Uncovering Latent Patterns in Ransomware Attack Behaviour and Recovery Outcomes Using Neural Network-Based Pattern Recognition in Healthcare Organizations

Rajesh Mojumder
School of Data and Science, Department of Computer Science and Engineering
BRAC University
Dhaka,Bangladesh
rajesh.mojumder@g.bracu.ac.bd

***Abstract*—Healthcare systems are now a major target of ransomware attacks. These attacks affect their data and daily operations. In this study, we use an unsupervised deep learning method. It combines an autoencoder with three clustering algorithms: KMeans, Agglomerative Clustering, and GMM. We use a dataset of 5000 ransomware cases. We convert the data into 8 features using an autoencoder. Then, we use clustering to group the organizations based on their response behavior. All three methods were found to be in the same groups: Resilient, vulnerable, and hybrid responders. KMeans gave the best result with a silhouette score of 0.2106. We also checked the reconstruction error. 5cases were found as anomalies. These are rare or very serious attacks. This method does not need labeled data. It helps hospitals and clinics understand risks better and plan safer recovery steps.**

***Index Terms*—Ransomware, Healthcare Cybersecurity, Autoencoder, Unsupervised Learning, Clustering, Anomaly Detection, KMeans, Agglomerative, GMM**

I. Introduction

Ransomware is a big threat to healthcare today. It can stop clinical work and steal medical records. It can also cost a lot of money to fix. Hospitals are at higher risk. This is because patient data is sensitive and services must run all the time. Normal detection systems need labeled data. However, labeled data is often missing or comes too late. So, these systems cannot detect new or unknown attacks well. To solve this, we use an unsupervised learning method. We train an autoencoder to learn the main features of incident data. Then, we use three clustering models: KMeans, Agglomerative, and GMM. These help find groups of organizations based on how they respond to attacks. Our method also helps detect rare cases. This helps healthcare systems prepare better and respond faster.

II. Related work

A. Rising Threat of Ransomware in Healthcare

Healthcare organizations have been under a significant increase in ransomware attacks, which affected 72.7% of formations in 2023 [1]. Driving through the weakness of legacy technology, these attacks use hefty demands on ransom and serious operational setbacks. With the evolution of ransomware linked to the increased RaaS trend, traditional detection procedures are floundering to keep up pace. Consequently, today, there are no less important simple technologies such as machine learning (ML) or deep learning (DL) for vital enhancement of detection and recovery methods.

B. Advancements in Deep Learning for Detection

Deep learning models include CNNs, RNNs, and LSTMs, which are used for ransomware detection. One of the important properties of these models is their effectiveness in processing high-dimensional information, which includes network traffic and system behaviours, and hence they are highly efficient in detecting malicious activities. A study by Alzahrani et al. (2025) indicates that CNNs and RNNs are the most commonly used architectures for real-time anomaly detection [2]. It is discovered that incorporating CNNs and LSTMs with hybrid approaches can speed up novel ransomware variant detection by up to 30% [3]. It has been observed that conventional approaches have been far surpassed by deep learning-based detection models. Research presented in Cyber Security and Applications (2025) proved that DNNs were able to have a detection accuracy of 95.4%, which far exceeded the standard methods, which only worked at 83.1% [4]. Moreover, a hybrid of DNNs and Support Vector Machines (SVM) achieved rates of detection of up to 98.7% [3].

C. Contrastive Learning in Action Recognition

Improved detection techniques have not sufficiently slowed the persistent problem of recovery. According to Sophos (2022), healthcare organisations could retrieve some data almost every time, yet only approximately 60% of their total database was not breached by paying the ransom. Over six and a half of every ten organisations paid a median ransom of $2.3 million, bringing home the complexity of the monetary impact ransomware may cause [2]. The recoveries from cases

of ransomware can add up to $10.5 million when considering downtime and reputational damage. Relevant cases, such as LockBit, that terrorised Royal Mail in 2023, unveil the significant operational impacts of these attacks. Utilisation of Threat Intelligence (TI) frameworks such as RDoS-CoAP, deployed in a manner such that it can leverage CNNs and LSTMs to support real-time network anomaly detection, offers a viable option in a battle against advanced IoT-targeting ransomware [5]. Such innovations need to be implemented to strengthen healthcare cybersecurity defence and mitigate the impacts of newly emerging ransomware threats [6].

