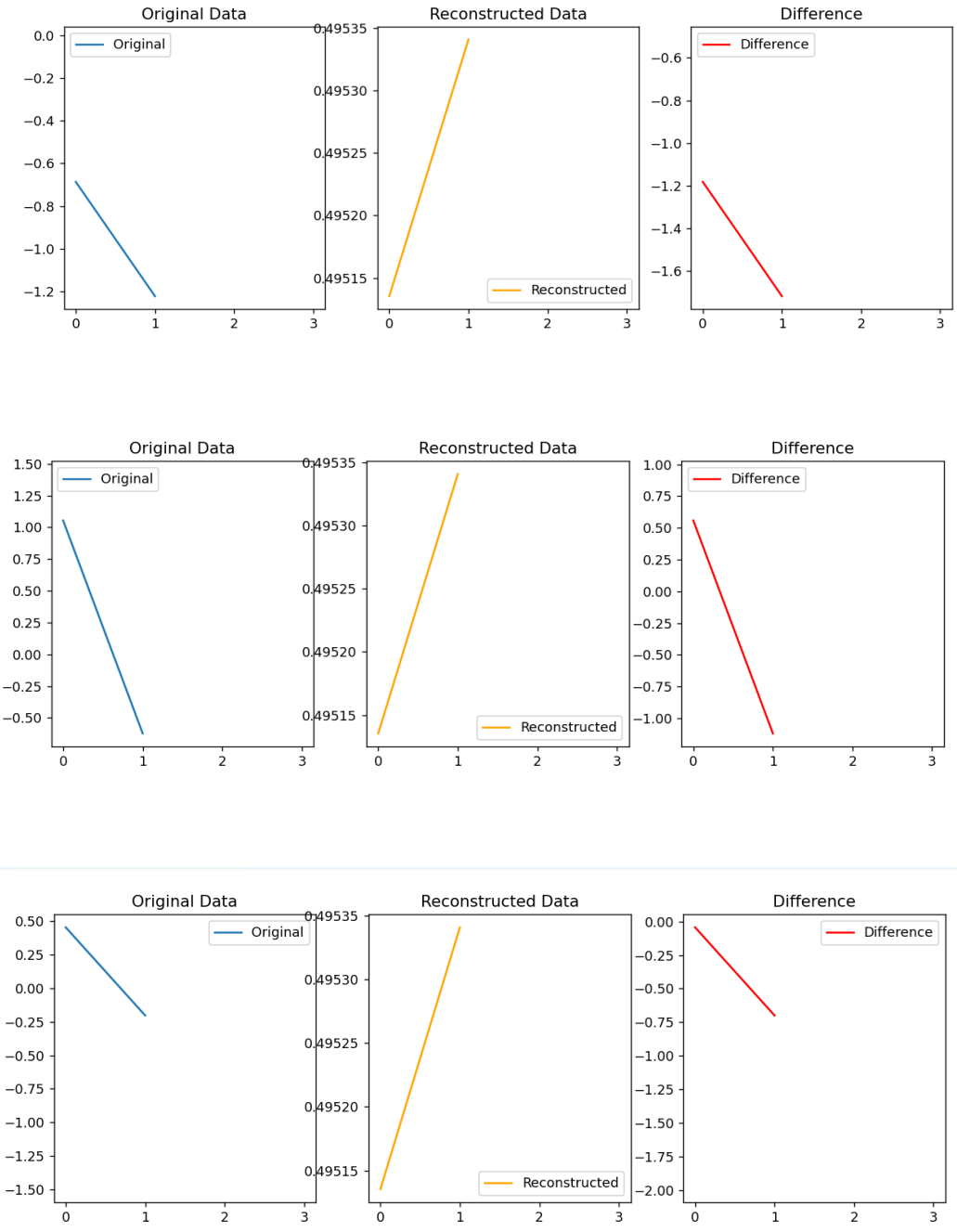

Figure 2. Anomaly detection data visualization and evaluation for example 2

5. CONCLUSIONS

This study provides a comprehensive framework for healthcare system security using generative artificial intelligence. This innovative technology detects and eliminates data poisoning attacks using federated learning, homomorphic encryption, and autoencoder-based anomaly detection. This paper discusses the growing risks of medical data security and how artificial intelligence can improve healthcare safety. Distributed modelling secures data and models. It can survive data poisoning assaults on simulated health data, proving its viability.

