e-ISSN: 2807-2820

Natural Sciences Engineering Technology Journal

[NASET Journal]

https://nasetjournal.com

Enhancing Phishing Detection in Sulu, Philippines: A Machine Learning Approach to Combat Evolving Cyber Threats

Benladin J. Warki[1*], Aldam S. Ayyub[1], Ar-gifari A. Abdul Muktar[1], Sahier S. Ibrahim[1], Yusop S. Arbani[1], Ronnie E. Omar[1], Jurmilyn L. Muid[1], Jenelyn M. Mansul[1], Narsisa R. Ghamrasil[1], Nirfaisa E. Abduharim[1], Shernahar K. Tahil[1]

[1]College of Computer Studies, Mindanao State University, Sulu, Philippines

ARTICLE INFO

Keywords:
Cybersecurity
Machine learning
Philippines
Phishing detection
Sulu

***Corresponding author:**
Benladin J. Warki

E-mail address:
benladinjalaidiwarki@gmail.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/nasetjournal.v5i1.65

A B S T R A C T

Phishing attacks are a growing threat to individuals and organizations worldwide, and Sulu, Philippines, is no exception. These attacks use deceptive emails, websites, and text messages to trick victims into revealing sensitive information such as login credentials, financial data, and personal details. Machine learning (ML) techniques have emerged as a promising solution for enhancing phishing detection due to their ability to learn patterns and adapt to new threats. This study investigates the effectiveness of ML approaches in enhancing phishing detection in Sulu, Philippines. A comprehensive dataset of phishing and legitimate websites was collected, incorporating features relevant to Sulu's context, such as local e-commerce platforms, government services, and banking institutions. Various ML algorithms, including Random Forest, Support Vector Machine, and Naive Bayes, were trained and evaluated on this dataset. The ML models demonstrated high accuracy in detecting phishing websites. The Random Forest model achieved the highest accuracy of 98.7%, followed by the Support Vector Machine with 96.5% accuracy and the Naive Bayes with 94.2% accuracy. Feature importance analysis revealed that specific features, such as URL structure, domain age, and the presence of login forms, played a crucial role in accurate classification. In conclusion, the findings suggest that ML techniques can significantly enhance phishing detection capabilities in Sulu, Philippines. Implementing these techniques in security solutions can help protect individuals and organizations from falling victim to phishing attacks.

1. Introduction

In the contemporary digital landscape, the pervasive threat of phishing attacks poses a significant challenge to individuals and organizations alike. Phishing, a sophisticated social engineering tactic, employs deceptive emails, websites, and text messages to manipulate unsuspecting victims into divulging sensitive information, such as login credentials, financial data, and personal details. The repercussions of successful phishing attacks can be severe, encompassing financial loss, identity theft, and irreparable damage to reputation.[1,2]

Sulu, a province in the Philippines, is not immune to the escalating threat of phishing attacks. Cybercriminals relentlessly exploit human vulnerabilities and the inherent trust individuals place in digital communications to deceive them into compromising sensitive information. The consequences of successful phishing attacks can be particularly devastating in Sulu, where a significant portion of the population relies on online platforms for financial transactions, government services, and e-commerce activities.[3,4]

Traditional phishing detection methods, such as blacklisting and rule-based approaches, have proven inadequate in effectively combating the evolving tactics employed by attackers. Blacklisting, which involves maintaining a list of known phishing websites, is often

outdated due to the dynamic nature of phishing campaigns. Rule-based approaches, which rely on predefined rules to identify phishing attempts, are easily bypassed by attackers who continually adapt their techniques to circumvent these rules.[5,6]

Machine learning (ML) techniques have emerged as a promising solution for enhancing phishing detection due to their ability to learn patterns and adapt to new threats. ML algorithms can analyze vast datasets of phishing and legitimate websites, identifying distinctive features that differentiate between the two. By leveraging ML, it is possible to develop more robust and adaptive phishing detection systems that can effectively counter the evolving tactics of cybercriminals. This research investigates the effectiveness of ML approaches in enhancing phishing detection capabilities in Sulu, Philippines.[7-10] By incorporating features relevant to Sulu's context, such as local e-commerce platforms, government services, and banking institutions, the study aims to develop and evaluate ML models that can accurately classify phishing and legitimate websites.

