EasyChair Preprint

№ 15938

Hacker Eye

M Suthar Kunta, Naga Surya Pakalapati, Priya Banik,
V L S Sai Sathwik Badam and G S Navaneeth Arumilli

March 24, 2025

HACKER EYE

Prof. Kunta M Suthar
Assistant Professor
Parul University
KUNTA.SUTHAR31680@paruluniversity.ac.in

Pakalapati Naga Surya
Parul University
Artificial Intelligence
pandusurya456@gmail.com

Priya Banik
Parul University
Computer Science
priyabanik222@gmail.com

Badam VLS Sai Sathwik
Parul University
Artificial Intelligence
ssathwik157@gmail.com

Arumilli G S Navaneeth
Parul University
Artificial Intelligence
210303124202@paruluniversity.ac.in

***Abstract*— Phishing sites are a dangerous cyber threat where users are tricked into revealing sensitive information. This paper proposes Hacker Eye, a machine learning-based phishing prediction tool. The paper discusses various classification techniques, such as Random Forest, XG Boost, and ensemble techniques, to ensure maximum detection accuracy. Feature extraction techniques, such as URL structure inspection, domain age, and SSL certificate verification, assist in making classification more precise. A new hybrid technique involving deep learning and natural language processing (NLP) is also proposed to achieve maximum phishing detection. Test results indicate that ensemble techniques, such as XG Boost and Voting Classifiers, provide maximum accuracy in phishing and genuine site detection. The paper concludes with the description of how machine learning can be utilized to achieve maximum cyber defense and recommends future improvements to adaptive learning and multidisciplinary approaches.**

Keywords: Phishing Detection, Machine Learning, XG Boost, Random Forest, Cybersecurity, Website Classification, Deep Learning, NLP, Adaptive Learning.

I. INTRODUCTION

Hacker Eye is one of the most prevalent cybersecurity threats nowadays. It entails fraudulent methods to pilfer sensitive user information such as login credentials and financial data. Cybercriminals are still modify their attack tactics, making it more difficult for conventional security measures to detect and successfully defend against phishing attacks. These fake sites usually replicate authentic sites, manipulating people's trust and enticing them into disclosing confidential information. Traditional rule-based detection systems based on predefined heuristics and blacklists, which become becomes outdated shortly as new phishing sites emerge. Machine learning-based approaches, on the other hand hand, offer a smarter and more agile solution by examining trends, habits, and characteristics of websites to determine whether they are authentic or phishing. The ability to learn lessons from past attacks and generalize with which to classify new threats requires machine learning an effective cybersecurity tool. This research explores many machine learning techniques for optimizing phishing detection effectiveness such as ensemble methods such as XG Boost and Voting Classifiers. Further, this studies explores how deep learning can be incorporated and natural language processing (NLP) techniques to enhance feature extraction and classification capabilities. By combining different methodologies, this study will help in adding to the current endeavors in creating strong and scalable phishing detection systems.

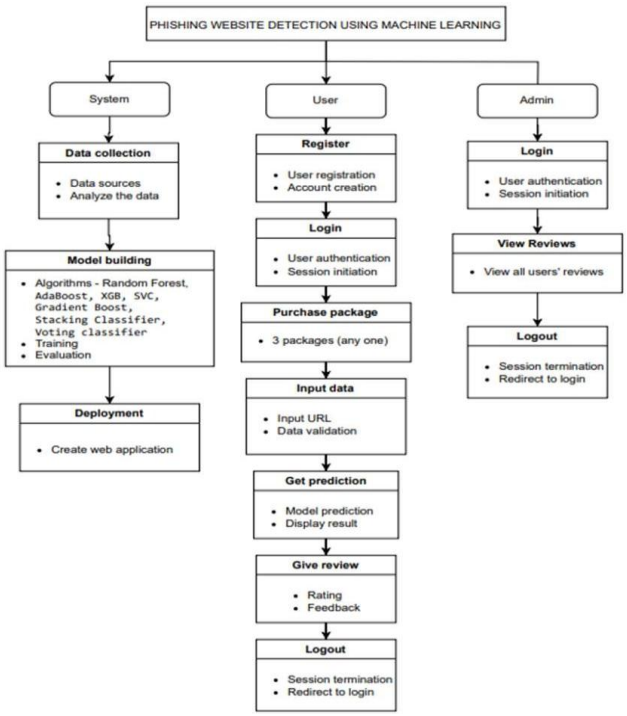

Figure 1.Project Flow Diagram

II. PROBLEM STATEMENT

The growing threat of phishing sites has emerged as a major threat in cybersecurity, and there is a need for more sophisticated detection methods. Rule-based systems, while still in use, are no longer very effective against the ever-changing techniques employed by cybercriminals. These conventional approaches rely on pre-defined rules and static attributes, making them inadequate in detecting newly designed phishing techniques that keep on changing to avoid security measures.