III. Methodology

A. Overview

This research uses an unsupervised deep learning model to identify ransomware response behaviour in healthcare organisations. The method follows a clear step-by-step process. Each step is designed to transform raw input data into meaningful behavioural clusters. The proposed framework uses an autoencoder to reduce the high-dimensional ransomware data into a small set of important features. These features are then used by 3 clustering algorithms. Then it was to group similar ransomware behaviours without using any labels. This method follows FAIR and CARE principles by ensuring that data is used responsibly, transparently, and without harming the rights of any group.

B. Workflow Diagram

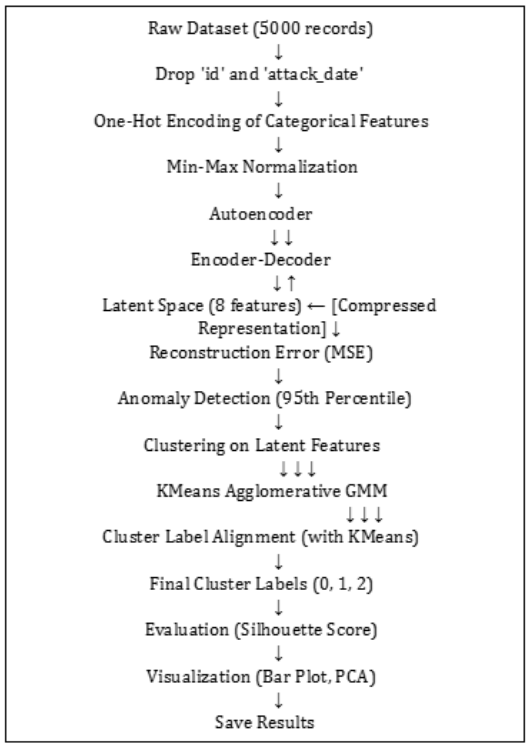

Fig. 1. **Workflow Diagram**

C. Data Acquisition

The dataset contains 5000 records. First, the ID and attack date columns are removed because they do not help in the analysis. Then, categorical features like org type, entry method, and paid ransom are converted into a numerical format using one-hot encoding. After that, all numerical values are scaled between 0 and 1 using Min-Max normalisation. This scaling helps in the stable training of neural networks.

D. Autoencoder Model and Architecture Diagram

An autoencoder is used to reduce the feature space. The input has 30 features. The encoder has four layers: 64, 32, 16, and 8 neurons. ReLU is used as the activation function. Dropout is used to prevent overfitting. Batch normalisation is added for stability. L2 regularisation is applied to improve generalisation. The decoder mirrors the encoder. The model is trained using the Adam optimiser. The loss function is a mean squared error. Early stopping is applied to avoid overfitting. The output of the encoder is an 8-dimensional latent feature for each record.

