

A Survey on Ransomware Detection and Prevention Using Machine Learning Models

Dr. Hemant N. Patel¹

¹ Assistant Professor, Computer Engineering, Sankalchand Patel College of Engineering, Sankalchand Patel University, Visnagar, India

ABSTRACT: Ransomware is one of the most serious cyber-security threats to both individuals and organizations, causing fairly serious financial losses and operational disruption. Ransomware attacks are getting increasingly sophisticated, making it more difficult for traditional signature-based security mechanisms to work effectively, which is why intelligent detection and prevention are needed. The predominant approach used for the detection of ransomware is via pattern recognition and behavior analysis, which is the domain of machine learning. This work provides an in-depth analysis of ransomware, its variants, attack cycles, and the idea of ransomware-as-a-service (RaaS). Furthermore, it discusses both traditional and deep learning approaches to ransomware detection, such as Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. Further, traditional and AI-based ransomware prevention methods, recovery procedures, and a comparative review of recent papers are discussed to identify the existing challenges and potential research directions in ransomware detection and prevention using machine learning.

KEYWORDS: Ransomware, Cybersecurity, Machine Learning, Deep Learning, Cyber Defense, Ransomware Prevention, Behavioral Analysis.

1. Introduction

Ransomware is a major concern for corporations when it comes to cyber security. Ransomware, a sort of malicious software that encrypts data and demands payment for a key to decrypt it, is being used more and more often by hackers as a means of financially extorting victims. From healthcare and banking to government and education, ransomware assaults have affected every sector [1]. The stakes are too high to ignore the need to learn about ransomware, its methods of distribution, and the possible repercussions of being a victim. Cryptographic malware known as ransomware encrypts user data secretly and without their knowledge or permission. It limits authorized access to user information. Ransomware stands apart from other malwares due to its permanent effect. After encryption is complete, the only method to access user data is to utilize the decryption key [2]. The attackers want payment in bitcoin, which cannot be traced, so they can decode the data. Organizations are also being targeted by ransomware attacks, which are extorting money from both individuals and businesses. The attackers take advantage of the fact that ransomware attacks leave no trace and that the money is untraceable.

The victim is put in danger since his data might be misused, lost, or revealed, including his search history [3].

Ransomware prevention aims to stop ransomware attacks before they can infect systems or encrypt valuable data. Security measures are used to limit vulnerabilities and the risk of infection thru frequent software upgrades, access control, user knowledge, and constant monitoring. Data confidentiality, integrity, and availability may be preserved by companies with the support of effective preventative techniques, which are the bedrock of robust cybersecurity defenses [4].

Ransomware recovery involves restoring systems, apps and data following a ransomware attack while also making sure the malware has been fully removed. The recovery process also involves restoring data from secure backups, reinstalling or reimaging infected systems, and verifying the integrity of the systems and returning to normal operations [5]. A robust recovery strategy reduces downtime and data loss, enabling the company to return to business operations as soon as possible. The ability to analyse data and trends, ML can help computers identify criminal activity, which makes it an excellent tool for

ransomware detection. Unlike the signature based-detection approach, ML-based detection can identify new and unknown ransomware variants by learning from past events. ML algorithms enhance ransomware detection precision and swiftness by examining attributes like file operations, process behavior, API uses, and network activity. In the ever-changing landscape of ransomware threats, ML has emerged as a valuable tool in the cybersecurity arsenal, particularly for improving ransomware detection and prevention [6]. Ransomware detection with ML relies on models trained to recognize ransomware characteristics from ransomware file behavior, process activity, API usage, network traffic, and more. Some of the common algorithms are LR, DT, RF, KNN and support vector [7]. RF is one of these, which is very well recognized for its detection strength and computational efficiency and hence it can be used for ransomware detection in real-time.

There are six parts to this paper. This section covers the basics of ransomware. Machine Learning Models is discussed in Section III. In Section IV, ransomware prevention and protection strategies are described. Section V critically summarizes the recent literature and Section VI concludes with limitations future research directions.

2. Classification and Types of Ransomware

Ransomware is a virus that has recently emerged as a major risk to users' data and computer resources. Its standard operating procedure is to encrypt user files, rendering them worthless, and then to demand payment in order to decrypt them [8]. Documents, databases, images, source codes, emails, compressed files, and many more kinds of data assets may all be found in this file format. Paying a ransom to decrypt and restore data or maintaining a good reputation are the sole choices for victims.

A. Categories of Ransomwares

There are several varieties of ransomware, but the two most common are computer locker and data locker. Computer locker protects the user's ability to use their computer or mobile device, while data locker protects their files and data from malicious encryption. There are two kind of ransomware that are still prevalent in today's digital world. Malware is defined by its ability to encrypt and decrypt data, either on a user's device or in their files and folders [9]. Although their goals are identical with respect to data or system access, the processes they employ are very different.

