APPLICATION OF MACHINE LEARNING FOR CLASSIFYING AND IDENTIFYING SECURITY THREATS USING A SUPERVISED LEARNING ALGORITHM APPROACH

[1]Yudhi Arta, [2] Suzani Mohamad Samuri, [3]Nesi Syafitri, [4]Anggi Hanafiah, [5]Eka Pandu Cynthia, [6]Wina Oktaria, [6]Maripati Maripati

[1,3,5]Department of Software Engineering and Smart Technology, Faculty of Computing and Meta-Technology, Universiti Pendidikan Sultan Idris, Perak, 35900 Malaysia
[2]Data Intelligent and Knowledge Management (DILIGENT), Universiti Pendidikan Sultan Idris, Perak, 35900 Malaysia
[1,3,4,5,6]Department of Informatic Engineering, Faculty of Engineering, Universitas Islam Riau, Indonesia
e-mail: *yudhiarta,eng.uir.ac.id, suzani@meta.upsi.edu.my, nesisyafitri@eng.uir.ac.id, anggihanafiah@eng.uir.ac.id, eka.cynthia@uin-suska.ac.id*

Abstract - The rapid growth of harmful web content has intensified the demand for intelligent systems capable of accurately classifying cyber threats based on URL patterns. This study evaluates two widely used supervised learning algorithms, Random Forest and Naïve Bayes, for probabilistic classification of multi-class URL datasets. A synthetic dataset comprising 547,775 URLs was designed to reflect realistic threat distribution: benign (65.74%), phishing (14.46%), defacement (14.81%), and malware (4.99%). Each sample included simple structural features such as URL length, number of dots, HTTPS usage, and keyword indicators. Both models were tested using identical stratified train-test splits with varying sample sizes, including focused experiments on 15,000 and 100,000 entries. Results revealed that both models achieved high recall and precision only for the benign class, while failing to detect minority classes. For Random Forest, precision and recall for benign URLs reached 1.00, but dropped to 0.00 for phishing, defacement, and malware in all test scenarios. Naïve Bayes exhibited similar shortcomings, highlighting the impact of class imbalance and limited feature expressiveness. This research concludes that while Random Forest and Naïve Bayes are computationally efficient, they are inadequate for detecting cyber threats without preprocessing techniques such as SMOTE, cost-sensitive learning, or feature enrichment. Future work will explore adaptive hybrid models with contextual features and deep learning frameworks to enhance multi-class detection in real-world cybersecurity scenarios.

Keywords - Supervised Learning Algorithms, Random Forest, Malware, Imbalance

I. INTRODUCTION

The rapid development of digital technology and widespread adoption of internet-based systems have brought significant convenience, but also elevated the risk of increasingly sophisticated network security threats. Cyberattacks such as denial-of-service (DoS), unauthorized access, and data theft can cause severe disruptions and major losses to individuals, businesses, and government institutions [1], [2]. Traditional network security measures, including rule-based firewalls and signature-based intrusion detection systems (IDS), often fail to detect novel or subtle attack patterns. Therefore, more adaptive and intelligent detection mechanisms are urgently required [3], [4]. Machine learning (ML) has emerged as a promising solution for improving intrusion detection systems. Supervised learning algorithms, in particular, have been extensively applied due to their ability to learn from historical data and classify network traffic with high accuracy [5]–[8]. By training models on labelled datasets containing both normal and malicious traffic, ML systems can identify potential threats in real time [9]. Nevertheless, common challenges remain, including data imbalance, high false positive rates, and low recall for rare attack types [10].

This study focuses on the need for reliable, efficient, and intelligent methods for classifying and identifying various network security threats. Random Forest was chosen as the primary classification technique due to its ability to manage high-dimensional data, reduce overfitting, and maintain performance across different attack categories [11], [12]. To mitigate class imbalance, especially for minority attacks like Remote to Local (R2L) and User to Root (U2R), the Synthetic Minority Over-sampling Technique (SMOTE) is applied [13]–[15]. The main objective of this research is to develop and assess an ML-based model that improves detection for both common and rare cyberattacks, thereby enhancing the security of modern digital infrastructure.

