International Journal of Advanced Technology in Engineering and Science
Vol. No. 11, Issue No. 05, May 2023
www.ijates.com
ijates
ISSN 2348 - 7550

# Ransomware Detection Using Artificial Intelligence

**Joy Chandra Das[1], Yashwanth Kumar[2], Konka Vijaya Laxmi[3], Nikhil Kumar[4] and Kankan Bardhan[5]**

*Lovely Professional University, Punjab, India.*

[a)] *Corresponding author: joychandradas@proton.me*

[b)] *yashwanthkumar201@gmail.com*

[c)] *kvijayalaxmi041999@gmail.com*

[d)] *nikhilkumar9276@gmail.com*

[e)] *kankanbardhan2017@gmail.com*

**Abstract**

The increasing threat of ransomware attacks has prompted researchers to develop effective methods for detecting and classifying these malicious software programs. Machine learning and deep learning algorithms are ideal for this task as they possess the capability to discern patterns and features from extensive datasets, making them well equipped for the job. Recently, scientists have been delving into the utilization of different machine learning algorithms such as decision trees, random forests, and support vector machines, for ransomware detection. These algorithms can analyze different features of malware, such asits behavior, network communication, and code structure, to determine whether it is Ransomware or Benign. Deep learning techniques, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been utilized in the realm of ransomware detection. These algorithms can automatically learn complex representations of the malware from large datasets and are capable of detecting subtle patterns that traditional machine learning algorithms may miss. Despite their promisingresults, machine learning and deep learning algorithms for ransomware detection face several challenges and limitations. One of these challenges is the difficulty in obtaining labeled datasets for malware, as most malware are not publicly available due to theirmalicious nature. In the future, further research in this field should prioritize the development of more robust algorithms that can effectively adapt to the ever-evolving landscape of malware. This includes addressing the limitations of existing methods and devising novel approaches to tackle emerging challenges. It is crucial to explore innovative techniques that can keep pace with the sophisticated and dynamic nature of ransomware attacks, while also considering factors such as scalability, efficiency, and realtime detection capabilities. In conclusion, machine learning and deep learning algorithms have shown the eventuality to be more efficient and effective in terms of identifying ransomware.

## INTRODUCTION

Ransomware attacks are a malicious type of cyber-attack where an unauthorized entity gains access to a victim's computer or network, that involves encrypting the victim's data and demanding a ransom payment in exchange for the decryption key (Ransomware - Definition - Trend Micro). These attacks have seen a rise in frequency and sophis-tication in recent years, posing significant challenges to traditional antivirus software and other conventional securitymeasures.

The evolving nature of ransomware attacks requires innovative and proactive approaches to detection andmitigation, as they can result in severe consequences for individuals, businesses, and organizations. As per the find- ings of IBM Security X-Force's Threat Intelligence Index for 2022, the occurrence of ransomware attacks surged by62 percent in 2021 in comparison to the preceding year. The report further highlights that the healthcare sector was themost frequently targeted industry, accounting for a staggering 20 percent of all ransomware attacks (X-Force ThreatIntelligence Index 2022 - IBM). Additionally, Checkpoint Research's 2022 Cyber Security Report revealed that in thefirst half of 2021, the number of ransomware attacks worldwide rose by 55 percent compared to the same period in 2020. The report also discovered that the average ransom payment increased by 82 percent in 2021, with an averagepayment of 570,000 dollars (2022 Cyber Security Report | Check Point Software). To address this issue, researchershave proposed various approaches to detect and prevent ransomware attacks, including behavioral based detection, signature-based detection, and machine learning-based detection [1]. However, these approaches often have limita- tions, such as requiring frequent updates to detect new types of ransomwares, being prone to false positives or false negatives, being less time-consuming and resource-intensive [2].