1) Locker Ransomware (Computer Locker)

Hackers utilize this ransomware to encrypt data and block people from accessing their systems. Users are

restricted to doing specified actions on a locked system, such as entering mathematical digits to confirm the payments authorized by the hacker. In order to get the victim to pay the ransom, the hacker then employ various psychological techniques. Like an intelligent automobile that can unlock and lock doors and switch on and off, locker ransomware might potentially increase in size as the user connects to the Internet. Once infected, it may become uncontrollable and require the user to pay a ransom to restore control.

2) Crypto Ransomware (Data Locker)

Internet and online connectivity are becoming more integral to every electronic device, including personal computers, due to the growing requirement of doing so. Because it may encrypt or lock data focused on specific files, crypto ransomware or data locker mostly targets personal data rather than the system itself [10][11]. During its design phase, ransomware implements a number of cryptographic techniques and encryption methods. Once the infection has encrypted all of the files on the system, the unknown hacker often sends a message across the system. This message may be a link to a website, a simple text file with instructions for the payment procedure, or it may be a demo file giving instructions to decrypt some random files.

- **Scareware:** This kind of ransomware takes use of pop-up advertising to present frightening messages that lure users into installing or purchasing harmful or worthless software. A key component of the fallacy is the idea that it may deceive people into downloading malicious software. That this ransomware inflict harm on the victim is not guaranteed.

- **Leak ware:** "Leak ware" is a new and terrifying sort of ransomware that demands payment in order to prevent the exposure of victims' data. The damage is irreversible because once the data is out in the open, anybody may access it.

B. Attack Lifecycle of Ransomware

The majority of ransomware attacks use asymmetric encryption methods. As shown graphically in Fig. 1. the ransomware attack workflow consists of separate stages.

1) Delivery and Reconnaissance

The first step for an attacker is to gather knowledge about their target in order to find its weak spots and possible ways to exploit them. Information such as IP addresses, types of networks, operating system versions, and types of infrastructure may be collected [12]. Ransomware may simultaneously use a wide variety of techniques to infiltrate its intended targets. In order to sneakily infiltrate computers without anybody noticing, ransomware uses a variety of infection vectors. There is a wide range in complexity and efficacy among these

vectors. Common entry points for malicious software include:

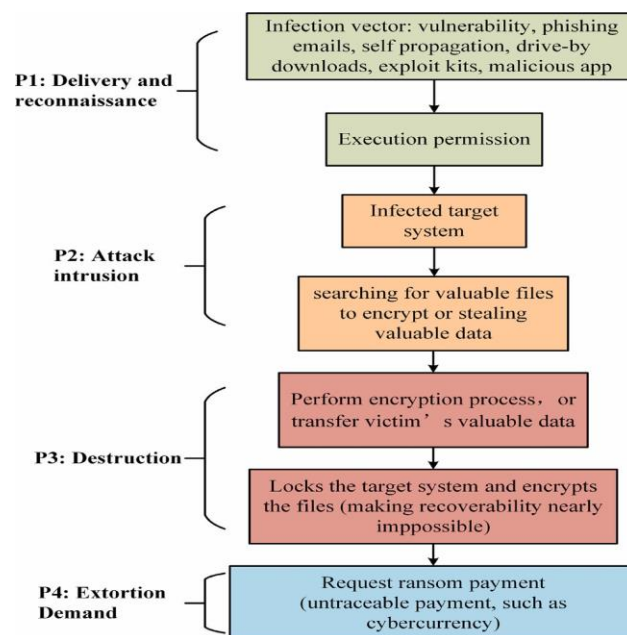


Fig. 1. Ransomware Attack Workflow

- **Security vulnerabilities:** Insecure passwords and remote code execution are examples of security weaknesses that allow attackers to access the victim's internal network and initiate attacks.

- **Phishing/spam email:** Hackers infect phishing emails with ransomware by either attaching it to documents or images or by including harmful links to the malware in the email itself. Ransomware infects victims' workstations and spreads throughout the whole network the moment a victim clicks on a malicious attachment or link. To illustrate the point, malicious macros inserted in Microsoft Word documents allowed the Locky ransomware to infiltrate target PCs.

- **Software supply chain:** Attackers execute extortion assaults using software supply chain methods like distribution and updates by taking advantage of the user's faith in the software source.

- **Remote desktop:** Hackers install ransomware utilizing a remote desktop login server after obtaining the target server's remote login information thru weak password brute force cracking. Malware detection and counter measurement by the target system's antivirus software may occur during the delivery phase. Ransomware uses obfuscation, packaging, and code alteration to evade detection by antivirus solutions that rely on signatures. These strategies can make the damage even worse and the ransom demands much higher.

2) Attack Instruction

Cybercriminals deploy resources in response to exploited vulnerabilities, which they then employ as a

foundation for more network attacks. One-way ransomware may be propagated is using malicious ledgers and DLLs [13]. The virus replicates itself and spreads laterally throughout infected systems thru file-sharing protocols or other methods [14]. The correct attack tools can also be used to obtain access. In order to maximize its success probability while avoiding discovery, ransomware often analyzes the target environment and uses stealth defenses to surreptitiously create and continue the assault.

