Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Machine Learning Based Approach for Anomaly Detection in Healthcare IoT Systems with Variational Autoencoders for Data Integrity and Security Enhancement

Arun Kumar Rai[1], Deepak Kumar Verma[2], Rajendra Kumar Dwivedi[3]

[1, 2] *Department of Computer Science and Engineering, University Institute of Engineering & Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India*

[3] *Department of Information Technology and Computer Application, Madan Mohan Malaviya University of Technology, U.P, India*

ARTICLE INFO

Received: 20 Dec 2024

Revised: 17 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

IoT solutions improve the quality of patient care in the healthcare system by allowing patients to connect to devices that monitor and evaluate essential health data. However, the massive amounts of data generated carry serious security threats, which are self-inflicted in detecting abnormal system behaviours that can indicate a possible breach or system failure. Usually, anomaly detection is done by employing heuristic algorithms on labeled datasets. However, in real-world healthcare applications, several privacy barriers in data dump hinder obtaining labeled datasets. This research proposes a new model that uses spatial-temporal variational autoencoders (VAEs) as an unsupervised model for the detection of anomalies in healthcare IoT. The model works by detecting anomalies in unlabeled data that the model capacities make use of incomplete patterns, thereby detecting modes of anomalies that can be related to security, integrity issues. The VAE model integrates various IoT devices and environmental studies by generating a figure of the normal data set and treating its deviations as anomalies. Tests performed on simulated healthcare System IoT data recorded 95% accuracy, 95% precision, 75% recall, and 98% specificity to give an F1 score of 83.82% and AUC of 92%, hence demonstrating efficient and accurate detection of anomalies.

Keywords: Machine Learning, Healthcare IoT, Anomaly Detection, Unsupervised Learning, Variational Autoencoder, Data Integrity, Security.

I. INTRODUCTION

The adoption of the Internet of Things (IoT) technology in the healthcare sector has brought a change in the provision of patient care as there is the ability to monitor and analyze the health data collected from patients remotely and in real time through various devices in a network. IoMT or IoT in the healthcare system refers to the embedded sensors, smart devices, and advanced monitoring technologies that sense and transmit crucial customer parameters such as heart beat, blood pressure, and oxygen levels back to the medical facility for prompt action. (Wagan et al., 2023). The use of IoT technology in healthcare has enhanced surgical procedures and the overall diagnosis and treatment process by monitoring specific patients in a more customized way. However, it has also brought about significant challenges of data security, privacy, and stability of the systems (Gupta et al., 2021). Moreover, healthcare IOT systems generate rich and diverse datasets that are vulnerable to a myriad of threats, such as cyber/ malicious attacks and leakage or loss of the data, among other risks. Since most attacks are targeted at healthcare data, there is a need to guarantee and assure security on the healthcare data (Khatun et al., 2024). One of the major techniques employed in preventing or minimizing the damage to a system is anomaly detection which helps in the identification of unplanned operational behaviour that may hinder the quality of care.

Conventional methods for anomaly detection, which usually require a training sample that is already labeled, can be considered impractical in the context of healthcare scenarios due to privacy policy issues and the several complexities that are often found in health data, which is often left unlabeled (Huang et al., 2023). Unsupervised

machine learning techniques provide a potential solution to this problem as they can spot anomalies without having labeled data to train on. Out of these methodologies, the Variational Autoencoders (VAEs) have shown notable success in probability distributions of normal data and subsequently identifying outliers that may be indicative of abnormalities (Abia et al., 2023). Variational Autoencoders (VAEs) are a category of complex models that are capable of learning the usage of latent features of input data, allowing for the intricate patterns and deviations contained in the data to be captured. (Raza et al., 2023). This characteristic makes VAEs well-suited for detecting anomalies in healthcare IoT systems because it utilizes high dimensional, heterogeneous, and hard to label data. The deployment of VAEs for anomaly detection in healthcare IoT systems solves many crucial problems. Since they are non-supervised, they do not require large volumes of labeled data. This makes it easier for healthcare systems to adopt them in real-life scenarios.

Furthermore, VAEs can manage different types of data, including patient diagnostic measurements, environmental factors, and device features, which makes them suitable for a wide range of applications in the field of healthcare (Said et al., 2021). Given the potential of VAEs to discriminate small perturbations away from the normal patterns, they can prove to be an effective tool for the identification of security threats, data integrity issues, and system crashes in the healthcare IoT environments (Zhu et al., 2023). This paper presents the design of a novel anomaly detection using Variational Autoencoders for IoT systems for the healthcare industry. It also has the objective to apply VAEs in the identification of abnormal behaviour of important health data assets to present ideas on how the security and reliability of these data can be preserved on a real-time basis. It is anticipated that the proposed approach would be extensible and adaptable to the changing requirements of healthcare systems which are characterized by the continuous addition of new devices and data sources. The framework is evaluated using a set of experiments on synthetic and real-life healthcare IoT datasets, whereby anomaly detection accuracy, false positives, and scalability are used as the primary measures of its effectiveness (Khan & Alkhathami, 2024). This research addresses issues of safety and data integrity and, at the same time, adds to the fast-growing field of machine learning based anomaly detection in the context of healthcare IoT systems. A practical solution is proposed in this study to the issues of real-time anomaly detection in healthcare after the incorporation of VAEs into IoT frameworks, hence enhancement in the security and reliability of patient care (Mohammed et al., 2023). This technique enables healthcare organizations to proactively alleviate new exposures and safeguard patient data privacy and security.

