JOURNAL
COMPUTER SCIENCES INSTITUTE

JCSI 38 (2026) 6-10
Received: 12 September 2025
Accepted: 18 November 2025

Behavioral analysis of ransomware threats to ESXi Hypervisors: a machine learning-based predictive model

Upakar Bhatta*

Department of Information Technology Management, Central Washington University, Ellensburg, WA, USA

Abstract
In today's virtualized world, ransomware threats to ESXi hypervisor are a significant and growing concern. Factors include lack of dedicated security tools, an expanding attack surface targeting virtualization infrastructure, the use of data encryption by attackers, and exploitation of known vulnerabilities. Evaluating the ESXi hypervisor is critical for security, performance, and cost efficiency. This paper leverages application of artificial intelligence, specifically machine learning, to assess ransomware threats. The experimental methodology applied in this paper leverages ESXi logs to construct a sample dataset containing5,000 labeled instances of observed attack outcomes across seven features, including vm_shutdowns, snapshot_deletions, high_entropy_files, shell_command_count, failed_login_attempts, suspicious_port_access, and data_exfil_volume. These features are used to train a model to enhance operational efficiency and maintain a robust virtualized environment.

Keywords: artificial intelligence; machine learning

*Corresponding author
Email address: **Upakar.Bhatta@cwu.edu** (U. Bhatta)

1. Introduction

With the rise of ransomware attacks specifically targeting ESXi environments, the cybersecurity landscape has been significantly transformed. These ransomware attacks allow attacker to gain full access to an ESXi hypervisor, enabling them to encrypt the file system and potentially disrupt hosted server functionality. In 2024, VMware ESXi servers targeted by ransomware attacks exposed 8,000 ESXi hosts directly to the internet [1].

As core infrastructure of organization's IT department, ESXi hypervisors have become a key target for attackers, and their targeting has increased over the years. This allows attackers to gain full administrative permission on an ESXi servers and exfiltrate sensitive data. The rise of ransomware attacks exploiting ESXi hypervisor vulnerabilities represents a growing crisis in today's virtualized world.

A growing number of ransomware actors targeting ESXi hypervisors are constantly innovating their attack techniques [2]. This research study explores machine learning techniques to assess the ransomware threats to ESXi environments, enabling organizations to secure their virtualized infrastructure. This study examines ESXi logs to construct a sample dataset with features such as failed login attempts, data exfiltration volume, suspicious ports, snapshots deletion, etc. These features are used to train a model that helps organizations maintain a secure virtualized environment.

2. Pre-requisite knowledge

2.1. Cloud computing

Cloud computing is the technology that connects virtual resources, such as computing power, storage, networking, and databases, over the internet to facilitate data sharing. Cloud technology is one of the fastest growing technologies relying on virtualization platforms like ESXi. Since ESXi plays a critical role in virtualization, securing it is essential to ensure a secure cloud environment. Protecting the virtualization infrastructure in a cloud environment requires comprehensive ESXi hardening measures, such as enabling TPM and secure boot, implementing memory protection, and using virtual machine encryption [3].

Cloud adoption includes service models such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). It also involves deployment models such as public, private, community, and hybrid clouds, each designed to meet specific organizational requirements and preferences [4]. The cloud computing services that the most organizations leveraging today are:

Infrastructure as a Service (IaaS): Provides highest level of flexibility and management control over the infrastructure compared to other services. In this service model, ESXi plays a crucial role in virtualizing physical hardware resources such as memory, CPU, network, and storage.

Platform as a Service (PaaS): Provides customer with an underlying infrastructure to deploy their own code. In this service model, ESXi delivers a scalable virtual infrastructure on which application can be deployed.

Software as a Service (SaaS): Offers complete software packages over the internet on a subscription basis. These services typically run on the virtual machines managed by ESXi.

2.2. Big data analytics

Big data analytics enables the preprocessing and analysis of ESXi log data. It involves examining large datasets generated by ESXi hypervisors, allowing machine learning developers to clean, transform, and normalize raw log information. Big data analytics deals with massive volumes of data that originative from divers sources.