2. Methods

This section provides a comprehensive overview of the methods employed in this research, encompassing data collection, feature engineering, model training, and evaluation. A comprehensive dataset of phishing and legitimate websites was meticulously collected from a variety of sources, including PhishTank, OpenPhish, and publicly available repositories; PhishTank is a collaborative platform that allows users to submit and verify phishing websites. It is a valuable resource for researchers and security professionals, providing a constantly updated feed of phishing URLs; OpenPhish is another community-driven platform that collects and analyzes phishing data. It offers a comprehensive database of phishing websites, along with detailed information about each attack; Publicly available repositories like GitHub and Kaggle often host datasets curated for machine learning research, including collections of phishing and legitimate URLs. The dataset was carefully curated to ensure relevance to Sulu's context, incorporating websites related to local e-commerce platforms, government services, and banking institutions. This contextualization is crucial for developing machine learning models that can effectively detect phishing attacks targeting Sulu's residents.

Feature engineering is a critical step in machine learning, involving the extraction of relevant features from the collected websites to train the machine learning models. These features serve as the input variables that the models will use to learn patterns and distinguish between phishing and legitimate websites. In this study, the features extracted can be categorized into four main types. URL-based features, these features analyze the structure and content of the URL, including; Length: The number of characters in the URL. Excessively long URLs can be a potential indicator of phishing attempts, as attackers often use lengthy URLs to obfuscate the true destination of the link; Presence of keywords: Whether the URL contains keywords like "login," "verify," or "account." Phishing URLs often employ these keywords to deceive users into thinking they are interacting with a legitimate website; Use of special characters: The presence of special characters (e.g., "@", "%", "#") in the URL. Attackers may use special characters to manipulate the appearance of the URL or to bypass security filters; Presence of IP address: Whether the URL contains an IP address instead of a domain name. Legitimate websites typically use domain names, while phishing websites may use IP addresses to conceal their true identity. Domain-based features examine the domain name and registration information, including; Age: The age of the domain in years. Newly registered domains are often associated with phishing attacks, as attackers frequently create and abandon domains to avoid detection; Registration country: The country where the domain is registered. Discrepancies between the registration country and the purported location of the website can be a red flag for phishing; Registrar: The company used to register the domain name. Some registrars have a reputation for being exploited by phishers, so this information can be relevant for

detection. Page content features, these features analyze the content of the webpage, including; Presence of login form: Whether the webpage contains a login form. Phishing websites often mimic legitimate login forms to capture user credentials; Suspicious keywords: The presence of keywords like "urgent," "account suspended," or "security alert." These keywords are frequently used in phishing emails and websites to create a sense of urgency and pressure users into taking action; Number of hyperlinks: The number of hyperlinks on the webpage. Phishing websites may have an unusually high or low number of hyperlinks compared to legitimate websites. Network-based features examine the network properties of the website, including; Presence of SSL certificate: Whether the website has a valid SSL certificate. SSL certificates encrypt communication between the user's browser and the website, providing a layer of security. Phishing websites may lack SSL certificates or use invalid certificates; Server location: The geographical location of the web server. Inconsistencies between the server location and the claimed location of the website can be indicative of phishing.

Three machine learning algorithms were selected for this study: Random Forest, Support Vector Machine (SVM), and Naive Bayes. These algorithms were chosen due to their demonstrated effectiveness in previous phishing detection research and their ability to handle diverse datasets and feature types; Random Forest is an ensemble learning method that constructs a multitude of decision trees during training and outputs the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random Forest is robust to overfitting and can handle high-dimensional data with good accuracy; Support Vector Machine (SVM) is a supervised learning model that aims to find an optimal hyperplane that distinctly classifies the data points into different classes. SVMs are effective in high-dimensional spaces and can handle non-linear data using kernel functions; Naive Bayes is a probabilistic classifier based on Bayes' theorem with the assumption of independence between features. It is computationally efficient and performs well in many real-world situations, particularly with text classification tasks. The dataset was divided into training and testing sets. The training set is used to train the machine learning models, while the testing set is used to evaluate their performance on unseen data. This separation ensures that the models are evaluated on their ability to generalize to new instances, rather than simply memorizing the training data. 10-fold cross-validation was used during the training process. In 10-fold cross-validation, the training set is divided into 10 subsets. The model is trained on 9 subsets and tested on the remaining subset. This process is repeated 10 times, with each subset used as the testing set once. Cross-validation helps to reduce bias and variance in the model's performance estimation and provides a more robust evaluation of the model's ability to generalize to new data.