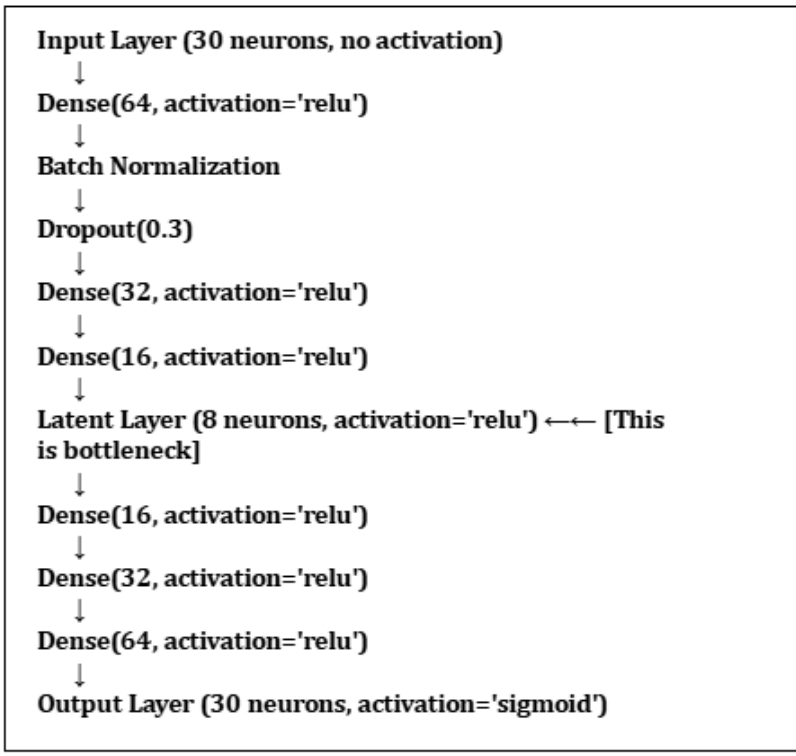

Fig. 2. **Autoencoder Architecture Diagram**

E. Anomaly Detection

After training the autoencoder, the input is reconstructed. Reconstruction error is calculated using mean squared error. The 95th percentile of reconstruction error is used as a threshold. Samples above this threshold are marked as anomalies. These may indicate severe or unique ransomware

attacks.

F. Latent Feature Extraction

The dataset first has 30 features after encoding the original categorical and numerical values. These include one-hot encoded columns like org type, entry method, and others. The autoencoder is trained to compress this 30-dimensional data. After training, the encoder part is used to extract the latent features. Each input of 30 features is passed through the encoder and becomes an 8-dimensional output. These 8 values are the compressed form of the original data. They carry the main information needed for clustering.

G. Dimensionality Reduction for Visualisation

Principal Component Analysis (PCA) is applied to reduce the latent space to two dimensions. This helps to visualise clusters in 2D space. It makes it easier to observe group separation.

H. Clustering Algorithms and Cluster-Level Adjustments

Three clustering algorithms are applied to the latent features. The first is KMeans clustering. It forms three clusters. The second is Agglomerative clustering. It uses Ward linkage to form clusters based on hierarchy. The third is the Gaussian Mixture Model (GMM). It uses probabilistic methods to assign clusters. Each algorithm is set to create three clusters. Each clustering algorithm gives different cluster labels. Cluster 0 in one method does not always mean the same group in another. To make them match, K-Means is used as the reference. First, the centroid of each cluster is calculated using the latent features. Then, the clusters from Agglomerative and GMM are compared with the centroids from K-Means. The closest centroids are matched using Euclidean distance. After matching, the cluster labels from Agglomerative and GMM are reassigned to match the labels from KMeans.
This is done using the function in Python:
def run-and-remap clustering(... reference-labels=kmeans-labels ...)
This function helps to keep the meaning of Clusters 0, 1, and 2 the same in all methods. After this step, Cluster 0 refers to the same group in K-Means, Agglomerative, and GMM.

I. Evaluation and Visualisation

Silhouette scores are calculated to measure cluster quality. A higher silhouette score means better cluster separation. Bar plots show the number of records in each cluster. PCA scatter plots show how clusters look in 2D. The same colour is used for the same cluster across all methods. This helps to track Clusters 0, 1, and 2 easily across K-Means, Agglomerative, and GMM.

IV. Dataset

The dataset contains synthetic data, 5,000 records describing ransomware events in healthcare organisations. The records mix categorical and numerical information to give a thorough understanding of the penetration and the countermeasures made by the organisations. At first, the dataset contained 16 main features, like the type of organisation, the method of ransomware, how often organisations monitored threats, whether a ransom was paid, and the consequent effects of data restoration. For the preparation of Deep Neural Network models, Categorical variables were converted using One-Hot encoding, while Numerical variables like organisation size, ransomware infection rate, and recovery were scaled. This preprocessing led to a feature set of 30 numeric variables per record, which was found beneficial for unsupervised learning approaches that emphasise pattern identification and anomaly detection.

A. Data Collection

The dataset, publicly sourced from Kaggle, contains 5,000 synthetic healthcare ransomware records with details like infection rate, recovery time, and ransom payment. It supports unlabeled pattern analysis while promoting reuse and transparency in line with FAIR principles.