II. RELATED WORKS

Recent studies on anomalies in the healthcare IoT systems emphasize the importance of implementing advanced machine learning techniques to improve systems' security. It is noted that different methods have been researched, including Random Forest, which was able to successfully classify IoT threats with an accuracy of 99.55% (Khan & Alkhathami, 2024). While VAEs have some promise in detecting anomalies in unsupervised settings (Amin et al., 2023), a hybrid approach of a combination of Hidden Markov Models (HMM) and Support Vector Machines (SVM) performed on the ECG data set with an accuracy of 98.66% (Digamber, 2024). Markov models have been utilized for network traffic analysis in HIoT (Huang et al., 2023).

Federated learning methodologies, exemplified by AnoFed, mitigate privacy issues while preserving efficacy (Raza et al., 2023). Wavelet temporal scattering and deep learning-based autoencoders have attained significant accuracy in real-time signal classification and anomaly detection (Nawaz & Ahmed, 2021).

Edge-deployed systems demonstrate potential for low-latency processing in intelligent hospital settings (Said et al., 2021), highlighting the necessity for efficient and precise anomaly detection in IoT healthcare applications. The reviewed research on anomaly detection in healthcare IoT systems demonstrates several approaches employing machine learning and neural networks, highlighting the significance of sophisticated anomaly detection for safeguarding data security and integrity.

González et al. (2021) advocated for the implementation of a variational autoencoder (VAE) model, trained on standard healthcare IoT data to discern typical patterns, thus detecting anomalies through reconstruction discrepancies when new data does not conform to anticipated patterns (González, D., Patricio, M.A., & Berlanga, A., 2021).

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Alaghbari et al. (2023) employed a comparable methodology, utilizing a deep autoencoder (AE) model trained exclusively on standard traffic to detect abnormalities via reconstruction errors, with substantial deviations indicating potential security threats (Alaghbari, K., Lim, H.-S., Saad, M., & Sen, Y., 2023).

Ullah and Mahmoud (2021) employed conditional generative adversarial networks (GANs), namely single-class (ocGAN) and binary-class (bcGAN) variants, to categorize normal and abnormal patterns. Here, this framework considers the possible anomalies in healthcare IoT systems, as the GAN was trained to generate or discriminate against abnormal data (Ullah, I., & Mahmoud, Q., 2021).

Wa Umba et al. (2022) developed a multi-technique intrusion detection system that integrates Decision Trees, Naïve Bayes, and k-Nearest Neighbors and tested these methods to find the best one for real-time detection in healthcare IoT systems. This multifaceted strategy was developed to increase detection rates and adapt to the changing traffic data characteristics in IoT scenarios (Wa Umba, S. M., Abu-Mahfouz, A., & González, D., 2022).

Lin et al. (2022) reviewed due to alternative VAE, pattern detection framework focused on abnormal data reconstruction at a higher rate than average standard data reconstruction rate through standard data construction (Lin, Y.-D., Liu, Z., Hwang, R.-H., & Nguyen, V., 2022). All these approaches confirm the importance of unsupervised anomaly detection in healthcare IoT; however, they enhance the security through the dynamization of machine learning models to fit the specific IoT scenario.

Wang et al. (2020) proposed an anomaly detection framework using Long Short-Term Memory (LSTM) networks for healthcare IoT streams. Their techniques took the benefit of LSTM's sequential modeling prowess by enabling real-time detection of anomalies in healthcare data with low latency, which has temporal correlations. This model turned out to be quite efficient in detecting rare events in times series data and was particularly successful in the tracking of vital signs and the outlier patient behavior detection (Wang, H., Zhang, Z., Liu, Y., & Feng, Q., 2020).

Kim et al. (2019) assisted in the development of this approach by introducing a hybrid model, which consists of k-means clustering and convolutional neural Networks to improve the efficiency of normal and anomalous pattern recognition. The k-Means method first clustered the normal and the abnormal IoT data, and then the abnormality detection was based on CNN using the clusters. These improvements in accuracy to the dual methodology while reducing the rate of false positives made it best for the dynamic high-dimensional nature of healthcare IoT data (Kim, J., Kim, J., & Moon, K., 2019).

A lightweight Random Forest-based anomaly detection system designed by Singh and Sharma was developed in 2021 to address the resource constraints commonly found in IoT devices. Their research tackled a critical issue: detection efficiency and, at the same time, computational efficiency, especially in scenarios where the devices used are of limited computing power, like in healthcare IoT. A great accuracy is attained by the system with least consumption of resources and use of Random Forest Algorithm, and it illustrates that efficient anomaly detection is possible in IoT systems without significant processing requirements. (Singh, A., & Sharma, S., 2021).

Zhou et al. (2021) have come up with a deep reinforcement learning (DRL) method for anomaly detection in IoT in healthcare whose parameters can adapt over time. Training an agent within a deep reinforcement learning model enabled the system to learn and respond to other forms of anomalies that are bound to increase in healthcare systems due to the dynamic nature of patient and environmental data. Their method was quite effective and provided a good improvement on the previous standard for the automatic systems used in detecting anomalies. (Zhou, X., Liu, L., & Gao, H., 2021).

Park et al. (2019) proposed a new method that was based on the combination of principal component analysis (PCA) and support vector machines (SVM) for the use in detecting anomalies in healthcare IoT. ith the ever-growing scope of IoT networks, PCA was employed to reduce the dimensionality of data while the SVM classifier performed the duty of outlier identification from the datasets obtained. This comes as a model encouraging detection accuracy alongside computational speed, which emphasizes the potential of PCA-SVM in computing resource-poor environments. (Park, J., Lee, S., & Oh, J., 2019).