For instance, the average mobile traffic was 6.2 Exabyte per month, and Netflix has over 86 million members globally [5]. By leveraging big data analytics, which encompasses data velocity, data variety, and data volume, organizations can uncover hidden patterns and insights. Companies can apply advanced analytical tools and technologies to study big data in depth and extract valuable business insights [6]. Without the implementation of big data analytics, machine learning-based predictive model would not be effective in predicting ransomware threats to ESXi hypervisors.

2.3. Machine Learning

Machine learning is a mathematical technique that uses algorithm to analyze the behavior of ESXi to identify ransomware threats. It enables systems to make predictions by analyzing hidden patterns in datasets. Machine learning introduces a new approach for training algorithms with real-time datasets to identify specific patterns andanomalies in network traffic [7]. There are three main categories of machine learning: supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning can be beneficial to predicting viruses on ESXi by training on labeled datasets. This approach is further divided into Classification and Regression [8]. In classification tasks, the output is a category, whereas in regression tasks, the goal is to predict a continuous value.

In unsupervised learning, the algorithm is trained on unlabeled data to detect ransomware attacks in ESXi environments. Since unsupervised learning techniques do not require pre-labeled data, they are beneficial for detecting unknown threats, although the learning process is complex. Due to its complexity, unsupervised learning is used less frequently than supervised learning [9].

In reinforcement learning, an AI agent is trained based on trial and error, receiving feedback in the form of rewards and punishments. The algorithm learns sequence of normal and malicious events, such as ransomware attacks on the hypervisor, over time. Through this process, the algorithm determines which steps yield the highest rewards [10].

3. Literature Review

Ransomware has emerged as an evolving threat in virtualized environments, such as VMware ESXi. Traditional signature-based mechanisms are not sufficient to prevent these threats. Recent research studies have explored machine learning (ML) approach to detect ransomware. Machine learning technique have been applied to identify complex pattern without reliance on static rules [11]. Unsupervised learning approaches have been applied to identify unique ransomware families based on shared behavioral characteristics [12]. While machine learning approaches have advanced detection capabilities, significant gaps persist in ransomware detection. The deployment of generative adversarial networks (GANs) has been purposed to introduce resilience against ransomware evasion tactics by generating synthetic ransomware samples [13]. However, deep learning models require substantial computational resources. Machine learning models also depend on a substantial training datasets to accurately predict ransomware behaviors. This research aims to address these gaps by applying experimental methodology that leverage supervised machine learning algorithms to assess Ransomware Threats to ESXi Hypervisors.

4. Methodology

In this paper, machine learning techniques are explored to analyze ransomware threats to ESXi Hypervisors. The study demonstrated various machine learning algorithms to assess ransomware threats to ESXi Hypervisors. The experimental methodology employed in this research includes the following steps:

Acquiring ESXi logs to construct a sample dataset, data Preprocessing, Exploratory Data Analysis (EDA), splitting the dataset into training and testing segments, applying SMOTE, selecting machine learning models, trainin the models, predicting the output, measuring the accuracy, and evaluating the models.

Acquiring Dataset: ESXi logs were used to construct a sample datasetconsisting of 5000 records.

Feature selection: A total of 7 features were included. The key features are vm_shutdowns, snapshot_deletions, data_exfil_volume, high_entropy_files, shell_command_count, failed_login_attempts, and suspicious_port_access.

Data preprocessing: Data cleaning, encoding, and normalization were performed.

Exploratory Data Analysis (EDA): To visualize patterns and identify relationship within the dataset, correlation heatmap was utilized to detect outliers, observe feature distributions, and identify relations between features.

Split the dataset: The dataset was divided into training (80%) and testing (20%) sets.

SMOTE: SMOTE was applied to address the class imbalance in the dataset.

Model selection: Appropriate algorithms were selected to train the machine learning models.

Model training: Five machine learning models were trained. Hyperparameter tuning and cross-validation were employed to optimize model performance and ensure robustness.

Model Accuracy and evaluation: Machine learning models were evaluated based on their accuracy.

5. Data Analysis

The sample dataset used in this research include numerical features, such as vm_shutdowns, snapshot deletions, failed login attempts, and suspicious port access.

The exploratory data analysis methods applied in this research include:

Correlation Heatmap: It visualizes the numerical features and helps identify redundant ones. Weakly correlated features which provides independent information, or moderate correlated features, which provide fairly independent information, such as failed login attempts, suspicious port, and snapshot deletions, are ideal for training the ML models.