3) Destruction

Ransomware run malicious binary code (such as a program) on the compromised machine at this stage. Malicious software goes a step further by scanning for and encrypting critical data files such as databases, images, Word documents, etc. In addition to infecting other systems (such the Internet of Things or industrial control systems), ransomware can utilize security holes in systems and networks to infiltrate other systems.

4) Extortion Demand

Ransomware operators often impose a deadline—often 24 or 48 hours—for victims to pay a ransom once the assault has been effective. Refusing to pay the ransom often leads to data loss or device breakdown. The attacker gives the victim the decryption key to get their data back if they pay the price for the ransom. The attacker may provide keys, but it doesn't guaranty they able to decode the data. As a result, victims run the danger of losing both their files and their financial information simultaneously.

3. Machine Learning for Ransomware Detection

ML is a branch of AI that employs algorithms on large datasets to discover patterns, spot outliers, and carry out

automated tasks; this makes it far better than traditional rule-based systems, which rely on predetermined rules to detect harmful acts. The problems of rule-based systems are lack of adaptability, high FP, and the difficulty in detecting modern malware while ML-based systems have their own problems such as the need for significant effort to train the system and the difficulty in creating the rule base. The three popular ML approaches used in ransomware analysis are supervised learning, unsupervised learning, and reinforcement learning [15]. Unsupervised ML does not need labeled data, as contrast to supervised ML, which necessitates labeled data, like ransomware or benign data. A few examples of supervised ML methods include ensemble learning, classification, and regression [16]. Namely, SVM, NN, KNN, and DT are some of the most famous supervised classification methods. Predicting and forecasting using one dependent variable and numerous variables is the purpose of regression analysis [17].

A. ML Workflow for Ransomware Detection

The process of detecting ransomware using ML is a systematic approach that involves several steps: The process starts with data collection and finds data from legitimate software and ransomware samples. These pieces of information can be static, such as the attributes of an executable file, such as the ones with the imported functions, or dynamic, including operations on the file system, API calls, behavior of the processes, registry changes and network traffic.

Data preparation follows data gathering to guaranty data quality and consistency. This step involves deduplication, imputation, normalization of features, and feature coding. Feature selection and/or feature extraction techniques are used to identify the most pertinent attributes in order to streamline the computation and improve detection results.

Separate sets of data are created from the processed data: training data and testing data. During the training phase, the ML algorithms learn the patterns associated with the selected features, thereby distinguishing ransomware from non-malicious applications. The supervised learning algorithms commonly used are DTs, RF, SVM, NB, KNN, LR, and others based on the problem.

The efficiency of the model in identifying ransomware is assessed after training using the testing dataset. Measures of performance often utilized include memory, accuracy, precision, F1 score, and FPR. The trained model can then be deployed in production systems to monitor system traffic and identify suspicious activity when the results are good. Because ransomware is always evolving, it is necessary to retrain the detection model with fresh data sets on a regular

basis to keep it effective and increase its capacity to identify new types of ransomware [18].

B. ML Models for Ransomware Detection

ML is a branch of AI that involves the computer learning to do a task better over time without explicit instructions. Ransomware is a type of malicious software that encrypts user data and then requests payment for the decryption key [19]. Ransomware attacks are becoming more common and more destructive, making ML methods crucial for detecting and preventing them. Some algorithms from the field of ML are used [20]. ML methods such as LR, KNN, DTs, RFs, and SVMs may all identify ransomware. Every technique has its own set of pros and cons; picking the right one is all a matter of context and evidence. A few algorithms for ML:

1) Decision Tree

Ransomware detection is only one use of DTs, a straightforward and easy-to-understand ML approach for classification tasks. DTs are a kind of framework for making decisions that iteratively split data into subsets according to feature values using a structure similar to a tree. Both continuous and categorical Big Data benefit from decision trees' intuitive design and ease of use, but these models are susceptible to overfitting and struggle to handle even minor data changes.

2) Random Forest

RFs are an extension of DTs that make them work better and stop them from overfitting. To make their forecasts, random forests generate several decision trees by randomly picking characteristics and data and then combining them. In addition to being less prone to overfitting, they excel at dealing with high-dimensional data. Nevertheless, they may provide computational challenges and be hard to understand.

3) Support Vector

Ransomware detection, classification, and regression can all benefit from support vector machines, which are dependable ML approaches [21]. Support vector machines are able to sort data into meaningful categories based on feature values by finding the hyperplane that does this as accurately as feasible. When it comes to high-dimensional data, support vector machines are up to the task. Depending on the kernel function and its settings, they may be able to handle linear or nonlinear boundaries.

4) K-Nearest Neighbor

The k-NN model is a non-parametric tool for regression and classification jobs. Predicting the input's label from the most frequent label among the k data points in the training set that are closest to the input (called "neighbors") is how it works. The algorithm has

many potential uses because it is both simple and effective.

5) Logistic Regression

LR is a type of parametric technique that is commonly employed in binary classification problems, where the output may be classified into one of two distinct groups. The connection between the input features and the likelihood of the output class is where the model gets its usefulness. Finding the settings that make the training data as likely as possible is the point of training the algorithm. To prevent overfitting, it can thereafter be fine-tuned.