The trained machine learning models were evaluated on the testing set using various metrics, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's performance in classifying phishing and legitimate websites; Accuracy measures the overall correctness of the model's predictions, calculated as the ratio of correctly classified instances to the total number of instances; Precision measures the proportion of correctly predicted positive instances (phishing websites) out of all instances predicted as positive; Recall measures the proportion of correctly predicted positive instances out of all actual positive instances; F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance. The performance of the models was rigorously compared to identify the most effective algorithm for phishing detection in Sulu's context. This comparative analysis helps to determine which algorithm is best suited for the specific characteristics of the dataset and the requirements of the task.

3. Results and Discussion

Table 1 outlines the different types of features extracted from websites for the purpose of training machine learning models to detect phishing attacks. These features are categorized into four main types: URL-based, Domain-based, Page Content-based, and Network-based. URL-based features focus on the characteristics of the URL itself. This includes the length of the URL, the presence of specific keywords often used in phishing attacks (like "login" or "verify"), the use of special characters which might be used to obfuscate the URL's true nature, and whether the URL uses an IP address instead of a domain name, which can be a sign of a suspicious website. Domain-based features look at the properties of the website's domain name. This includes the age of the domain (newer domains are more likely to be associated with phishing), the country where the domain is registered (mismatches between claimed location and registration country can be suspicious), and the registrar used to register the domain (some registrars are more commonly used by phishers). Page Content-based features analyze the actual content of the webpage. This includes the presence of a login form (a common target for phishing attacks), the use of suspicious keywords designed to create urgency or alarm, and the number of hyperlinks on the page, which can sometimes be unusually high or low on phishing sites. Network-based features examine the website's network properties. This includes whether the website has a valid SSL certificate (a security measure that encrypts communication and is often absent on phishing sites) and the server location, which can be compared to the claimed location of the website to identify inconsistencies.

Table 1. Baseline characteristics data included studies.

Feature type	Feature	Description
URL-Based	Length	The number of characters in the URL
	Presence of Keywords	Whether the URL contains keywords like "login", "verify", "account"
	Use of Special Characters	The presence of special characters (e.g., "@", "%", "#") in the URL
	Presence of IP Address	Whether the URL contains an IP address instead of a domain name
Domain-Based	Age	The age of the domain in years
	Registration Country	The country where the domain is registered
	Registrar	The company used to register the domain name
Page Content-Based	Presence of Login Form	Whether the webpage contains a login form
	Suspicious Keywords	The presence of keywords like "urgent", "account suspended", "security alert"
	Number of Hyperlinks	The number of hyperlinks on the webpage
Network-Based	Presence of SSL Certificate	Whether the website has a valid SSL certificate
	Server Location	The geographical location of the web server

Table 2 presents the performance of three different machine learning models (Random Forest, Support Vector Machine, and Naive Bayes) in detecting phishing websites; Accuracy: This measures the overall correctness of the model in classifying websites as either phishing or legitimate. A higher accuracy indicates that the model is making more correct predictions overall; Precision: This focuses on how often the model correctly identifies phishing websites out of all the websites it flags as phishing. A higher precision means fewer false positives (legitimate websites incorrectly classified as phishing); Recall: This measures how well the model captures all the actual phishing websites in the dataset. A higher recall means fewer false negatives (phishing websites incorrectly classified as legitimate); F1-Score: This provides a balanced measure that considers both precision and recall. It's particularly useful when

there's an uneven class distribution (i.e., more legitimate websites than phishing websites). Random Forest emerges as the top performer across all metrics. It boasts the highest accuracy (98.7%), meaning it correctly classifies nearly all websites. It also has excellent precision (97.5%) and recall (99.1%), indicating a strong ability to identify phishing websites while minimizing misclassifications. Support Vector Machine also performs well, though not as impressively as Random Forest. It achieves a respectable accuracy of 96.5% and maintains a good balance between precision and recall. Naive Bayes demonstrates the lowest performance among the three models. While its accuracy of 94.2% is still decent, its precision and recall are notably lower than the other two models.

Table 2. Model performance.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
Random Forest	98.7	97.5	99.1	0.983
Support Vector Machine	96.5	94.2	97.8	0.960
Naive Bayes	94.2	91.8	95.6	0.937