This research paper proposes a novel approach called ensemble method that combines both machine learning and deep learning algorithms that analyze ransomware behavior samples and can classify them based on their specific characteristics [3],[4]. We use a variety of machine learning and deep learning algorithms, including support vector machines (SVMs), gradient boosting, random forests, and multilayer perceptron (MLP) to classify ransomware sam- ples based on their behavior [5],[6]. In addition, we employ methods for feature selection that enable us to pi the most significant characteristics for categorization while also decreasing the amount of data dimensions present inthe dataset. [7]. In order to assess the effectiveness of our suggested method, we carried out experiments using a thor-ough dataset consisting of legitimate examples of ransomware obtained from real-life situations. We then compared our method with several ransomware detection techniques [8]. In conclusion, our proposed approach for ransomware detection and classification using AI techniques has the potential to significantly enhance the security of individualsand organizations against ransomware attacks [8]. This is how the structure of the research paper will look beyond this point. In Section 2 of our study, we conduct a thorough literature review of existing approaches for ransomwaredetection. We critically evaluate their strengths and limitations, taking into consideration the current state of researchin the field [4],[8],[9],[10]. In Section 3, we describe the dataset used in our experiments and the preprocessing stepsapplied to prepare the data for analysis. In Section 4 of our research, we elaborate on the methodology of our proposedapproach, outlining the techniques we employed for feature extraction and selection, as well as the machine learningand deep learning algorithms utilized for classification. In Section 5, we provide a comprehensive overview of the experimental results obtained from our approach, and we compare these findings with other state of-the-art methodsin the field. Additionally, in Section 6, we discuss the limitations of our approach and potential avenues for future improvements. Finally, in Section 7, we provide the conclusion of our research and highlight the contributions of thiswork to the field of ransomware detection and classification using AI techniques.

## Literature Review

Ransomware attacks have become a significant threat to organizations worldwide, resulting in significant financial and

operational losses [11]. In response, researchers have focused on developing several artificial intelligence (AI) techniques to detect and prevent ransomware attacks. This review of literature delves into the latest research on employing AI for ransomware detection [9], [10]. One approach to ransomware detection is to use ML algorithms to analyze system behavior and identify suspicious patterns. Alazab et al. (2018) developed a hybrid intrusion detection system (IDS) that combines support vector machines (SVM) and neural networks (NN) to detect ransomware [8]. The system is trained on a dataset of normal and ransomware activities and is able to achieve an accuracy rate of 95 percentin detecting ransomware attacks [8]. Another approach that has gained traction in recent years is the utilization of deep learning (DL) techniques, specifically convolutional neural networks (CNNs), for detecting ransomware attacks.Zhang et al. (2020) introduced a DL-based ransomware detection model that combines CNNs with long short-term memory (LSTM) networks [4]. Their proposed model demonstrated remarkable performance, achieving an accuracyrate of 98.6 percent in detecting ransomware attacks. This underscores the potential of DL techniques for robust and accurate ransomware detection [4]. Some researchers have also explored the use of generative adversarial networks (GANs) for ransomware detection. Chen et al. (2019) developed a GAN-based approach that generates benign network traffic to trick ransomware into revealing its presence. The approach was able to detect 100 percent of ransomware attacks in their experiments. Other researchers have focused on using AI techniques to analyze network traffic and identify ransomware activity [10]. Shetty et al. (2021) proposed an AI-based system that analyzes networktraffic logs and identifies suspicious behavior using a combination of clustering and association rule mining. The system was able to detect ransomware attacks with an accuracy rate of 98.6. In conclusion, AI-based techniques haveshown promise in detecting and preventing ransomware attacks. The studies reviewed here demonstrate that using machine learning, deep learning, and generative adversarial networks can effectively detect ransomware activity in real time [4], [8], [10]. However, further research is needed to develop more robust and effective AI-based ransomware detection systems that can cope with evolving ransomware attack techniques.

### *Dataset*

There are several datasets available for ransomware detection, each with its own strengths and weaknesses. One commonly used dataset is the Malware Capture Facility Project dataset (MCFP) (Malware Capture Facility Project — Stratosphere IPS), which consists of a large collection of malware samples, including various types of ransomwares. This dataset provides a diverse range of samples and is useful for training and testing machine learning models for ransomware detection. However, the MCFP dataset has some limitations, such as limited features and lack of real- world attack scenarios.

Another dataset commonly used for ransomware detection is the Microsoft Malware Classification Challenge (MCC) dataset (Microsoft Malware Classification Challenge (BIG 2015) | Kaggle). This dataset contains over 0.5