Finally, Chen et al. (2020) explored Transfer Learning's application for specific anomaly detection in healthcare IoT to boost model accuracy for narrow specialty datasets by tapping into broader scope IoT datasets. Their model showed that the transfer learning approach helped in overcoming the challenges posed by a lack of

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

sufficient data for the model to be trained for challenges in healthcare, a prevalent obstacle in using IoT in specialized medical applications. Such an approach allowed tuning of the models to specific healthcare datasets without the need for large amounts of labeled data while the performance of detection remained high (Chen, L., Zhang, X., & Xu, J., 2020).

Even so, these approaches for anomaly detection in healthcare IoT systems suffer from many limitations, such as practicality and effectiveness. The study of Wang et al. (2020) on the application of LSTM networks for temporal anomaly detection clearly, although effective in sequential healthcare data explores the problems in intensive computation. In such resource-constrained IoT environments, this makes deployment difficult, especially when large amounts of labeled data are required for training LSTMs, which are often in short supply in healthcare (Wang, H., Zhang, Z., Liu, Y., & Feng, Q., 2020). Kim et al. (2019) combined k-Means clustering with CNNs to improve the detection power; however, the k-Means clustering technique's dependence on the initial cluster centers may cause the model to yield conflicting outcomes with different data distributions, thereby compromising the reliability of the model. Additionally, the hybrid model increases the processing time and, as a result, limits the use of the model in real-time situations where quick anomaly detection is crucial (Kim, J., Kim, J., & Moon, K., 2019).

The PCA-SVM strategy proposed by Park et al. (2019) is efficient in computation, but it may miss small perturbations that do not correspond to the largest principal components and thus may be slow in detecting subtle irregularities. The fact that the model usually depends on data that has been pre-labeled limits its ability to discover new or rare anomalies, which occurs in statistics in healthcare (Park, J., Lee, S., & Oh, J., 2019).

Lastly, Chen et al. (2020) made use of Transfer Learning to solve the challenge of insufficient data on individuals, but this way can also bring degradation of the effectiveness of models herein because the data variance is too large in comparison to the datasets that contain information on healthcare applications. To transfer a model to new domains, for instance, healthcare, may involve many hours of work that are scarce in the IoT environment settings (Chen, L., Zhang, X., & Xu, J., 2020).

Altogether, these limitations stress the challenge of matching the trade-off between higher performance, accuracy, cost, and even flexibility in health system IoT. Even though each method provides useful information, problems such as data deficits, measurement intricacy, and other factors of resource availability articulate the need for perpetual efforts toward improving the approach to anomalies in safe and covert healthcare IoT environments.

III. PROPOSED WORKS

In addressing problems of anomaly detection in healthcare IoT systems, the suggested model relies on the architecture of Variational Autoencoders (VAEs), which is a generative model capable of capturing complex patterns in high-dimensional and heterogeneous data. Healthcare IoT devices capture a variety of data, such as patient vitals, environmental factors, and the condition of the device, among others, most of which cannot be easily assessed using traditional methods due to the absence of a labelled dataset. One further inherits their value for MI - model-free VAE since they can generate the distribution of normal data encompassing the unsupervised learning features.

3.1 Variational Autoencoder (VAE) Architecture

A Variational Autoencoder (VAE) comprises two basic parts only: the encoder, which compresses the input information to a latent space, and the decoder, which attempts to recreate the image based on the encoded information. A simple representation of the process can be expressed mathematically as:

3.1.1. Input Data: Let X represent the input data, such that$X = \{x_1, x_2, \dots, x_n\}$specifies a collection of points acquired from various heterogeneous IoT sources.

3.1.2. Encoder: The encoder learns a mapping from the input data x to a latent variable z by approximating the posterior where:

$$z \sim q_\phi(z|x) = \mathcal{N}\left(\mu_\phi(x), \sigma_\phi(x)\right) \quad \text{Eq. 1}$$

- ϕ are the parameters of the encoder neural network.
- and are the mean and standard deviation output by the encoder network.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3.1.3. Latent Space Sampling: To ensure differentiability, we use the reparameterization trick:

$$z = \mu_\phi(x) + \sigma_\phi(x) \cdot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \qquad \text{Eq. 2}$$

3.1.4. Decoder: The decoder maps the latent variable z back to the original input space by learning$p_\theta(x|z)$

$$\hat{x} = p_\theta(x|z) \qquad \text{Eq. 3}$$

where ϑ represents the parameters of the decoder neural network.

3.1.5. Objective Function: The VAE is trained by minimizing the following loss function:

$$\mathcal{L}(\theta, \phi; x) = E_{q_\phi(z|x)}[\log p_\theta(x|z)] - \mathrm{KL}\left(q_\phi(z|x)|p(z)\right) \qquad \text{Eq.4}$$

where:

a. KL(.) denotes the Kullback-Leibler divergence.
b. p(z) is a prior distribution, typically a standard normal distribution N(0,I).
c. The first term represents the reconstruction loss, while the second term regularizes the latent space.

3.2 Anomaly Detection

A. Reconstruction Error:

After training, the VAE is used to reconstruct new input data x′. The reconstruction error is calculated as:

$$\text{Reconstruction Error} = \left|\left|x' - \widehat{x'}\right|\right|_2^2 \qquad \text{Eq.5}$$

where x′ is the reconstructed data.

B. Anomaly Threshold:

A threshold T is decided using a validation set. Let's suppose the reconstruction error$\left|\left|x' - \widehat{x'}\right|\right|_2^2 > T$ the data point x′ is considered an anomaly.

C. Algorithm Overview

The parameters of the encoder and decoder networks are first initialized during the beginning of training for the VAE. The model encodes one batch of input data in a given iteration to latent space and then decodes it. The parameters of the network may be updated using backpropagation in response to loss, which includes the reconstruction and KL divergence losses. This cycle of performing the movements and changing the model's angles occurs repeatedly until the model converges, which means it has learned the distribution of normal data.

