IPS Journal of Physical Sciences
ISSN: 3115-5715
IPS J. Phys. Sci., 3(1): 176-183 (2026)
DOI: https://doi.org/10.54117/ijps.v3i1.27

Hybrid Cloud Security Framework for E-Health Data Protection using Machine Learning and Advanced Encryption Techniques

Mgbeafulike, I. J.[1], Okeke, O. C.[1], Ugbor, I. C.[2*], Uba, C. P.[1], Emeto, I. C.[2], Ewunonu, T. C.[2], Ibeneme-Sabinus, I. L.[2], and Amaka, E. N.[2]

[1] Department of Computer Science, Faculty of Physical Sciences, Chukwuemeka Odumegwu Ojukwu University, P.M.B. 02, Uli, Anambra State, Nigeria.
[2] Department of Cyber Security, School of Information and Communication Technology, Federal University of Technology, Owerri, Imo State.

*Corresponding author- ugbor.ihechiluru@futo.edu.ng; Tel: +234 8037294485

Abstract

This study presents a hybrid cloud security framework for E-Health data protection that integrates machine learning-based behavioural analysis with advanced encryption techniques. With the use of Agile methodology, the system was developed to iteratively refine the proposed encryption algorithms in order to ensure adaptive functionality of the technique. A primary dataset made up of 15,001 user activity logs was collected from a cloud-based healthcare platform (Anderson Hospital) capturing both legitimate, suspicious or malicious behaviours. Furthermore, the dataset was then pre-processed using missing value imputation, min-max normalization and Principal Component Analysis (PCA) so as to optimize model training process. A Multilayer Perceptron (MLP) neural network was trained for the prediction of user sessions into three categories such as Legitimate, Suspicious or Malicious. The model achieved strong predictive continuous threat score performance with $R^2 = 0.9946$, MAE = 0.0689, and MSE = 0.0188, demonstrating a high predictive accuracy. For data protection, AES-128-bit encryption was used for routine access, while a hybrid Advanced Encryption Standard- Rivest–Shamir–Adleman (AES-256+RSA) approach secured high-risk scenarios. Then, the Experimental results show that the hybrid system provides robust security with acceptable processing overhead, ensuring confidentiality, integrity, and secure access control to sensitive health records. It significantly enhances security against key exchange vulnerabilities and interception attacks.The framework demonstrates the feasibility of real-time cloud-based E-Health data protection and provides a practical solution for safeguarding sensitive healthcare information.

Keywords: Cloud Security; Machine Learning; Multilayer Perceptron (MLP); AES Encryption; RSA Encryption; E-health.

Article History
Received: 20 Apr 2026
Accepted: 11 May 2026
Published: 21 May 2026

Scan QR code to view•

How to cite this paper: Mgbeafulike, I., Okeke, O. C., Ugbor, I. C., Uba, C. P., Emeto, I. C., Ewunonu, T. C., Ibeneme-Sabinus, I. L., & Amaka, E. N. (2026). Hybrid Cloud Security Framework for E-Health Data Protection using Machine Learning and Advanced Encryption Techniques. *IPS Journal of Physical Sciences*, *3*(1), 176–183. https://doi.org/10.54117/ijps.v3i1.27

1. Introduction

Cloud computing has quickly become one of the most desired platforms in the current computing and has provided organizations with flexibility, scalability and cost-effective solutions. Cloud technology is an emerging technology that came after the achievements of grid computing, utility computing, and distributed computing, with the technology transforming data storage and service delivery (Singh and Chatterjee, 2020). It has been very appealing in all industries due to its simplicity and efficiency. Nevertheless, there are also serious issues regarding security presented by cloud storage, especially in the safety of sensitive data. Threats of malicious insiders, external attacks and insufficient access controls are some of the major challenges in the cloud environments (Kumar et al., 2022). Although the cybersecurity is widely in place, the number of data breaches is increasing, which proves the inefficiency of the traditional security mechanisms. Traditional encryption methods can be ineffective, particularly in the process of protecting dynamic data, because of excessive computational cost, ineffective key management, and restricted flexibility (Zobaed and Amini, 2023).