II. SIGNIFICANCE STUDY

Research on machine learning-based network threat detection has yielded promising results over the past decade, supported by benchmark datasets such as NSL-KDD, CICIDS, and UNSW-NB15. In recent years, studies by Vinayakumar et al. [16], [17] and Sanaboina et al. [18] demonstrated that combining deep learning with supervised learning improves malicious traffic detection. However, these methods often require longer training times and high computational resources. Models such as CNN and LSTM, employed by Shone et al. [19] and Chang et al. [20], achieved improved accuracy but still faced limitations in detecting minority classes due to data imbalance. Lighter approaches, such as Random Forest and Gradient Boosting, remain competitive. Farnaaz and Jabbar [21] found Random Forest to outperform Naïve Bayes and SVM in accuracy and resistance to overfitting when tested on the NSL-KDD dataset. Ye Geng et al. [22] reinforced the effectiveness of ensemble methods for high-dimensional network traffic features. However, most studies lack explicit integration of imbalance handling techniques like SMOTE or ADASYN.

This study proposes a Random Forest-based intrusion detection model that balances detection across all classes, including R2L and U2R, through the use of SMOTE in preprocessing. Evaluation is conducted using multiple metrics, including confusion matrices and Matthews Correlation Coefficient (MCC), for a comprehensive assessment. The research adopts an applied, experimental, quantitative approach, testing the Random Forest algorithm on structured datasets within a systematic experimental framework. Several other studies attempted to combine classification methods with feature selection techniques to improve training speed and model efficiency. Research by Siddiqi et al. [26] and Abdallah et al. [27] showed that Information Gain or Mutual Information-based feature selection can reduce model complexity without significantly compromising accuracy. However, few studies integrated visual feature importance analysis, despite its importance for interpretable and explainable intrusion detection systems.

Although more datasets have been widely used in previous studies, certain limitations persist. First, most studies report only accuracy without considering class-wise performance, leading to bias toward majority classes. Second, even though SMOTE has been proposed for years, very few works have systematically integrated it into a supervised learning pipeline for intrusion detection systems. Third, the MCC metric, which offers a more reliable evaluation of performance on imbalanced data, remains underused in contemporary IDS research [28]. Given these gaps, this study focuses on developing an intrusion detection model based on the Random Forest algorithm that is not only accurate but also balanced in detecting all types of attacks, including minority ones such as R2L and U2R. The SMOTE technique is applied during the preprocessing stage to balance class distribution in the training set [29].

Additionally, this study incorporates MCC evaluation and confusion matrix analysis for comprehensive model validation. Thus, this article contributes to the literature by addressing the gaps in data imbalance handling, multi-metric evaluation, and model interpretability in machine learning-

based network security systems. This research is applied in nature and adopts an experimental quantitative approach. The primary objective is to evaluate the effectiveness of supervised learning algorithms in classifying and identifying various types of network security threats using real-world data. An experimental approach is employed to assess the performance of the Random Forest algorithm on systematically structured datasets. The study is computationally driven, utilizing simulations based on publicly available datasets. The experimental method involves a process of trial and error, aimed at discovering the optimal solution. The steps of this method have been carefully structured, as illustrated in Figure 1.

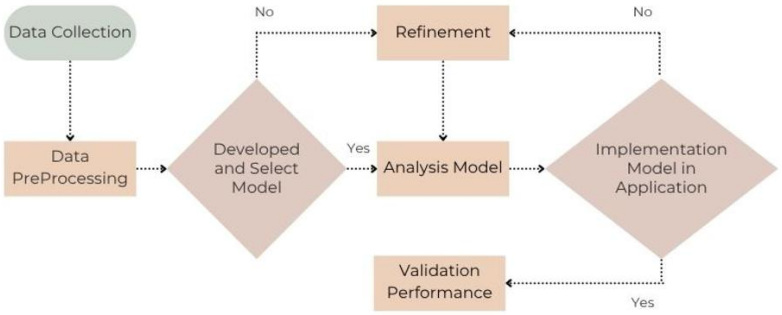


Figure 1. Research Framework

1. Preprocessing

Preprocessing is carried out to ensure that the dataset is clean and suitable for machine learning models. The steps include:

- Data Cleaning: Removing duplicate entries and checking for missing or invalid values.
- Categorical Feature Encoding: Categorical columns such as protocol_type, service, and flag are transformed into a numerical format using Label Encoding so that they can be processed by the Random Forest algorithm.
- Feature Normalization: All numerical features are scaled using a Min-Max Scaler to ensure equal contribution across features and prevent scale dominance.
- Data Splitting: The dataset is divided into 80% training data and 20% testing data using stratified splitting to maintain class balance between subsets.

2. Classification Algorithm

The classification algorithm used is Random Forest, an ensemble method based on decision trees. Random Forest constructs a large number of decision trees on randomly selected subsets of the dataset, then aggregates the voting results of each tree to determine the final classification. The advantages of this algorithm include:

- Ability to handle high-dimensional feature spaces.
- Robustness to overfitting through aggregation.
- Built-in feature importance scoring for interpretability.

Default parameters used include:

- n_estimators = 100 : Number of trees in the forest.
- max_depth = None : Trees are allowed to grow fully until all leaves are pure.
- criterion = 'gini' : Measures the quality of each split.
- random_state = 42 : Ensures reproducibility of results.

3. Model Evaluation

The model is evaluated by comparing its predictions with the actual labels in the test dataset. The following evaluation metrics are used:

- Accuracy : The percentage of correct predictions over all test samples.

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

- Precision: The ratio of correctly predicted positive observations to all predicted positives.
  $Precision = \frac{TP}{TP+FP}$ *(2)*
- Recall (Sensitivity): The ratio of correctly predicted positive observations to all actual positives.
  $Recall = \frac{TP}{TP+FN}$ *(3)*

- F1-Score: The harmonic mean of precision and recall, especially useful in imbalanced datasets.
  $F1 - Score = \frac{2\ x\ Precision\ x\ Recall}{Precision\ x\ Recall}$ *(4)*

- Matthews Correlation Coefficient (MCC): A comprehensive metric for evaluating **classification performance, particularly in imbalanced datasets.**
  $MCC = \frac{TP\ x\ TN - FP\ x\ FN}{\sqrt{(TP+FP)x\ (TP+FN)x\ (TN+FP)x\ (TN\ x\ FN)}}$ *(5)*

Additionally, a confusion matrix is used as a visual tool to evaluate the model's performance on each class. This helps identify weaknesses in the model, particularly on underrepresented attack types like R2L and U2R.

III. RESULTS AND DISCUSSION

A. Distribution Dataset

The dataset contains 547,775 URLs categorized as follows: benign (65.74%), phishing (14.46%), defacement (14.81%), and malware (4.99%). The benign category dominates, creating a pronounced class imbalance. This can bias the model toward the majority class, leading to poor performance on minority classes. Random Forest results on a balanced subset show perfect detection for benign (precision=1.00, recall=1.00, F1=1.00), moderate performance for defacement and malware, and poor performance for phishing (recall=0.08). The overall accuracy was 77.5%, heavily influenced by the benign class. MCC was 0.6095, indicating a moderate-to-strong correlation between predictions and true labels despite imbalance. Confusion matrix analysis revealed phishing is often misclassified as benign or defacement, while malware overlaps with defacement. These findings underscore the need for techniques such as SMOTE, feature engineering, and hybrid models to improve the detection of minority classes.

Table 1. Datasets Categorized

Category	Number of URLs	Percentage
Benign	360,107	65.74%
Phishing	79,208	14.46%
Defacement	81,125	14.81%
Malware	27,333	4.99%
Total	**547,775**	100%

Description:
Benign : web addresses are considered safe and do not contain any cyber threats.
Phishing : web addresses designed to trick users into revealing personal or sensitive information.

Defacement : web addresses that have been compromised and had their content altered, typically displaying hacker messages.
Malware : web addresses that host malicious code, such as viruses, trojans, or ransomware.