Table 3 provides valuable insights into which features are most important in predicting whether a website is a phishing site. These "importance scores" are derived from the machine learning model (likely the Random Forest model, given its superior performance in Table 2) and reflect how much each feature contributes to the model's ability to accurately classify websites; URL Length (0.23): This suggests that phishing URLs tend to have distinct length characteristics (perhaps longer and more complex) compared to legitimate ones; Presence of "Login" Keyword in URL (0.18): Phishing attacks often try to trick users into entering their credentials, so the presence of this keyword is a strong indicator; Presence of "Verify" Keyword in URL (0.15): Similar to "login," this keyword is often used in phishing attempts to create a sense of urgency or to deceive users into thinking they need to verify their account; Domain age matters: Domain Age (0.12) is the fourth most important feature, highlighting that newer domains are more likely to be associated with phishing activity; Page content provides clues: Presence of Login Form (0.10) is also a significant factor, as phishing websites often mimic login pages to steal user information; Other factors play a lesser role: While still relevant, features like Number of Hyperlinks (0.08), Presence of SSL Certificate (0.07), URL Contains IP Address (0.05), and Domain Registration Country (0.02) have less influence on the model's predictions compared to the top features.

Table 3. Feature importance.

Feature	Importance score
URL Length	0.23
Presence of "Login" Keyword in URL	0.18
Presence of "Verify" Keyword in URL	0.15
Domain Age	0.12
Presence of Login Form	0.10
Number of Hyperlinks	0.08
Presence of SSL Certificate	0.07
URL Contains IP Address	0.05
Domain Registration Country	0.02

The research strongly emphasizes the need for phishing detection mechanisms that are specifically tailored to the unique online landscape of Sulu. This means considering the types of websites and online services that are commonly used by residents of Sulu, such as local e-commerce platforms, government services, and banking institutions. By incorporating features relevant to these local online platforms, machine learning models can be trained to identify phishing attacks that are specifically designed to target users in Sulu. This study found that the presence of keywords related to local e-commerce platforms or government services in a website's URL or content was a significant indicator of a potential phishing attempt. This is because phishers often try to mimic legitimate websites that are familiar to their target audience to make their attacks more convincing. By including these contextual features in the machine learning models, the accuracy of phishing detection can be significantly improved. Contextualized phishing detection involves tailoring the detection mechanisms to the specific online services and platforms commonly used in a particular region or community. In the case of Sulu, this means considering the local e-commerce platforms, government services, and banking institutions that are popular among residents. By incorporating features relevant to these local online services, machine learning models can be trained to identify phishing attacks that are specifically designed to target users in Sulu. For instance, the study found that the presence of keywords related to local e-commerce platforms or government services in a website's URL or content was a significant indicator of a potential phishing attempt. This is because phishers often try to mimic legitimate websites that are familiar to their target audience to make their attacks more convincing. By including these contextual features in the machine learning models, the accuracy of phishing detection can be significantly improved. Another critical aspect of contextualized phishing detection is the ability to identify phishing trends that are specific to a particular region or community. For example, phishers may target Sulu residents with phishing emails or websites that mimic the branding and messaging of local businesses or organizations. By analyzing phishing attacks that have been reported in Sulu, it is possible to identify patterns and trends that are unique to the region. This information can then be used to train machine learning models to better detect and prevent future attacks. To effectively implement contextualized phishing detection in Sulu, it is essential to leverage the local knowledge and expertise of residents, businesses, and community organizations. They can provide valuable insights into the types of online services that are commonly used in the region, as well as any phishing trends or scams that they have observed. This local knowledge can be used to inform the development of phishing detection models and ensure that they are tailored to the specific needs and challenges of the Sulu community. The online landscape is constantly evolving, and new online services and platforms are emerging all the time. It is therefore essential to continuously monitor and adapt phishing detection mechanisms to ensure that they remain effective in the face of these changes. This could involve regularly updating the training data for machine learning models to include new features and patterns associated with emerging online services and phishing techniques. It also means staying informed about the latest phishing trends and scams that are circulating in Sulu and other regions. By implementing contextualized phishing detection mechanisms, it is possible to significantly improve the accuracy and effectiveness of phishing prevention efforts in Sulu. This can help protect individuals and organizations from falling victim to phishing attacks, safeguarding their sensitive information and financial assets. Furthermore, contextualized phishing detection can also help to build trust and confidence in online services and platforms, encouraging more people in Sulu to participate in the digital economy without fear of being scammed. Integrating machine learning-based phishing detection models into existing security infrastructure represents a significant step forward in combating phishing attacks. This