B. Data Preprocessing

Data was processed in Google Colab using TensorFlow, Scikit-learn, Pandas, and Seaborn. Categorical variables were one-hot encoded, and numerical features were scaled using Min-Max normalisation. The dataset was split into 80% for training and 20% for validation.

V. Results and Discussion

This study used three clustering algorithms to group ransomware behaviors. These were KMeans, Agglomerative Clustering, and Gaussian Mixture Models (GMM). The input for these models was 8-dimensional data. This data came from an autoencoder. The autoencoder compressed 30 original features into 8 important ones. These features included infection rate, backup condition, and recovery time. All three clustering methods found the same types of behaviour. Each algorithm grouped the organisations into three main response types.

A. Autoencoder Reconstruction Error

After training, the autoencoder rebuilt the input records with low error. Then, the reconstruction error was calculated for each record. The 95th percentile value was used as the threshold. Any record above this value was marked as an anomaly. These records may have unusual or rare behaviour. Reconstruction Error: 0.0448

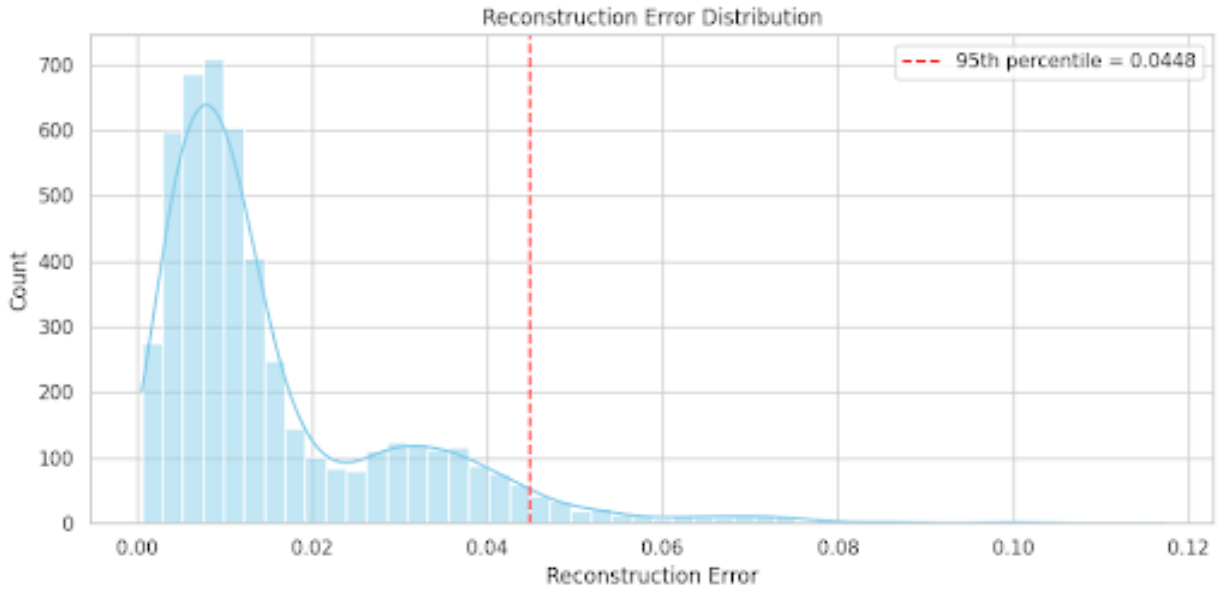

Fig. 3. **Reconstruction Error**

B. Cluster Alignment Insight and Cluster Behaviour Patterns

After aligning the cluster labels across all three algorithms, the same group behaviours were found in each method. Cluster 0 showed fast recovery, strong backups, and almost no ransom paid. This group is called Resilient Responders. Cluster 1 showed slow recovery, high ransom payments, and poor backup systems. This group is called Vulnerable Responders. Cluster 2 showed mixed behaviour with medium recovery time, partial ransom payments, and some backup issues. This group is called Hybrid Responders. The cluster alignment was done using centroid similarity, which helped match the labels in GMM and Agglomerative to the KMeans cluster structure. This interpreted all models consistently.