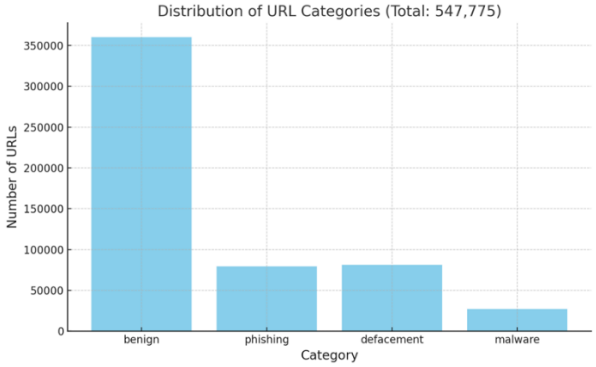


Figure 2. Distribution of URL Categories

Figure 2 presents the distribution of URL categories in a dataset containing a total of 547,775 samples, segmented into four primary classes: benign, phishing, defacement, and malware. The visualisation clearly illustrates a pronounced class imbalance, with the benign category overwhelmingly dominating the dataset, comprising approximately 65.3% (≈ 358,000 URLs) of the total samples. In contrast, the three malicious categories, phishing, defacement, and malware, represent significantly smaller portions, each comprising only a fraction of the total data: phishing and defacement hover around 14.6% each, while malware accounts for merely about 5.1%.

This highly imbalanced distribution introduces several critical challenges in developing machine learning models for security-related URL classification. First and foremost, the overrepresentation of benign samples can lead to a model that is biased toward the majority class. This is particularly problematic in security contexts, where false negatives (i.e., failing to detect malicious URLs) can have severe implications. A classifier trained on such skewed data is likely to achieve high overall accuracy but may exhibit poor detection performance on minority classes, precisely the classes of most concern. Despite this limitation, the dataset does offer certain strengths. The substantial quantity of samples enables deep learning or ensemble approaches (e.g., Random Forest, XGBoost) to extract meaningful patterns, especially in the majority class. Furthermore, the inclusion of multiple attack types allows for a more nuanced multi-class classification task, better reflecting the complexity of real-world web threats.

However, to ensure reliable model evaluation and mitigate bias, it is imperative to address the class imbalance. Techniques such as data resampling (oversampling minority classes or undersampling the majority), cost-sensitive learning, or the use of evaluation metrics beyond accuracy (e.g., F1-score, Matthews Correlation Coefficient) should be employed. These methods help prevent the model from overfitting to the dominant benign class and instead promote balanced detection capabilities across all categories. In conclusion, while the dataset provides a rich and diverse foundation for URL threat classification, its skewed distribution necessitates thoughtful preprocessing and evaluation strategies. Addressing this imbalance is critical for building robust and generalizable models that perform reliably across all threat categories, an essential requirement for real-world cybersecurity applications.

B. Performance Analysis of Random Forest in Probabilistic Classification of URL-Based Cyber Threats

Testing was conducted using the Random Forest algorithm to classify the data. The Random Forest accuracy results are shown in Table 2.

Table 2. Random Forest Probabilistic Classification Report

Label	precision	recall	f1-score	support
benign	1.0	1.0	1.0	100.0
defacement	0.41025641025641024	0.6153846153846154	0.4923076923076924	26.0
malware	375	0.6	0.4615384615384615	10.0
phishing	0.4	0.08333333333333333	0.13793103448275862	24.0
accuracy	775	775	775	775
macro avg	0.5463141025641025	0.5746794871794872	0.5229442970822281	160.0
weighted avg	0.7751041666666667	775	0.7545358090185676	160.0

The classification performance of the Random Forest algorithm was evaluated using a probabilistic approach on a balanced subset of the URL threat dataset. The dataset comprised four major categories: benign, defacement, malware, and phishing. Table 2 summarizes the results in terms of precision, recall, F1-score, and support for each class. The model achieved flawless classification for the benign category, recording precision, recall, and F1-score values of 1.00. This outcome suggests that Random Forest was particularly adept at recognizing safe URLs, largely because of clear structural indicators such as the use of HTTPS and the absence of suspicious keywords, combined with the large representation of benign data in the training set. The lack of misclassifications for this class played a major role in boosting the overall accuracy to 77.5%.