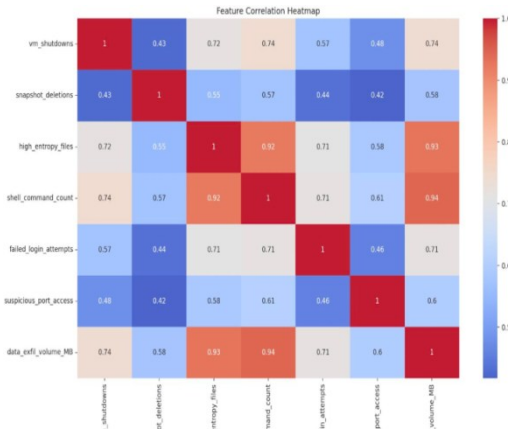

Figure 1: Correlation heatmap.

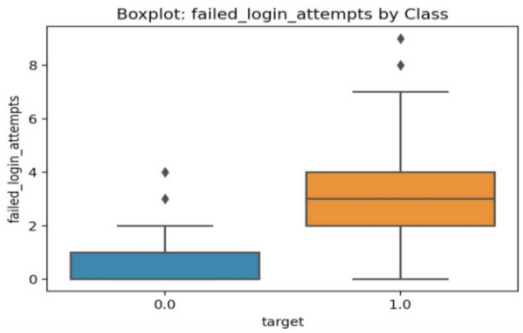

Figure 2: Boxplot failed login attempts by class.

Boxplot: The above boxplot shows the distribution of failed login attempts by target class either 0 or 1. The target on X_axis labeled 0 represent normal users, while 1 represents compromised users. Failed login attempts on Y-axis represents the total number of failed login attempts. Target class 1 shows users with the highest number of failed logging attempts.

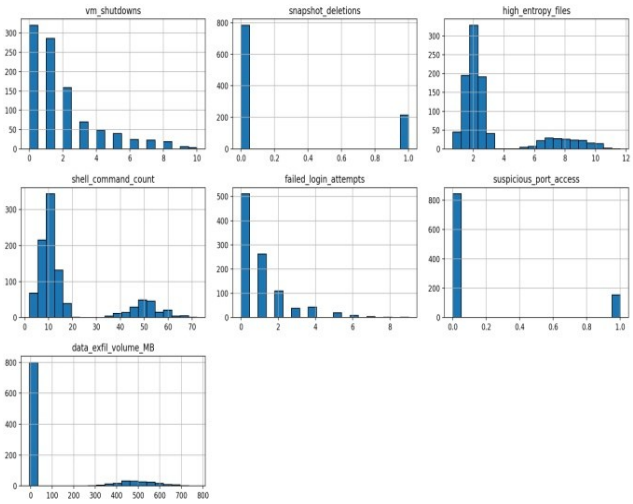

Figure 3: Histograms of various features.

Histogram: The above diagram shows multiple histograms of various features extracted from ESXi log data, analyzing whether the data is normal or anomalous. The histogram of vm_shutdowns shows that most VMs were shut down 0-2 times, indicating most systems are operating normally. The histogram of snapshot deletions shows that most snapshots were not deleted, as most entries are 0. The histogram of high entropy files indicates abnormal activities, representing suspicious behavior. The histogram of failed login attempts shows the number of failed logins, with most systems having 0-2 failed login attempts. The histogram of suspicious port access indicated a small number of systems accessed suspicious ports.

Synthetic Minority Over-Sampling Technique(SMOTE): SMOTE technique has been applied to address class imbalance issues. It helps machine learning models avoid bias toward the majority class and improve overall performance.

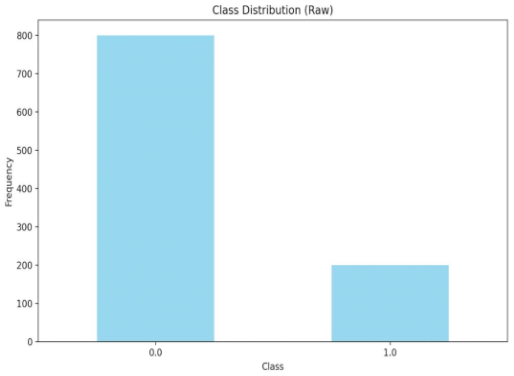

Figure 4: Class distribution before SMOTE.

