Machine Learning for Cybersecurity (Ransomware)

A Study

Sharaj Jagadeesan
MSC *Data Analytics*
University of Galway
Student ID: 23100668
s.jagadeesan1@universityofgalway.ie

Jason Martin
MSC *Data Analytics*
University of Galway
Student ID: 23101224
j.martin29@universityofgalway.ie

Estella Roberts
MSC *Data Analytics*
University of Galway
Student ID: 23102324
e.roberts10@universityofgalway.ie

Shayna Tuscano
MSC *Data Analytics*
University of Galway
Student ID: 23104937
s.tuscano1@universityofgalway.ie

Abstract— **A type of malicious software called Ransomware demands a ransom by encrypting the victim's data in return for its decryption. Ransomware attacks can cause significant damage to individuals and organizations, such as data loss, financial losses, operational disruptions, and reputational harm. Attackers continuously adapt their ransomware techniques, making it challenging for detection systems to keep up. This is especially true for traditional systems that incorporate detection via signature-based methods. As soon as defenders recognize and shut down an attack vector, another one is opened. This calls for the need for a dynamic system that analyses the communications in the system and constantly checks for anomalies or variations in the hash keys or other security protocols.**

However, this cannot be achieved using traditional models and it calls for a more advanced, dynamic model which can constantly monitor the system and check for any blips in the system. This dynamic model will require a lot of computation power and development resources which can get tedious over time considering the evolution speed of Ransomware. The reaction time to shut down ransomware is also crucial, and it should be shut down as soon as it is detected.

Fortunately, machine learning is well suited for this task. Machine learning is a branch of artificial intelligence that enables computers to learn from data and make predictions or decisions. Machine learning can be applied to various aspects of ransomware defence. Machine learning algorithms like Random Forest (RF), Naïve Bayes (NB), Decision Trees (DT), and Logistic Regression (LR) can be used in Ransomware detection. Alternatively, we can also use Neural Network (NN) based architectures.

In this paper, we will be discussing the potential ransomware detection techniques using Machine Learning, the metrics of different machine learning algorithms, advantages, limitations, and how we can improve our defences against future ransomware attacks.

II. Introduction

Securing systems, networks, and programs from digital threats defines cybersecurity. These attacks often target sensitive data, exploit ransomware for financial gain, or disrupt business operations. Ransomware is a type of cyberattack. This form of attack is aimed at extorting money from individuals, businesses, or organisations. Ransomware, a type of malicious software, encrypts a victim's data, files, devices, or systems, making them inaccessible until a ransom is paid to the attacker. Cybercriminals exploit this malware for financial gain, with some even offering ransomware-as-a-service, making it more accessible to potential attackers. Luckily, there's a smart way to catch ransomware using something called machine learning. Traditional detection techniques are inadequate against advanced threats, highlighting the need for cutting-edge security measures. Machine learning (ML) presents a promising approach for ransomware detection, leveraging dynamic behaviors for automatic malware detection. Algorithms like Decision Trees, Random Forests, Naïve Bayes, Logistic Regression, and Neural Networks show potential for ransomware classification and detection. Deep learning techniques offer a promising solution to the challenges faced by traditional supervised ransomware detection tools, aiming to improve the accuracy and reliability of ransomware detection. These algorithms leverage automatic feature generation and excel in handling unstructured datasets, making them well-suited for detecting various forms of ransomware, including those involving audio, text, and images. It is not always possible to access labeled data Machine learning (ML) clustering is applicable here (unsupervised learning). It analyses the behavior patterns, grouping entities like endpoints, container images, or users based on similarities. This helps identify outlier behaviors that might not stand out individually but are significant when compared with similar entities. Applications of clustering in threat detection include spotting irregular endpoint configurations, detecting abnormal behaviors in users or containers compared to similar ones, and identifying changes in activity patterns over time. We'll explore these applications further below. Also, Semi-supervised learning can be applied to ransomware detection by leveraging both labeled and unlabeled data to improve the accuracy of classifiers. This paper aims to compare the performance of machine learning algorithms across various categories to determine their efficacy in ransomware detection. The study evaluates the performance of algorithms belonging to different categories, including supervised, unsupervised, and semi-supervised learning, in detecting ransomware based on static and dynamic features.