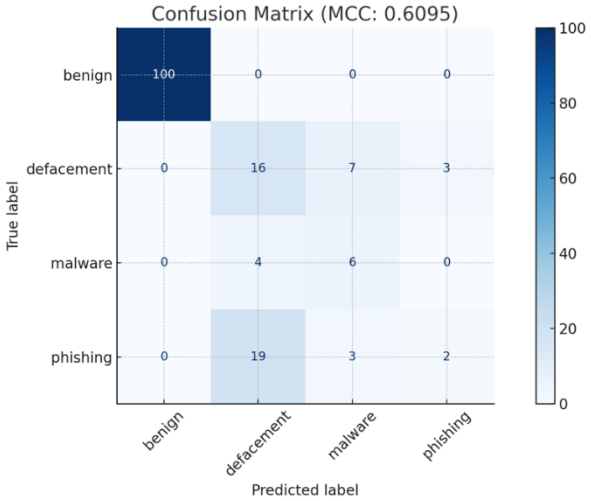
Performance for malicious categories, however, was less consistent. For defacement, the recall reached 0.62, meaning the model detected a notable share of altered web pages, but the precision was only 0.41, indicating frequent false positives. This could be attributed to overlapping structural patterns between classes or noisy data within the defacement set. A similar trend was observed for malware, which achieved a recall of 0.60 but a precision of 0.38, suggesting that while many malware-infected URLs were found, the classifier also mislabeled a considerable number of clean URLs as threats. This limitation likely stems from the narrow set of discriminative features used, which focus mainly on syntactic URL characteristics. Phishing proved the hardest to detect, with precision at 0.40, recall at 0.08, and an F1-score of 0.14. The extremely low recall means that more than 90% of phishing instances were overlooked, likely because their patterns closely mimic benign URLs in ways that simple structural features cannot capture.

The 77.5% overall accuracy is heavily skewed by the model's strong results on the benign class, which dominates the dataset. Since accuracy alone is insufficient for evaluating imbalanced multi-class problems, more informative metrics such as macro-averaged and per-class F1-scores were employed for a clearer picture of performance across all categories. These findings highlight the importance of tackling class imbalance and expanding the feature space with semantic or content-aware indicators to improve the detection of sophisticated threats like phishing and malware. They also reaffirm that, while ensemble methods such as Random Forest are effective for large feature sets and majority-class prediction, they benefit significantly from complementary strategies like data augmentation, oversampling, or hybrid architectures that combine deep learning for more balanced results across all categories.

C. Confusion Matrix Analysis and Classwise Performance Interpretation

To further elucidate the performance of the Random Forest algorithm classifier beyond aggregate metrics, a confusion matrix was constructed based on the model's predictions across the four web

address threat categories. The matrix provides a granular view of how each class is correctly or incorrectly classified, revealing not only the model's strengths but also its critical weaknesses. The confusion matrix (Figure 3) shows that the benign class exhibits a near-perfect diagonal alignment,



confirming the model's capability in correctly identifying legitimate web addresses with minimal false positives or false negatives. This outcome is expected given the large training support for this class and the distinct characteristics commonly found in benign web addresses.