To effectively manage these issues, this study also suggests a new hybrid cryptographic model that will be used to protect dynamic data, which is regularly updated, edited, or erased, i.e., real-time databases and collaborative systems, and streams of data. The hybrid structure combines the symmetric and asymmetric encryption techniques, pooling their advantages together to offer a high level of security, flexibility and performance. This guarantees the confidentiality, integrity and, availability of sensitive data in the cloud. Due to the dynamic nature of cloud security, an intelligent and flexible

system is essential to success in protection (El-Attar et al., 2021).

The present paper is dedicated to design and implementation of a hybrid cryptography system designed specifically to provide dynamic data security in cloud applications. The proposed system overcomes the current security drawbacks by providing powerful mechanisms of keeping data safe without impairing the performance. It facilitates smooth adaptation to the regular changes in data, offers scalable encryption policies, and implements robust cryptographic keys management (Ahirwar and Tyagi, 2024). Moreover, it is equipped with real-time threat detection capabilities to prevent unauthorized access and respond to cyberattacks efficiently (Nwatuzie et al., 2025).

Contribution to Knowledge

1. This research contributed to knowledge by demonstrating a novel approach to securing cloud environments through a dual-layered framework that combines traditional cryptographic methods with intelligent machine learning algorithms.
2. The study showed how real-time learning can enhance encryption systems beyond static protection, offering dynamic security responses.
3. The study provided practical implementation guidance and performance validation, serving as a reference model for future developments in secure and intelligent cloud computing solutions.

2. Methodology

2.1 Development Approach

The methodology adopted for this research is the Agile software development methodology. Agile methodology promotes an iterative and incremental approach to software development, allowing for frequent evaluations, stakeholder feedback, and adaptive planning. This was particularly valuable in refining the framework's encryption algorithms, dynamic data operation support, and key lifecycle management. Agile sprints enabled the continuous integration and testing of security features, such as encryption/decryption efficiency, integrity verification, and user authentication modules. This methodology supported the goal of delivering a robust, secure, and practically implementable cryptographic framework for cloud data protection.

2.2 System Overview

The proposed framework consists of three major components:

1. **Data Processing Layer**
2. **Machine Learning Threat Detection Layer**
3. **Adaptive Encryption Layer**

2.3 Data Collection

The data used for this work is a primary dataset collected directly from a cloud- based healthcare system(Anderson Hospital). The sample size consists of 15,001 records representing user activity log behaviours, captured over several operational cycles involving both legitimate and simulated malicious users. The data reflects various patterns of access to sensitive cloud-stored electronic health records (EHRs), enabling effective training of the machine learning threat detection models. Each record contained 15 distinct attributes, including login stamps, user roles, device types, IP addresses, session durations, access pattern, encryption protocols applied and machine learning threat prediction outputs. These attributes represent key features indicative of user behaviour and access risk. The dataset also includes a target attribute labelled `threat_level`, predicting each session as either Legitimate, Suspicious, or Malicious, which serves as the ground truth for supervised learning. All data were securely collected and anonymized using de-identification techniques to maintain patient confidentiality and comply with data privacy regulations and hospital policies. Access is restricted through role-based authentication, allowing only authorized personnel (doctors, admins) to decrypt data. This historical log forms the foundation for training, evaluating, and validating the proposed intelligent cryptographic protection framework for cloud-based E-Health systems.

2.4 Data Preprocessing

To prepare the dataset for machine learning analysis, a structured data processing pipeline was implemented, comprising missing value imputation, min-max normalization, and Principal Component Analysis for dimensionality reduction. The collected access log data from the cloud testbed had occasional missing entries due to session timeouts or incomplete data transfers. These missing values were handled using statistical imputation. For numerical attributes such as session duration, encryption latency, and packet size, missing values were filled using the mean of each respective column. For categorical features like user role or device type, the most frequent value (mode) was used. After handling missing values, all numerical features were scaled using min-max normalization to ensure uniformity and remove bias caused by varying feature ranges. To reduce dimensionality and improve model efficiency, PCA was applied after normalization. PCA transformed the original feature space into a smaller set of uncorrelated principal components, retaining over 95% of the dataset's variance. This step helped eliminate redundant or correlated features, speeding up the training process while preserving essential behaviour patterns required for threat prediction. Through these preprocessing steps, the dataset was optimized for the training of machine learning, ensuring high model accuracy, faster convergence, and improved generalization on unseen cloud activity logs.