C. Deep Learning Models for Ransomware Detection

One way that DL approaches might improve upon the shortcomings of older supervised ransomware detection methods is by making the detection process more precise and trustworthy. The algorithms are self-learning and are good for unstructured data, needing little or no human involvement [22]. Textual and picture-based ransomware data may be effectively detected by using their audio, text, and image data classification capabilities. But, general-purpose applications, especially those with limited datasets or sizes, may not be a good fit for DL algorithms because of the massive amounts of data needed to train them. Ransomware detection makes use of a few DL models.

- **Deep Neural Network (DNN):** DNNs consist of

multiple interconnected layers that learn hierarchical representations of ransomware behavior. Binary classification tasks, like differentiating ransomware from safe software, are a good fit for them, and they can handle a variety of properties borrowed from system operations.

- **Convolutional Neural Network (CNN):** The CNNs were first designed for image analysis, they are now capable of analyzing structured security data including memory patterns, network traffic, API call sequences, and more [23]. These methods can automatically extract significant qualities, so ransomware detection doesn't need human feature engineering.

- **Long Short-Term Memory (LSTM):** LSTM networks understand the long-term relationships between events in order to interpret sequential input. This makes them especially useful for a ransomware detection based on API call sequences, file operations and process behaviors where the sequence of the events matters in identifying malicious activity.

Generally speaking, DL models can be trained to recognize complex behavioral patterns, detect changes in ransomware methods, and enhance ransomware detection [24]. They can also automatically extract features, which eliminates the need for manual feature selection and enhances the detection of known and unknown ransomware families. In the next Table I, compare the ML and DL models about the ransomware detection.

TABLE I. COMPARISON OF MACHINE LEARNING AND DEEP LEARNING MODELS FOR RANSOMWARE DETECTION

Model	Strengths	Limitations	Suitable For
Decision Tree	Simple, easy to interpret	Overfitting	Small datasets
Random Forest	High accuracy, robust	Higher computational cost	General ransomware detection
Support Vector Machine	Effective for high-dimensional data	Sensitive to kernel selection	Binary classification
K-Nearest Neighbors	Simple implementation	Slow on large datasets	Small-scale detection
Logistic Regression	Fast and lightweight	Limited for complex patterns	Baseline classification
Deep Neural Network (DNN)	Learns complex features	Requires large datasets	Behavioral analysis
Convolutional Neural Network (CNN)	Automatic feature extraction	High computational requirements	API and memory pattern analysis
Long Short-Term Memory (LSTM)	Captures sequential behavior	Longer training time	API call and file

			operation sequences
--	--	--	---------------------

4. Ransomware Prevention and Defense Mechanisms

The complexity of ransomware attacks highlights the importance of taking a proactive and strategic cybersecurity stance. Companies can reduce the risks incurred by these attacks and mitigate potential consequences through a comprehensive preventative approach [25]. The Following draws attention to the need for a proactive defense strategy to strengthen the digital protections of irreplaceable data and systems. As shown in Fig 2, there are two types of ransomware prevention techniques: traditional and AI-based.

A. Traditional Ransomware Prevention Techniques

Reducing the risk of ransomware attacks by strengthening the security of systems, networks, and user data. These preventive measures have been crucial to cybersecurity and have helped reduce the threat and damage that can be caused by the ransomware attack. The following sections discuss the most popular traditional countermeasures used to defend against ransomware. AI-powered malwares are a growing concern for cybersecurity globally. Unlike conventional malware, AI-driven ones can be adaptive, evade detection, and scale attacks with minimal man power. Additionally, the existence of generative models like LLMs allows criminals to generate polymorphic code, which can appear as if it were normal code to avoid detection by conventional security methods [26].

- **Anticipatory Security Measures:** Advanced threat detection systems and proactive monitoring are key elements of anticipatory security measures. These steps can help organizations detect and deter threats before they can gain a foothold, minimizing the risk of ransomware infiltration.
- **Vulnerability Mitigation:** Software/ systems should be updated and patched regularly, to close vulnerability holes that are frequently a target of ransomware attacks. Security and integrity of IT ecosystems depend on patch management.
- **Continual Improvement:** Technology has advanced and cyber threats have too. Security over provided should also be changed. It is updated and patched regularly and is based on the most current threat intelligence, which ensures that defenses are robust in the midst of an ever-changing threat landscape.
- **Human Firewall:** One excellent first line of defense is the workforce. Employees who receive regular

training on ransomware avoidance, safe surfing, and phishing may become proactive and watchful in the battle against threats.

- **Culture of Security:** Culture of security is not an IT responsibility. An eye for security team is a priceless resource for ransomware attacks prevention.
- **Data Resilience:** Backup Data regularly and securely, this is part of robust ransomware defense. The benefit of clean backups is that they may help ensure business continuity during an attack, while also minimizing potential gains from data encryption by attackers.
- **Disaster Recovery Planning:** A disaster recovery plan ought to contain a recovery plan and should indicate that stakeholders have confidence that the organization prepared for any disaster [27].