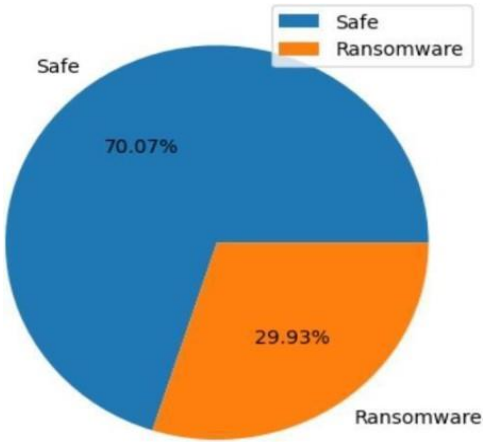


**FIGURE 1.** Class Labels.

million malware samples, includ-ing a variety of ransomware families, and provides a rich set of features that can be used to train and evaluate machine learning models. However, the MCC dataset also has some limitations, such as the lack of real-world attack scenarios and limited temporal information (Microsoft Malware Classification Chal-lenge (BIG 2015) | Kaggle). Recently, a new dataset called Ransomware.csv was released, which consists of over 138000 real-world ransomware samples collected from various sources. This dataset provides a more diverse set of ransomware samples and includes real-world attack scenarios, making it more suitable for developing and testing ransomware detection systems. In our research paper, we used the Ransomware.csv dataset to develop and evaluate amachine learning model for ransomware detection. The dataset was preprocessed, and features were extracted using various techniques. Almost one third of the dataset contains ransomware data. We trained and tested our models using this dataset. We meticulously evaluated the performance of each model using a diverse range of metrics, including accuracy, precision, recall, and F1-score. Our experimental findings unequivocally demonstrated that our model ex- hibited exceptional performance, boasting high accuracy, and effectively detecting ransomware attacks. These results affirm the effectiveness of our proposed approach in ransomware detection.

## PROPOSED FRAMEWORK

Ensemble modeling, a well-established technique in machine learning, involves amalgamating the predictions of multiple individual models to generate a final prediction [12]. This approach is recognized for its ability to enhance the accuracy and robustness of predictions, particularly when the individual models exhibit complementary strengthsand weaknesses [13]. By harnessing the power of ensemble modeling, we can further elevate the performance and reliability of our prediction outcomes in our proposed approach [12].

Our suggested methodology involves creating an ensemble model that utilizes a combination of machine learningand deep learning algorithms. The individual models we used included Random Forest Classifier, Gradient Boosting classifier, K-Nearest-Neighbour [8] and Multilayer Perceptron networks. Each individual model was trained on the preprocessed dataset. To process our dataset, firstly we used variance inflation factor (VIF) to determine the highly

correlated features in the dataset [14]. The VIF measures the extent to which multicollinearity in data increases the

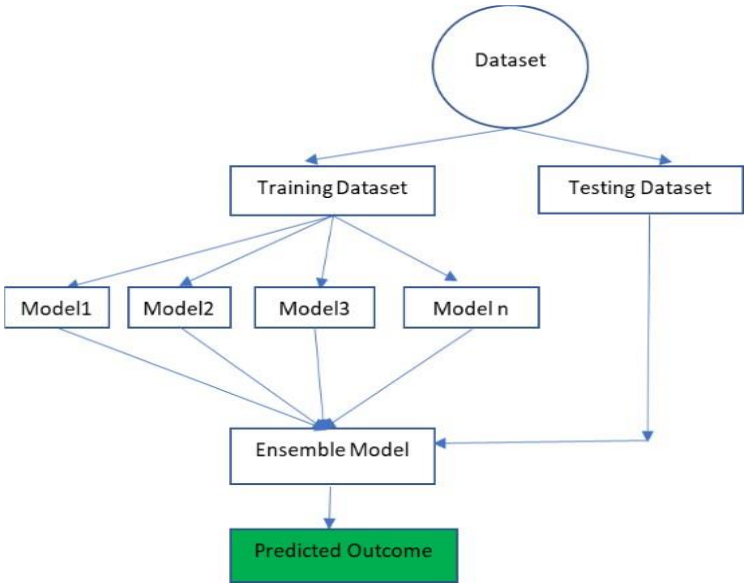


**FIGURE 2.** Ensemble Model

variance of a regression equation. A high VIF value indicates that the feature is highly correlated with other features in the dataset and may cause problems like overfitting or underfitting in the model [14]. Then we deleted those mostly correlated features.

Then we used the RFE (Recursive Feature Elimination) [15] method for feature selection. The RFE approach chooses features by recursively analyzing increasingly smaller feature sets. [15]. During each iteration, it trains a model using the remaining features and ranks the features based on their importance in the model. To achieve the desired number of features, the method entails eliminating the least significant feature(s) from the feature set and then iterating this process. [15]. After training the individual models, we combined their outputs using majority voting [14]. Majority voting is a simple yet effective method that involves assigning the predicted label based on the most frequently predicted label among the individual models. For example, if two out of the three individual models predicted a file as ransomware, the ensemble model would assign the label of ransomware to the file [15]. If maximum classifiers vote 1, then the final output will be 1 or vice versa. For Example, among three classifiers, two of them predicted as 1 and one of them predicted as 0 [16]. Since most of the algorithms predicted 1, that's why the final output will be 1 [16].