2.5 The Machine Learning Model for Behavioural Analytical Model

The machine learning algorithm used for this work is the multi-layered neural network. A Multilayer Perceptron (MLP) is a class of feed-forward artificial neural networks consisting of at least three layers: an input layer, one or more hidden layers, and an output layer. Each node (neuron) in a layer is connected to every node in the next layer and performs a weighted sum followed by a non-linear activation (Elmezughi, et al., 2022). The proposed model is based on ANN architecture that includes two hidden layers of feed-forward neural networks. The first hidden layers contain 96 neurons, and the last hidden layers have 32 neurons, with each hidden layer followed by the ReLU activation function. The input for the network accepts all features from the processed data. The output layer is a single neuron with a linear activation function as the transfer function that leads to the predicted value. The

output result is a real value representing the threat and malicious packet. The hyper-parameter values of the proposed ANN (number of layers, number of neurons, activation functions, learning rate = 0.001and gives the best hyper-parameters that fit the measured data during the training process using gradient descent back-propagation algorithm. The MLP model learns to move up or down depending on the trend feature extraction from the data, giving the fit curve. The parameter weight at every epoch is adjusted using the gradient descent with momentum to reach the global minimum error. The proposed model uses a comprehensive hyper-parameter set up to identify the best weights of the parameters for the prediction of threat.

Mathematical Formulation

$$y = W_2 \cdot f(W_1 X + b_1) + b_2$$

Where:

- y= predicted threat score
- f= ReLU activation
- W, b= model parameters

2.6 The Proposed Advanced Encryption for Security of cloud Data

This work proposed two different encryption methods for data protection in the cloud network. The first method is the Advanced Encryption Standard–128-bit, while the second is a hybrid approach which integrated the Rivest Shamir Adleman (RSA) and AES-256 as an improved encryption solution.

2.6.1 Advanced Encryption Standard (AES)-128-bit

AES-128bit encrypts and decrypts data using the same 128-bit key, making it efficient and fast for handling large volumes of data. AES-128 works by dividing the input plaintext into 128-bit blocks and then processing each block through 10 rounds of transformation. Each round includes four key operations: SubBytes (a nonlinear substitution using a predefined S-box), ShiftRows (which shifts the rows of the state array to achieve diffusion), MixColumns (a matrix multiplication over a finite field to further spread the bits), and AddRoundKey (which combines the block with a round-specific key derived from the original key). The first and final rounds are slightly modified to improve cryptographic strength.

Algorithm 1: AES-128bit

1. A 128-bit secret key is generated using a secure pseudo-random number generator
2. Load patient data
3. Convert patient data to binary and divided into 128-bit blocks.
4. Apply 10 rounds of transformations
5. Encrypt data
6. Upon data request = true
7. Apply decryption key
8. Decrypt data
9. End.

2.6.2 Rivest–Shamir–Adleman (RSA)

RSA is an asymmetric cryptographic algorithm that uses a pair of mathematically linked public key for encryption and a private key for decryption. The security of RSA relies on the computational difficulty of factoring the product of two large prime numbers. It uses large key pairs between 1024 and 4096 bits, to secure data through encryption and decryption. To generate RSA keys, two large primes p and q are chosen, and their product $n = pq$ forms part of both the public and private keys.

Algorithm 2: Rivest-Shamir-Adleman

1. Key generation phase
2. Select two large prime numbers: p and q
3. Compute modulus: n = p x q as modulus for keys.
4. Compute Euler's totient: $\varphi(n) = (p - 1)(q - 1)$
5. Choose public exponent e:
6. $1 < e < \varphi(n)$ such that $\gcd(e, \varphi(n)) = 1$
7. Compute private exponent d:
8. $d \equiv e^{-1} \bmod \varphi(n)$ (modular inverse)
9. Publish public key: (e, n)
10. Keep private key secret: (d, n)
11. Encryption Phase
12. Obtain receiver's public key (e, n)
13. Convert plaintext message M to integer m such that $0 < m < n$
14. Compute ciphertext: $c \equiv m^e \bmod n$
15. Send ciphertext c to receiver
16. Decryption Phase
17. Use private key (d, n)
18. Decrypt ciphertext: $m \equiv c^d \bmod n$
19. Convert integer m back to original message M
20. End

2.7 System Flowchart

The system flowchart was presented in Figure 1. In the diagram the user activities input to the network are identified and the trained MLP model predicted the logs to detect threat and normal user. For threat the hybrid encryption algorithm is applied to generate a more complex and secured decryption key. when the data access request is from legitimate user, the system generate AES key to encrypt the data. In the Figure 1, to access data, the authenticate user decrypt only if authorized to decrypt the data and view health records.