To address such issues, the present work explores the application of machine learning methods for enhancing phishing web site identification. Machine learning methods excel in identifying intricate patterns, acquiring relevant features, and making better predictions with time. By combining various solutions like Natural Language Processing (NLP), behavior monitoring, and ensemble learning, the present work tries to develop a highly robust and versatile phishing detection framework.

This study takes into account different machine learning classifiers and tests their performance. It uses important metrics like accuracy, precision, recall, F1-score, and confusion matrix. The main aim is to create a detection system that can evolve and function without failure, enabling improved cybersecurity. This will reduce the harm inflicted by phishing attacks on individuals and organizations.

III. LITERATURE REVIEW

Several studies for machine learning techniques for phishing detection, highlighting the effectiveness

A. Traditional Machine Learning Techniques

Traditional classifiers such as Support Vector Machines (SVM) and Decision Trees have been experimented with for phishing detection with varied success (Patil & Patil, 2018). These have some disadvantages including:

Limited feature extraction: Traditional classifiers utilize manual feature extraction, which is computationally costly and might not identify intricate patterns.

Poor handling of imbalanced data: Phishing data is skewed in nature, with a large number of legitimate websites and limited phishing websites. Traditional classifiers can struggle with such imbalances.

B. Ensemble Methods

Ensemble models, including Random Forest and XG Boost, are more effective in identifying phishing (Abbasi et al., 2019). The methods:

Ensemble the models: Ensemble methods blend the predictions of several models, minimizing the risk of overfitting and enhancing overall performance.

Handle unbalanced data: Ensemble methods handle unbalanced data more effectively than the default classifiers.

C. Deep Learning and Natural Language Processing (NLP)

Recent developments merge deep learning and NLP to attain the highest detection rate. Deep learning methods, including:

Convolutional Neural Networks (CNN): Optimally suited for image-based phishing detection.

Recurrent Neural Networks (RNN): Ideal for text-based phishing detection.

NLP methods have also been incorporated to search web pages and enhance detection rates (Brown et al., 2021).

D. Behavioural Analysis

Adaptive Learning Behaviour analysis techniques have been researched to detect phishing attacks on user behaviour patterns (Liu & Wang, 2022). Adaptive learning models have also been introduced to:

Learn from emerging data: Adaptive learning models learn from emerging data and evolve along with changing phishing methods.

Improve detection accuracy: Adaptive learning models improve detection accuracy with time.

E. Interdisciplinary Approaches

Interdisciplinary approaches incorporating cybersecurity, psychology, and human behavior analytics have been suggested in order to advance phishing detection (Zhang et al., 2020).

F. Gaps and Future Directions

Although research in phishing detection has improved, there are a few limitations to existing studies. Future studies need to investigate:

Real-time detection: Creating real-time phishing detection tools that can evolve with changing attacks. Adversarial attacks: Investigating the effect of adversarial attacks on phishing detection models and designing countermeasures to counter these attacks. Interdisciplinary approaches: Researching interdisciplinary approaches combining cybersecurity, psychology, and human behavior analysis to improve phishing detection.

IV. METHODOLOGY

The data used in this study includes labeled phishing and normal websites, which were collected from public repositories. The following features were collected for enhancing phishing detection: content features, URL length, domain age, and the presence of the SSL certificate. For the

purpose of feature selection, Natural Language Processing (NLP) techniques were employed for content analysis, and user behavior features like click patterns were added to enhance the performance of the model.

To effectively detect phishing websites, several machine learning algorithms were utilized and compared:

1)Random Forest Classifier – An ensemble algorithm that creates a collection of decision trees and aggregates their predictions to produce improved accuracy.