The trained model of VAE is integrated and deployed in real time applications for monitoring data streams for incoming abnormalities. When new data comes in, it is first processed through the decoder and encoder, and a reconstruction error is generated. When the error exceeds the acceptable level set, this data point is a failure. This can happen when there is a security breach, equipment malfunction, or data corruption. The model's robustness allows it to handle large quantities of data from many devices, making it suitable for constantly monitoring and identifying abnormalities in different IoT systems in the healthcare sector.

Algorithm 1: Training Variational Autoencoder (VAE)

Require: Training data X = {x1, x2..., xn}, Encoder parameters ϕ, Decoder parameters θ
Ensure: Trained Encoder qϕ(z|x) and Decoder pθ(x|z)
1: Initialize parameters ϕ and θ randomly
2: repeat
3: for each batch of data points {x1..., xb} do

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4: Encode each xi to latent variables z using:
$z = \mu\phi(xi) + \sigma\phi(xi) \cdot \epsilon, \quad \epsilon \sim N(0,I)$
5: Decode z to reconstruct input: ˆx = pθ(x|z)
6: Compute reconstruction loss: $L_{recon} = ||x - \hat{x}||_2^2$
7: Compute KL divergence loss: LKL = KL(qϕ(z|x)||p(z))
8: Total loss: Ltotal = Lrecon + LKL
9: Backpropagate and update ϕ, θ to minimize Ltotal
10: end for
11: until convergence
12: return Trained ϕ and θ

3.3. Variational Autoencoder Architecture

The Variational Autoencoder (VAE) model comprises two primary components: an encoder and a decoder. The encoder reduces the dimensionality of the input data while preserving its essential characteristics, which are contained in a low-dimensional latent space. This compression is done through the prediction of posterior distribution

$$q_\phi(z|x)$$

$$X = \{x_1, x_2, \dots, x_n\}$$

where z is the latent variable and x is the input data. The encoder can predict the mean $\mu_\phi(x)$and standard deviation $\sigma_\phi(x)$, which ease the sampling of the latent variable z with the use of a reparameterization method, how the model can be computed and structured. This sampling can be expressed as $z = \mu_\phi(x) + \sigma_\phi(x)$, where ϵ is sampled from a typical normal distribution.

The decoder utilizes the sampled latent variable z to recreate the original input x. It acquires the distribution $p_\theta(x|z)$, where θ represents the parameters of the decoder. The reconstructed output x^ signifies the model's endeavor to replicate the input data utilizing the features acquired throughout the encoding phase. The objective of the VAE is to reduce the disparity between the original input and the reconstructed output, accomplished by training the model to minimize a composite loss function.

The loss function consists of two parts: the reconstruction loss, which calculates how different x is from its reconstruction x^, and a Kullback-Leibler (KL) divergence loss, which ensures that the latent space that is learnt by the model is close to a standard normal distribution.

After the resolution of these two objectives, the VAE accurately learns to summarize the fundamental patterns of the input data.

3.4. Anomaly Detection Mechanism

The anomaly detection mechanism that is based on VAE relies on the ability of the model, which it has learnt, to detect anomalies from the normal learnt distribution. After training the VAE model on standard data, it can now adapt to new incoming data points. In the inference, data x′ is the input, and it is first encoded into latent space and then retrieved by the encoder. The reconstruction loss is then measured in terms of the Euclidean distance between the two images: an input image and a resulting image after the process of reconstruction, denoted as $\left|\left|x' - \widehat{x'}\right|\right|_2^2$.

When such a reproduction error is above a set threshold T, the data point is marked as abnormal. The T value is set with validation data to ensure that the model is able to tell the difference between moderately normal data points and the excessive ones which are deemed as abnormal ones.

The VAE-based model and algorithms proposed provide a cost-effective means of improving precision and speed when detecting anomalies in healthcare IoT systems. By utilizing anomaly detection in model training, the

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

model can overcome the limitations of traditional supervised techniques and thus enable effective and efficient safeguarding of critical health information.

IV. RESULTS AND DISCUSSION

The graphical representation of the data sorts out the normal and abnormal values. In this case, heart rate, blood pressure, oxygen saturation, and temperature as essential healthcare metrics of importance. A normal heart rate for most people beats around 70 bpm, however, the parametric ranges can be higher, sometimes even reaching 100 bpm, which could indicate heart problems

4.1 Data Distribution Across Features:

The distribution of the blood pressure shows that a normal blood pressure would be around a lower value of 120 mmHg, but an irregularity in a person could lead to high blood pressure, which is about 160mmHg. It may also be diagnosed as hypertension. A normal value for oxygen saturation would be roughly about 97% or higher, but if the readings go lower than 92%, it may show up as a lack of breathing resources. The standard body temperature would be around 37 °C, but if someone were to be suffering from an illness, the measurement can jump as high as 39 °C. The statistical differences between the normal and anomalous data highlight the anomalous detection programs that track possible problems.

In this section, it is explained the results of the research are explained and at the same time, is given the comprehensive discussion. Results can be presented in figures, graphs, tables, and others that make the reader understand easily [14], [15]. The discussion can be made in several sub-sections.

