Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 9, No.1, September : 2025

EVALUATE THE EFFECTIVENESS OF HYBRID META VOTE CLASSIFIER IN ENHANCING CYBERSECURITY FOR SMALL AND MEDIUM ENTERPRISES

Bhanwar Lal Patel, Research Scholar, Faculty of Computer Science, Pacific Academy of Higher Education & Research University, Udaipur, Rajasthan, India
Dr. Rajesh Kanja, Assistant Professor, Faculty of Computer Science, Pacific Academy of Higher Education & Research University, Udaipur, Rajasthan, India
Corresponding Email: patelbharatudr@gmail.com; rajesh.kanja@gmail.com

Abstract:
The study evaluates the effectiveness of machine learning predictive models in enhancing cybersecurity for small and medium enterprises (SMEs). Given the increasing cyber threats faced by SMEs, traditional models—including REP Tree, J48, and Naive Bayes—are compared with a hybrid meta vote classifier predictive model. The findings indicate that the hybrid model achieved an accuracy of 67.87%, outperforming the traditional models. The REP Tree and J48 also demonstrated strong performance, while simpler models like Naive Bayes showed lower efficacy. Both null hypotheses were rejected, indicating a significant difference in the effectiveness of various models and confirming that the hybrid model is more effective than traditional models. This research highlights the potential of hybrid approaches to strengthen cybersecurity resilience in SMEs against evolving threats.

Keywords: J48, Accuracy, SMEs, Machine Learning

1. Introduction:

In today's digital landscape, small and medium enterprises (SMEs) play a vital role in the global economy, driving innovation, employment, and growth. However, their increasing dependence on digital technologies exposes them to a myriad of cybersecurity threats. Unlike larger organizations, SMEs often lack the financial resources, technical expertise, and comprehensive security frameworks necessary to defend against sophisticated cyber-attacks. This vulnerability makes them attractive targets for cybercriminals, who exploit the inadequacies in their security measures to carry out data breaches, ransomware attacks, and other malicious activities.

Cybersecurity incidents can have devastating consequences for SMEs, including financial losses, reputational damage, and legal ramifications. Studies indicate that a significant percentage of SMEs that experience a major cyber-attack cease operations within six months due to the resulting financial strain and loss of customer trust. Consequently, there is an urgent need for effective cybersecurity strategies tailored to the specific challenges faced by SMEs. This necessity has spurred interest in leveraging advanced technologies, particularly machine learning, to enhance cybersecurity defences.

Machine learning (ML) offers the capability to analyse vast amounts of data, recognize patterns, and detect anomalies that could indicate potential security threats. Traditional cybersecurity approaches often rely on static rules and signature-based detection methods, which are increasingly ineffective against new and evolving threats. In contrast, machine learning algorithms can adapt and learn from new data, enabling them to identify previously unseen attack vectors. Hybrid machine learning models, which combine multiple algorithms to leverage their complementary strengths, represent a promising direction for improving the effectiveness of cybersecurity measures.

The objective of this research is to evaluate the effectiveness of hybrid machine learning predictive models in enhancing cybersecurity for SMEs. By integrating various machine learning techniques, such as decision trees, support vector machines, and neural networks, this study aims to create a robust predictive framework capable of identifying vulnerabilities and predicting cyber-attack patterns. Such models can provide SMEs with actionable insights, enabling them to take proactive measures to mitigate risks before they result in significant harm.

Furthermore, this research will explore how hybrid models can be tailored to the unique operational contexts of SMEs. Understanding the specific factors that influence the effectiveness of these models—such as the type of data being analysed, the nature of the threats faced, and the operational constraints of SMEs—will be critical for developing practical solutions.
The integration of hybrid machine learning models into the cybersecurity strategies of SMEs can not only enhance their ability to detect and respond to threats but also contribute to building a more resilient cybersecurity infrastructure. By empowering SMEs with the tools and knowledge necessary to defend against cyber threats, this research seeks to foster a safer digital environment, ensuring that these essential contributors to the economy can thrive in an increasingly interconnected world.
Through a comprehensive evaluation of the effectiveness of these models, this study aims to provide insights and recommendations for SMEs looking to strengthen their cybersecurity posture and protect their vital assets in the face of growing cyber threats.