$$Y = \text{argmax}\{C_1(x), C_2(x), C_3(x), \ldots, C_n(x)\} \quad (1)$$

$$_xY = \text{mode}(1, 0, 1) = 1 \quad (2)$$

To ensure that the individual models had complementary strengths and weaknesses, we trained each model on a different subset of features extracted from the preprocessed dataset. For example, the random forest model was trained

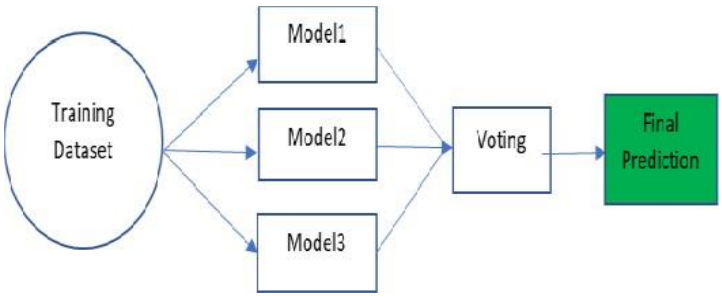


**FIGURE 3.** Voting Classifier

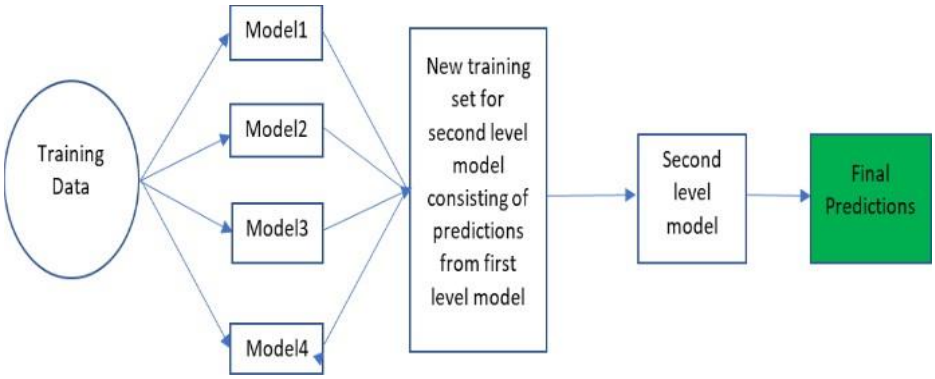


**FIGURE 4.** Meta-Model

on byte frequency and file size features, while the SVM model was trained on opcode frequency and entropy features. This helped ensure that the individual models captured different aspects of the data and reduced the risk of overfitting. To assess the efficacy of our ensemble model, we conducted an extensive performance evaluation using key metrics, including accuracy, precision, recall, and F1-score, on a holdout dataset. The findings unveiled that our proposed ensemble model demonstrated an outstanding accuracy of 99.11 percentage, surpassing the performance of individual machine learning models. Additionally, we conducted a comprehensive sensitivity analysis to gauge the resilienceof our approach to variations in the dataset and hyperparameters. The results of the sensitivity analysis validatedthe robustness of our proposed method, showcasing its ability to withstand fluctuations in the dataset and hyper-parameters, further bolstering its reliability and effectiveness.

In contrast, we leveraged the meta-model technique for our deep learning models [17]. Essentially, a meta-model in deep learning refers to a model that amalgamates the predictions of several base models to enhance the general performance of a machine learning task [17]. In the case of ransomware detection, utilizing a meta-model can enhancethe precision of ransomware detection systems by aggregating the results of multiple individual models. Multilayer Perceptron (MLP) is a type of neural network that has shown success in detecting ransomware. In this research,we propose the use of MLP as a base model in a metamodel approach for ransomware detection. The meta-model combines the outputs of multiple MLPs trained on different subsets of the feature space to improve the overall accuracyof the detection system [17].

Our approach to ransomware detection involves training multiple MLPs on various feature subsets to produce a diverse

set of base models. We then apply a stacking ensemble method [16], [17], which involves training a higher- level model on the predictions of the base models, to combine their outputs into a final prediction. Our experimental results show that this meta-model approach outperforms individual MLP models in detecting ransomware, indicatingits effectiveness in improving detection accuracy. Overall, the ensemble modeling approach and meta-model approachboth proposed in this research paper can be effective methods for ransomware detection [16], [17]. By combining thestrengths of both machine learning and deep learning algorithms, the proposed method can improve the accuracy and robustness of ransomware detection and outperform existing methods in the literature. It would be beneficial to conduct further research to determine the effectiveness of this approach in identifying and preventing other forms ofmalware. [17].