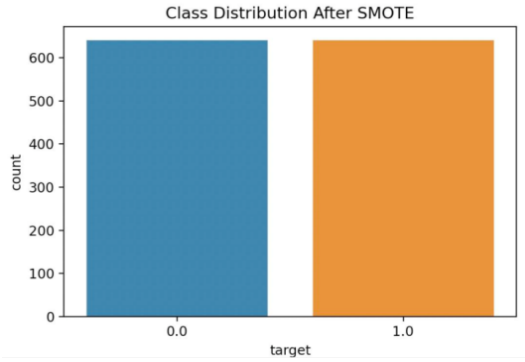

Figure 5: Class distribution after SMOTE.

6. Results

This paper leverages a sample ESXi logs to construct a sample dataset, containing5,000 labeled instances of observed attacks outcomes across seven features, including such as vm_shutdowns, snapshot deletions, high entropy files, data exfil volume, failed login, shell command count, suspicious port access, and data exfil volume. These features are used to train a model that helps organizations enhance their operational efficiency. The machine learning model demonstrated strong performance, achieving an accuracy of 100% in assessing ransomware threats to ESXi hypervisor.

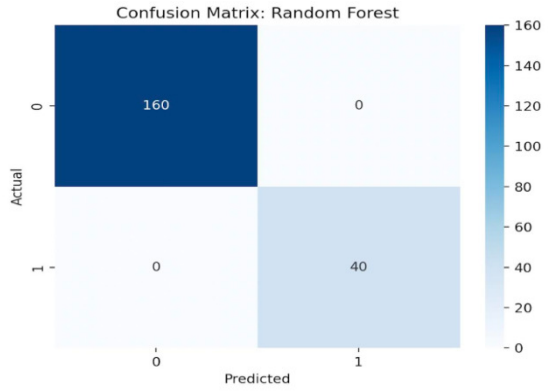

Figure 6: Confusion Matrix Random Forest.

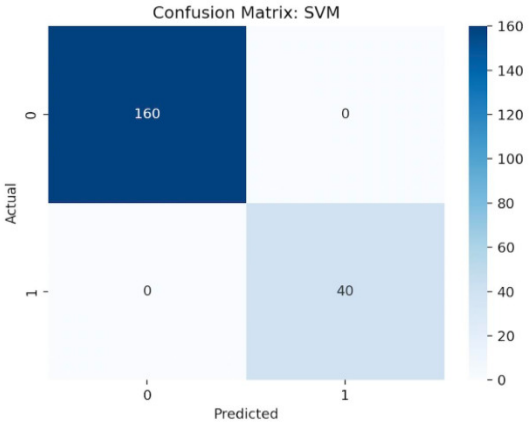

Figure 7: Confusion Matrix SVM.

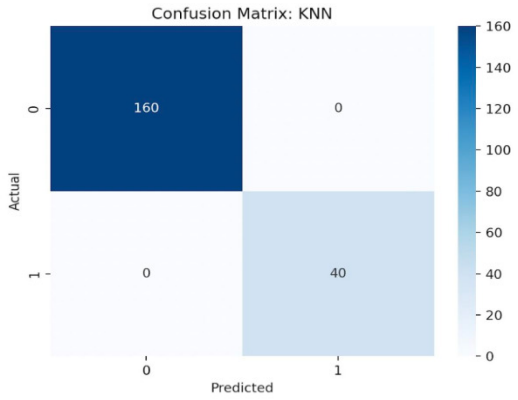

Figure 8: Confusion Matrix KNN.

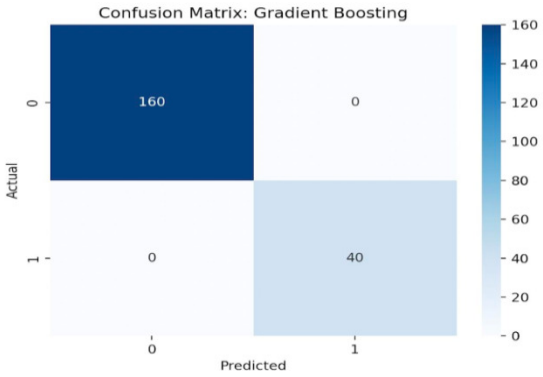

Figure 9: Confusion Matrix Gradient Boosting.