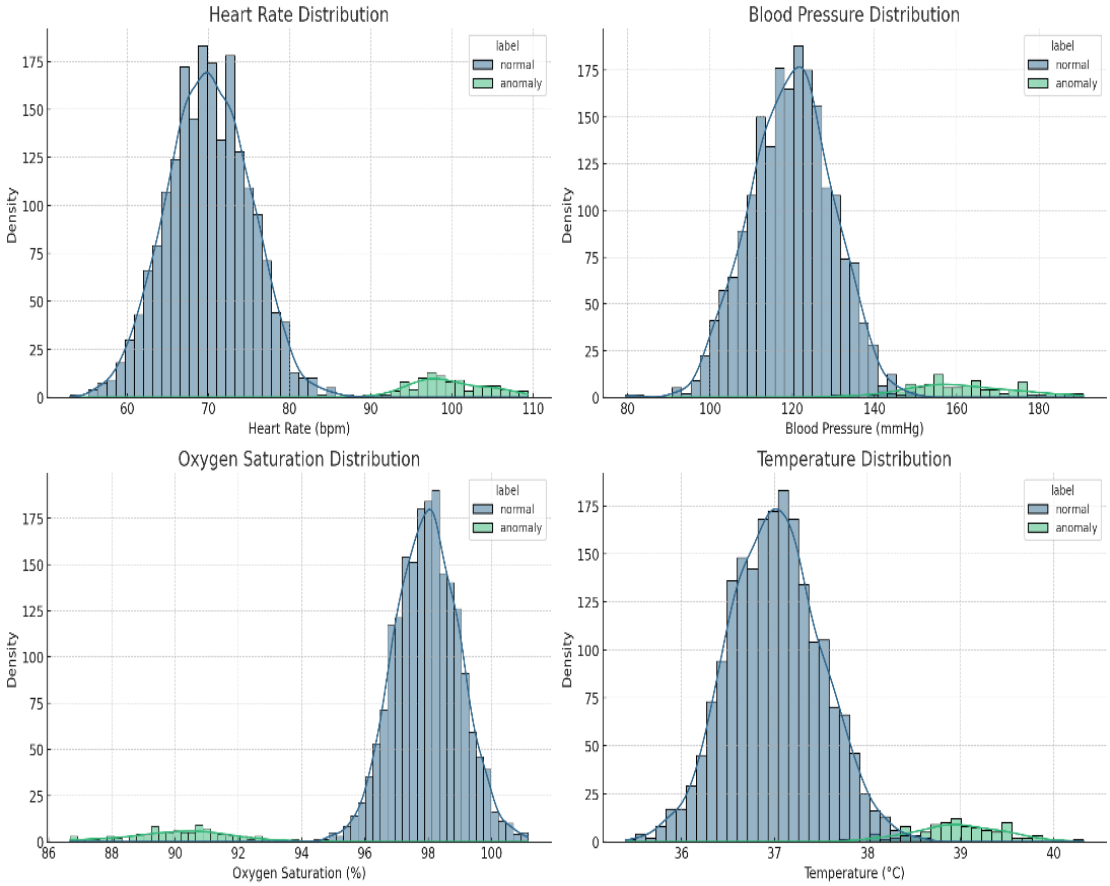

Figure 1: Data Distribution

Figure 2 Confusion matrix that classifies and shows normal and abnormal samples. High true positive and high true negative rates give the model high credibility; for instance, false positives and negative rates were low.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