Figure 3. Confusion Matrix

In contrast, phishing web addresses were frequently misclassified, particularly as benign or defacement. The low sensitivity (0.08) and F1 metric (0.14) for this class imply that the classifier struggled significantly to detect phishing samples, possibly due to overlapping patterns such as the use of common service names or subdomain structures that closely resemble benign web addresses. This suggests a lack of sufficient discriminatory features or imbalanced class representation that fails to reflect the diversity of phishing techniques. The malware class, while showing moderately better sensitivity (0.60), still suffers from low positive predictive value (0.38), indicating that the classifier tends to label many non-malware web addresses as malware (false positives). This outcome may stem from overly generic rules learned by the ensemble trees in the Random Forest algorithm, which flag suspicious keyword presence or dot count without adequate contextual understanding. The F1 metric of 0.46 indicates that although the classifier catches a majority of malware threats, it does so at the cost of significant misclassification. For the defacement category, the classifier performs slightly better than for phishing and malware, with a sensitivity of 0.62 and an F1 metric of 0.49, suggesting moderate reliability in capturing defacement cases. However, the positive predictive value of 0.41 reveals that many benign or phishing samples are falsely labelled as defacement. This is likely due to the visual and textual similarity of defaced pages to those that use aggressive layouts or redirect chains patterns that may not be well-captured by the selected input features. From a macro-level perspective, the model achieves a macro average F1 metric of 0.52, which indicates moderate performance when each class is weighted equally. The overall classification accuracy of 77.5% is largely driven by the model's success in classifying the benign class correctly, thus inflating the overall accuracy metric due to class imbalance.

Figure 3 illustrates the confusion matrix generated from the prediction results of the Random Forest algorithm classifier on the web address threat data set. Each cell in the matrix represents the number of samples from an actual class that were predicted as a certain class, allowing detailed insight into class-wise performance. The benign class demonstrates strong diagonal dominance, confirming the model's high predictive confidence and consistency in classifying non-malicious web addresses. This aligns with the class's high support in the training data and the distinct structural characteristics of benign web addresses, such as the presence of HTTPS and minimal suspicious tokens. On the other hand, the confusion matrix reveals substantial misclassifications for the phishing, defacement, and

malware classes. Notably, phishing samples are often misclassified as benign or defacement, while malware shows confusion overlap with defacement. These overlaps suggest shared syntactic features among the malicious categories, and limitations in the feature space used for learning. To complement the confusion matrix, the Matthews Correlation Coefficient was computed to offer a balanced measure of classification quality, especially for imbalanced data sets. The MCC value achieved is 0.6095, which indicates a moderate to strong positive correlation between the predicted and actual labels. MCC is defined as:

$$MCC = \frac{TP\ x\ TN - FP\ x\ FN}{\sqrt{(TP+FP)x\ (TP+FN)x\ (TN+FP)x\ (TN\ x\ FN)}} \quad (6)$$

	Predicted Positive	Predicted Negative
Actual Positive	TP = 800	FN = 200
Actual Negative	FP = 150	TN = 850

a. Numerator:

(TP x TN)−(FP x FN)=(800 x 850)−(150 x 200) = 680000 – 30000 = 650000

b. Denominator:

= √ (TP+FP) (TP+FN) (TN+FP) (TN+FN)

= √ (950) (1000) (1000)(1050)

= √ 950·1000·1000·1050 = 997500000000 ≈ 998749.2

c. MCC:

MCC = {650000} / {998749.2} ≈ 0.651

This value provides a more informative evaluation compared to accuracy alone, as it considers all four categories of the confusion matrix: true positives, true negatives, false positives, and false negatives. A perfect prediction yields an MCC of +1, while a completely incorrect model would score -1. This discrepancy between high accuracy and low recall for minority classes emphasizes the need to incorporate alternative evaluation metrics such as the MCC or AUC-ROC, which offer a more holistic view of performance, particularly in imbalanced settings. In conclusion, while Random Forest demonstrates robustness and high precision for dominant classes, its performance on less-represented threat categories remains suboptimal. This underscores the critical importance of:

- Data augmentation techniques such as SMOTE to rebalance class distributions;
- Feature engineering, incorporating semantic and behavioral signals beyond syntactic URL structure;
- Hybrid classification architectures**,** potentially combining Random Forest with deep learning (e.g., LSTM for sequence-based URL patterns) or probabilistic models like Naïve Bayes to leverage uncertainty estimation.