III. Literature Review

After an in-depth analysis of the research papers and relevant links, a comprehensive literature review on machine learning algorithms for ransomware detection is as follows:

Alraizza and Algarni's survey offers a comprehensive look at present issues and potential future discussions surrounding automated ransomware detection. It addresses current practices, innovations, and prospective approaches. In their work,

Ransomware Classification and Detection with Machine Learning Algorithms primarily focuses on feature selection-based frameworks that leverage neural network architectures for classification. A novel framework for ransomware detection is examined with a basis in classification utilizing various machine learning algorithms and neural network-based classifiers. The primary focus centered on feature selection to develop models that effectively identify ransomware threats [3].

The study by Abu Saad most likely looks into supervised machine learning methods for identifying ransomware assaults. Using a data-centric methodology, Vehabovic et al. explore early detection and attribution. Smyth studies the effect of semi-supervised feature selection on ransomware identification.

In the survey paper titled Cybersecurity Data Science: An Overview from Machine Learning Perspective, the authors discuss the use of machine learning in the fortification of cybersecurity.

Greater emphasis was placed on core principles of confidentiality, availability, and integrity. Critical analysis points out the limitations of signature-based intrusion detection systems and advocates for anomaly-based systems due to their detection capabilities. The challenges of cybersecurity data science are considered including the need for quality up-to-date datasets and the development of dynamic security models. The advocation of a multi-tiered framework that initiates proactive cybersecurity responses while adapting to evolving threats is also discussed [1].

A publication titled Classification Model for Accuracy and Intrusion Detection using Machine Learning Approach presents a study on the effectiveness of three machine learning classification algorithms in detecting network intrusions using the UNSW-NB15 dataset. The comparison consisted of Naïve Bayes, Support Vector Machines, and K-Nearest Neighbor. Out of the three machine learning algorithms Support Vector Machines performed with the highest accuracy. The purpose of this study was to emphasize the significance of selecting an appropriate algorithm to strengthen network security measures against the progressive advancement of cyber threats [2].

A diversified survey on Ransomware Detection using Machine Learning discusses the evolution, detection methodologies, and challenges pertaining to ransomware. This article provides in-depth background knowledge of the existing types of ransomware and distribution methods. Manual ransomware detection and automated ransomware are compared in addition to artificial intelligence-based approaches [4].

IV. Review on Machine Learning for Ransomware

Attackers continuously adapt their ransomware techniques, making it challenging for detection systems to keep up [6]. This is especially true for traditional systems that incorporate detection via signature-based methods. As soon as defenders recognize and shut down an attack vector, another one is opened. This calls for the need for a dynamic system that analyses the communications in the system and constantly checks for anomalies or variations in the hash keys or other security protocols.

However, this cannot be achieved using traditional models and it calls for a more advanced, dynamic model which can constantly monitor the system and check for any blips in the system. This dynamic model will require a lot of computation power and development resources which can get tedious over time considering the evolution speed of Ransomware. The reaction time to shut down ransomware is also crucial, and it should be shut down as soon as it is detected.

Thankfully, machine learning is well suited for this task. Machine learning is a branch of artificial intelligence that enables computers to learn from data and make predictions or decisions. Machine learning can be applied to various aspects of ransomware defence. Machine learning techniques have been increasingly used in the detection of ransomware due to their ability to learn behavior patterns and detect anomalies. Machine learning algorithms like Random Forest (RF), Naïve Bayes (NB), Decision Trees (DT), and Logistic Regression (LR) can be used in Ransomware detection. Alternatively, we can also use Neural Network (NN) based architectures [3].

Evaluating ransomware behavior and traits, including its encryption type, attack vector, file modifications, and network traffic, can be accomplished using machine learning. Machine learning models are capable of classifying files or processes as ransomware or not by identifying pertinent features from the data [11]. For instance, we can group and detect ransomware attacks using Random Forest algorithms by looking at the frequency and pattern of API calls.

Malicious emails, websites, and downloads are examples of potential ransomware sources that can be found and stopped using machine learning. Machine learning models can identify suspicious or unusual behaviors and notify users or administrators by evaluating the context or substance of the data [11]. Utilizing machine learning to monitor email attachments for the identification of potential malware and ransomware is one example.

