URL-BASED PHISHING DETECTION USING HYBRID ENSEMBLE TECHNIQUE

H. Pannirchelvam[1], N. Mohd Salleh[1]
M. F. Abdollah[1], S. R. Selamat[1] and A. L. Amir[2]

[1]Fakulti Kecerdasan Buatan dan Keselamatan Siber,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

[2]Faculty of Computer and Mathematical Sciences,
Universiti Teknologi Mara, Cawangan Melaka, Kampus Jasin, Jalan Lembah Kesang 1/1-2, Kampung Seri Mendapat, 77300 Melimau, Melaka, Malaysia.

Corresponding Author's Email: [1]nurhashikin@utem.edu.my

Article History: Received 25 November 2025; Revised 16 December 2025; Accepted 17 December 2025

ABSTRACT: Phishing attacks pose a significant and growing threat to cybersecurity by deceiving users into disclosing sensitive information through malicious websites. Most existing URL-based phishing detection studies rely on individual classifiers or traditional ensemble techniques, which often struggle to generalize against evolving phishing patterns. To overcome this limitation, this study proposes a hybrid ensemble learning approach for URL-based phishing detection by integrating a Random Forest Classifier with AdaBoost, Bagging, and Stacking strategies. Experiments were conducted using a publicly available benchmark dataset from the UCI Machine Learning Repository consisting of 11,055 URLs and 30 features. Model performance was evaluated using 10-fold cross-validation. The results show that the Random Forest–Stacking hybrid model achieved the highest accuracy of 97.21%, outperforming other hybrid configurations. The findings demonstrate that stacking-based hybrid ensemble learning enhances generalization and robustness in phishing URL detection.

KEYWORDS: *Phishing Attacks; Hybrid Ensemble Technique, Uniform Resource Locator (URL), WEKA (Waikato Environment for Knowledge Analysis).*

1.0 INTRODUCTION

The rise of internet usage has brought about significant advancements and conveniences, but it has also given rise to various cyber threats, with phishing attacks being one of the most pervasive. Phishing attacks deceive users into divulging sensitive information, such as passwords and credit card numbers, by masquerading as legitimate websites. This threat undermines trust in online transactions and poses serious risks to both individuals and organizations. In the field of cybersecurity, the detection of phishing websites is critical to safeguarding user data and maintaining the integrity of digital interactions. Traditional methods for detecting phishing websites, such as blacklists, heuristics, and individual machine learning classifiers, have been employed with varying degrees of success. However, the dynamic and evolving nature of phishing tactics often renders these methods insufficient. This necessitates the exploration of more advanced detection techniques.

Although the widespread adoption of machine learning methods for phishing detection, many existing approaches depend on individual classifiers or basic ensemble methods, which may be inadequate for capturing complex and evolving URL manipulation patterns. While traditional ensemble techniques such as bagging and boosting improve performance by reducing variance or bias, they fail to fully leverage the complementary strengths of multiple classifiers. Hybrid ensemble learning, which integrates ensemble strategies through meta-learning, has received limited attention in URL-based phishing detection. This study aims to bridge this gap by conducting a systematic evaluation of hybrid ensemble combinations using Random Forest as a base model combined with AdaBoost, Bagging, and Stacking. The goal is to determine the most effective hybrid configuration for enhancing detection accuracy and generalization performance in URL-based phishing classification.

The rest of this paper is structured as follows. Section II is the related work of the previous literature review. Section III outlines the methodology of this study. Section IV presents the results from the study, while Section V discusses implementation and implications of the results, as well as the limitations of the study. Section VI concludes the paper and provides several suggestions for future research.

2.0 RELATED WORK

2.1 Phishing Attack

The term phishing attack is becoming growing common issue nowadays, yet a strong definition for it and what it creates is subjective and encloses a broad area. Phishing is an online scam where scammers send malicious links to user accounts to obtain private, sensitive, and financial data [1]. While according to [2], phishing is the term for online frauds that employ websites, social media, and email to get personal information. [3] said phishing is one of the most popular ways to gain access to a website. This tactic entails deceiving particular people or companies into revealing confidential information. It has been supported by [4] claimed phishing is a fraudulent tactic that takes use of technological and social tactics to get bank credentials and customer identity. Its aims to expose the weakness of the users by sending them fake websites in an attempt to obtain their sensitive or private information [3]. Characteristics of phishing attacks introduce common principles, signs or indicators that can guide to determine a phishing attack. [5] stated the characteristics of phishing attacks consist of indirect greeting, text wording, personal information, urgency, irresistible offers, domain, link and attachments.