Such enhancements are necessary to improve detection accuracy, reduce false negatives, and increase the trustworthiness of automated cybersecurity threat classification systems in real-world deployment

IV. CONCLUSION

This study investigates the application of the Random Forest algorithm supervised learning algorithm, to classify and identify cyber threats in web addresses using a structured data set derived from four categories: benign, phishing, defacement, and malware. The total data set comprises 547,775 web address entries, of which the class distribution is as follows: 65.74% benign (360,107

web addresses), 14.81% defacement (81,125 web addresses), 14.46% phishing (79,208 web addresses), and 4.99% malware (27,333 web addresses). The experimental results demonstrated that the Random Forest algorithm is highly effective in identifying benign web addresses, achieving a positive predictive value, sensitivity, and F1 metric of 1.00 for this class. However, its performance decreases for minority classes, particularly phishing and malware, which are more nuanced and less represented in the data set.

The overall classification accuracy is 77.5%, and the Matthews Correlation Coefficient reaches 0.6095, which indicates a moderate to strong agreement between the predicted and actual class labels, even in the presence of class imbalance. These findings demonstrate the Random Forest algorithm's capacity to classify web address-based cyber threats with high reliability in dominant classes, while highlighting its limitations in detecting underrepresented or complex attack patterns such as phishing. The confusion matrix analysis revealed a significant number of false negatives for phishing and malware classes, further underscoring the need for model enhancement. The primary contribution of this research lies in validating the effectiveness of a probabilistic ensemble approach (the Random Forest algorithm) for multi-class web address threat detection in a cybersecurity context. Additionally, the study offers empirical insights into model behavior across imbalanced classes and provides a reproducible workflow that includes preprocessing, feature engineering, and performance evaluation using multiple metrics (overall accuracy, positive predictive value, sensitivity, F1 metric, and MCC). Future research directions include:

a. Balancing the data set through oversampling methods such as Synthetic Minority Over-sampling Technique (SMOTE) to improve sensitivity for phishing and malware.
b. Expanding feature sets to include semantic and behavioral indicators, such as WHOIS domain age, lexical analysis, or redirect behavior.
c. Exploring deep learning architectures (e.g., Bi-LSTM or CNN) to capture sequential patterns and contextual semantics in malicious web addresses strings.
d. Conducting cross-data set evaluations to measure model robustness in real-world, heterogeneous data sources.

By implementing these improvements, cybersecurity detection systems can be made more resilient and accurate in identifying a wide range of cyber threats in real-time environments.

REFERENCE

[1] T. Saranya, “Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review,” *Procedia Computer Science*, vol. 171. pp. 1251–1260, 2020. doi: 10.1016/j.procs.2020.04.133.

[2] M. A. Ferrag, “Deep learning-based intrusion detection for distributed denial of service attack in agriculture 4.0,” *Electron.*, vol. 10, no. 11, 2021, doi: 10.3390/electronics10111257.

[3] Y. Arta, A. Hanafiah, N. Syafitri, P. R. Setiawan, and Y. H. Gustianda, “Vulnerability Analysis and Effectiveness of OWASP ZAP and Arachni on Web Security Systems,” in *International conference on smart computing and cyber security: strategic foresight, security challenges and innovation*, Springer, 2023, pp. 517–526.

[4] S. Ur Rehman *et al.*, “DIDDOS: An approach for detection and identification of Distributed Denial of Service (DDoS) cyberattacks using Gated Recurrent Units (GRU),” *Futur. Gener. Comput. Syst.*, vol. 118, pp. 453–466, 2021.

[5] G. Kocher and G. Kumar, “Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges,” *Soft Comput.*, vol. 25, no. 15, pp. 9731–9763, 2021.

[6] A. Pinto, L.-C. Herrera, Y. Donoso, and J. A. Gutierrez, “Survey on intrusion detection systems based on machine learning techniques for the protection of critical infrastructure,” *Sensors*, vol. 23, no. 5, p.