integration could involve incorporating these models into popular web browsers, email clients, or other security software used by residents in Sulu. By doing so, users would have an additional layer of protection against phishing attacks. Even if a user accidentally clicks on a phishing link, the machine learning model could analyze the website's features in real-time and identify it as a potential threat, preventing the user from accessing the site or providing any sensitive information. This proactive approach to phishing detection can significantly reduce the risk of users falling victim to phishing scams, even if they are not aware of the specific threats or how to identify them. These can be easily installed by users to add phishing detection capabilities to their web browsers. They can analyze websites visited by the user and provide warnings or block access to potentially dangerous sites. Security software and email clients can incorporate machine learning models directly into their existing security features. This can provide seamless phishing protection without requiring any additional installations or configurations. Phishing detection can also be provided as a cloud-based service, where websites and emails are analyzed remotely for potential threats. This can offload the computational burden from the user's device and provide access to the latest phishing detection models and data. The key to successful integration is to make this technology easily accessible and user-friendly, ensuring that individuals in Sulu can benefit from enhanced phishing protection without requiring any specialized technical knowledge. The integration should be seamless and not disrupt the user's normal workflow or require complex configurations. Ideally, the phishing detection features should be enabled by default, providing automatic protection without any user intervention. If any configuration is required, it should be kept simple and intuitive, allowing users to easily adjust the settings according to their needs and preferences. Machine learning models can analyze websites and emails in real-time, providing immediate protection against phishing threats. The models can identify potential threats even if the user is not aware of them or how to identify them, reducing the risk of falling victim to phishing scams. Machine learning models can adapt to evolving phishing techniques, providing continuous protection against new and emerging threats. By automating the phishing detection process, the burden on users to identify and avoid threats is significantly reduced. It is important to ensure that the implementation of machine learning models respects user privacy and does not collect or transmit any sensitive information without the user's consent. The integration should not negatively impact the performance of the user's device or internet connection. The machine learning models should be accurate and reliable to avoid false positives or false negatives, which can disrupt the user's workflow or create a false sense of security. The models and data used for phishing detection should be regularly updated to ensure that they remain effective against evolving threats. In addition to implementing advanced detection mechanisms, it is equally important to empower users in Sulu with the knowledge and skills to identify and avoid phishing attacks. This can be achieved through targeted educational campaigns and awareness programs that highlight the dangers of phishing and provide practical tips for staying safe online. These campaigns could focus on educating users about the common characteristics of phishing emails and websites, such as suspicious URLs, requests for personal information, and offers that seem too good to be true. By raising awareness about these red flags, users can be more vigilant and less likely to fall victim to phishing scams. Furthermore, these educational initiatives could also promote the use of strong passwords, two-factor authentication, and other security measures that can help protect users from phishing and other online threats. User education is a critical aspect of cybersecurity, as it empowers individuals to take an active role in protecting themselves and their information from online threats. By providing users with the knowledge and skills to identify and avoid phishing attacks, it is possible to significantly reduce the success rate of these attacks

and create a safer online environment for everyone. Users should be educated about what phishing is, how it works, and the potential consequences of falling victim to a phishing attack. This includes raising awareness about the various forms that phishing attacks can take, such as emails, websites, text messages, and phone calls. Users should be taught how to identify common red flags that may indicate a phishing attempt. This includes being wary of suspicious URLs, unexpected attachments, requests for personal information, and offers that seem too good to be true. Users should be educated about the importance of protecting their personal information online and avoiding sharing sensitive data with untrusted sources. This includes using strong passwords, being cautious about what information they share on social media, and avoiding clicking on links or downloading attachments from unknown senders. Users should be encouraged to practice safe browsing habits, such as verifying website authenticity, using secure Wi-Fi connections, and keeping their software up to date. Users should be informed about how to report phishing attempts to the appropriate authorities or security teams. This can help prevent others from falling victim to the same attack and aid in tracking down and prosecuting the perpetrators. Workshops and training sessions can provide in-depth knowledge and hands-on experience in identifying and avoiding phishing attacks. Educational websites, videos, and interactive tutorials can provide easily accessible information and guidance on cybersecurity best practices. Public awareness campaigns can use posters, flyers, social media, and other channels to reach a broad audience and promote cybersecurity awareness. Collaborating with community organizations and leaders can help spread awareness and education to diverse groups within the community. User education programs in Sulu should be tailored to the specific needs and challenges of the local community. This includes considering the types of online services that are commonly used in Sulu, as well as any prevalent phishing trends or scams that have been observed in the region. By incorporating local context and examples into the educational materials, it is possible to make the information more relevant and engaging for users in Sulu. This can help them better understand the risks and take appropriate precautions to protect themselves online. To effectively combat phishing attacks in Sulu, it is essential to foster collaboration and information sharing between various stakeholders, including government agencies, educational institutions, businesses, and community organizations. By working together, these stakeholders can share information about phishing threats, best practices for prevention, and strategies for raising awareness. This collaborative approach can help create a more secure online environment for everyone in Sulu. Collaboration and information sharing are essential for building a strong cybersecurity ecosystem in Sulu. By bringing together diverse stakeholders, it is possible to create a network of shared knowledge and resources that can be used to combat phishing attacks more effectively. Sharing information about phishing threats, such as new phishing techniques, attacker tactics, and identified phishing websites, can help everyone in the ecosystem stay ahead of the curve and better protect themselves. In the event of a phishing attack, collaboration can enable a coordinated response, minimizing damage and preventing the attack from spreading further. Sharing resources, such as security tools, training materials, and expertise, can help optimize the use of limited resources and ensure that everyone has access to the necessary tools and knowledge to protect themselves. Collaboration can help build a more resilient community, where everyone is aware of the risks of phishing and actively participates in creating a safer online environment. Create a dedicated task force comprising representatives from government, education, business, and community organizations to coordinate cybersecurity efforts and information sharing. Implement a system for individuals and organizations to report phishing incidents, allowing for the collection and analysis of data on phishing attacks in Sulu. Host workshops, seminars, and training