2. Literature Review:
The application of machine learning algorithms for malware detection and cybersecurity incident prediction has been the focus of extensive research efforts aimed at identifying and enhancing the performance of various algorithms in diverse contexts. In their study, Akhtar and Feng (2023) performed a thorough evaluation of machine learning models, highlighting the necessity of employing multiple performance metrics, such as accuracy, precision, recall, and area under the ROC curve (AUC-ROC). Their findings indicate that sophisticated machine learning techniques can substantially enhance the accuracy of malware detection, although they also point out the challenges associated with adapting these models to effectively combat real-world adversarial conditions.
Ali et al. (2019) introduced a proactive methodology for malware identification through the development of a framework tailored for digital forensic examiners. This framework integrates both static and dynamic analysis techniques to strengthen malware detection capabilities. By addressing critical factors such as detection time and accuracy, the proposed approach is particularly relevant in the realm of digital forensics, where prompt and precise identification is crucial for supporting legal proceedings. Additionally, Alqahtani (2021) provided a comprehensive overview of various machine learning techniques applicable to malware detection, underscoring existing challenges and potential future directions. The study emphasizes the need for adaptive and resilient models capable of responding to the continuously evolving landscape of cyber threats.
The specific roles of algorithms in malware detection have also been rigorously explored. Research by Bhavsar and Panchal (2012) and Biggio et al. (2014) investigated Support Vector Machines (SVMs), emphasizing their robustness and adaptability for data classification tasks in cybersecurity contexts. SVMs have shown significant promise, particularly when paired with data augmentation strategies. Catak et al. (2021) demonstrated this potential by applying Convolutional Neural Networks (CNNs) combined with data augmentation techniques, resulting in improved malware detection performance. Collectively, these studies reveal both the strengths and limitations of machine learning algorithms—particularly SVMs, decision trees, and neural networks—in enhancing cybersecurity defenses against increasingly sophisticated threats.
The evolution of malware classification has been significantly influenced by decision tree algorithms. Bashari Rad et al. (2011) conducted research on the statistical features of opcodes for classifying morphed virus families using decision trees, showcasing the effectiveness of these classifiers in analyzing intricate malware behaviors through feature extraction methods. This research not only improves classification accuracy but also offers insights into the distinguishing characteristics of various malware types, which is vital for formulating targeted detection strategies.
The challenges associated with existing algorithms are also well-documented in the literature. Cobian (2022) performed an in-depth efficiency analysis of machine learning algorithms aimed at developing hardware-based cybersecurity countermeasures. This analysis highlights the trade-offs between computational efficiency and detection accuracy, particularly in resource-constrained environments

where rapid response times are critical. As cyber threats continue to evolve, there is an increasing demand for adaptive models that can adjust to new malware variants and effectively address emerging challenges.

Additionally, the integration of artificial intelligence (AI) techniques in cybersecurity is gaining momentum, as explored by Camacho (2024). This study emphasizes the role of AI in anticipating and mitigating cyber threats by enabling real-time data analysis and decision-making support. As the prevalence of machine learning models increases within cybersecurity frameworks, continuous evaluation and refinement of these models will be essential to ensure they remain effective against emerging threats. Ongoing research in this area is crucial for developing future strategies for malware detection and predicting cybersecurity incidents, equipping organizations to defend against an ever-evolving landscape of cyber threats.

Furthermore, the issue of cyberbullying in the digital age has emerged as a significant concern, with detection techniques categorized into text-based and multimodal approaches. According to Ojha, Patil, and Joshi (2024), these techniques are analysed for their respective strengths and weaknesses, while also addressing current challenges in detection methods. The authors explore potential future enhancements to these models, aiming to adapt to the continually evolving nature of cyberbullying.