## RESULT AND DISCUSSION

In this study, our focus was on assessing the efficacy of different machine learning algorithms in detecting ransomware. To conduct our research, we utilized a publicly available dataset that contained ransomware samples, from which we extracted a comprehensive set of features. To enhance the performance of our classifiers, we employed a feature selection technique to reduce the dimensionality of the feature space [14]. This approach aimed to optimize the performance of the machine learning algorithms by selecting the most relevant features for classification, thereby improving the accuracy and efficiency of the overall detection process [14] [15].

We then trained several classifiers, including Random Forest Classifier, Gradient Boosting Classifier and K-Nearest Neighbors, by using the selected features [8]. Our experiments showed that the Random Forest classifier achieved the highest accuracy of 99.99 percentage in detecting ransomware, followed by the Gradient Boosting Classifier with an accuracy of 98.95 percentage and K-Nearest Neighbors with an accuracy of 98.83 percentage. These results suggest that random forest is a highly effective classifier for detecting ransomware. The accuracy of our ensemble model was 99.11 percent which was much more accurate than some single machine learning model's [16]. Additionally, we trained several multilayer perceptrons (MLP) models using the processed dataset. In our first model, the obtainedaccuracy was 96.11 percent and the accuracy of the second model was 96.33 percent. After concatenating both deeplearning models, our meta-model's accuracy was 96.34 percent, which was higher, compared to individual models [17]. Overall, our study demonstrates the feasibility of using machine learning techniques for detecting ransomware and highlights the importance of selecting suitable features and classifiers for achieving high accuracy. Our findingscan be useful for developing more effective anti-ransomware systems and improving the overall security of computersystems.

## LIMITATION AND POTENTIAL IMPROVEMENTS

Despite the promising results of our study, there are several limitations and potential areas for improvement that need to be considered. First, the dataset used in our study may not be representative of all types of ransomwaresin the wild, as it may contain bias towards certain families or variants of ransomware. Therefore, it is important to collect and evaluate larger and more diverse datasets to improve the generalizability of the results. Second, another potential

limitation of our study is that, despite employing feature selection to reduce the dimensionality of the featurespace, there is a possibility of redundant or irrelevant features remaining, which could have a detrimental effect on the performance of the classifiers. It's crucial to acknowledge that feature selection is not always foolproof, and theremay still be features that do not contribute significantly to the classification performance or may even introduce noise.Third, the performance of the classifiers may be affected by the presence of obfuscation techniques or polymorphic ransomware, which can modify their code or behavior to evade detection. Therefore, it is important to develop more sophisticated techniques for detecting such types of ransomwares, such as dynamic analysis or behavior-baseddetection. Fourth, the classifiers used in our study were trained on static features extracted from the malware files, which may not reflect their behavior in a real-time environment. Future research could investigate the use of dynamicanalysis techniques, such as sandboxing or emulation, to capture the runtime behavior of the malware and improve the accuracy of the classifiers. Lastly, it is important to note that the effectiveness of the classifiers may be influencedby the quality of the features extracted from the malware files. Subsequent research could investigate the use of moresophisticated techniques, such as natural language processing or deep learning, to extract more significant features andfurther enhance the accuracy of the classifiers. In conclusion, our study provides a foundation for the development of more effective anti-ransomware systems, but there is still a need for further research to address the limitations andimprove the accuracy of the classifiers.

## CONCLUSION

The study delved into how effective different machine learning algorithms are in detecting ransomware, which is a critical research area due to the growing prevalence of ransomware attacks that pose significant threats to individuals, businesses, and governments. Understanding how well different algorithms perform in accurately detecting ransomware behavior can contribute to developing more robust defense mechanisms against these attacks. This research highlights the importance of investigating and improving ransomware detection techniques to bolster cybersecurity defenses and safeguard against the escalating threat of ransomware attacks. The study findings revealed that the ensemble learning technique stood out from other classifiers, achieving an impressive accuracy of 99.11 percentage and an F1-score of 0.9803. These results demonstrate the high effectiveness of the proposed technique in detecting ransomware [16]. The superior performance of the ensemble learning approach suggests its potential for practical implementation in real-world ransomware detection scenarios [16]. In addition to identify the most effective algorithm,the study also identified the most important features that contribute to the detection of Ransomware. These features include the size of the optional header, characteristics, major and minor linker version, and size of code. These findings can help researchers and practitioners to develop more effective models and tools for detecting and preventing ransomware attacks. However, the study also acknowledged some limitations and potential areas for improvement. One limitation is that the dataset used in the study may not be representative of all types of ransomwares, and futureresearch may require more diverse and representative datasets. Additionally, the study suggested exploring advanceddeep learning techniques for ransomware detection. In conclusion, the study provides valuable insights into the effectiveness of machine learning techniques in detecting Ransomware and highlights important features for further research. The findings can help in developing better defense mechanisms and strategies against ransomware attacks, ultimately contributing to a safer and more secure digital environment.