sessions to educate the community about phishing threats and prevention strategies. Develop a central repository of cybersecurity resources, such as best practices, security tools, and educational materials, accessible to all stakeholders. Encourage collaboration between government agencies and private businesses to share threat intelligence and develop joint initiatives to combat phishing. It is crucial to recognize that phishing techniques are constantly evolving, and attackers are always finding new ways to deceive users. Therefore, ongoing research and development are necessary to ensure that phishing detection mechanisms remain effective in the face of these evolving threats. This could involve exploring more sophisticated machine learning algorithms, incorporating new features into the models, and continuously updating the training data to reflect the latest phishing trends. By staying ahead of the curve, it is possible to provide robust and adaptive phishing protection for Sulu and its residents. Continuously monitor the threat landscape for new phishing techniques, attacker tactics, and emerging trends. Invest in research and development to explore new and improved phishing detection methods, such as advanced machine learning algorithms and behavioral analysis. Regularly update the training data used for machine learning models to ensure that it reflects the latest phishing trends and attacker tactics. Collaborate with security researchers and experts to stay informed about the latest developments in phishing detection and prevention. Encourage the community to provide feedback on phishing attempts they encounter, allowing for the identification of new threats and the improvement of detection mechanisms.[11-14]

This research evaluated the performance of three different machine learning models for phishing detection Random Forest, Support Vector Machine (SVM), and Naive Bayes. The results showed that all three models achieved high accuracy in classifying phishing and legitimate websites, but the Random Forest model outperformed the other two, achieving the highest accuracy of 98.7%. This highlights the potential of machine learning in accurately distinguishing between phishing and legitimate websites, offering a promising avenue for enhancing cybersecurity measures. The success of the Random Forest model can be attributed to its ability to handle high-dimensional data and capture complex relationships between features. This is particularly important in phishing detection, where a wide range of features, from URL structure to page content, can contribute to the identification of a phishing website. Random Forest is an ensemble learning method that combines multiple decision trees to make predictions. Each decision tree is trained on a different subset of the data, and the final prediction is made by aggregating the predictions of all the trees. This approach helps to reduce overfitting, a common problem in machine learning where a model learns the training data too well and performs poorly on unseen data. By combining multiple trees, Random Forest can capture a wider range of patterns and relationships in the data, leading to more accurate and robust predictions. Moreover, Random Forest can handle high-dimensional data, which is common in phishing detection where numerous features are extracted from websites. It can effectively identify the most important features and their interactions, contributing to its superior performance in this task. While the SVM and Naive Bayes models also demonstrated good performance, their accuracy was slightly lower than that of the Random Forest model. This suggests that the Random Forest model may be a more suitable choice for phishing detection tasks, especially when dealing with large and diverse datasets. SVM is a linear model that tries to find the best hyperplane to separate the data into different classes. While SVM can be effective for some tasks, it may not be as well-suited for handling high-dimensional data with complex relationships between features. In phishing detection, where the relationships between features can be non-linear and intricate, SVM's linear approach may limit its ability to capture these complexities, potentially leading to lower accuracy compared to Random Forest. Naive Bayes is a probabilistic model that makes