Finally, the literature demonstrates a robust commitment to improving machine learning algorithms for cybersecurity applications. Significant strides have been made in enhancing detection accuracy and operational efficiency. However, the dynamic nature of cyber threats necessitates ongoing research and development efforts to keep pace with new malware strategies, ensuring that machine learning models remain adaptable and effective in addressing emerging challenges.

3. Research Methodology:

The study employs a comparative analysis of various machine learning algorithms for predicting and detecting cybersecurity incidents, utilizing the WEKA software tool for implementation and evaluation. The research focuses on six classification techniques: Naive Bayes, SMO, Logistic, Multi Class Classifier, J48 and REP Tree. Performance measures, including accuracy, precision, recall, and ROC, are assessed using 10-fold cross-validation to ensure robust and reliable results.

Objective:

1. To evaluate the effectiveness of various machine learning predictive models in enhancing cybersecurity for small and medium enterprises.
2. Comparative Analysis of hybrid meta vote classifier predictive model and traditional cybersecurity models for Small and Medium Enterprises (SMEs).

Hypothesis:

$H_0$1: There is no significant difference in the effectiveness of various machine learning predictive models in enhancing cybersecurity for small and medium enterprises (SMEs).

$H_0$2: The hybrid meta vote classifier predictive model is not effective than traditional cybersecurity models in enhancing cybersecurity for small and medium enterprises (SMEs).

4. Data Analysis & Interpretation:

4.1 Analysis Machine Learning Models used for Enhancing Cybersecurity:

The performance of various machine learning algorithms was evaluated based on several metrics to enhance cybersecurity for small and medium enterprises (SMEs). Among these algorithms, the REP Tree model achieved the highest accuracy at 99.30%, followed closely by the J48 model at 99.04%. These models significantly outperformed Naive Bayes, which recorded an accuracy of 89.69%. The high accuracy rates of REP Tree and J48 indicate that these models are capable of correctly classifying a large proportion of instances, making them effective tools for cybersecurity applications.

In terms of incorrectly classified instances, the REP Tree model had the lowest rate at 0.70%, demonstrating its reliability. On the other hand, Naive Bayes exhibited the highest rate of incorrectly classified instances at 10.31%, which suggests that it may not be as effective in accurately detecting

cybersecurity threats. This discrepancy in classification accuracy underscores the varying effectiveness of different algorithms in threat detection.

Table 4.1: Analysis Machine Learning Algorithms used for enhancing cybersecurity for small and medium enterprises (SMEs) at Configuration Setting Percentage Split: 30%

Performance Measures	Naive Bayes	SMO	Logistic	Multi Class Classifier	J48	REP Tree
Correctly Classified Instances: Accuracy	89.69%	97.22%	97.53%	97.53%	99.04%	99.30%
Incorrectly Classified Instances	10.31%	2.78%	2.47%	2.47%	0.96%	0.70%
Kappa statistic	0.7922	0.944	0.9502	0.9502	0.9806	0.986
Mean absolute error	0.1	0.0278	0.0356	0.0356	0.0125	0.0103
Root mean squared error	0.3085	0.1667	0.1364	0.1364	0.0954	0.079
Precision	0.897	0.972	0.975	0.975	0.99	0.993
Recall	0.897	0.972	0.975	0.975	0.99	0.993
F-Measure	0.897	0.972	0.975	0.975	0.99	0.993
ROC Area	0.966	0.971	0.994	0.994	0.994	0.998
Average Execution Time (Model Building)	0.14 seconds	88.55 seconds	6.49 seconds	11.21 seconds	1.07 seconds	0.47 seconds