B. Ai-Based Ransomware Prevention

AI has become an essential component of modern ransomware protection, enabling security solutions to identify, understand, and respond to threats in real-time. AI security systems analyze massive amounts of data pertaining to user, network, and system activity using ML and DL approaches, as opposed to conventional security systems that usually depend on pre-defined signatures and attack patterns. This enables them to detect suspicious activities, predict potential ransomware attacks, and initiate automated responses before significant damage occurs.

- **Behavioral pattern:** One of the most important uses of AI in ransomware prevention is behavioral monitoring. AI systems monitor system's activity and events, including the access and usage of files, processes, memory, registry modifications, API calls, and network communication. By learning the normal behavior of a system, AI models can be used to detect abnormal activity, which could help detect the early signs of a ransomware attack. The security system could break the attack by detecting it early, if the ransomware is going to encrypt files or get to other machines.
- **Anomaly Detection:** Another important prevention technique is anomaly detection. ML algorithms build a profile of typical system and user activity for comparison – Anomaly detection works by detecting behavior, not malware signatures, and as such can identify ransomware and attack variants that were never seen before. This adaptability is essential for AI-driven ransomware prevention, as ransomware families are known to adjust their code to bypass traditional detection techniques.

- **Honeypots, Honey files, and Honey folders:** The use of honeypots, honey files, and honey folders has also become an effective AI-assisted prevention mechanism. These decoy resources are designed to look valuable to ransomware, but are devoid of anything that is important. AI-powered monitoring systems are able to detect ransomware attacks on these decoy files, alerting users and taking necessary protective measures like stopping the malicious process or isolating the infected device. These decoy resources are an effective option for an early warning system due to their low probability of false alerts, and are often not engaged by legitimate users.

- **Email and Network Security:** Phishing emails, malicious attachments, compromised hyperlinks and websites are all likely to be the source of many ransomware infections. Email filtering software powered by AI can detect phishing emails based on email content, sender's reputation, attachment type, and

email communication patterns before the user accesses them. Likewise, AI-driven network monitoring tools constantly analyze network traffic for abnormal communication patterns, unusual remote connections, or ransomware delivery or communication with command-and-control servers. Blocking malicious traffic from accessing the network greatly lowers the risk of a ransomware attack succeeding.

- **File Protection and Access Control:** Monitoring of attempts to modify sensitive files and directories. Smart access control systems can identify legitimate user activities from unauthorized modifications and prevent ransomware from encrypting protected files. Intelligent monitoring and secure backup technologies are also integrated into other methods of prevention, ensuring that clean backups of important data are available even if attacked. All of these measures make the data more resilient and easier to rebuild after a ransomware attack is attempted.

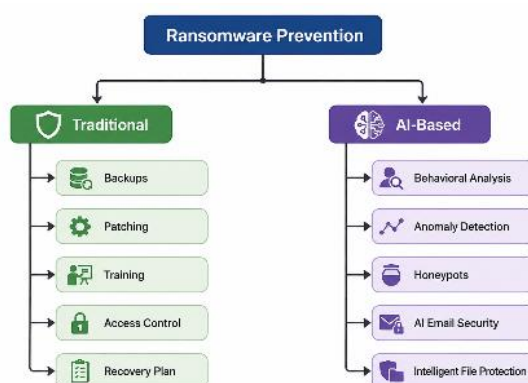


Fig. 2. Traditional and AI-Based Ransomware Prevention Techniques.

C. Ransomware Recovery Mechanism

Ransomware recovery systems are designed to restore ransomware-affected devices, applications, and data to get the business up and running. Organizations require well-thought-out recovery strategies that enable them to minimize downtime and data loss, as modern ransomware can quickly spread across networks, can become persistent and can disable recovery features. Good recovery not only entails recovering encrypted files but also ensuring that compromised systems are thoroughly sanitized before putting back online [28].

- **Full System Backup Recovery:** Recovery of the entire operating system, applications and data to a previously created backup image. This is a full system recovery, but requires a lot of storage and back-up administration.

- **System Reimaging:** The affected system re-imaged and re-installed on a fresh OS image. Secure backups or cloud storage are then used to restore user data. All traces of malwares are eliminated in this way.

- **Modern Cloud-Based Recovery:** Modern endpoint management platforms automatically install operating systems, security updates, and enterprise applications after reinstalling [29]. Cloud sync restores user files and reduces recovery time and complexity in the event of a large-scale recovery.

- **Automated Recovery Mechanisms:** There are some organizations that use automated recovery mechanisms that can reliably rebuild an operating system and check recovery images. These mechanisms minimize human involvement and accelerate recovery from ransomware attacks.

5. Literature Review

This section summarizes the recent researches on ML based ransomware detection and prevention. It highlights key approaches, key findings, and uncertainties that have contributed to the effectiveness of ransomware defenses.