To recover from ransomware attacks, machine learning can be used to decrypt data, restore the system, or track down the attackers. Through the examination of data patterns and trends, machine learning models can produce solutions or insights for the victims [13]. For instance, based on the ransomware family and the victim's profile, we can use machine learning to estimate the ransom amount and the chance of decryption.

One of the research projects conducted by Masum et al. compared multiple machine learning algorithms. In the research, they used RF, LR, NB, DT algorithms, and a NN-based architecture. The dataset used had 138,047 samples with 54 features where 70% are variants of ransomware and the remaining 30% are legit observations. Below is a histogram which shows the distribution of the dataset.

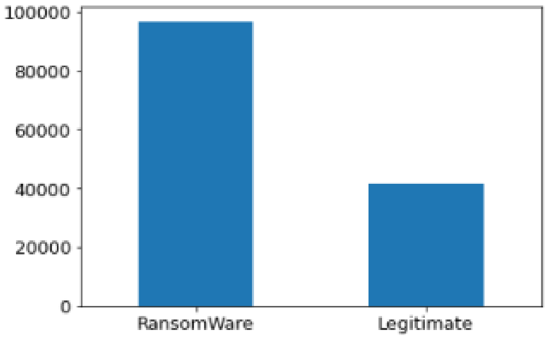
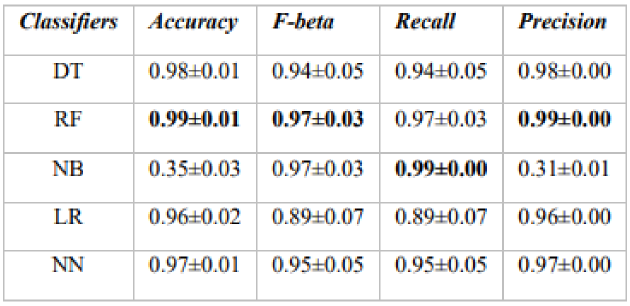


Figure 1 Distribution of the Dataset.

Researchers applied DT, RF, NB, LR, and NN classifiers to classify between legitimate and ransomware samples. The table below demonstrates the results of each model in terms of its performance metrics which includes accuracy, recall, precision, and F-beta score. The highest accuracy and precision are achieved by Random Forest model (around 0.99), also F-beta (0.97) is high compared to other models; there by outperforming all the other models. Although the NB model has the highest recall, its accuracy and precision are extremely low. The performance of DT and NN classifier is reasonable when compared to RF. Although the accuracy of LR is around 0.96 its recall and F-beta score are significantly lower compared to the other models [3].



Classifiers	*Accuracy*	*F-beta*	*Recall*	*Precision*
DT	0.98±0.01	0.94±0.05	0.94±0.05	0.98±0.00
RF	**0.99±0.01**	**0.97±0.03**	0.97±0.03	**0.99±0.00**
NB	0.35±0.03	0.97±0.03	**0.99±0.00**	0.31±0.01
LR	0.96±0.02	0.89±0.07	0.89±0.07	0.96±0.00
NN	0.97±0.01	0.95±0.05	0.95±0.05	0.97±0.00

Figure 2 Results of the experiment.

For classification tasks, such as ransomware detection, decision trees are an easy-to-understand machine learning algorithm. To create a tree-like structure that represents the decision-making process, decision trees recursively partition the data into subsets based on the values of the features. Decision trees, which are easy to understand but sensitive to even small changes in the data, can handle both continuous and categorical components.

This experiment shows the Random Forest algorithm performs better overall, considering that it is an ensemble method. Random forests are an extension of decision trees that improve performance and reduce overfitting.

By randomly selecting features and data, random forests create multiple decision trees and combine their predictions. They are better equipped to handle high-dimensional data and are less likely to overfit. However, they can be computationally demanding and difficult to interpret.

ML classifier performance is heavily dependent upon input training data. Hence Feature extraction is extremely important since ML classifier's performance is dependent on input training data. Feature extraction helps model to train on most important features. This feature selection process can be executed in multiple ways. However, Supervised feature extraction may not prepare the model to pick up new variants of the Ransomware, Unsupervised feature extraction may not yield the expected results. Therefore, we are left with Semi-Supervised learning.