2)XG Boost Classifier – A highly optimized speed and efficiency gradient-boosting algorithm that can handle large datasets efficiently.

3)Support Vector Classifier (SVC) – Supervised classification model that calculates a best-fit hyperplane with better discrimination between phishing and legitimate sites.

4)Gradient Boosting Classifier – An iterative boosting technique that progressively enhances weak learner errors to optimize classification performance.

5)Voting and Stacking Classifiers – Ensemble learning methods that involve voting among several classifiers for increased reliability and accuracy.

6)Hybrid Model with NLP – A model based on deep learning that applies NLP to analyze website content to help in the detection of phishing attempts.

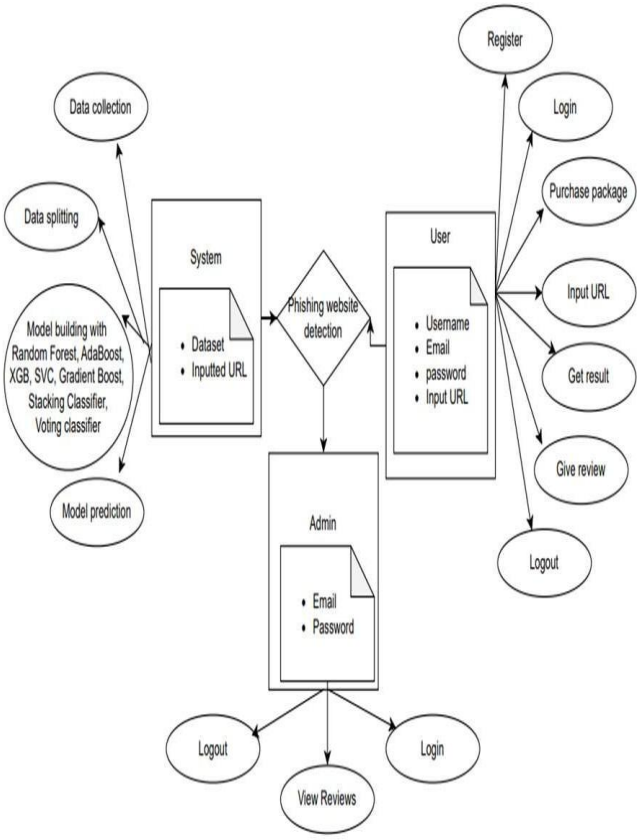

Figure 2. Entity-Relationship Diagram

V. MODEL EVALUATION

The Hacker Eye project conducted a comprehensive comparison of multiple machine learning models for the classification of phishing websites. The models analyzed included Random Forest, XGBoost, AdaBoost**,** Support Vector Machine (SVM), Gradient Boosting, Stacking, and Voting Classifiers. The dataset used for this evaluation consisted of a mix of legitimate and phishing websites, allowing for an in-depth performance assessment based on various evaluation metrics.

To measure the effectiveness of each model, five key performance metrics were utilized: accuracy, precision**,** recall, F1-score, and confusion matrix. The accuracy of the models ranged between 78% and 83%, with ensemble methods such as Voting and Stacking Classifiers achieving the best results. These ensemble models effectively combined multiple classifiers to enhance prediction reliability and mitigate errors.

Among the individual models, XGBoost and Random Forest demonstrated strong detection capabilities, with F1-scores exceeding **0.80**, indicating a high balance between precision and recall. This suggests that these models performed well in distinguishing between legitimate and phishing websites. However, the confusion matrix revealed that certain models struggled with classification. Specifically, SVM and Gradient Boosting were proficient in identifying legitimate websites but encountered challenges in detecting some phishing sites. This limitation may stem from the complexity of phishing website features, which can be highly deceptive and adaptive.

On the other hand, ensemble learning methods such as Stacking and Voting Classifiers proved to be more effective by aggregating the strengths of multiple models. These methods reduced misclassification rates and improved overall detection accuracy. The success of ensemble techniques highlights the importance of model diversity and integration in building robust phishing detection systems.

Despite these promising results, there remains room for improvement. Future work should focus on fine-tuning model parameters, implementing real-time adaptability, and integrating reinforcement learning techniques to enhance detection performance. By refining hyperparameters and incorporating adaptive learning mechanisms, phishing detection systems can become more responsive to evolving threats.