D. Manivannan (2026) covered a wide range of topics in ransomware research, including: storage and

hardware aided defenses, deception defenses, platform specific defenses for mobile and IoT, ML and DL based defenses, network and SDN based defenses, backup and recovery, and storage and hardware assisted defenses. Economic and policy ramifications, operational repercussions of adversarial ransomware attacks for vital sectors such as healthcare, adversarial evasion tactics, and blockchain analysis of ransomware payments are all included in the report [30].

R. Kakadiya et al. (2026) proposed an AI-based malware and ransomware detection system based on LSTM networks in order to address this evolving threat profile. The model is developed using DL using databases containing static and behavioral characteristics of executable files, both synthetic and real-world. However, AI-powered cybersecurity is a formidable force—a proactive, potent solution that lets businesses automate, anticipate, and maintain resilience to safeguard critical assets and activities [31].

V. S. Vivek et al. (2025) provide promising outcomes for the detection and prevention of ransomware attacks using advanced algorithms. The validation of the effectiveness of the model is done by using the evaluation metrics. The findings offer a proactive approach to protecting against new ransomware attacks, improving cybersecurity. It provides a sound method for detecting ransomware attacks, whilst also inspiring further research in the area, as ongoing ransomware attacks prove, in order to continuously improve ransomware security [32].

M. Rele et al. (2025) suggests a novel method for detecting ransomware that makes advantage of AI and ML. A new method has been created that combines cutting-edge feature extraction from system logs, network traffic, and file metadata with anomaly detection and classification methods. For high accuracy and low false positive rates, it makes use of autoencoders, RF, isolated forests, and SVM for anomaly detection. Overall, the study offers a strong foundation for the proactive identification and prevention of ransomware, emphasizing the potential of AI and ML in the field of cybersecurity [33].

W. F. Elersy et al. (2024) enters the cybersecurity world and talks about the correlation between the ML algorithms and the malware detection. The main goal is to improve ransomware detection capabilities using binary file signature analysis. The tasks include data set exploration, segmentation, visualization, data cleaning, feature selection, and model creation, using a wide range of datasets including legit binaries and malwares and a comprehensive methodology. In this research, three machine learning models, namely, RF Classifier, LR, and NN were used to classify a binary data. As a first approach, use Logistic Regression, which is reasonably effective but reveals some complexity in the data [34].

Table II summarizes the objectives, methodology, conclusions, shortcomings, and research gaps found in the reviewed literature on the topic of ML based ransomware detection and prevention.

TABLE II. COMPARATIVE ANALYSIS OF EXISTING RANSOMWARE DETECTION STUDIES

Author (Ref.)	Focus Area	Methodology / Approach	Key Findings	Challenges / Limitations	Research Gap
D. Manivannan (2026)	Comprehensive ransomware survey	Reviews behavioral detection, ML/DL approaches, SDN, IoT, mobile, storage protection, deception techniques, backup and recovery strategies	Provides a broad overview of ransomware research and identifies major challenges in modern ransomware defense	Survey-based work without implementation or experimental validation	Need for integrated, scalable, and cross-layer ransomware defense architectures with improved adversarial robustness
R. Kakadiya et al. (2026)	AI-driven ransomware detection	LSTM model trained on static and behavioral executable features	Demonstrates that deep learning effectively detects ransomware using executable characteristics	Vulnerable to adversarial attacks, privacy concerns, and requires large datasets	Develop secure, explainable, and adversarially robust deep learning models
V. S. Vivek et al. (2025)	Machine learning-based ransomware detection	Advanced machine learning algorithms evaluated using classification metrics	Achieves reliable ransomware detection and supports proactive cybersecurity defense	Evaluation limited to selected ransomware samples and datasets	Improve generalization to zero-day and emerging ransomware variants
M. Rele et al. (2025)	AI and ML-based ransomware detection	Combines feature extraction with Autoencoders, Isolation Forest, Random Forest,	Achieves high detection capability with low false-positive rates through hybrid anomaly	Detection performance depends on feature	Design scalable real-time systems capable of adapting to continuously evolving ransomware

		and Support Vector Machine	detection and classification	engineering and dataset diversity	
W. F. Elersy et al. (2024)	Malware-based ransomware detection	Binary file analysis using RF, LR, and Neural Network classifiers	Random Forest provides the best overall classification performance among evaluated models	Relies primarily on static binary features and limited classifier diversity	Incorporate behavioral analysis and advanced deep learning techniques for improved detection accuracy

6. Conclusion and Future Work

The healthcare, retail, financial, and manufacturing sectors are all vulnerable to ransomware assaults, which are among the most expensive and destructive forms of cyberattacks. This survey provided an overview of ransomware, including different types, attack lifecycle, and the rise of RaaS. It also explored how ML and DL technologies are applied for the detection of ransomware, outlining the merits of the models used widely including DT, RF, SVM, LR, KNN, DNN, CNNs, and LSTM networks. Moreover, traditional and AI-driven ransomware prevention and recovery methods were analyzed, along with current ransomware research avenues. The studies reviewed showed that ML brings substantial benefits to ransomware detection, specifically by detecting abnormal behavior, improving detection rates, and facilitating the early detection of threats, compared to the traditional signature-based method.