2.2 URL-Based Phishing Detection Using Machine Learning

URL-based phishing detection has been widely studied due to its efficiency and independence from webpage content, which allows for faster detection and reduced computational overhead [6]. Early studies primarily relied on traditional machine learning classifiers such as Support Vector Machines (SVM), Decision Trees, Naïve Bayes, and k-Nearest Neighbors (k-NN) to classify URLs as legitimate or phishing [7]. These approaches typically utilize lexical and structural URL features, including URL length, the presence of special characters, abnormal subdomains, and suspicious domain patterns [7]. While these models demonstrated reasonable detection accuracy, their performance often degrades when confronted with sophisticated or previously unseen phishing URLs. Table 1 summarizes representative studies on URL-based phishing detection using machine learning techniques. The

comparison highlights the datasets employed, the best-performing algorithms, and reported classification accuracy. Although high accuracy values are reported across different studies, most approaches rely on individual classifiers or single ensemble methods, with limited exploration of hybrid ensemble configurations.

Table 1: Phishing detection techniques using machine learning

Author	Dataset used	Best algorithm	Accuracy
[7]	UCI Machine Learning Repository	Pruned Decision Tree	91.5%
[1]	CatchPhish_D3	Random Forest	83%
[8]	Kaggle dataset (Public)	Ensemble model	96%
[9]	i. Kaggle dataset for benign Websites URLs ii. Phishtank dataset for phishing website URLs	XGBoost	93.73%
[10]	UCI machine learning repository	K-NN+RFC	97.33%

As shown in Table 1, reported phishing detection accuracies span from 83% to 97.33% across existing studies. Ensemble approaches consistently demonstrate better performance than individual classifiers, reflecting their effectiveness in handling sophisticated phishing patterns. However, inconsistencies in datasets and evaluation methodologies limit fair comparison, while systematic analysis of hybrid ensemble models remains limited.

2.3 Ensemble and Hybrid Learning Approaches for Phishing Detection

Ensemble learning techniques such as Bagging, Boosting, and Random Forest have been extensively applied in phishing detection in order to enhance generalization performance [8], [9], [10]. Bagging mitigates model variance by training multiple classifiers on different subsets of the data, while Boosting emphasizes misclassified instances to reduce bias [9], [10]. Although these techniques enhance performance compared to standalone classifiers, they remain limited in capturing complex interactions among diverse model predictions [10]. Recent

studies have explored hybrid ensemble approaches that combine multiple classifiers or ensemble methods to improve detection accuracy [9], [11]. Some works integrate Random Forest with other classifiers such as k-NN or neural networks, demonstrating performance improvements over individual models. Nevertheless, many of these studies lack systematic evaluation of different hybrid configurations and often focus solely on accuracy without deeper analysis of ensemble integration strategies [8], [10].

Stacking-based ensemble learning, which employs a meta-learner to optimally combine predictions from base classifiers, has shown promise in other cybersecurity domains [11]. Nevertheless, its application to URL-based phishing detection remains limited. Existing studies rarely compare stacking-based hybrids with other ensemble approaches such as Bagging and Boosting under consistent experimental settings [9], [11]. Despite the extensive use of machine learning and ensemble techniques for phishing detection, several research gaps remain. First, most existing studies focus on individual classifiers or single ensemble methods, with limited exploration of hybrid ensemble configurations [8], [10]. Second, there is a lack of comprehensive comparative analysis that evaluates different ensemble integration mechanisms such as Bagging, Boosting, and Stacking—using the same dataset and evaluation metrics [9], [11]. Third, limited attention has been devoted to explaining why certain ensemble strategies outperform others in the context of URL-based phishing detection [11]. To address these gaps, this study systematically evaluates hybrid ensemble learning approaches by integrating Random Forest as a base model with AdaBoost, Bagging, and Stacking techniques. By conducting a comparative analysis, this work aims to provide deeper insights into the effectiveness of hybrid ensemble strategies for improving phishing URL detection performance.