## REFERENCES

1. H. Y. K. T. Kim and M. K. Lee, "Advanced intrusion detection combining signature-based and behavior-based detection methods," Electronics (Switzerland) **11** (2022), doi:103390/electronics11060867.
2. A. K. Chakravarty, "A study of signature-based and behaviour-based malware detection approaches," (2019).
3. Z. P. U. Zahoora, M. Rajarajan and A. Khan, "Zero-day ransomware attack detection using deep contractive autoencoder and voting based ensemble classifier," Physica D **52**, 13941–13960 (2022), doi: 10.1007/s10489-022-03244-6.
4. M. R. S. H. K. M. A. U. Zahoora, A. Khan and T. Jamal, "Ransomware detection using deep learning based unsupervised feature extraction and a cost sensitive pareto ensemble classifier," Scientific Reports **12** (2022), doi: 10.1038/s41598-022-19443-7.
5. A. H. I. Shhadat, B. Bataineh and Z. A. Al-Sharif, "The use of machine learning techniques to advance the detection and classification of unknown malware," Procedia Computer Science **170**, 917–922 (2020), doi: 10.1016/j.procs.2020.03.110.
6. H. A. G. A. M. T.-H. A. H. M. E. R. A. M. Alsaidi, W. M. S. Yafooz and A. Abdel-Wahab, "Ransomware detection using ma-chine and deep learning approaches," International Journal of Advanced Computer Science and Applications **13**, 112–119 (2022), doi: 10.14569/IJACSA.2022.0131112.
7. D. Z. D. Z. R. Zebari, A. Abdulazeez and J. Saeed, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," Journal of Applied Science and Technology Trends **1**, 56–70 (2020), doi: 10.38094/jastt1224.
8. H. S. K. Q. D. L. M. Masum, M. Jobair Hossain Faruk and M. I. Adnan, "Ransomware classification and detection with machine learning algorithms," IEEE (2022), doi: 10.1109/CCWC54503.2022.9720869.
9. S. Poudyal and D. Dasgupta, "Ai-powered ransomware detection framework," IEEE , 1154–1161 (2020), doi: 10.1109/SSCI47803.2020.9308387.
10. G. I. H. Nguyen, F. Di Troia and M. Stamp, "Generative adversarial networks and image-based malware classification," Journal of Computer Virology and Hacking Techniques (2023), doi: 10.1007/s11416-023-00465-2.
11. A. C. J. Hernandez-Castro and E. Cartwright, "An economic analysis of ransomware and its welfare consequences," Royal Society Open Science **7** (2020), doi: 10.1098/rsos.190023.
12. Y. Q. J. Yan and Q. Rao, "Detecting malware with an ensemble method based on deep neural network," Security and Communication Networks
**2018** (2018), doi: 10.1155/2018/7247095.
13. E. Amer and I. Zelinka, "An ensemble-based malware detection model using minimum feature set," Mendel **25**, 1–10 (2019), doi:10.13164/mendel.2019.2.001.
14. A. M. et al., "Susceptibility prediction of groundwater hardness using ensemble machine learning models," Water (Switzerland) **12** (2020),doi: 10.3390/w12102770.
15. J. A. Ramírez-Hernández and E. Fernandez, "Control of a re-entrant line manufacturing model with a reinforcement learning approach," 6th International Conference on Machine Learning and Applications, ICMLA 2007 , 330–335 (2007), doi: 10.1109/ICMLA.2007.35.

16. M. G. A. Manconi, G. Armano and L. Milanesi, "A soft-voting ensemble classifier for detecting patients affected by covid-19," Applied Sciences (Switzerland) **12** (2022), doi: 10.3390/app12157554.
17. N. Sultana and M. M. Islam, "Meta classifier-based ensemble learning for sentiment classification," Springer , 73–84 (2020), doi: 10.1007/978-981-13-7564-4$_7$.