The goal of future research should be to develop ML models that can identify zero-day and rapidly developing ransomware assaults in real-time while being lightweight, scalable, and interpretable. Behavioral analysis should also be given more attention and hybrid ML and DL models should be utilized, and larger and more diverse datasets should be used to create improved models with increased model generalization. Furthermore, by combining ML with cloud security, the IoTs, edge computing, and automated incident response, ransomware protection can be enhanced. The creation of standardized benchmarking datasets and evaluation frameworks will continue to aid in the design of practical, efficient, and robust ransomware D&D systems for future cybersecurity applications.

REFERENCES

1. V. Rohilla, K. Rohilla, P. Kumar, and N. Jain, "Intrusion Detection in Network Traffic Using Feature Selection and Optuna-Based Machine Learning Optimization," in 2026 3rd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India: IEEE, Apr. 2026, pp. 1–6. doi: 10.1109/RMKMATE69073.2026.11518834.
2. Monika, P. Zavorsky, and D. Lindskog, "Experimental Analysis of Ransomware on Windows and Android Platforms: Evolution and Characterization," *Procedia Comput. Sci.*, vol. 94, pp. 465–472, 2016, doi: 10.1016/j.procs.2016.08.072.
3. U. Urooj, B. A. S. Al-rimy, A. Zainal, F. A. Ghaleb, and M. A. Rassam, "Ransomware Detection Using the Dynamic Analysis and Machine Learning: A Survey and Research Directions," *Appl. Sci.*, vol. 12, no. 1, 2022, doi: 10.3390/app12010172.
4. J. A. Abraham and S. M. George, "A Survey on Preventing Crypto Ransomware Using Machine Learning," in 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), IEEE, Jul. 2019, pp. 259–263. doi: 10.1109/ICICT46008.2019.8993137.
5. V. Pahuja, A. Khanna, and I. Sharma, "RansomShield: Novel Framework for Effective Data Recovery in Ransomware Recovery Process," in 2024 IEEE International Conference on Big Data & Machine Learning (ICBDML), IEEE, Feb. 2024, pp. 240–245. doi: 10.1109/ICBDML60909.2024.10577365.
6. M. Aljabri et al., "Ransomware detection based on machine learning using memory features," *Egypt. Informatics J.*, vol. 25, p. 100445, Mar. 2024, doi: 10.1016/j.eij.2024.100445.
7. R. E. Asma A. Alhashmi, Abdulbasit A. Darem, Ahmed B. Alshammari, Laith A. Darem, Huda K. Sheatah, "Ransomware Early Detection Techniques," *Eng. Technol. Appl. Sci.*, vol. 14, no. 3, pp. 14497–14503, 2024.
8. M. Cen, F. Jiang, X. Qin, Q. Jiang, and R. Doss, "Ransomware early detection: A survey," *Comput. Networks*, vol. 239, p. 110138, Feb. 2024, doi: 10.1016/j.comnet.2023.110138.
9. R. T. Da-Yu KAO, Shou-Ching HSIO, "Analyzing WannaCry Ransomware Considering the Weapons and Exploits," *ICACT Trans. Adv. Commun. Technol.*, vol. 7, no. 2, pp. 1098–1108, 2018.
10. A. Zimba, L. Simukonda, and M. Chishimba, "Demystifying Ransomware Attacks: Reverse