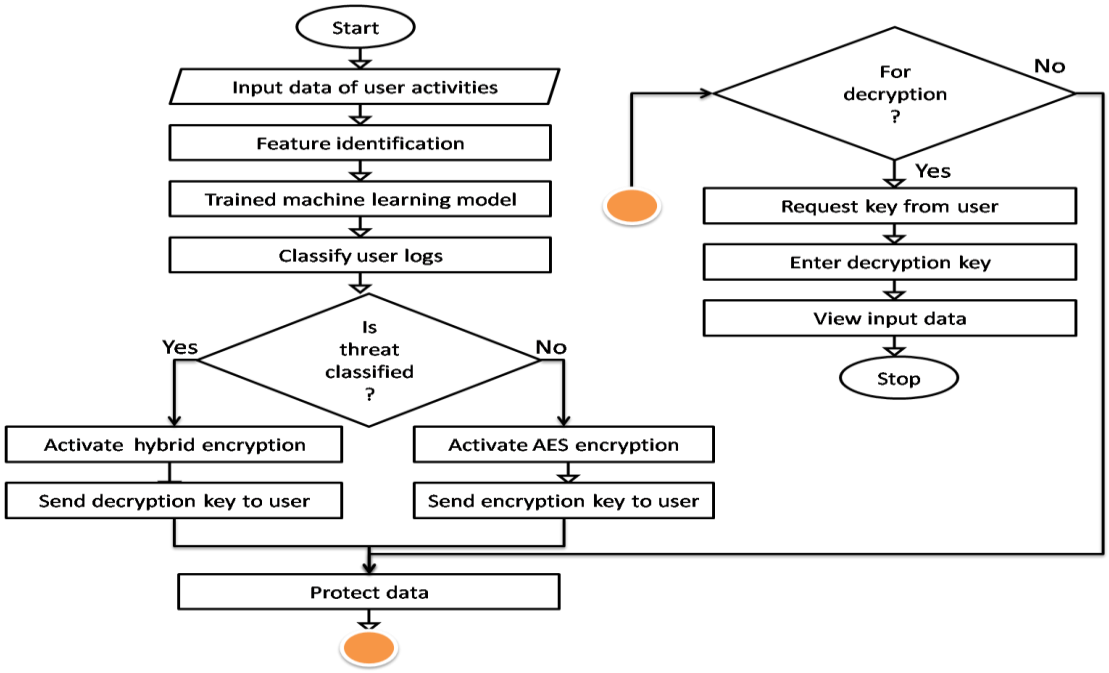

Figure 1: Flowchart of the hybrid encryption solution for cloud security

2.8 System Implementation

The hybrid cloud encryption system with machine learning was implemented using a layered approach that combined cryptographic techniques with intelligent pattern recognition. This was achieved by integrating a hybrid encryption mechanism that leveraged both symmetric and asymmetric encryption methods to ensure data confidentiality, integrity, and secure access control within the cloud environment. The MLP was trained to recognize patterns associated with secure and insecure data activities. Inputs were passed through multiple hidden layers, where features were processed using nonlinear activation functions to capture complex relationships within the data. The output layer provided predictive decisions, supporting the overall security posture of the system. The integration of encryption and MLP allowed for continuous monitoring of data behaviour in real time. Encrypted data, user access patterns, or related metadata were analysed by the MLP to flag anomalies or unauthorized activities. The system was designed to respond adaptively to potential threats, improving overall reliability and responsiveness.

2.9 Performance Evaluation

The metrics used for the evaluation of the new system performance are Mean Square Error (MSE), coefficient of determination R^2, Mean Absolute Error (MAE),

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y - \tilde{y}| \quad (2.1)$$

$$MSE = \sum_{i=1}^{n}\frac{(y-\tilde{y})^2}{N} \quad (2.2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - Y)^2}{\sum_{i=1}^{n}(y_i - Y_i)^2} \quad (2.3)$$

Where y is the dataset, $\tilde{y}$ is the predicted malicious data, N is the overall observation applied to train the model. y_i is the actual value for the i-th data point, Y is the predicted value, Y_i is the mean of all actual values and n s the number of samples.

3. System Results

This section presents the various results attained from implementing the various techniques before integration as a hybrid protection system. The results gotten from training the MLP and execution of the AES technique are presented before the final outcome from the execution of the hybrid system model.