The approach described herein should make medical AI systems safer via autoencoder anomaly detection, homomorphic encryption, and federated machine learning. These methods prevent security threats to sensitive data and models. Future tasks are listed in the report:

Develop extensive medical data and model protection techniques for potential dangers.

- Innovate healthcare privacy and security.
- Simulate data poisoning resistance and medical data anomaly detection [42].
- Detect anomalies with autoencoders, homomorphic encryption, and federated learning.
- Encrypt homomorphically.
- Use AI to detect and stop data pollution.
- Explain increasing hazards and offer flexible security.
- Research on secure multi-site model training using federated learning.

Generative artificial intelligence, federated learning, and homomorphic encryption are used to improve healthcare system security in this study. The approach detects and mitigates data poisoning threats, protecting sensitive health data. Future work will include developing more efficient coding techniques and investigating the application of this framework in a variety of healthcare settings.

Ethical Considerations

Institutions offered data handling ethical authorization guidelines. This study used only synthetic data to replicate common health datasets without patient data.

Acknowledgement

The authors thank Mustansiriyah University (https://uomustansiriyah.edu.iq/) and Al-Iraqia University in Baghdad, Iraq, for their support in the present work.

REFERENCES

1. Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implementation Science 2024; 19(27):1-15
2. Bekbolatova M, Mayer J, Ong C W, Toma M. Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. Healthcare 2024; 12(12):125
3. Huang J, Neill L, Wittbrodt M, Melnick D, Klug M, Thompson M, Bailitz J, Loftus T, Malik S, Phull A, Weston V, Heller A, Etemadi M. Generative Artificial Intelligence for Chest Radiograph Interpretation in the Emergency Department. JAMA Network Open 2023; 6(10):e2336100
4. Yang Z, Mitra A, Liu W, Berlowitz D, Yu H. TransformEHR: transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records.

Nature Communications 2023; 14(7857):1-10
5. Rodriguez-Almeida A J, Fabelo H, Ortega S, Deniz A, Balea-Fernandez F J, Soguero-Ruiz C, Wägner A M, Callico G M. Synthetic Patient Data Generation and Evaluation in Disease Prediction Using Small and Imbalanced Datasets. IEEE Journal of Biomedical and Health Informatics 2022; 27(6):2670-2680
6. Data Poisoning: The Newest Threat to Generative AI. (2023, November 8). Forcepoint. https://www.forcepoint.com/blog/x-labs/data-poisoning-gen-ai
7. Pashkova T V. To the question of the ways of naming the disease "hernia" (based on the Karelian language). Bulletin of Ugric Studies 2022; 12(2):291-299
8. Olsen, E. (2023, December 13). Google reveals new generative AI models for healthcare. Healthcare Dive. https://www.healthcaredive.com/news/google-cloud-medlm-generative-ai-in-healthcare/702345/
9. Sallam M, Al-Farajat A, Egger J. Envisioning the Future of ChatGPT in Healthcare: Insights and Recommendations from a Systematic Identification of Influential Research and a Call for Papers. Jordan Medical Journal 2024; 58(1):95-108
10. de Aguiar E J, Traina C, Traina A J M. Security and Privacy in Machine Learning for Health Systems: Strategies and Challenges. Yearbook of Medical Informatics 2023; 32(01), 269–281
11. Nass S J, Levit L A, Gostin L O. Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research. In National Academies Press, 2009.
12. Ali J J, Shati N M, Gaata M T. Abnormal activity detection in surveillance video scenes. TELKOMNIKA (Telecommunication Computing Electronics and Control) 2020; 18(5):2447-2453
13. Wang C, Chen J, Yang Y, Ma X, Liu J. Poisoning attacks and countermeasures in intelligent networks: Status quo and prospects. Digital Communications and Networks 2022; 8(2):225-234
14. Sattar I A, Gaata M T. Image steganography technique based on adaptive random key generator with suitable cover selection. In proceedings of Annual Conference on New Trends in Information & Communications Technology Applications (NTICT), 07-09 March 2017:208-212
15. Khaleel M K, Ismail M A, Yunan U, Kasim S, Review on Intrusion Detection System Based on The Goal of The Detection System. International Journal of Integrated Engineering 2018; 10(6):197-202
16. Al-Qaysi Z T, Suzani M S, Rashid N A, Ismail R D, Ahmed M A, Aljanabi R A, Gil-Costa V. Generalized Time Domain Prediction Model for Motor Imagery-based Wheelchair Movement Control. Mesopotam
17. Alqaraghuli S M, Karan O. Using Deep Learning Technology Based Energy-Saving For Software Defined Wireless Sensor Networks (SDWSN) Framework. Babylonian Journal of Artificial Intelligence 2024; 2024: 34–45.
18. Ali M, Naeem F, Tariq M, Kaddoum G. Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. IEEE Journal of Biomedical and Health Informatics 2023; 27(2): 778 – 789
19. Nawaz A, Khan S S, Ahmad A. Ensemble of Autoencoders for Anomaly Detection in Biomedical Data: A Narrative Review. IEEE Access 2024; 12: 17273 – 17289
20. Gopi R S, Suganthi R, Hephzipah J J, Amirthayogam G, Sundararajan P N, Pushparaj T. Elderly People Health Care Monitoring System Using Internet of Things (IOT) For Exploratory Data Analysis. Babylonian Journal of Artificial Intelligence 2024; 2024: 54–63.
21. Ibraheem H R, Zaki N D, Al-mashhadani M I. Anomaly detection in encrypted HTTPS traffic using machine learning: a comparative analysis of feature selection techniques. Mesopotamian Journal of Computer Science 2022; 2022: 18–28.
22. Chen Y, Esmaeilzadeh P. Generative AI in Medical Practice: In-Depth Exploration of Privacy and Security Challenges. Journal of Medical Internet Research 2024; 26:e53008
23. Psychogyios K, Velivassaki T, Bourou S, Voulkidis A, Skias D, Zahariadis T. GAN-Driven Data Poisoning Attacks and Their Mitigation in Federated Learning Systems. Electronics 2023; 12(8):1-17