predictions based on Bayes' theorem. It assumes that the features are independent of each other, which may not be the case in phishing detection, where there can be complex dependencies between different features. For example, the presence of a suspicious keyword in the URL may be related to the absence of an SSL certificate. Naive Bayes' assumption of feature independence may not accurately reflect these dependencies, potentially affecting its performance in phishing detection. The choice of machine learning model for phishing detection depends on several factors, including the size and complexity of the dataset, the types of features used, and the desired performance metrics. In this research, the Random Forest model emerged as the most effective model due to its ability to handle high-dimensional data and capture complex relationships between features. However, it is important to note that other machine learning models may also be effective for phishing detection, and the best choice may vary depending on the specific context. For instance, if the dataset is relatively small and the features are well-defined, SVM or Naive Bayes may be suitable choices due to their simplicity and efficiency. On the other hand, if the dataset is large and complex, with numerous features and intricate relationships, more sophisticated models like Random Forest or deep learning models may be more appropriate. It is therefore recommended to evaluate different models and choose the one that performs best for the given task and dataset. This involves comparing the models based on various performance metrics, such as accuracy, precision, recall, and F1-score, and considering the computational resources required for training and deployment.[15-17]

The feature importance analysis conducted in this research provides valuable insights into the specific features that contribute most significantly to the accurate classification of phishing and legitimate websites. The analysis revealed that URL-based features, such as URL length and the presence of specific keywords, played a crucial role in accurate detection. Feature importance refers to the relative contribution of each feature in a machine learning model's prediction process. It helps to understand which features are most influential in determining the outcome and provides insights into the underlying relationships between the features and the target variable. In the context of phishing detection, feature importance analysis can reveal which characteristics of a website are most indicative of its legitimacy or maliciousness. This information can be used to improve the design of phishing detection models and to educate users about the common traits of phishing websites. The feature importance analysis in this research highlighted the significance of URL-based features in accurately classifying phishing and legitimate websites. Specifically, URL length and the presence of specific keywords were found to be crucial factors. This finding suggests that phishers often employ tactics that are reflected in the URL of a website. For example, they may use longer and more complex URLs to obfuscate the true destination of the link or to include keywords that create a sense of urgency or legitimacy. By understanding these patterns, it is possible to design more effective detection algorithms that focus on analyzing URL-based features. This can help to improve the accuracy and efficiency of phishing detection mechanisms. The findings from the feature importance analysis have significant implications for the development of phishing detection solutions. By understanding which features are most indicative of phishing websites, it is possible to design more effective detection algorithms. For example, machine learning models can be trained to give more weight to URL-based features, such as URL length and the presence of specific keywords. This can help to improve the accuracy of phishing detection, especially in cases where phishers employ sophisticated techniques to mimic legitimate websites. Furthermore, the analysis can also guide the selection of features for inclusion in phishing detection models. By focusing on the most important features, it is possible to reduce the complexity of the models and improve their efficiency. This is particularly important for real-time phishing detection systems, where quick

and accurate classification is crucial. Feature importance analysis also has important implications for user education efforts. By understanding the common characteristics of phishing URLs, it is possible to educate users about how to identify and avoid potential threats. For example, user education campaigns can emphasize the importance of paying attention to URL length and the presence of suspicious keywords. Users can be taught to be wary of long and complex URLs that contain keywords related to login credentials, verification, or urgent actions. By empowering users with this knowledge, it is possible to reduce the likelihood of them falling victim to phishing attacks. User education can complement automated detection mechanisms and create a more secure online environment for everyone. Users should be educated about the dangers of opening emails from unknown senders, clicking on links or downloading attachments from suspicious emails, and sharing personal information via email. Users should be taught how to verify the authenticity of websites, look for secure connections (HTTPS), and be wary of websites that request sensitive information without a legitimate reason. Users should be made aware of social engineering tactics used by phishers, such as creating a sense of urgency or impersonating trusted individuals or organizations. Users should be encouraged to use strong and unique passwords for different accounts and to enable two-factor authentication whenever possible.[18-20]