Fig. 4. **K-Mean cluster distribution**

C. Clustering Performance Summary

Clustering performance are displayed in this section.

1) Cluster Performance for K-Means Algorithm:

TABLE I
PERFORMANCE TABLE OF K MEANS ALGORITHMS

Silhouette Score	Cluster Label	Records	Avg. recovery time(Days)	Paid Ransom(%)	Backup Compromised(%)
0.2106	Cluster 0(Resilent)	2013	2.3	3	11
	Cluster 1(Vulnerable)	1473	7.9	87	79
	Cluster 2(Hybrid)	1514	5.1	42	35

2) Cluster Performance for Agglomerative:

TABLE II
PERFORMANCE TABLE OF AGGLOMERATIVE

Silhouette Score	Cluster Label	Records	Avg. recovery time(Days)	Paid Ransom(%)	Backup Compromised(%)
0.1686	Cluster 0(Resilent)	1385	2.1	4	9
	Cluster 1(Vulnerable)	1855	8.3	85	82
	Cluster 2(Hybrid)	1760	5.3	39	31

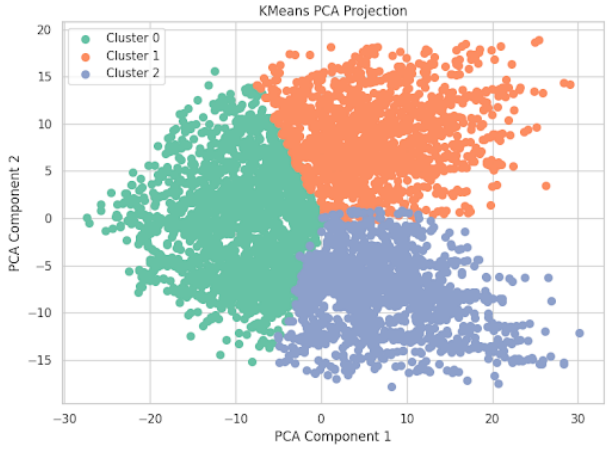

Fig. 5. **K Means Clustering in PCA 2D**

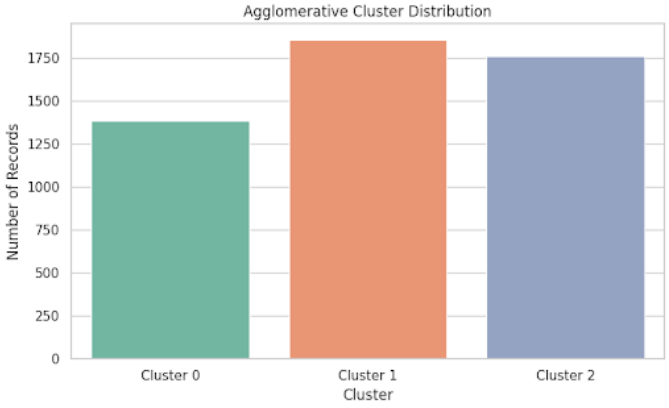

Fig. 6. **Agglomerative Cluster distribution**

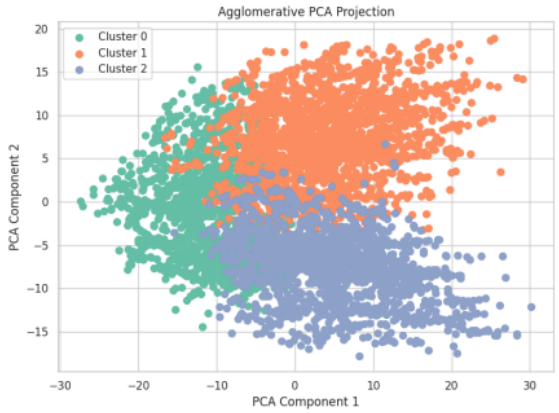

Fig. 7. **Agglomerative Clustering in PCA 2D**

3) Cluster Performance for Gaussian Mixture Model:

D. Performance Evaluation

Silhouette score tells how good the clusters are. The value can be between –1 and 1. A score close to 1 means the clusters are well separated. A score near 0 means the clusters are overlapping. A score below 0 means the clustering is poor and