V. Understanding

There has been a substantial growth of ransomware due to the availability and ease of use created by modern cryptographic tools paired with the anonymity of cryptocurrencies creating a robust and profitable platform for criminals to prosper [5].

The three primary types of ransomware attacks include locker, crypto, and scareware. Locker ransomware denies access to the device and crypto ransomware prevents data and files from being accessed. Scareware deceives the individual by generating alarming pop-ups or creating a cause for concern that leads to the installation of malicious software. Research communities must adopt futuristic technologies such as machine learning to create optimal levels of protection and enhance security measures [6].

Ransomware attacks are generally conducted through email attachments or malicious websites. Once access to the system has been gained through the elevation of system commands the ransomware begins iterating through the infected machine's system files. Then the ransomware begins to encrypt files with a private key encryption rendering the system or data useless for authorized users. Once the encryption has occurred, the hackers will require a ransom to regain system access via a symmetric key or will alternatively threaten to release private information gained in the attack if the ransom is not paid [4]. Even if the ransom is paid there is no guarantee the decryption will be successful, and that the stolen information won't still be released.

Ransomware detection consists of automated and manual methodologies. Manual techniques include routine scans and devices to monitor and indicate possible anomalies or attacks. This method is labour-intensive because it requires highly trained human intervention to continuously analyse and evaluate scan results or flagged incidents. Manual scanning processes can contribute to a higher rate of false positives, increasing the possibility of disruption to system operations [4].

A. *Behaviour-Based Detection Models*

Several innovative approaches in cybersecurity have been employed to combat ransomware attacks with the usage of machine learning algorithms. These models are trained to observe file access patterns, monitor system file changes, and oversee network activity to distinguish between benign and malicious behaviour. The following machine learning algorithms are used to analyse behaviours, anomalies, and patterns that manual security practices might not recognize.

Machine Learning Algorithms

Algorithm Name	***Description***
Decision Tree	A simplistic classification method that distinguishes between ransomware and benign software behaviours by training on data on file modifications, system calls, and network traffic [4].
Random Forest	Trained with the same features as Decision Tree for the detection of ransomware but creates an advancement to the Decision Tree algorithm by creating multiple decision trees and combining the predictions of each tree to improve performance and reduce the likelihood of overfitting [4].
Support Vector Machines (SVM)	Known for its ability to handle highly dimensional data. In terms of ransomware detection. Trained on features consisting of system calls, file modifications, and network traffic. In ransomware detection, the data can be non-linearly separable and have a high dimensionality which makes Support Vector Machines a viable option when using machine learning to detect ransomware [4].
K-Nearest Neighbours (KNN):	Uses a non-parametric approach for classification and regression. Can be used to detect ransomware hidden in executable files. The simplicity of this algorithm also assists in identifying behavioural patterns of ransomware [4].
Extreme Gradient Boosting (XGBoost)	Utilizes a combination of gradient boosting and decision trees to increase predictive accuracy. This algorithm is known for its powerful computational capabilities when analysing large complex datasets [4].
Logistic Regression	Operates as a binary classifier to help distinguish between benign or malicious files, behaviours, or network traffic patterns by modeling the output class as a function of the input features. When this algorithm is trained properly it is possible to find the optimal parameters for training and prevent overfitting [4].

B. *Semi-Supervised Learning*

In the field of machine learning, semi-supervised learning focuses on carrying out specific learning tasks with both labeled and unlabeled data. It enables the combination of often smaller sets of labeled data with the enormous volumes of unlabeled data available in many use cases. It's situated conceptually between learning that is supervised and unsupervised.

To enhance an unsupervised feature selection's performance, semi-supervised feature selection incorporates a small amount of labeled data into unlabeled data as extra information. Since the study of semi-supervised feature selection has received more attention recently, there are a lot of semi-supervised feature selection techniques available.