- Engineering and Dynamic Malware Analysis of WannaCry for Network and Information Security," *Zambia ICT J.*, vol. 1, no. 1, pp. 35–40, Dec. 2017, doi: 10.33260/zictjournal.v1i1.19.
11. M. S. Hosain, "Ransomware: Are We Safe on the Web?," *Inf. Secur. Technol.*, vol. 1, no. 103, pp. 1–9, 2022.
 12. S. C. Pallaprolu, "Zero-day Attack Identification in Streaming Data: Nearest Neighbor Heuristics and Dynamic Semantic Network Generation in the Spark Eco-system," University of Maryland, Baltimore County, 2017.
 13. G. Nagar, "The evolution of ransomware: tactics, techniques, and mitigation strategies," *Int. J. Sci. Res. Manag.*, vol. 12, pp. 1282--1298, 2024.
 14. S. P. Sivashanmugam, S. Kumara, A. Mohile, and D. R. Suram, "Improving Detection Accuracy of Phishing Attacks with Feature Selection and Machine learning Techniques in Cybersecurity," in *2026 1st International Conference on Emerging Trends in Advancements and Applications of Computational Intelligence Techniques (ETA-CT)*, Bhubaneswar, India: IEEE, Apr. 2026, pp. 1–6. doi: 10.1109/ETA-CT69135.2026.11542138.
 15. M. Najafabadi, F. Villanustre, T. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, 2015, doi: 10.1186/s40537-014-0007-7.
 16. S. Debnath, S. A. Devanira Poovaiah, N. Khurramov, S. A. Chikop, and S. Sadullaeva, "KEDenNet-MEO: A Knowledge-Driven Deep Intrusion Detection Framework for VANETs with Realistic Mobility Simulation," in *2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/ISAC364032.2025.11156856.
 17. J. Ferdous, R. Islam, A. Mahboubi, and M. Zahidul Islam, "AI-Based Ransomware Detection: A Comprehensive Review," *IEEE Access*, vol. 12, pp. 136666–136695, 2024, doi: 10.1109/ACCESS.2024.3461965.
 18. R. Soni, A. Kumar, A. Vashisth, and G. Kaur, "Next-Generation AI-Driven Ransomware Detection: Trends, Taxonomy, and Their Limitations," in *2026 3rd International Conference on Emerging Trends in Engineering and Medical Sciences (ICETEMS)*, 2026, pp. 1–6. doi: 10.1109/ICETEMS66917.2026.11469261.
 19. A. Azmoodeh, A. Dehghantanha, M. Conti, and K.-K. R. Choo, "Detecting crypto-ransomware in IoT networks based on energy consumption footprint," *J. Ambient Intell. Humaniz. Comput.*, vol. 9, no. 4, pp. 1141–1152, 2018, doi: 10.1007/s12652-017-0558-5.
 20. M. Mittal, "Quantum Machine Learning: Harnessing Quantum Algorithms for Supervised and Unsupervised Learning," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 11, no. 09, Sep. 2022, doi: 10.15680/IJIRSET.2022.1109004.
 21. A. Alraizza and A. Algarni, "Ransomware Detection Using Machine Learning: A Survey," *Big Data Cogn. Comput.*, vol. 7, no. 3, p. 143, Aug. 2023, doi: 10.3390/bdcc7030143.
 22. M. S. Akhtar and T. Feng, "Detection of Malware by Deep Learning as CNN-LSTM Machine Learning Techniques in Real Time," *Symmetry (Basel)*, vol. 14, no. 11, 2022, doi: 10.3390/sym14112308.
 23. S. Kakkar, S. Janarthanan, B. Makkena, and A. Kakkar, "Deep Learning Approaches for Predicting Attack Vulnerabilities in Quantum-Secure Financial Networks," in *2026 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, IEEE, Feb. 2026, pp. 780–785. doi: 10.1109/ESIC68176.2026.11496277.
 24. M. Davidian, M. Kiperberg, and N. Vanetik, "Early Ransomware Detection with Deep Learning Models," *Futur. Internet*, vol. 16, no. 8, 2024, doi: 10.3390/fi16080291.
 25. T. Jabid et al., "Pre-Print Ransomware Prevention Strategies: Building Robust Cyber Defenses," 2025.
 26. K. S. Anjani Gupta, "Malware Analysis on AI Technique," 2022.
 27. M. Jabid, Taskeed and Rashid, Mohammad Rifat Ahmmad and Ferdous, Md Hasanul and Ali, Md Sawkat and Islam, Mohammad Manzurul and Hasan, Mahamudul and Islam, "Ransomware prevention strategies: Building robust cyber defenses," *Ransomware Evol.*, pp. 141–171, 2024.
 28. K. J. Hole, "Robustness to Malware Reinfections," in *Anti-fragile ICT Systems*, Cham: Springer International Publishing, 2016, pp. 93–98. doi: 10.1007/978-3-319-30070-2_9.
 29. D. P. Marius-Constantin Ilau, Adriam Baldwin, Tristan Caulfield, "Modelling and simulating organizational ransomware recovery: structure, methodology, and decisions," *J. Cybersecurity*, vol. 11, no. 1, p. 45, 2025.
 30. D. Manivannan, "A Comprehensive and Critical Analysis of Ransomware Detection, Prevention,

- Mitigation, and Recovery Approaches," J. Cybersecurity, vol. 8, p. 397, 2026.
31. R. Kakadiya, K. Vora, A. Bambharoliya, and H. Kag, "AI-Driven Cyber Defense for Malware and Ransomware Protection," in 2026 18th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), IEEE, Jul. 2026, pp. 1–7. doi: 10.1109/ECAI69016.2026.11613710.
 32. V. S. Vivek, V. Lokesh, T. Kaushik, U. Kumaran, B. Umah, and G. G. Devarajan, "Ransome Ware Detection Using Machine Learning and Deep Learning Models," in Proceedings of International Conference on Recent Trends in Computing, R. P. Mahapatra, S. Roy, and P. Parwekar, Eds., Singapore: Springer Nature Singapore, 2025, pp. 105–117.
 33. M. Rele, J. Samuel, D. Patil, and U. Krishnan, "Exploring Ransomware Detection Based on Artificial Intelligence and Machine Learning," Procedia Comput. Sci., vol. 252, pp. 548–556, 2025, doi: 10.1016/j.procs.2025.01.014.
 34. W. F. ElSersy, A. ElShamy, and M. Samy, "Ransomware Detection Using Machine Learning Algorithms," in 2024 Intelligent Methods, Systems, and Applications (IMSA), IEEE, Jul. 2024, pp. 186–192. doi: 10.1109/IMSA61967.2024.10652659.