3.0 METHODOLOGY

In the detection of phishing attack using a hybrid ensemble technique, this project proposedthe framework of detection process as shown in Figure 1.

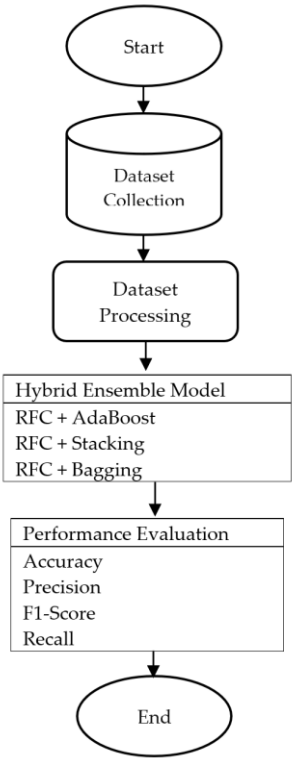


Figure 1: Proposed framework of phishing detection

The details of each process will be explained in the following subsections.

3.1 Dataset Collection

The dataset used in this study was obtained from the UCI Machine Learning Repository and consists of 11,055 URL samples with 4,898 legitimate URLs and 6,157 phishing URLs. Each sample includes 30 manual features representing lexical and structural characteristics of URLs [12]. The dataset shows moderate class imbalance, which was addressed through robust ensemble learning rather than resampling techniques.

3.2 Data Preprocessing

Data preprocessing involved checking for missing values and ensuring all features were in a numerical format suitable for classification. Since the dataset had no missing value, no imputation was required. All features were retained to preserve discriminative information. No feature scaling was applied, as tree-based models are generally insensitive to the magnitude of features.

3.3 Hybrid Ensemble Model

After data preprocessing, hybrid ensemble learning models were developed to enhance phishing URL detection performance. Random Forest Classifier (RFC) was selected as the base model due to its robustness, ability to handle high-dimensional feature spaces, and resistance to overfitting. Three hybrid ensemble configurations were evaluated by integrating RFC with AdaBoost, Bagging, and Stacking techniques. In the AdaBoost-based hybrid model, RFC was employed as the base classifier to iteratively emphasize misclassified instances and reduce model bias. In the Bagging-based hybrid model, multiple RFC models were trained on bootstrap samples to reduce variance and improve stability. For the Stacking-based hybrid model, predictions from multiple RFC-based learners were combined using a meta-learner to capture higher-level relationships among model outputs. All models were evaluated using 10-fold cross-validation under consistent experimental conditions to ensure a fair comparison. The hybrid ensemble framework was designed to systematically evaluate the effectiveness of different ensemble integration strategies for URL-based phishing detection.

3.4 Performance Evaluation

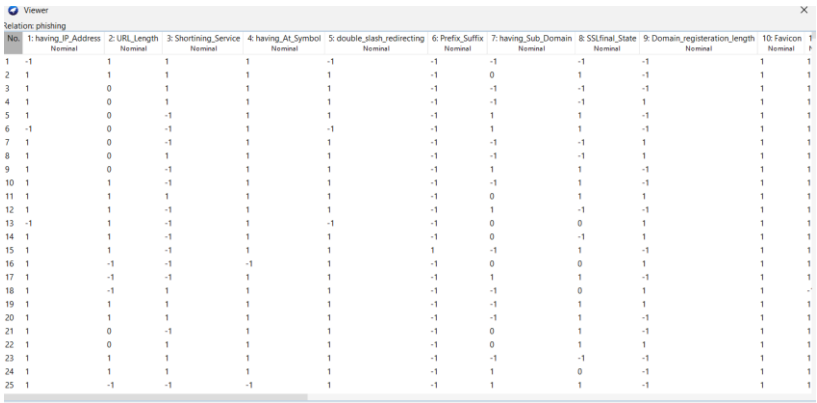
Model performance was evaluated using 10-fold cross-validation to ensure reliable and unbiased estimation. This approach ensures that each instance contributes to both training and testing, minimizing potential performance bias. Classification performance was evaluated using accuracy, precision, recall, and F1-score.

4.0 RESULT & ANALYSIS

The implementation for data collection, data preprocessing, and implementing hybrid ensemble model and performance evaluation are all covered in this chapter. The evaluation of model accuracy will be measured for this project to get the expected output.