TABLE III
PERFORMANCE TABLE OF GAUSSIAN MIXTURE MODEL

Silhouette Score	Cluster Label	Records	Avg. recovery time(Days)	Paid Ransom(%)	Backup Compromised(%)
0.1418	Cluster 0(Resilent)	1631	2.5	41	10
	Cluster 1(Vulnerable)	1464	8.1	88	77
	Cluster 2(Hybrid)	1905	5.2	41	34

Fig. 8. **Gaussian Mixture Model distribution**

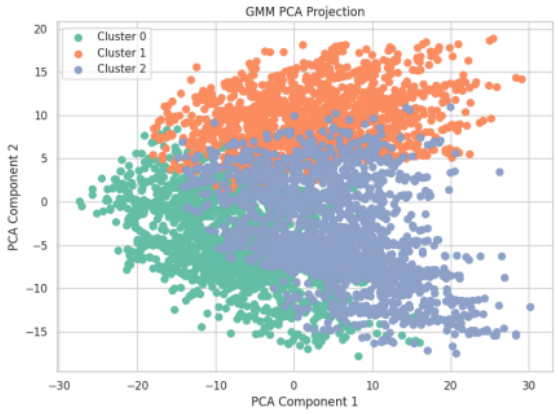

Fig. 9. **Gaussian Mixture Model PCA 2D**

points are in the wrong group. KMeans gave the best result. Its silhouette score was 0.2106. This means the clusters were clear and well separated. Agglomerative had a score of 0.1686. It also worked, but the clusters were a little closer and had more overlap. GMM had the lowest score of 0.1418. The groups were soft and less clear. This shows that KMeans worked the best out of the three.

E. Anomaly Insight

The autoencoder also detected anomalies. This was done using reconstruction error. Records with errors above 0.0448 were marked as outliers. These may be special ransomware

cases. They may include: Rare attacks Advanced threats Unusual organizational behavior These records need further analysis. They may help improve early detection in the future.

VI. Conclusion and Future Work

This study shows that autoencoders and clustering can work well together. They can find ransomware behavior patterns without using labeled data. All three clustering models gave similar groupings. These are Resilient, Vulnerable, and Hybrid Responders. KMeans gave the best result with a silhouette score of 0.2106. It made clear and separate groups. We also used reconstruction error to find outliers. 5% of the records were marked as anomalies. These could be extreme or rare attacks. This method helps healthcare groups find their weak points. It also helps them improve their response plans. They can use this system to stay ready for future ransomware threats. This research can be improved in many ways. We can add time-based data. This will help track how attacks change over time. We can make the clusters more explainable. Tools like SHAP or LIME can help tell us why a group behaves a certain way. We can test this on real-world ransomware datasets. These can come from threat reports or incident logs. We can also use this in real-time. The autoencoder can run on live data and catch new attacks early. Lastly, this method can be used in other fields. Sectors like finance, energy, and government also face ransomware problems.

References

[1] Alzahrani, S., Xiao, Y., Asiri, S., Zheng, J., & Li, T. (2025). A Survey of Ransomware Detection Methods. IEEE Access.

[2] Kirubavathi, G., Regis Anne, W., & Sridevi, U. K. (2024). A recent review of ransomware attacks on healthcare industries. International Journal of System Assurance Engineering and Management, 15(11), 5078-5096.

[3] Alam, M., Ashraf, Z., Singh, P., Pandey, B., Rehman, K., & Aldasheva, L. (2025, March). Deep Learning Techniques for Intrusion Detection Systems in Healthcare Environments. In 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 105-111). IEEE.

[4] Kritika, E. (2024). A comprehensive literature review on ransomware detection using deep learning. Cyber Security and Applications, 100078.

[5] Al-Hawawreh, M., Moustafa, N., & Slay, J. (2024). A threat intelligence framework for protecting smart satellite-based healthcare networks. Neural Computing and Applications, 36(1), 15-35.

[6] Khatun, M. A., Memon, S. F., Eising, C., & Dhirani, L. L. (2023). Machine learning for healthcare-iot security: A review and risk mitigation. IEEE Access, 11, 145869-145896.