24. Mohialden Y M, Hussien N M, Salman S A, Aljanabi M. Secure Federated Learning with a Homomorphic Encryption Model. International Journal Papier Advance and Scientific Review 2023; 4(3):1-7
25. Salman S A, Al-Janabi S, Sagheer A M. Security Attacks on E-Voting System Using Blockchain. Iraqi Journal for Computer Science Mathematics 2023; 4(2):179-188
26. Paul W, Cao Y, Zhang M, Burlina P. Defending Medical Image Diagnostics Against Privacy Attacks Using Generative Methods: Application to Retinal Diagnostics. In Clinical Image-Based Procedures, Distributed and Collaborative Learning, Artificial Intelligence for Combating COVID-19 and Secure and Privacy-Preserving Machine Learning 2021; 2021:174–187
27. Liu H, Crespo R G, Martínez O S. Enhancing Privacy and Data Security across Healthcare Applications Using Blockchain and Distributed Ledger Concepts. Healthcare 2020; 8(3):1-17
28. Generative AI Security Issues and how to mitigate them? - Securiti. (2023, December 20). Securiti. https://securiti.ai/generative-ai-security
29. Sallam M, Khalil R, Sallam M. Benchmarking Generative AI: A Call for Establishing a Comprehensive Framework and a Generative AIQ Test. Mesopotamian Journal of Artificial Intelligence in Healthcare 2024; 2024: 69–75.
30. Mijwil M M, Naji A S, Doshi R, Hiran K K, Bala I, Ali G. The Effect of Human-Computer Interaction on New Applications: A ChatGPT Use-Case Analysis in Healthcare Services. In Modern Technology in Healthcare and Medical Education: Blockchain, IoT, AR, and VR 2024; 1, 74-84
31. Bala I, Pindoo I A, Mijwil M M, Abotaleb M, Yundong W. Ensuring Security and Privacy in Healthcare Systems: A Review Exploring Challenges, Solutions, Future Trends, and the Practical Applications of Artificial Intelligence. Jordan Medical Journal 2024; 58(2):250-270
32. Mohammed S Y, Aljanabi M, Gadekallu T R. Navigating the Nexus: A systematic review of the symbiotic relationship between the metaverse and gaming. International Journal of Cognitive Computing in Engineering 2024; 5:88-103
33. Khan M A, Ullah S, Ahmad T, Jawad K, Buriro A. Enhancing Security and Privacy in Healthcare Systems Using a Lightweight RFID Protocol. Sensors 2023; 23(12): 1-15
34. Seh A H, Al-Amri J F, Subahi A F, Agrawal A, Kumar R, Khan R A. Machine Learning Based Framework for Maintaining Privacy of Healthcare Data. Intelligent Automation & Soft Computing 2021; 29(3):697-712
35. Marulli F, Marrone S, Verde L. Sensitivity of Machine Learning Approaches to Fake and Untrusted Data in Healthcare Domain. Journal of Sensor and Actuator Networks 2022; 11(2): 1-12
36. Aceto G, Giampaolo F, Guida C, Izzo S, Pescapè A, Piccialli F, Prezioso E, Synthetic and privacy-preserving traffic trace generation using generative AI models for training Network Intrusion Detection Systems. Journal of Network and Computer Applications 2024; 229:103926
37. Khalid N, Qayyum A, Bilal M, Al-Fuqaha A, Qadir J. Privacy-preserving artificial intelligence in healthcare: Techniques and applications. Computers in Biology and Medicine 2023; 158:106848
38. Li L, Fan Y, Tse M, Lin K. A review of applications in federated learning. Computers & Industrial Engineering 2020; 149:106854
39. Wood A, Najarian K, Kahrobaei D. Homomorphic Encryption for Machine Learning in Medicine and Bioinformatics. ACM Computing Surveys 2020; 53(4): 1-35
40. Chen Z, Yeo C K, Lee B S, Lau C T. Autoencoder-based network anomaly detection. In 2018 Wireless Telecommunications Symposium (WTS) 2018; 1-5
41. Hussain Z F, Ibraheem H R, Alsajri M, Ali A H, Ismail M A, Kasim S, Sutikno T. A new model for iris data set classification based on linear support vector machine parameter's optimization. International Journal of Electrical and Computer Engineering 2020; 10(1):1079-1084
42. Aljanabi M, Ismail M A. Improved Intrusion Detection Algorithm based onTLBO and GA Algorithms. International Arab Journal of Information Technology 2021; 18(2):170-179