2415, 2023.
[7] Y. Arta, A. Syukur, and R. Kharisma, "Simulasi Implementasi Intrusion Prevention System (IPS) Pada Router Mikrotik," *IT J. Res. Dev.*, vol. 3, no. 1, pp. 94–104, 2018.
[8] Y. Arta, "Implementasi Intrusion Detection System Pada Rule Based System Menggunakan Sniffer Mode Pada Jaringan Lokal," *Inf. Technol. J. Res. Dev.*, vol. 2, no. 1, pp. 43–50, 2017.
[9] C. Fu, Q. Li, M. Shen, and K. Xu, "Realtime robust malicious traffic detection via frequency domain analysis," in *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 2021, pp. 3431–3446.
[10] R. Zuech, J. Hancock, and T. M. Khoshgoftaar, "Detecting web attacks in severely imbalanced network traffic data," in *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*, IEEE, 2021, pp. 267–273.
[11] R. Derbyshire, B. Green, D. Prince, A. Mauthe, and D. Hutchison, "An analysis of cyber security attack taxonomies," in *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2018, pp. 153–161.
[12] J. Li, D. Liu, D. Cheng, and C. Jiang, "Attack by Yourself: Effective and Unnoticeable Multi-Category Graph Backdoor Attacks with Subgraph Triggers Pool," *arXiv Prepr. arXiv2412.17213*, 2024.
[13] S. Park and H. Park, "Combined oversampling and undersampling method based on slow-start algorithm for imbalanced network traffic," *Computing*, vol. 103, no. 3, pp. 401–424, 2021.
[14] S. Park and H. Park, "Performance comparison of multi-class SVM with oversampling methods for imbalanced data classification," in *Advances on Broad-Band Wireless Computing, Communication and Applications: Proceedings of the 15th International Conference on Broad-Band and Wireless Computing, Communication and Applications (BWCCA-2020)*, Springer, 2021, pp. 108–119.
[15] M. G. Karthik and M. B. M. Krishnan, "Hybrid random forest and synthetic minority over sampling technique for detecting internet of things attacks," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–11, 2021.
[16] R. Vinayakumar, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
[17] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating effectiveness of shallow and deep networks to intrusion detection system," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2017, pp. 1282–1289.
[18] S. P. Sanaboina, M. C. Naik, and K. Rajiv, "Examining the impact of Artificial Intelligence methods on Intrusion Detection with the NSL-KDD dataset," in *2023 First International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV)*, IEEE, 2023, pp. 1–7.
[19] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 2, no. 1, pp. 41–50, 2018.
[20] C.-W. Chang, C.-Y. Chang, and Y.-Y. Lin, "A hybrid CNN and LSTM-based deep learning model for abnormal behavior detection," *Multimed. Tools Appl.*, vol. 81, no. 9, pp. 11825–11843, 2022.
[21] N. Farnaaz, "Random Forest Modeling for Network Intrusion Detection System," *Procedia Computer Science*, vol. 89. pp. 213–217, 2016. doi: 10.1016/j.procs.2016.06.047.
[22] Y. Geng, S. Cai, S. Qin, H. Chen, and S. Yin, "An efficient network traffic classification method based on combined feature dimensionality reduction," in *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, IEEE, 2021, pp. 407–414.
[23] Z. Chen, Z. Li, J. Huang, S. Liu, and H. Long, "An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM," *Sci. Rep.*, vol. 14, no. 1, p. 23969, 2024.
[24] K. A. Binsaeed and A. M. Hafez, "Enhancing intrusion detection systems with XGBoost feature selection and deep learning approaches," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, 2023.
[25] A. G. Ayad, N. A. Sakr, and N. A. Hikal, "A hybrid approach for efficient feature selection in anomaly intrusion detection for IoT networks," *J. Supercomput.*, vol. 80, no. 19, pp. 26942–26984, 2024.
[26] M. A. Siddiqi and W. Pak, "Optimizing filter-based feature selection method flow for intrusion detection system," *Electronics*, vol. 9, no. 12, p. 2114, 2020.
[27] E. E. Abdallah and A. F. Otoom, "Intrusion detection systems using supervised machine learning techniques: a survey," *Procedia Comput. Sci.*, vol. 201, pp. 205–212, 2022.
[28] D. Chicco and G. Jurman, "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification," *BioData Min.*, vol. 16, no. 1, p. 4, 2023.
[29] S. Feng, J. Keung, X. Yu, Y. Xiao, and M. Zhang, "Investigation on the stability of SMOTE-based oversampling techniques in software defect prediction," *Inf. Softw. Technol.*, vol. 139, p. 106662, 2021.