3.1 Results MLP training

The MLP training performance was measured and reported in this section. The results showed the correlations matrix between the multi variables in the dataset. This result is necessary for the data analysis to reveal correlation patterns in the data, their alignment and how they perfectly fit to train a model. From the results, it was observed that the diagonal matrix which related each attributes scored 1.00 which means that each of the data attributes perfectly fits and make the dataset good for further analysis. In the Figure 2, the class distribution was analysed. The reason was to identify those classes which are bias that might affect the quality of classification output. From the results recorded, it was observed that the target class without abnormally behaviour recorded 9000 features, while the classes which model normal client behaviour recorded 6000 features. These results implied that the class is bias and then SMOTE was applied to balance the outcome and the results presented in Figure 3.

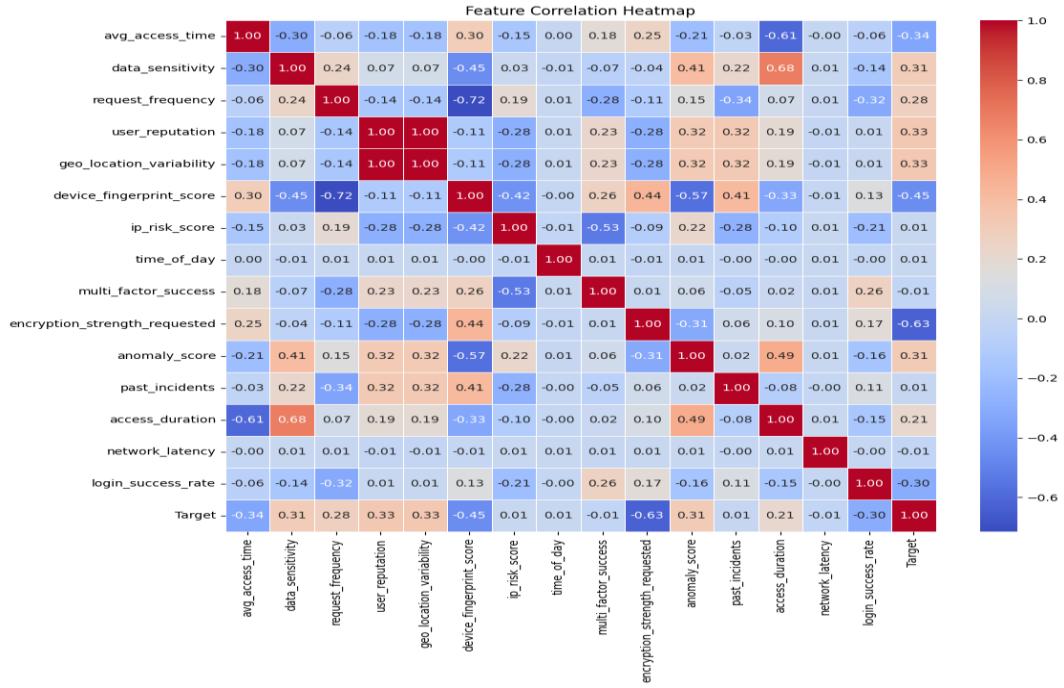

Figure 2: Result of the data correlation matrix

Figure 3: Result of the target class distribution

Figure 2 reported the data correlation matrix, while Figure 3 reported the class distribution, showing that the class with abnormally recorded lower distribution and need to be optimized. The data was then applied to train MLP for the prediction of user log activities. The results of the training were reported in figure 4 for MAE, figure 5 for MSE and coefficient of determination result in figure 6 respectively. The MAE measures the error by calculating the average of the absolute differences between the predicted and actual values. A lower MAE value, as shown in Figure 4, indicates that the model's predictions are consistently close to the actual outputs. Figure 5 presents the MSE, which penalizes larger errors more heavily by squaring the differences before averaging. This makes MSE particularly useful for identifying models that make large prediction errors. Finally, the coefficient of determination shown in figure 6 quantifies how well the model explains the variance in the target variable, with higher values closer to 1 indicating a better fit.