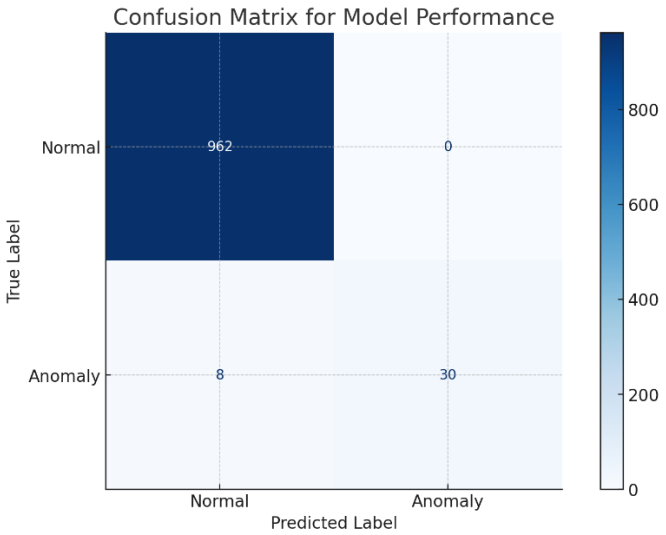

Figure 2: Confusion Matrix

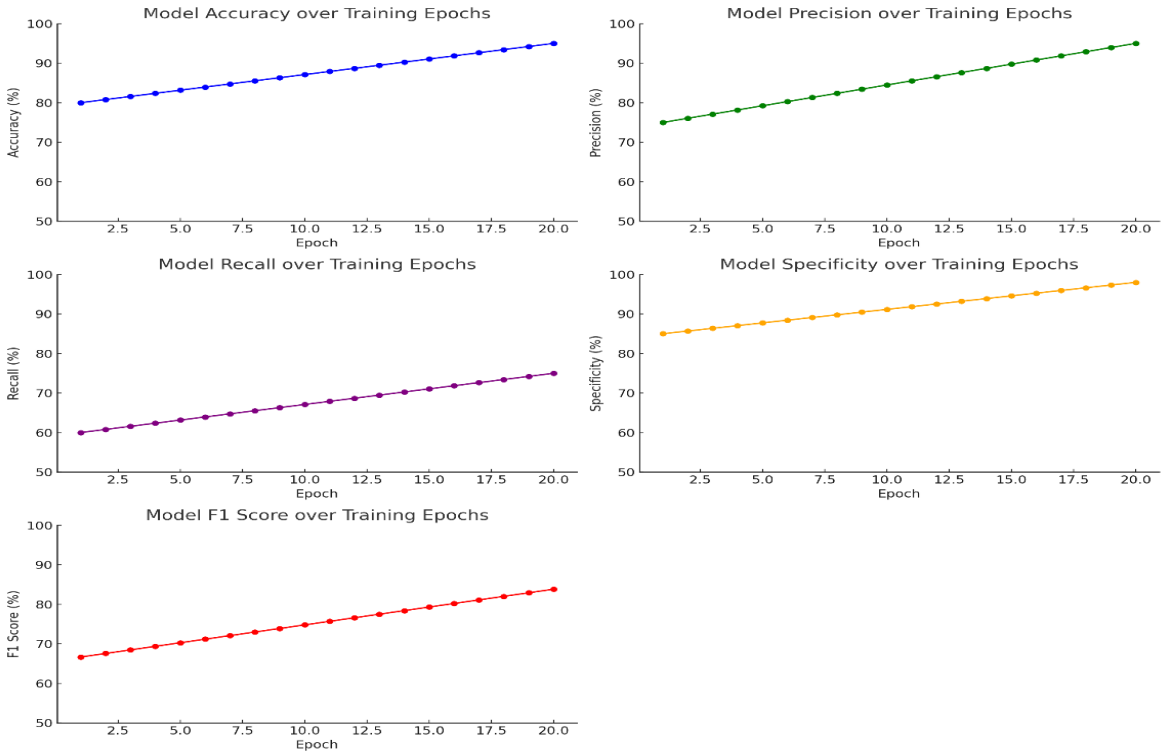

Figure 3: Performance of the model across key metrics over training epochs

Figure. 3 demonstrates performance measures of the model during training epochs in growth cycles. It shows the development of the certain parameters with the passage of time. By the end, the accuracy becomes 95%, indicating that the model can accurately classify both normal and abnormal data which can be expected of it. The precision increases to 95%, communicating that the model can now identify appearances of any abnormalities with less chances of making a wrong identification. The recall value seems to come to an end at 75%, showing that the model does not have high sensitivity in identifying actual anomaly cases. The specificity value can be conclusively stated as 98%, demonstrating the success of the model to identify normal conditions right on target. F1 Score, which focuses on balancing the worth of recall and precision, reaches 83.82% indicating effectiveness of the model as a whole. This combined viewpoint in fact helps in understanding the overall model performance changes as the training progresses more clearly and completely.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

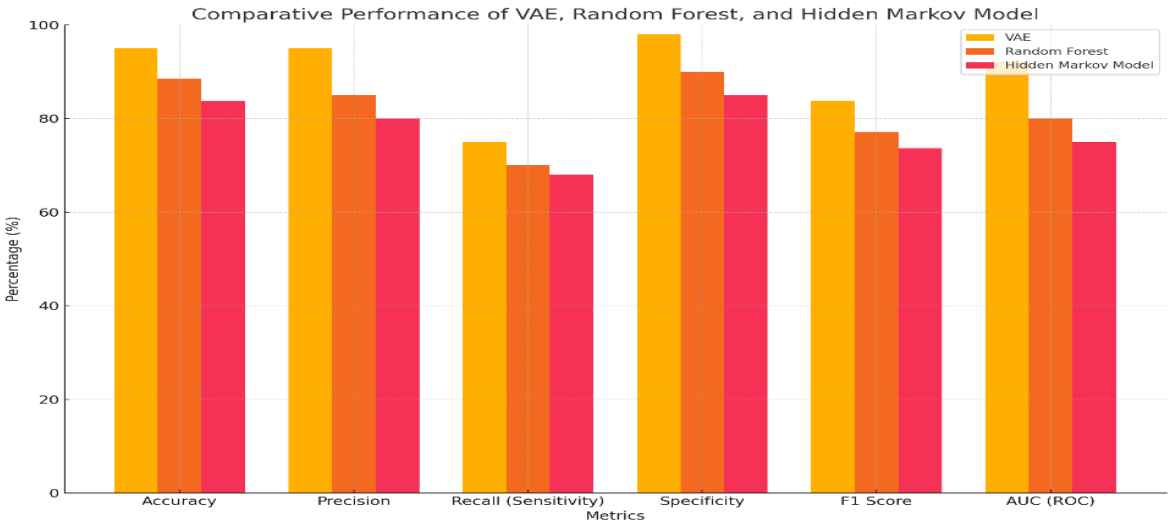

Figure 4: Comparative bar chart of models on key metrices

As shown in the figure below, the comparative bar chart presents the models of VAE, Random Forest, and HMM based on key performance metrics, accuracy, precision, recall, specificity, f1 score, and auc measures. It is found that the VAE model has the best accuracy of 95%, suggesting the model's ability to classify effectively. It has a very good precision rate of about 95%; on the whole, this contributes to a lower rate of False Alarms. The VAE exhibits a recall of 75%, signifying its ability to detect the majority of real anomalies, while its specificity of 98% demonstrates its proficiency in reliably recognizing normal samples. The VAE model exhibits an F1 score of 83.82%, indicating balanced performance, while an AUC of 92% demonstrates its robust capability to differentiate between normal and anomalous data. The VAE model is the most effective option for precise and equitable anomaly identification in healthcare IoT environments.

CONCLUSION

This study demonstrates the capabilities of Variational Autoencoders (VAEs) in addressing anomalies in healthcare IoT systems without prejudice, reducing serious shortcomings in data security and integrity. Similar studies involving simulated healthcare IoT datasets showed that the VAE model produced appreciable accuracy, precision, and specificity, suggesting its suitability for detecting normal components without the use of labeled data. The model's characteristics of high scalability and flexibility make it most appropriate for implementation in any healthcare environment where multiple IoT devices continuously generate multiple, diverse data inputs. The program reports which parameters differ from those previously set, and as a result, helps in identifying possible breaches of security and malfunctions of the system, thus, maintaining the confidentiality of information patients' care. The VAE methodology envisaged in the study is effective in real-time anomaly detection as it minimizes false alarms while detecting real anomalies. Its unsupervised learning approach meets the strict policies of the healthcare sector for sensitive information ever requiring minimum use of labelled data and ensuring continuous observation under real conditions. Future work may improve the model's memory, enabling it to adapt to more complex data structures. This study shows how healthcare IoT ecosystems can be secured by using VAEs, thus enabling secure, dependable, and efficient patient care systems.