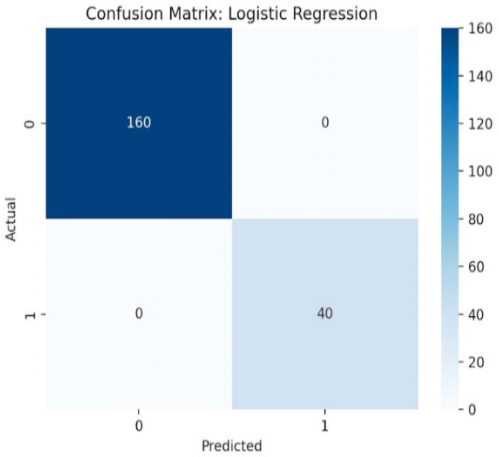

Figure 10: Confusion Matrix Logistic Regression.

Table 1: Model metrics

Random forest	SVM	KNN	Gradient boosting	Logistic regression
TP= 40	TP= 40	TP= 40	TP= 40	TP= 40
TN= 160	TN= 160	TN= 160	TN= 160	TN= 160
FP= 0	FP= 0	FP= 0	FP= 0	FP= 0
FN= 0	FN= 0	FN= 0	FN= 0	FN= 0
A= 1.0	A= 1.0	A= 1.0	A= 1.0	A= 1.0
P= 1.0	P= 1.0	P= 1.0	P= 1.0	P= 1.0
R= 1.0	R= 1.0	R= 1.0	R= 1.0	R= 1.0
F1-Score= 1.0	F1-Score= 1.0	F1-Score= 1.0	F1-Score= 1.0	F1-Score= 1.0

$$Accuracy\ (A) = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision\ (P) = \frac{TP}{TP+FP} \quad (2)$$

$$Recall\ (R) = \frac{TP}{TP+FN} \quad (3)$$

$$F1-Score = \frac{2*(P*R)}{P+R} \quad (4)$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

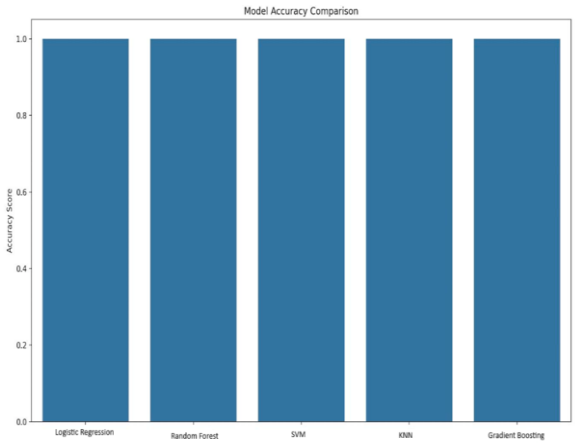

Figure 11: Model accuracy comparison.

Based on the confusion matrix metrics, the machine learning model perfectly predicted all the samples. The model accuracy 100% because ESXi logs were generated using an algorithm to mimic the pattern of real logs for machine learning testing. The synthetic dataset was created from ESXi log with features such as vm_shutdowns, snapshot_deletions, high_entropy_files, shell_command_count, failed_login_attempts, suspicious_port_access, and data_exfil_volume. This experimental research evaluated the model using a single train-test split and SMOTE (Synthetic Minority Over-Sampling Technique), which was applied to address class imbalance issues. To evaluate the model performance, the study employed confusion matrix along with other key metrics such as precision, recall, and F1 score. The results indicated that the model perfectly predicted all the samples.

7. Discussion results

Precision (0.1): An ML model produced no false positive and is correct 100% of the time. Recall (0.1): An ML model correctly identified all actual positive instances. F1 Score (0.1): The F1 score indicated that the model accurately detected the positive class.

Justification: Random Forest was chosen for assessing ransomware threats to ESXi hypervisors in this research paper for following reasons:

Random forest is known for its high accuracy and works well with the network feature.