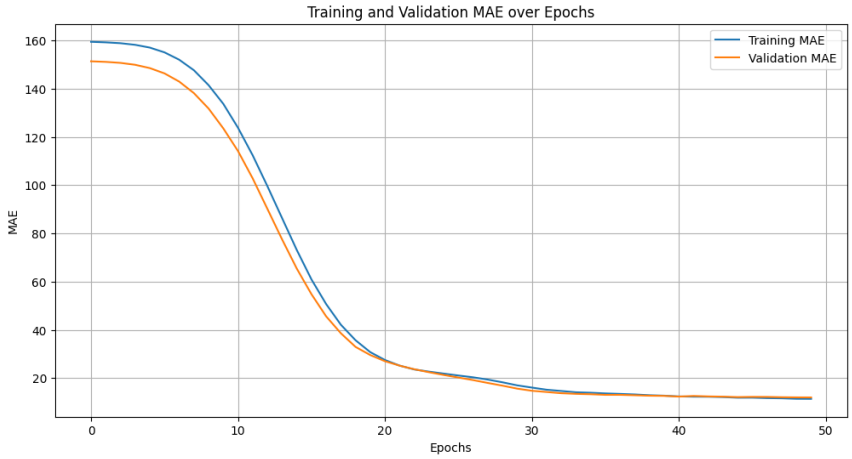

Figure 4: Result of the MAE

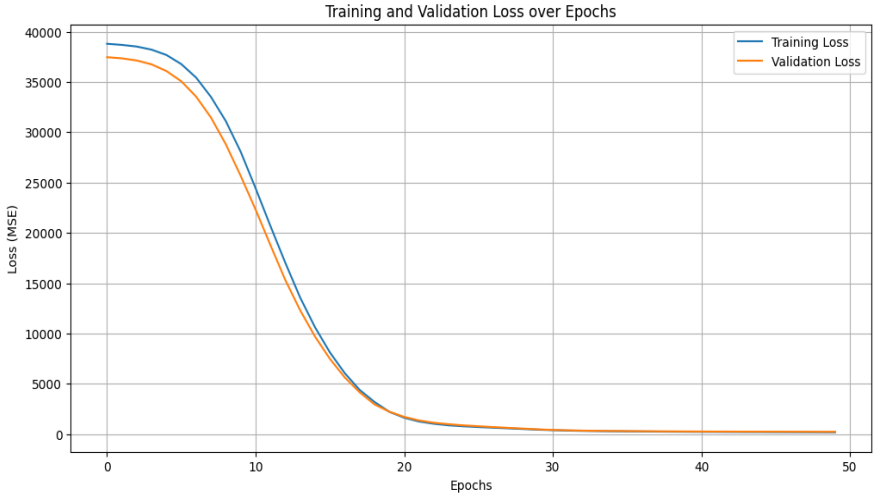

Figure 5: Result of the MSE

The MAE curve in Figure 4 shows a consistent decline over epochs, indicating that the model's absolute prediction error reduced as training progressed and recorded 0.0689. Similarly, the MSE trend in Figure 5 confirms that the model's overall prediction error decreased steadily and recorded 0.0188.

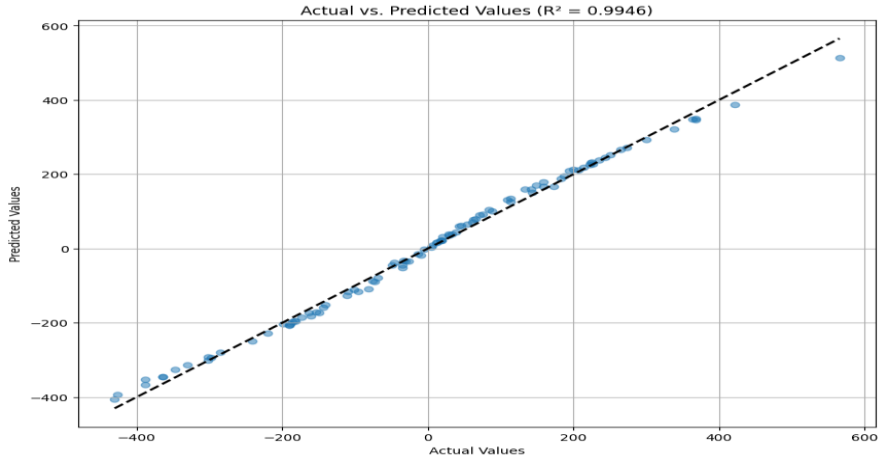

Figure 6: Result of the R^2

Furthermore, the coefficient of determination (R^2) presented in Figure 6 demonstrates a strong fit between the predicted and actual values, signifying the model's effectiveness in capturing the underlying patterns in the data. This model explains ~99.46% of variance. Therefore low MAE = 0.0689; low MSE = 0.0188 and High R^2 = 0.9946 strongly indicates an excellent model fit and predictive reliability.