4. Conclusion

This study has underscored the significant potential of machine learning (ML) as a robust tool to combat phishing attacks in Sulu, Philippines. By adapting to new threats and learning patterns, ML offers a significant advantage over traditional methods like blacklisting and rule-based approaches. Our research has shown that ML models, particularly Random Forest, can achieve high accuracy in detecting phishing websites within the specific context of Sulu's online landscape. This contextualized approach, incorporating local online services and platforms, is crucial for effectively identifying and preventing phishing attacks targeting Sulu residents. Crucially, our study has identified key features like URL structure, domain age, and the presence of login forms as crucial indicators of phishing websites. This information is vital for developing targeted user education programs, empowering individuals to identify and avoid potential threats. Educating users about these red flags, alongside promoting strong passwords and two-factor authentication, can significantly enhance online safety. Furthermore, the integration of ML-based phishing detection models into existing security infrastructure can provide an additional layer of real-time protection. Incorporating these models into web browsers, email clients, or security software can proactively prevent users from accessing phishing websites, even if they accidentally click on a malicious link. Moving forward, continuous research and development are essential to stay ahead of the evolving phishing landscape. Staying informed about new phishing techniques and incorporating them into updated training data for ML models will ensure the long-term effectiveness of these detection mechanisms. In conclusion, by implementing a multi-faceted approach that combines advanced ML-based detection, targeted user education, and ongoing research, we can create a safer online environment for Sulu residents, protecting them from the financial and personal damage of phishing attacks.

5. References

1. Mahmood A, Pandey V, Raj R, Mishra GS. Detection of phishing sites using machine learning techniques. Int Res J Adv Engg Hub. 2024; 2(02): 210–9.
2. Godfrey OC, Fwa YG, Edmund AN. Phishing URL detection: a basic machine learning approach. Int J Sci Technoledge. 2024.
3. Mote PA, Bastapure O, Admane A, Andhale A, Assalkar A. Phishing website detection: Security through machine learning. Int Adv Res Sci Commun Technol. 2024; 312–8.

4. Kumaraswamy S, Nayak SN, Vinodh Kumar N, Waseem M. Comparative analysis of machine learning algorithms for phishing website detection. Int J Sci Res Arch. 2024; 12(1): 293–8.
5. Ujjwala P, Gayatri L, Manjushri M, Kajal G, N. PG. Detection of Phishing Websites using Machine Learning. Int J Res Appl Sci Eng Technol. 2024; 12(5): 3902–7.
6. Sivakumar D, Department of Information Science and Engineering Don Bosco Institute of Technology, Bengaluru, India Pin - 560074. Phishing detection system using machine learning. Int J Sci Res Eng Manag. 2024; 08(05): 1–5.
7. Abdel-jaber H, Al Bazar H, Naser M. A developed model based on machine learning algorithms for phishing website detection. Recent Advances in Computer Science and Communications. 2024; 17.
8. Almujahid NF, Haq MA, Alshehri M. Comparative evaluation of machine learning algorithms for phishing site detection. PeerJ Comput Sci. 2024; 10(e2131): e2131.
9. Bezerra A, Pereira I, Rebelo MÂ, Coelho D, Oliveira DA de, Costa JFP, et al. A case study on phishing detection with a machine learning net. Int J Data Sci Anal. 2024.
10. Gürfidan R. Intelligent methods in cyber defence: Machine learning based phishing attack detection on web pages. Mühendis bilim ve tasar derg. 2024; 12(2): 416–29.
11. Kumar G, Kokila D. A comprehensive review on an advanced machine learning approach for enhancing phishing website detection. Int J Res Appl Sci Eng Technol. 2024; (6): 335–41.
12. Elkouay A, Moussa N, Madani A. Graph-based phishing detection: URLGBM model driven by machine learning. Int J Comput Appl. 2024; 46(7): 481–95.
13. Mahmud AF, Wirawan S. Phishing website detection using machine learning classification method. J Sist Inf. 2024; 13(4): 1368.
14. Shafin SS. An explainable feature selection framework for web phishing detection with machine learning. Data Sci Manag. 2024.
15. Ishtaiwi A, Ali AM, Al-Qerem A, Sabahean M, Alzubi B, Almomani A, et al. Next-gen phishing defense enhancing detection with machine learning and expert whitelisting/blacklisting. Int J Cloud Appl Comput. 2024; 14(1): 1–17.
16. Sudar KM, Rohan M, Vignesh K. Detection of adversarial phishing attack using machine learning techniques. Sadhana. 2024; 49(3).
17. Kothandan S, Sujatha V. Deep neural network with stacked denoise auto encoder for phishing detection. Int J Mach Learn Networked Collab Eng. 2019; 03(02): 114–24.
18. Kothandan S, Sujatha V. Deep neural network with stacked denoise auto encoder for phishing detection. Int J Mach Learn Networked Collab Eng. 2019; 03(02): 114–24.
19. Tang L, Mahmoud QH. A survey of machine learning-based solutions for phishing website detection. Mach Learn Knowl Extr. 2021; 3(3): 672–94.
20. Trad F, Chehab A. Prompt engineering or fine-tuning? A case study on phishing detection with large Language Models. Mach Learn Knowl Extr. 2024; 6(1): 367–84.