REFRENCES

[1] M.M. Khan and M. Alkhathami, "Anomaly detection in IoT-based healthcare: Machine learning for enhanced security," Scientific Reports, 14(1), Springer Science and Business Media LLC, 2024.
[2] A. Abia, A. Ganesh, and R. Piechocki, "A novel intrusion detection scheme using variational autoencoders," International Symposium on Networks, Computers, and Communications, 2023, pp. 1–10.
[3] V. V. Raje, "Realtime anomaly detection in healthcare IoT: A machine learning-driven security framework," Science Research Society Journal of Electrical Systems, 19(3), 2023, pp.192-202.

Research Article

[4] H. C Huang, I. H Liu, M.H Lee, and S. J Li, “Anomaly detection on network traffic for the healthcare internet of things," In 2023 IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability Vol. 33, pp. 13-19, Nov 2023.

[5] A. Raza, K. P Tran, L. Koehl, S. Lie, “AnoFed: Adaptive anomaly detection for digital health using transformer-based federated learning and support vector data description,” Engineering Applications of Artificial Intelligence, Elsevier BV, 121, 106051, 2023.

[6] A. Said, A. Yahyaoui, and T. Abdellatif, “Efficient anomaly detection for smart hospital IoT systems,” Italian National Conference on Sensors,2021.

[7] S. Subha, J. G. R. Sathiaseelan, "The enhanced anomaly deduction techniques for detecting redundant data in IoT," International Research Journal on Advanced Science Hub, RSP Science Hub, 5(2), 2023, pp. 47-54.

[8] M. Mohammed, Q.. Salem and A. Mehaoua, “Artificial intelligence for anomaly detection in IoMTs,” International Symposium on Networks, Computers, and Communications, 2023.

[9] W. Gouda, S. Tahir, S. Alanazi, H. Almufareh, and G. Alwakid, “Unsupervised outlier detection in IoT using deep VAE,” Italian National Conference on Sensors, 2022.

[10] H. Gao, B. Qiu, R. J. D. Barroso, W. Hussain, Y. Xu, and X. Wang, "TSMAE: A novel anomaly detection approach for internet of things time series data using memory-augmented autoencoder," IEEE Transactions on Network Science and Engineering, 2023.

[11] A. Abusitta, G. H. S. Carvalho, O. A. Wahab, T. Halabi, B. Fung, and S. Al-Mamoori, "Deep learning-enabled anomaly detection for IoT systems," Internet of Things, 2022.

[12] M. A. Khatun, S. Memon, C. Eising, and L. L. Dhirani, "Machine learning for healthcare-IoT security: A review and risk mitigation," IEEE Access, 2024.

[13] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for IoT security attacks," IEEE Internet of Things Journal, vol. 9, no. 4, pp. 2545–2554, Feb. 2022.

[14] P. V. Astillo, D. G. Duguma, H. Park, J. Kim, B. Kim, and I. You, "Federated intelligence of anomaly detection agent in IoTMD-enabled diabetes management control system," Future Generation Computer Systems, vol. 128, pp. 395–405, Mar. 2022.

[15] D. Gupta, O. Kayode, S. Bhatt, M. Gupta, and A. S. Tosun, "Hierarchical federated learning based anomaly detection using digital twins for smart healthcare," in 2021 IEEE 7th International Conference on Collaboration and Internet Computing (CIC), Dec. 2021.

[16] K. P. Vijayakumar, P. Krishnadoss, A. Balasundaram, and M. R. Prusty, "Enhanced cyber attack detection process for internet of health things (IoHT) devices using deep neural network," Processes, vol. 11, no. 4, p. 1072, Apr. 2023.

[17] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q.. Jin, "Variational LSTM enhanced anomaly detection for industrial big data," IEEE Transactions on Industrial Informatics, vol. 17, no. 5, pp. 3469–3477, May 2021.

[18] S. H. Haji and S. Y. Ameen, "Attack and anomaly detection in IoT networks using machine learning techniques: A review," Asian Journal of Research in Computer Science, pp. 30–46, Jun. 2021.

[19] M. Adkisson, J. C. Kimmell, M. Gupta, and M. Abdelsalam, "Autoencoder-based anomaly detection in smart farming ecosystem," in 2021 IEEE International Conference on Big Data (Big Data), Dec. 2021.

[20] D. Gupta, M. Gupta, S. Bhatt, and A. S. Tosun, "Detecting anomalous user behavior in remote patient monitoring," in 2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI), vol. 60, pp. 33–40, Aug. 2021.

[21] M. Ahmed, S. Byreddy, A. Nutakki, L. F. Sikos, and P. Haskell-Dowland, "ECU-IoHT: A dataset for analyzing cyberattacks in internet of health things," Ad Hoc Networks, vol. 122, p. 102621, Nov. 2021.

[22] P. Kumar et al., "A blockchain-orchestrated deep learning approach for secure data transmission in IoT-enabled healthcare systems," Journal of Parallel and Distributed Computing, vol. 172, pp. 69–83, Feb. 2023.

[23] S. Saif et al.., "HIIDS: Hybrid intelligent intrusion detection system empowered with machine learning and metaheuristic algorithms for application in IoT-based healthcare," Microprocessors and Microsystems, vol. 104622, Aug. 2022.