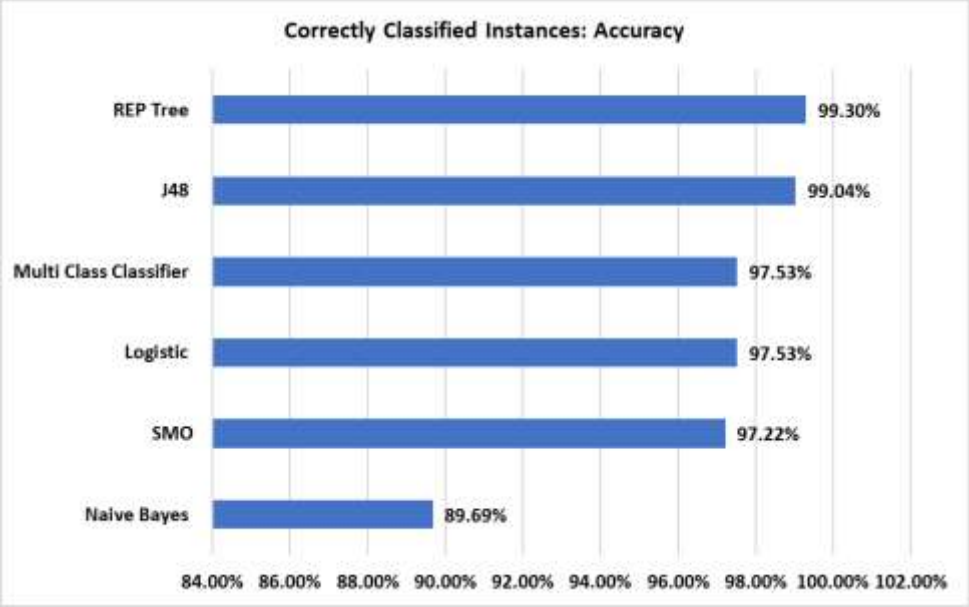

Figure 4.1: Evaluation based on Performance Measure: Accuracy

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 9, No.1, September : 2025

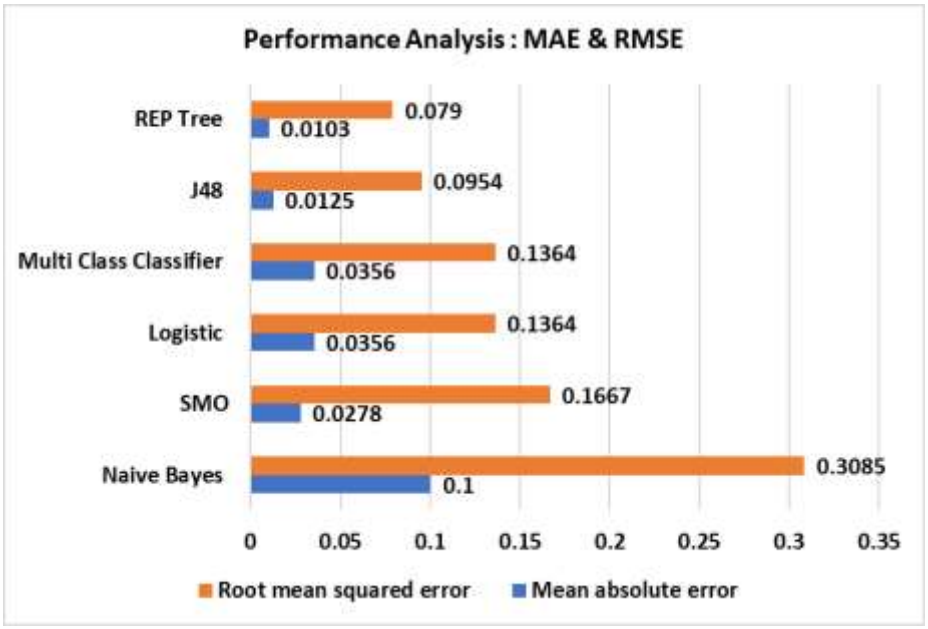

Figure 4.2: Evaluation based on Performance Measure: MAE & RMSE

When assessing predictive performance through Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), the REP Tree model excelled with the lowest MAE at **0.0103** and RMSE of **0.079**. These low error rates signify minimal average error in its predictions, reinforcing its effectiveness. Conversely, Naive Bayes had the highest MAE (**0.1**) and RMSE (**0.3085**), highlighting its inferior predictive accuracy compared to other models.