4.1 Dataset Collection

In this phase, the dataset has been downloaded from the UCI Machine Learning Repository. The dataset has been taken as an input and identify features. It contains 11,055 instances and 30 unique features. This is a publicly available dataset that has been used in past research. Figure 2 below shows the training dataset in WEKA viewer.


Viewer

Relation: phishing

No.	1: having_IP_Address Nominal	2: URL_Length Nominal	3: Shortining_Service Nominal	4: having_At_Symbol Nominal	5: double_slash_redirecting Nominal	6: Prefix_Suffix Nominal	7: having_Sub_Domain Nominal	8: SSLfinal_State Nominal	9: Domain_registeration_length Nominal	10: Favicon Nominal	1
1	-1	1	1	1	-1	-1	-1	-1	-1	1	1
2	1	1	1	1	1	-1	0	1	-1	1	1
3	1	0	1	1	1	-1	-1	-1	-1	1	1
4	1	0	1	1	1	-1	-1	-1	1	1	1
5	1	0	-1	1	1	-1	1	1	-1	1	1
6	-1	0	-1	1	-1	-1	1	1	-1	1	1
7	1	0	-1	1	1	-1	-1	-1	1	1	1
8	1	0	1	1	1	-1	-1	-1	1	1	1
9	1	0	-1	1	1	-1	1	1	-1	1	1
10	1	1	-1	1	1	-1	-1	1	-1	1	1
11	1	1	1	1	1	-1	0	1	1	1	1
12	1	1	-1	1	1	-1	1	-1	-1	1	1
13	-1	1	-1	1	-1	-1	0	0	1	1	1
14	1	1	-1	1	1	-1	0	-1	1	1	1
15	1	1	-1	1	1	1	-1	1	-1	1	1
16	1	-1	-1	-1	1	-1	0	0	1	1	1
17	1	-1	-1	1	1	-1	1	1	-1	1	1
18	1	-1	1	1	1	-1	-1	0	1	1	-1
19	1	1	1	1	1	-1	-1	1	1	1	1
20	1	1	1	1	1	-1	-1	1	-1	1	1
21	1	0	-1	1	1	-1	0	1	-1	1	1
22	1	0	1	1	1	-1	0	1	1	1	1
23	1	1	1	1	1	-1	-1	-1	-1	1	1
24	1	1	1	1	1	-1	1	0	-1	1	1
25	1	-1	-1	-1	1	-1	1	1	-1	1	1

Figure 2: Training dataset

4.2 Data Preprocessing

In this step, the dataset has been divided into two classification which is legitimate and phishing URLs. In this project, it has 4898 legitimate samples and 6157 phishing samples.

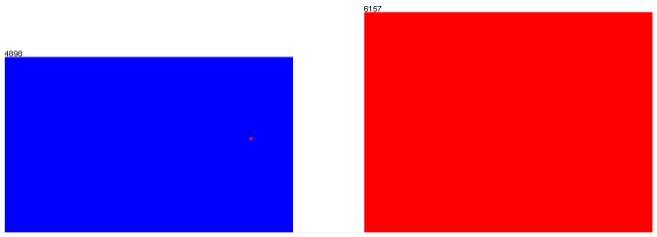


Figure 3: Legitimate and phishing sample

Figure 3 shows 4898 legitimate sample and 6157 phishing sample. After that, make sure to check the features to see whether it has any missing data. If no missing data has been found, then proceed to the next step.