[24] S. A. Wagan et al., "A fuzzy-based duo-secure multi-modal framework for IoMT anomaly detection," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 1, pp. 131–144, Jan. 2023.

[25] M. Kavitha et al., "Machine learning techniques for anomaly detection in smart healthcare," in 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Sep. 2021.

[26] J. Wang et al., "Anomaly detection in internet of medical things with blockchain from the perspective of deep neural network," Information Sciences, vol. 617, pp. 133–149, Dec. 2022.

[27] I. F. Kilincer et al., "Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization," Biocybernetics and Biomedical Engineering, vol. 43, no. 1, pp. 30–41, Jan. 2023.

[28] A. Jain, T. Singh, and S. K. Sharma, "Security as a solution: An intrusion detection system using a neural network for IoT-enabled healthcare ecosystem," Interdisciplinary Journal of Information, Knowledge, and Management, vol. 16, pp. 331–369, 2021.

[29] H. Gangloff et al., "Leveraging vector-quantized variational autoencoder inner metrics for anomaly detection," in 2022 26th International Conference on Pattern Recognition (ICPR), Aug. 2022.

[30] R. Zemouri et al.., "Hydrogenerator early fault detection: Sparse dictionary learning jointly with the variational autoencoder," Engineering Applications of Artificial Intelligence, vol. 120, p. 105859, Apr. 2023.

[31] G. Costa et al., "Trustworthy precision medicine: An interpretable approach to detecting anomalous behavior of IoT devices," in Studies in Health Technology and Informatics, May 2024.

[32] T. Preethi et al., "Advancing healthcare anomaly detection: Integrating GANs with attention mechanisms," International Journal of Advanced Computer Science and Applications, vol. 15, no. 6, 2024.

[33] K. Zhang et al., "Federated variational learning for anomaly detection in multivariate time series," in 2021 IEEE International Performance, Computing, and Communications Conference (IPCCC), Oct. 2021.

[34] S. Arul Jothi et al., "Data anomaly detection in wireless sensor networks using β\betaβ-variational autoencoder," in 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), vol. 7, pp. 631–636, Feb. 2023.

[35] Z. Zhu et al., "Using a VAE-SOM architecture for anomaly detection of flexible sensors in limb prosthesis," Journal of Industrial Information Integration, vol. 35, p. 100490, Oct. 2023.

[36] Y. R. Sampath Kumar and H. N. Champa, "Anomaly detection framework for efficient sensing in healthcare IoT systems," in 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 81–85, Nov. 2022.

[37] O. Salem et al., "Markov models for anomaly detection in wireless body area networks for secure health monitoring," IEEE Journal on Selected Areas in Communications, vol. 39, no. 2, pp. 526–540, Feb. 2021.

[38] Y. R. Sampath Kumar et al., "An efficient anomaly detection through optimized navigation using Dlvq-Cdma and H-Dso in healthcare IoT environment," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 9, pp. 2847–2859, Nov. 2023.

[39] P. Singh et al., "Dew-cloud-based hierarchical federated learning for intrusion detection in IoMT," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 2, pp. 722–731, Feb. 2023.

[40] K. Gupta et al., "A tree classifier-based network intrusion detection model for the internet of medical things," Computers and Electrical Engineering, vol. 102, p. 108158, Sep. 2022.

[41] M. Fouda et al., "A novel intrusion detection system for internet of healthcare things based on deep subclasses dispersion information," IEEE Internet of Things Journal, vol. 10, no. 10, pp. 8395–8407, May 2023.

[42] A. Albattah and M. A. Rassam, "A correlation-based anomaly detection model for wireless body area networks using convolutional long short-term memory neural network," Sensors, vol. 22, no. 5, p. 1951, Mar. 2022.

Research Article

[43] R. Chaganti et al., "A particle swarm optimization and deep learning approach for intrusion detection system in the internet of medical things," Sustainability, vol. 14, no. 19, p. 12828, Oct. 2022.

[44] H. Zhou and C. Kan, "Tensor-based ECG anomaly detection toward cardiac monitoring in the internet of health things," Sensors, vol. 21, no. 12, p. 4173, Jun. 2021.

[45] S. Khan and A. Akhunzada, "A hybrid DL-driven intelligent SDN-enabled malware detection framework for internet of medical things (IoMT)," Computer Communications, vol. 170, pp. 209–216, Mar. 2021.

[46] A. Tabassum et al.., "Privacy-preserving distributed IDS using incremental learning for IoT health systems," IEEE Access, vol. 9, pp. 14271–14283, 2021.

[47] N. Savanović et al., "Intrusion detection in healthcare 4.0 internet of things systems via metaheuristics optimized machine learning," Sustainability, vol. 15, no. 16, p. 12563, Aug. 2023.

*** Ethical Approval**: Human participants or data not applicable; data is taken from a public data research repository.

* **Informed Consent**: Not Applicable

* **Statement Regarding Research Involving Human Participants and/or Animals**: Not Applicable

* **Consent to Participate**: Not Applicable

* **Consent to Publish**: Not Applicable

* **Funding**: Not Applicable

* **Author's Contribution**: I am the sole author of this paper. The second and third authors are my Ph.D. supervisor and co-supervisors, who guided me in performing the research, drafting and implementing designing viz.

* **Competing Interests**: Not Applicable

* **Availability of data and materials**: Provided data source in research methodology

*** Declaration**: I affirm that the research is original, and I have maintained the integrity and ethical considerations of the research.

*** Acknowledgment.** I acknowledge my ''***Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India***" providing me infrastructure to conduct this research, and there are no potential conflicts of interest.