Random Forest ensemble learning method is beneficial for this research as it improves performance compared to other model in detecting anomalies effectively.

Compare to other models, Random Forest can handle both numerical and categorical features without any extensive preprocessing, making it ideal for analyzing the diverse datasets used in this research.

8. Conclusions

In modern virtualized world, evaluating the ESXi hypervisor is critical for ensuring security, performance, and costs efficiency. The experimental method applied in this paper leverages machine learning to assess ransomware threats to ESXi hypervisors. It leverages a sample ESXi logs to construct a dataset, containing 5,000 labeled instances of observed attacks outcomes across seven features, including vm_shutdowns, snapshot_deletions, high_entropy_files, data_exfil_volume, failed_login, shell_command_count, suspicious_port_access, and data_exfil_volume to train a model that helps organizations enhance its operational efficiency and maintain a robust virtualized environment. With an accuracy of 100 percent, this paper highlights how effective ML technique can be in assessing ransomware threats to ESXi hypervisor and ensuring security in a virtualized environment.

References

[1] The Hacker News, Ransomware on ESXi: The Mechanism of Virtualized Attacks, https://thehackernews.com/2025/01/ransomware-on-esxi-mechanization-of.html, [16.11.2025].

[2] Microsoft Threat Intelligence, Ransomware Operators Exploit ESXi Hypervisor Vulnerability for Mass Encryption, https://www.microsoft.com/en-us/security/blog/2024/07/29/ransomware-operators-exploit-esxi-hypervisor-vulnerability-for-mass-encryption/, [09.25.2025].

[3] Cloud Flow Technologies, ESXi Hardening: A Comprehensive Security Guide for System Administrators, Cloud Flow Tech, https://cloudflowtech.com/f/esxi-hardening-a-comprehensive-security-guide-for-system-admin, [10.09.2025].

[4] Q. Zhang, L. Cheng, R. Boutaba, Cloud computing: State-of-the-art and research challenges, J. Internet Serv. Appl. 1 (2010) 7–18, https://doi.org/10.1007/s13174-010-0007-6.

[5] C.L.P. Chen, C.-Y. Zhang, Big Data: Survey, Technologies, Opportunities, and Challenges, Information Sciences 275 (2014) 314–347, https://doi.org/10.1016/j.ins.2014.01.015.

[6] P. Russom, TDWI Best Practices Report, Big Data Analytics, https://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx, [10.11.2025].

[7] V. Kumar, D. Sinha, A.K. Das, S.C. Pandey, R.T. Goswami, An integrated rule-based intrusion detection system: Analysis on UNSW-NB15 data set and the real time online dataset, Cluster Computing 23 (2020) 1397–1418, https://link.springer.com/article/10.1007/s10586-019-03008-x.

[8] A. St-Aubin, Introduction to Supervised Machine Learning,https://www.math.mcgill.ca/gsams/drp/papers/papers2024/Alexandre-St-Aubin.pdf, [10.11.2025].

[9] S. Mishra, Unsupervised Learning and Data Clustering Explained, https://towardsdatascience.com/unsupervised-learning-and-data-clustering-eeecb78b422a/, [09.25.2025].

[10] K. Murphy, Reinforcement Learning: An Overview, arXiv preprint arXiv:2412.05265 (2024), https://arxiv.org/abs/2412.05265.

[11] S. Wasoye, M. Stevens, C. Morgan, D. Hughes, J. Walker, Ransomware classification using BTLS algorithm and machine learning approaches, https://www.researchsquare.com/article/rs-5131919/v1, [10.11.2025].

[12] K. Schmaltz, S. Thompson, D. Mendes, J. Carvalho, Robust defense mechanisms against adversarial ransomware attacks: Implementing a universal network-level detection filter, https://www.researchsquare.com/article/rs-5123680/v1, [10.11.2025].

[13] V. Miranem, G. Petrescu, D. Schelling, and A. Vasiliev, Ransomware detection on Windows systems using file system activities and a hybrid machine learning approach, https://files.osf.io/v1/resources/27neh/providers/osfstorage/670c4c13fcb08c88b8b9392e?action=download&direct&version=1, [10.11.2025].