C. *Hybrid Model*

A hybrid model consists of a combination of machine learning techniques and deep learning to overcome the limitations of missing features or unsatisfactory profiling to increase target feature extraction to provide insight and make quick decisions in the event of a ransomware attack. Deep learning approaches can alleviate the need for manual feature engineering processes by passing the data through a neural network consisting of multiple layers to perform classification and feature learning simultaneously. Once the data has been processed through the neural network it is passed to a machine learning algorithm like gradient boosting for training to create enhanced processing performance and predictive accuracy [7].

VI. Limitations

To effectively train a machine learning algorithm to detect ransomware, large amounts of labeled data are needed. This leads us to the most impactful limitation hindering the ability to defend and prevent ransomware attacks due to a lack of relevant, timely, and quantifiable data. Many of the trainable datasets available are generated by researchers from synthetic datasets and sandbox environments. Having access to real-world ransomware data samples would greatly improve the ability to train the ML algorithm and allow for better analysis of modern real-world ransom attacks. Although many victims are unwilling to report the occurrence of the ransomware attack, it would aid in the global prevention and mitigation of ransomware attacks [4,8]

The corporate shift to an increased remote workforce has amplified security vulnerabilities depicting a relationship between the increase of remote workers and the upsurge of ransomware attacks. The rise of ransomware attacks during this time can be attributed to the lack of security present on home networks in comparison to corporate networks which allowed for vulnerabilities to be exploited. Not only was the lack of security on home networks an issue but also the increase in social engineering and phishing attacks assisted in the rapid spread of ransomware [9]. Due to these occurrences, companies must implement enhanced security measures and training procedures if they are to continue a hybrid or remote workforce employment style.

VII. Suggestions/Modifications

When data privacy is of concern, federated learning can alleviate the need for raw data to be shared. By utilizing distributed ransomware analysis, decentralized training local models can train on real-time data without the need for data transmission to a separate location [10].

As the model trains, updates will be sent to a central server to create a global model. This method seemingly bodes well in the theory of continuous updates to ransomware detection systems while simultaneously ensuring data privacy [11]. Although this is an idealistic solution there are many hurdles to overcome such as participation from organizations, regulatory restrictions, and adequate communications protocols. For these reasons, more research is needed.

To keep up with the continuous advancement of ransomware attack techniques it is important to invest more research into transfer learning methods that adapt quickly when new ransomware variants are released. Transfer learning provides models with knowledge derived from pre-existing models that have been previously trained on large datasets. This allows the model to increase predictive performance on a new task when the scarcity of labeled data for that task is present [12]. To ensure the best pre-trained models are utilized for existing model training more research will need to be conducted and a continuous iterative process should be implemented to remain valid in the detection of advancing ransomware attack methods.

In conjunction with ransomware detection methodologies; preventative measures should also be implemented to halt, moderate, or roll back any damage caused by ransomware. These preventative measures include stringent user access control, comprehensive system backups, and extensive user awareness training. Often a privileged user account is leveraged during an attack, but this can be moderated by implementing role-based access controls and routine permission audits. Although email filtering can efficiently prevent the delivery of malicious emails, it is still of the utmost importance to ensure individuals have been properly trained and educated to detect and report suspicious incidents [13].

VIII. Conclusion

During this study, the role of machine learning in the advancement of cybersecurity defences against ransomware attacks was explored. In much dismay, traditional cybersecurity mechanisms have fallen short in their abilities to combat the rapid evolution of ransomware tactics. Not only do ransomware attacks create substantial financial damage, but they also cause major disruptions to business operations and significantly harm the reputations of companies or individuals. Thus, creating heightened importance on generating a better understanding of ransomware behaviours to effectively prevent future attacks. Through a detailed analysis of relevant literature, algorithms, models, and techniques we have explored various ways machine learning can aid cybersecurity practices in the prevention, detection, and mitigation of ransomware. An examination of various machine learning algorithms ranging from decision trees to neural networks provides valuable insight into the untapped potential of cybersecurity defences.

Although, machine learning provides us with a hopeful outlook for the advancements of cybersecurity some limitations need to be addressed. Not only is there a high computational demand for complex machine-learning algorithms, but these algorithms also require sizeable quantities of labeled data for training. A unified approach to increase the posture of cybersecurity can consist of various adaptations such as transfer learning and federated learning which can be employed with the integration of machine learning and deep learning methodologies.