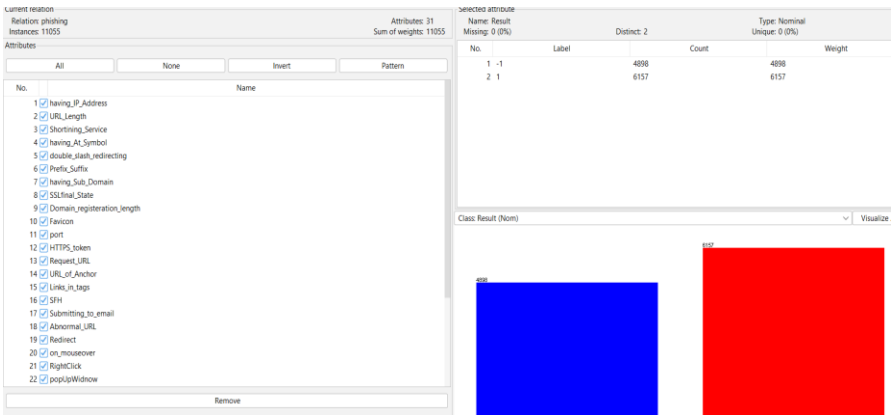


Figure 4: Preprocess tab in WEKA

Figure 4 shows the preprocess tab in WEKA. This is place where handling of missing data, encoding of categorial data and any removal of unnecessary features. In this dataset, there is no missing data and no removal of the attribute because we are using all the features available.

4.3 Hybrid Ensemble Model

After performing data preprocessing, the implementation of the

combination of the ensemble classifiers. Before we begin the implantation of the classifier, we must split the dataset into training and testing set. Figure 5 shows a model evaluation that was conducted using 10-fold cross-validation to ensure reliable and unbiased performance estimation. Cross-validation allows each data instance to be used for both training and testing, which is particularly suitable for phishing detection datasets where generalization performance is critical.

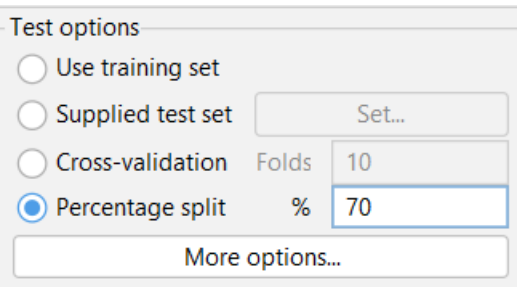


Figure 5: Setting percentage of split for training and testing

After setting split percentage, we need to set the number of cross-validation. In this project, we are using 10 folds of cross validation.

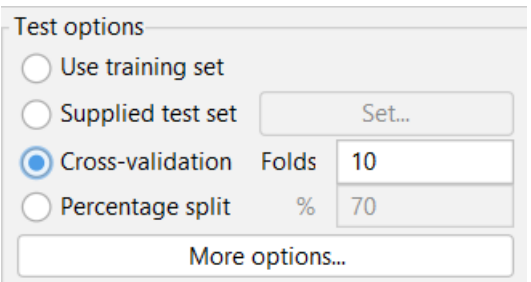


Figure 6: Setting cross-validation

Figure 6 shows cross-validation which is set at 10 folds. After that, the testing of machine learning classifiers is done as shown in Figure 7, Figure 8, and Figure 9.