تعزيز الأمن والخصوصية في الرعاية الصحية باستخدام برامج الكشف عن هجمات التسمم بالبيانات والتخفيف منها القائمة على الذكاء الاصطناعي التوليدي

ياسمين مكي محي الدين [1*]، صبا عبد الباقي سلمان [2]، معد محسن مجول [3،4]، نادية محمود حسين [5]، محمد الجنابي [6]، مصطفى أبوطالب [7]، كلوديان دوسكا [8]، براديب ميشرا [9]

[1] قسم علوم الحاسوب، كلية العلوم، جامعة المستنصرية، بغداد، العراق

[2,5] قسم علوم الحاسوب، كلية التربية، الجامعة العراقية، بغداد، العراق

[3] كلية الإدارة والاقتصاد، الجامعة العراقية، بغداد، العراق

[4] قسم هندسة تقنيات الحاسوب، كلية العلوم الاقتصادية، بغداد، العراق

[6] جامعة الإمام جعفر الصادق، بغداد، العراق

[7] قسم برمجة النظم، جامعة جنوب الأورال الحكومية، تشيليابينسك، روسيا

[8] قسم الإنتاج والإدارة، جامعة البوليتكنيك في تيرانا، ألبانيا

[9] كلية الزراعة، ريوا، JNKVV، (M.P.) الهند

Received: May 29, 2024
Accepted: October, 13, 2024
DOI:
https://doi.org/10.35516/jmj.v58i3.2712

الملخص

بحثت هذه الدراسة في نهج متقدم لتعزيز الأمن والخصوصية في الرعاية الصحية من خلال دمج استراتيجيات تعتمد على الذكاء الاصطناعي للكشف عن هجمات التسمم بالبيانات والتخفيف منها. جمعت الطريقة المقترحة بين التعلم الموحد والتشفير المتماثل واكتشاف الشذوذ القائم على الترميز التلقائي. وضمنت تدريب النماذج في أماكن متنوعة وحماية البيانات وتحسين أمان النموذج. وتم التحقيق في تحديد الشذوذ والتخفيف منه ومقاومة التسمم بالبيانات باستخدام بيانات محاكاة طبية. النتائج الرئيسية. قام هذا النهج بتصور وتقييم أداء النموذج. تقدم هذه الدراسة حلاً كاملاً لتأمين البيانات والنماذج الطبية ضد التهديدات الجديدة.

الكلمات الدالة: التعمية مثلية الشكل، التعلم المُتَّحِد، هجمات التسمم بالبيانات، محاكاة أمن الرعاية الصحية، اكتشاف الشذوذ.