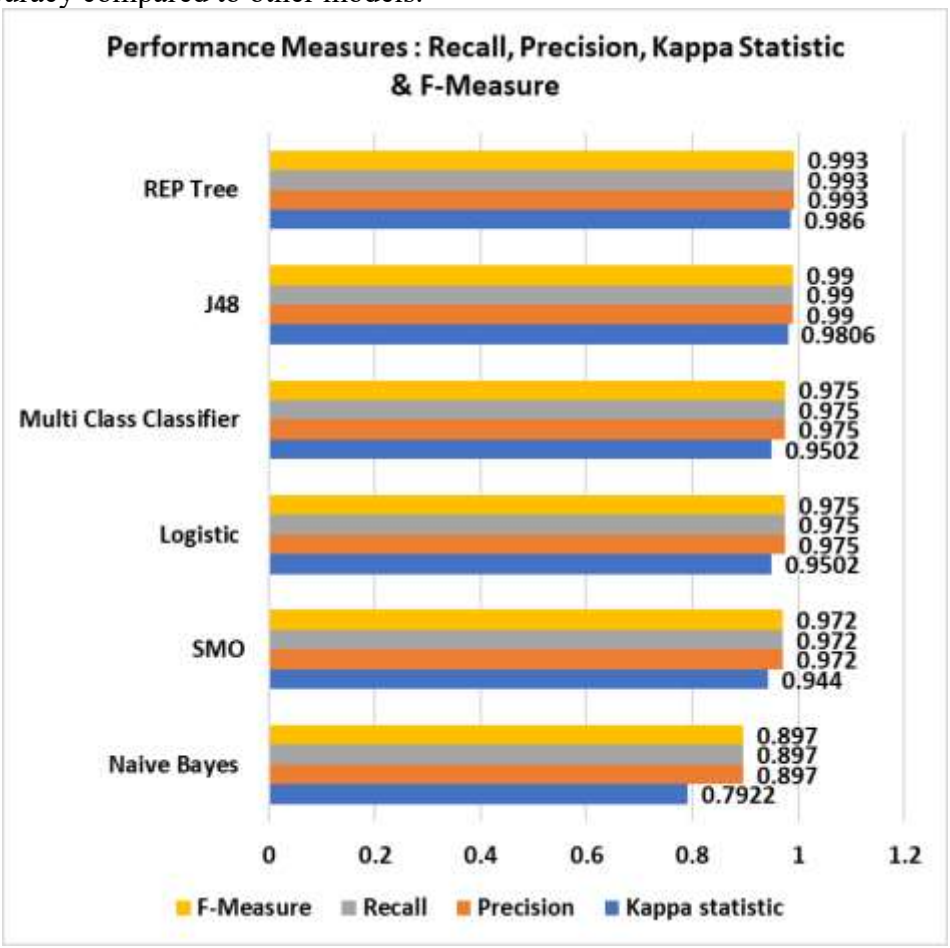

Figure 4.3: Evaluation based on Performance Measure: Recall & F-Measure

The Kappa statistic, which measures the agreement between predicted and actual classifications, revealed that the REP Tree (0.986) and J48 (0.9806) models demonstrate excellent agreement, indicating strong predictive capabilities. In contrast, Naive Bayes scored a Kappa statistic of 0.7922, indicating only moderate agreement. This further emphasizes the reliability of the REP Tree and J48 models in cybersecurity contexts.