3.2 Result of the AES encryption technique

This section presents the performance of the AES technique applied for the encryption of data. The encryption result was measured considering time of encryption, decryption, throughput, memory usage and CPU utilization factor as shown in Figure 7.

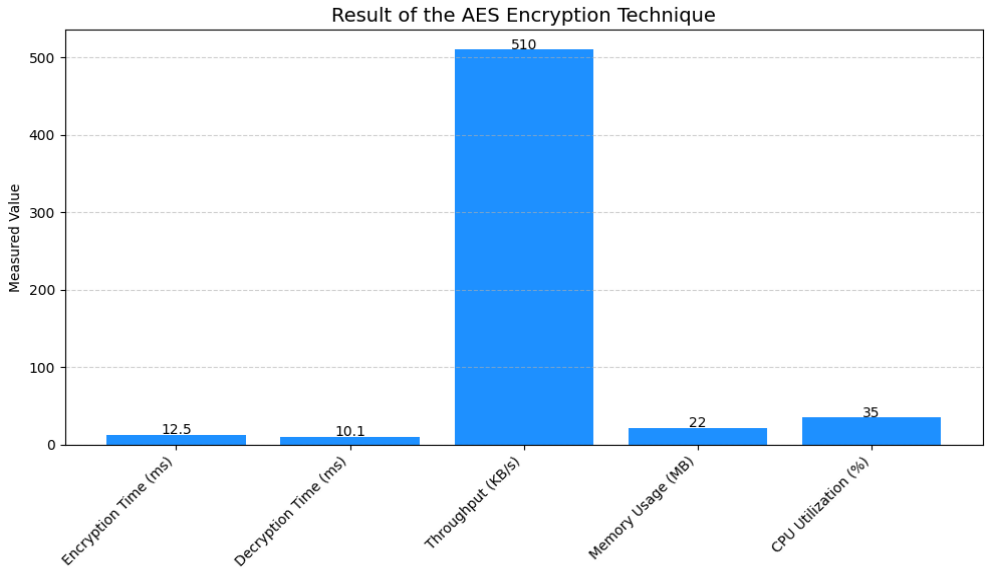

Figure 7: Result of the AES technique when applied for data encryption

The Figure 7 presents the result of the AES when applied for the data encryption. The results showed that the time it takes to encrypt file when the user is normal packet. The encryption key sent to the user is generated with the AES approach and then applied to encrypt the file, while decryption key is applied to decrypt the file. From the result, it was observed that to encrypt file, the AES takes 12.5ms, while to decrypt file the AES takes 10.1ms. The throughput which is the amount of data that can be managed by the AES is 510KB/s which is good, while the memory capacity consumed is 22MB, the CPU utilization factor during the data processing is 35% which is also good and showed limited energy consumption while encrypting and decryption of files. In the next results, the performance of the system when applying hybrid encryption approach or the security is data is reported in Figure 8.

3.3 Result of the Hybrid Encryption technique

This section presents the performance of the system when the hybrid encryption solution was applied for the security of data. From the results, it was observed that while the machine learning model was able to correctly classify abnormal behaviour from user activity log, the model then applied the hybrid encryption solution which combines AES and RSA which are the encryption of files. Figure 8 presents the results of the hybrid solution performance.