In conclusion, machine learning algorithms and deep learning models have great potential to enhance cybersecurity defences against ransomware attacks, but this does not negate the need for continuous improvements in network security and increased security training to increase defences and aid in the prevention of ransomware attacks. To remain vigilant against cybercriminals a collaborative effort must be made between industry professionals and members of academia to enhance cyber defences and combat the advancements of ransomware.

References

[1] Sarker, Iqbal & Kayes, A. S. M. & Badsha, Shahriar & Alqahtani, Hamed & Watters, Paul & Ng, Alex. (2020). Cybersecurity data science: an overview from machine learning perspective. *Journal of Big Data*. 7. 10.1186/s40537-020-00318-5.*)*

[2] Agarwal A, Sharma P, Alshehri M, Mohamed AA, Alfarraj O. 2021. Classification model for accuracy and intrusion detection using machine learning approach. *PeerJ Computer Science* 7:e437 https://doi.org/10.7717/peerj-cs.437 I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[3] K Masum, Mohammad & Hossain Faruk, Md Jobair & Shahriar, Hossain & Qian, Kai & Lo, Dan & Adnan, Muhaiminul. (2022). Ransomware Classification and Detection With Machine Learning Algorithms.

[4] A. Alraizza and A. Algarni, "Ransomware Detection Using Machine Learning: A Survey," *Big Data and Cognitive Computing*, vol. 7, no. 3, p. 143, 2023. [Online]. Available: https://doi.org/10.3390/bdcc7030143

[5] S. Sharmeen, Y. A. Ahmed, S. Huda, B. Ş. Koçer and M. M. Hassan, "Avoiding Future Digital Extortion Through Robust Protection Against Ransomware Threats Using Deep Learning Based Adaptive Approaches," in *IEEE Access*, vol. 8, pp. 24522-24534, 2020, doi: 10.1109/ACCESS.2020.2970466.

[6] Beaman, Craig et al. “Ransomware: Recent advances, analysis, challenges and future research directions.” *Computers & security* vol. 111 (2021): 102490. doi:10.1016/j.cose.2021.102490

[7] D. Gibert, J. Planes, C. Mateu, and Q. Le, "Fusing feature engineering and deep learning: A case study for malware classification," *Expert Systems with Applications*, vol. 207, p. 117957, 2022, Elsevier BV. DOI: 10.1016/j.eswa.2022.117957. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417422011927

[8] Cremer, F., Sheehan, B., Fortmann, M. *et al.* Cyber risk and cybersecurity: a systematic review of data availability. *Geneva Pap Risk Insur Issues Pract* **47**, 698–736 (2022). https://doi.org/10.1057/s41288-022-00266-6

[9] Beaman, Craig et al. “Ransomware: Recent advances, analysis, challenges and future research directions.” *Computers & security* vol. 111 (2021): 102490. doi:10.1016/j.cose.2021.102490

[10] A. Vehabovic et al., "Federated Learning Approach for Distributed Ransomware Analysis," in *Applied Cryptography and Network Security Workshops. ACNS 2023. Lecture Notes in Computer Science*, vol. 13907, J. Zhou et al., Eds. Cham: Springer, 2023. https://doi.org/10.1007/978-3-031-41181-6_33

[11] P. Liu, X. Xu, and W. Wang, "Threats, attacks and defenses to federated learning: issues, taxonomy and perspectives," *Cybersecurity*, vol. 5, no. 4, 2022. https://doi.org/10.1186/s42400-021-00105-6

[12] A. Singh, Z. Mushtaq, H. A. Abosaq, S. N. F. Mursal, M. Irfan, and G. Nowakowski, "Enhancing Ransomware Attack Detection Using Transfer Learning and Deep Learning Ensemble Models on Cloud-Encrypted Data," *Electronics*, vol. 12, no. 18, p. 3899, 2023. https://doi.org/10.3390/electronics12183899

[13] S. Parkinson, "Use of access control to minimise ransomware impact," *Network Security*, vol. 2017, no. 7, pp. 5-8, 2017. ISSN: 1353-4858. [Online]. Available: https://doi.org/10.1016/S1353-4858(17)30069-7.