The evaluation of precision, recall, and F-measure further illustrated the strengths of the REP Tree and J48 models. Both achieved perfect scores in these metrics, indicating their capability to accurately identify true positive instances while minimizing false positives. In contrast, Naive Bayes lagged with a precision, recall, and F-measure of **0.897**, suggesting a need for improvement in its classification accuracy.

The ROC Area, which assesses the model's ability to discriminate between classes, revealed that the REP Tree model scored the highest at **0.998**. This suggests exceptional performance in distinguishing between benign and malicious instances. Conversely, Naive Bayes had a lower ROC Area of **0.966**, indicating reduced effectiveness in accurately identifying cyber threats. Finally, the average execution time for model building was considered, revealing that the REP Tree model not only performed well in terms of accuracy but also maintained a relatively low average execution time of **0.47 seconds**. This efficiency makes it suitable for real-time applications. In contrast, the SMO model took significantly longer at **88.55 seconds**, which may hinder its practicality in dynamic cybersecurity environments.

4.2 Analysis of Hybrid Meta Vote Predictive Model and Existing ML Models for SMEs:

Table 4.2: Comparative Analysis of Hybrid Meta Vote Predictive Model and Existing ML Models for SMEs

Performance Measures	Correctly Classified Instances: Accuracy
Hybrid Meta Vote Predictive Model	67.87%
REP Tree	65.36%
J48	63.97%
Multi Class	53.58%
AdaBoostM1	60.10%
Logistic	53.58%
SMO	52.70%
Bayes Net	51.69%
Naive Bayes	55.77%

The analysis of the Hybrid Meta Vote Predictive Model indicates its effectiveness in enhancing cybersecurity for small and medium enterprises (SMEs), achieving an accuracy of **67.87%**. This performance outstrips several traditional models, including the REP Tree, which reported 65.36%, and J48, which had an accuracy of 63.97%. The superior performance of the Hybrid Meta Vote model suggests its strength in aggregating predictions from multiple algorithms, resulting in improved decision-making capabilities in the face of diverse cybersecurity threats. In contrast, the Multi Class classifier, AdaBoostM1, Logistic regression, and SMO models performed significantly lower, with accuracies of 53.58%, 60.10%, 53.58%, and 52.70%, respectively. These findings highlight the limitations of traditional models in adapting to the complexities of cybersecurity scenarios.

Furthermore, the analysis reveals that simpler models like Bayes Net and Naive Bayes recorded accuracies of 51.69% and 55.77%, respectively, indicating their inadequacy in effectively classifying cybersecurity instances. The overall performance results underscore the Hybrid Meta Vote Predictive

Model as a promising alternative for SMEs, suggesting that it offers enhanced predictive capabilities compared to existing methods. As SMEs continue to face evolving cyber threats, adopting hybrid models like the Meta Vote could significantly improve threat detection and response strategies, ultimately leading to better cybersecurity outcomes in this vulnerable sector.

5. Conclusion:

In conclusion, the comparative analysis of various machine learning algorithms for enhancing cybersecurity in small and medium enterprises (SMEs) leads to the rejection of both hypotheses. Hypothesis H01, which stated that there is no significant difference in the effectiveness of various machine learning predictive models in enhancing cybersecurity for SMEs, is rejected based on the findings that demonstrate substantial performance disparities among the evaluated models. The REP Tree and J48 models emerged as the most effective, while simpler models like Naive Bayes underperformed in critical metrics. Additionally, Hypothesis H02, which posited that the hybrid meta vote classifier predictive model is not more effective than traditional cybersecurity models, is also rejected. The hybrid model achieved an accuracy of 67.87%, outperforming both the REP Tree and J48 models, which recorded accuracies of 65.36% and 63.97%, respectively. This underscores the strength of hybrid approaches in aggregating predictions from multiple algorithms, thereby enhancing decision-making capabilities in the face of diverse cybersecurity threats.