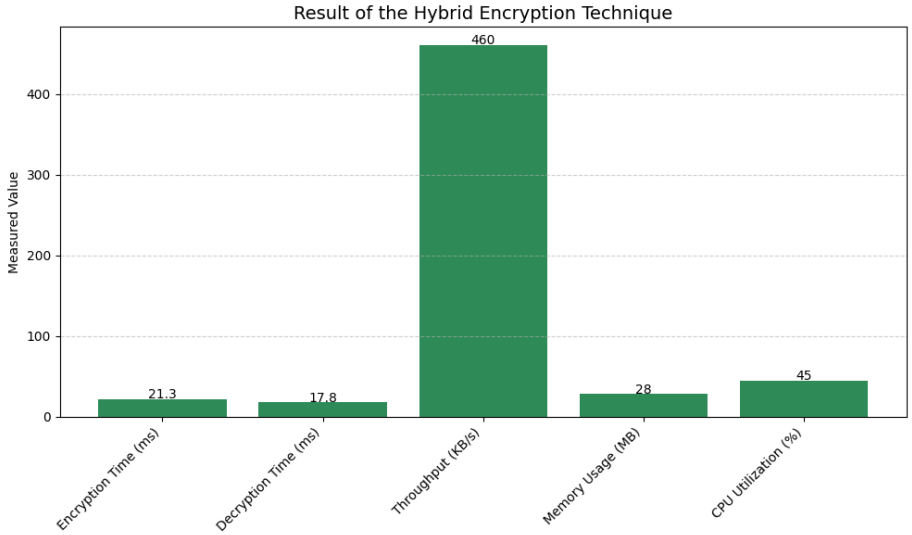

Figure 8: Result of the hybrid encryption solution

Figure 8 shows the time it takes to encrypt a file when the user is under threat. The encryption key sent to the user is generated using the AES+RSA hybrid approach, combining the speed of AES for data encryption with the secure key exchange of RSA. In this process, a unique symmetric key is created and encrypted with the recipient's RSA public key before being

transmitted. Upon receiving the encrypted file and key, the user applies their RSA private key to decrypt the symmetric key, which is then used to decrypt the file content. This dual-layer approach ensures both the confidentiality of the data and secure delivery of the encryption key. Although the hybrid method introduces slightly higher encryption and decryption times compared to standalone AES, it significantly strengthens the overall security of the system, particularly in high-risk communication scenarios.

From the result, it was observed that to encrypt file, the AES takes 21.3ms, while to decrypt file the AES takes 17.8ms. The throughput which is the amount of data that can be managed by the AES is 460KB/s which is good, while the memory capacity consumed is 28MB, the CPU utilization factor during the data processing is 45% which is also good and showed limited energy consumption while encrypting and decryption of files.

Comparative Analysis

Method	Security	Speed	Key Security
AES	Medium	High	Low
RSA	High	Low	High
Proposed Hybrid	Very High	Moderate	Very High

4. Conclusion

This paper suggested a hybrid cloud security model to protect E-Health data using a combination of machine learning-based behavioural analysis with superior encryption methods. The evaluation of performance revealed that the hybrid system was a bit slower than the AES alone but offered greater security with acceptable throughput, memory and CPU usage. To sum up, the framework proposed provides a strong, safe, and effective solution to the cloud-based E-Health data protection. The combination of proper behavioural threat identification and adaptive encryption will provide confidentiality, integrity, and secure access control to sensitive health records. This strategy shows the possibility of real-time deployment, and the future work which may be aimed at the expansion of the system to bigger datasets and adding more machine learning methods to increase threat detection further.

References

Ahirwar, M., & Tyagi, M. (2024). Hybrid cryptographic framework for cloud data security: Integration of AES-OTP, RSA, and temporal access control mechanisms. *International Journal of Telecommunications & Emerging Technologies, 10*(2), 1-6.

El-Attar, N. E., El-Morshedy, D. S., & Awad, W. A. (2021). A new hybrid automated security framework to cloud storage system. *Cryptography, 5*(4), 37. https://doi.org/10.3390/cryptography5040037

Elmezughi, M. K., Salih, O., Afullo, T. J., & Duffy, K. J. (2022). Comparative Analysis of Major Machine-Learning-Based Path Loss Models for Enclosed Indoor Channels. *Sensors*, *22*(13), 4967. https://doi.org/10.3390/s22134967

Kumar, R., Singh, M., & Sharma, A. (2022). A review on cloud computing security challenges and solutions. *International Journal of Computer Applications, 184*(20), 1–5.

Nwatuzie, G. A., Enyejo, L. A., & Umeaku, C. (2025). Enhancing cloud data security using a hybrid encryption framework integrating AES, DES, and RC6 with file splitting and steganographic key management. *International Journal of Innovative Science and Research Technology, 10*, 1555-1569.

Singh, A., & Chatterjee, K. (2020). Cloud computing: Security issues and research challenges. *International Journal of Computer Science and Information Technologies, 11*(2), 123–129.

Zobaed, M. A., & Amini, M. (2023). Securing dynamic cloud data using adaptive cryptographic models. *Journal of Cloud Computing Advances, 12*(1), 45–58.

• *Thank you for publishing with us.*