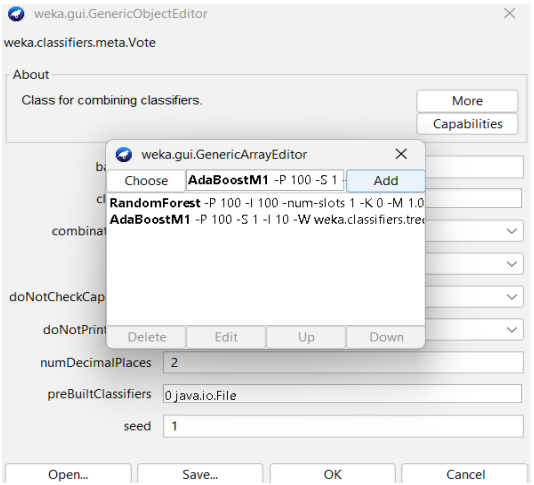


Figure 7: Random Forest combines with AdaBoost

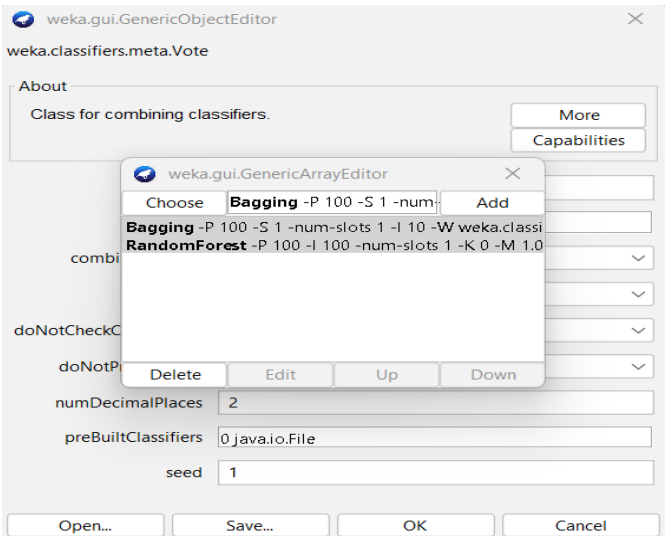


Figure 8: Random Forest combines with Bagging

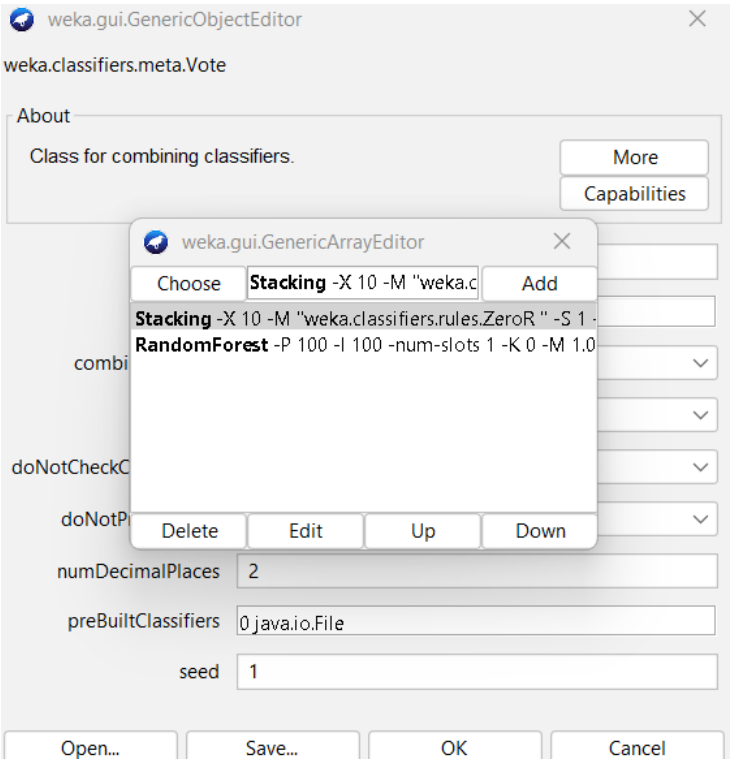


Figure 9: Random Forest combines with Stacking

4.4 Performance Evaluation

The final step of this project is to conduct performance evaluation. The trained models will be put to the test using the test or unseen dataset during this phase. Figures 10, Figure 11, and Figure 12 show the confusion matrix for implemented algorithm.

```
=== Confusion Matrix ===

    a    b   <-- classified as
 4651  247 |    a = -1
  185 5972 |    b = 1
```

Figure 10: Confusion Matrix for RFC + AdaBoost

```
=== Confusion Matrix ===

   a    b   <-- classified as
4676  222 |    a = -1
 113 6044 |    b = 1
```

Figure 11: Confusion Matrix for RFC + Bagging

```
=== Confusion Matrix ===

   a    b   <-- classified as
4676  222 |    a = -1
  86 6071 |    b = 1
```

Figure 12: Confusion Matrix for RFC + Stacking

Table 2 shows the performance of the three Hybrid Ensemble Techniques which contains the data of TP rate, FP rate, Precision, Recall, F-Measure(F1-score), Accuracy and ROC area.

Table 2: Performance evaluation results of three hybrid ensemble techniques

Metrics	RFC + AdaBoost	RFC + Bagging	RFC + Stacking
Accuracy	96.09%	96.97%	97.21 %
F1-score	96.5%	97.30%	97.5%
Precision	96.0%	96.75%	96.5%
Recall	97.0%	98.2%	98.6%

Table 2 shows the results of three types of hybrid ensemble techniques using the same number of attributes. From the table, it can be observed that the hybrid version of RFC with stacking achieved the highest accuracy and F1 score, at 97.21% and 97.5%, respectively. In terms of precision and recall, RFC with stacking obtained the second-highest values, with 96.5% and 98.6%, respectively. The superior performance of the RFC with Stacking can be attributed to its ability to learn optimal combinations of base classifier predictions through a meta-learner.