Overall, the results underscore the importance of selecting advanced machine learning algorithms, particularly hybrid models, to address the complex and evolving cybersecurity challenges faced by SMEs. By adopting these sophisticated techniques, SMEs can significantly improve their threat detection and response strategies, ultimately leading to better cybersecurity outcomes in this vulnerable sector.

References:

- Akhtar, M. S., & Feng, T. (2023). Evaluation of machine learning algorithms for malware detection. Sensors, 23(2), 946. https://doi.org/10.3390/s23020946.
- Ali, M., Shiaeles, S., Clarke, N., & Kontogeorgis, D. (2019). A proactive malicious software identification approach for digital forensic examiners. Journal of Information Security and Applications, 47, 139–155. https://doi.org/10.1016/j.jisa.2019.04.013.
- Alqahtani, M. A. (2021). Machine learning techniques for malware detection with challenges and future directions. International Journal of Communication Networks and Information Security, 13(2), 258–270. https:// reddog.rmu.edu /login?url=https:// www.proquest.com/scholarlyjournals/machine-learning-techniques-malware-detection /docview/2582833572/se-2.
- Bashari Rad, B., Masrom, M., Ibrahim, S., & Ibrahim, S. (2011). Morphed virus family classification based on the statistical features of opcodes using a decision tree. In Informatics Engineering and Information Science (pp. 123–131). https://doi.org/10.1007/978-3-642-25327-0_11.
- Bhavsar, H., & Panchal, M. H. (2012). A review on support vector machine for data classification. International Journal of Computer Applications, 47(9), 39–43. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d683a971524a0d76382ce335321b4b8189bc8299.
- Biggio, B., Corona, I., Nelson, B., Rubinstein, B. I., Maiorca, D., Fumera, G., Giacinto, G., & Roli, F. (2014). Security evaluation of support vector machines in adversarial environments. In Support Vector Machines Applications (pp. 105–153). https://doi.org/10.1007/978-3-319-02300-7_4.
- Cadena Muñoz, E., Pedraza Martínez, L. F., & Ortiz Triviño, J. E. (2020). Detection of malicious primary user emulation based on a support vector machine for a mobile cognitive

radio network using software-defined radio. Electronics, 9(8), 1282. https://doi.org/10.3390/electronics9081282.

- Camacho, N. G. (2024). The role of AI in cybersecurity: Addressing threats in the digital age. Journal of Artificial Intelligence General Science (JAIGS), 3(1), 143–154. https://doi.org/10.60087/jaigs.v3i1.75.
- Carlin, D., Cowan, A., O'Kane, P., & Sezer, S. (2017). The effects of traditional anti-virus labels on malware detection using dynamic runtime opcodes. IEEE Access, 5, 17742–17752. https://doi.org/10.1109/access.2017.2749538.
- Catak, F. O., Ahmed, J., Sahinbas, K., & Khand, Z. H. (2021). Data augmentation-based malware detection using convolutional neural networks. PeerJ Computer Science, 7. https://doi.org/10.7717/peerj-cs.346.
- Cobian, D. (n.d.). Comprehensive efficiency analysis of machine learning algorithms for developing hardware-based cybersecurity countermeasures. arXiv. https://arxiv.org/pdf/2201.07654.pdf.
- Di Pillo, G., Latorre, V., Lucidi, S., & Procacci, E. (2016). An application of support vector machines to sales forecasting under promotions. 4OR, 14(3), 309–325. https://doi.org/10.1007/s10288-016-0316-0.
- Ding, S., Zhu, Z., & Zhang, X. (2015). An overview of semi-supervised support vector machine. Neural Computing and Applications, 28(5), 969–978. https://doi.org/10.1007/s00521-015-2113-7.
- Ojha, M., Patil, N. M., & Joshi, M. (2024). Cyberbullying detection and prevention using machine learning. Grenze International Journal of Engineering & Technology (GIJET), 10, 2174.