Unlike Bagging, which primarily reduces variance, and AdaBoost, which emphasizes misclassified instances, Stacking integrates multiple model outputs and captures higher-level relationships among predictions. This enables the stacked model to better distinguish subtle patterns in phishing URLs, particularly in cases involving complex URL structures and obfuscation techniques. As a result, the stacking-based hybrid ensemble demonstrates improved generalization and higher classification accuracy.

5.0 DISCUSSION

This research focuses on advanced machine learning techniques to develop a competent phishing detection system. Phishing is an online scam in which attackers impersonate legitimate websites to trick users into disclosing sensitive personal information. To address this issue, we propose a hybrid ensemble model that combines multiple classifiers. The dataset contains 11,055 instances, including 4,898 legitimate and 6,157 phishing URLs, characterized by 30 features. This model has gone through 10-fold cross-validation which contributes to improved accuracy and reliability.

In the context of URL-based phishing detection, it is critical to identify phishing attacks that can compromise individuals or organizations. Our proposed technique integrates Random Forest Classifier (RFC) with three other ensemble methods: AdaBoost, Bagging, and Stacking. Among these integrations, the combination of RFC with Stacking achieves the highest accuracy. Despite its strong performance, the proposed hybrid ensemble approach has several limitations. The study relies on a single benchmark dataset, which may limit generalizability to real-world phishing scenarios involving zero-day or adversarial URLs. Additionally, URL-based features alone may not capture content-level deception techniques. Future work should evaluate the model across multiple datasets and integrate content-based or behavioral features to enhance robustness.

ACKNOWLEDGMENTS

The authors would like to thank Universiti Teknikal Malaysia Melaka

(UTeM) and Universiti Teknologi Mara, Cawangan Melaka, Kampus Jasin for the support.

REFERENCES

[1] G.Buket, K. Erensoy and E. Kocyigit, "Detection of phishing websites from URLs by using classification techniques on WEKA." *2021 6th International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 2021.

[2] G. Sonowal, "Introduction to phishing" In *Phishing and Communication Channels: A Guide to Identifying and Mitigating Phishing Attacks*, pp. 1-24. Berkeley, CA: Apress, 2021.

[3] Z.Alkhalil, C.Hewage, L.Nawaf and I. Khan, "Phishing attacks: A recent comprehensive study and a new anatomy. *Frontiers in Computer Science*, *3*, 563060, 2021.

[4] M. S. Kheruddin, M. A. E. M. Zuber and M. M. M. Radzai, "Phishing attacks: Unraveling tactics, threats, and defenses in the cybersecurity landscape" *Authorea Preprints*, 2024.

[5] M.Madleňák and K.Kampová, "Phishing as a cyber security threat", In *2022 20th international conference on emerging elearning technologies and applications (ICETA)*, IEEE, pp. 392-396, 2022.

[6] A.Karim, M.Shahroz, K. Mustofa, , S. B. Belhaouari and S. R. K. Joga, "Phishing detection system through hybrid machine learning based on URL", *IEEE Access*, *11*, 36805-36822, 2023.

[7] K.Omari, "Comparative study of machine learning algorithms for phishing website detection", *International Journal of Advanced Computer Science and Applications*, vol.14, no.9, 2023.

[8] A.Kulkarni and L. L. Brown, "Phishing Websites Detection using Machine Learning", International Journal of Advanced Computer Science and Applications, vol. 10, 2019.

[9] A. Soni and J. Tiwari, "Phishing Website Detection Using Ensemble Learning", vol. 11, 2023.

[10] Y. S.Tambe, S. S.Mhangore and K. S. Desai, "Phishing URL Detection Using Machine Learning", *Journal of Advanced Research in Production and Industrial Engineering*, vol.10, no.1, pp. 1-5, 2023.

[11] A.Basit, M.Zafar, , A. R.Javed, and Z.Jalil, "A novel ensemble machine learning method to detect phishing attack", In *2020 IEEE 23rd International Multitopic Conference (INMIC),* pp. 1-5, 2020.

[12] M. H. Mumu, and T.Aishy, "Malicious URL detection using machine learning and deep learning algorithms", *Ph.D. dissertation, East West University*, Bangladesh, 2023.