Overall, this study underscores the effectiveness of ensemble-based machine learning approaches in phishing detection. The findings suggest that a combination of multiple classifiers can significantly improve accuracy and resilience against deceptive phishing tactics. The proposed methodology not only enhances detection precision but also provides a scalable solution for securing online environments against cyber threats

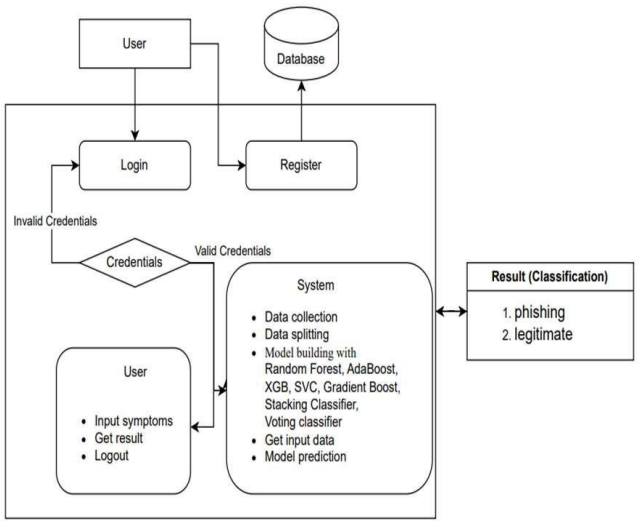

Figure 3. Architectural Diagram

VI. RESULT DISCUSSION

The experiment results indicate that the ensemble of numerous models performs better than a single model in phishing website detection. The top-performing individual model was XGBoost at 83%, followed by Random Forest and Gradient Boosting at 82%. The lowest performing was the Support Vector Classifier (SVC) at 78%.

Both Stacking and Voting classifiers were successful, both achieving 83% accuracy, which indicates that having different models combined together gives power to the detection. The application of Natural Language Processing (NLP) also helped to make it simpler to scan website pages and detect phishing patterns more effectively.

Main Outcomes:

XGBoost Accuracy: 83%

Random Forest Accuracy: 82%

Gradient Boosting Accuracy: 82%

SVC Accuracy: 78%

Stacking and Voting Classifiers Accuracy: 83%

Hybrid NLP Model Accuracy is 85%.

The outcome of the confusion matrix shows that ensemble techniques minimize false negatives and false positives, making them more accurate. The Hybrid NLP model had the highest performance with an accuracy of 85%, which shows that it is highly efficient in detecting phishing websites with clever text instead of suspicious website designs.

VII. CONCLUSION

This research indicates that machine learning approaches, particularly those that utilize combined multiple models, are extremely effective in detecting phishing attacks. These approaches utilize several models to enhance the accuracy with which they predict, reduce false positives, and make phishing detection systems more stable overall. Utilizing ensemble learning, in which several classifiers are combined, enables these systems to learn more effectively in response to new phishing attempts.

Also, applying Natural Language Processing (NLP) techniques to extract features has significantly contributed to enhancing phishing detection. Through the examination of text content, email headers, URLs, and website information, NLP techniques assist in identifying patterns and language features that are commonly associated with phishing attacks. This approach enhances the detection models, making them improved and more precise in distinguishing between real and malicious entities.

In addition to conventional machine learning techniques, observing user behavior has also been a key factor in detecting phishing scams. Through observation of how users react and interact, security systems can identify anomalous behavior indicative of phishing activity. Real-time analysis assists in fortifying defenses that reduce the likelihood of phishing attacks.

In addition, reinforcement learning approaches provide a strong direction for future research. Unlike conventional supervised learning models, which rely on historic data, reinforcement learning allows the system to learn during run-time from experience and adapt to defend against evolving phishing techniques. Continual learning has the potential to significantly enhance the efficacy of phishing detection systems against evolving threats.

VIII. FUTURE WORK

1)Extending the hybrid model to deep

Learning and Reinforcement Learning for Self-Adaptive

phishing detection.

2)Enhancing real-time detection systems with

streaming data analysis techniques.

3)Identification of interdisciplinary techniques that combine

cybersecurity, psychology, and human behaviour

analytics for detecting phishing

4)Analyzing the adversarial effect on

phishing detection models and countermeasures to avert

their impact.

REFERENCES

[1] Abbasi, A., Zhang, Z., & Chen, H. (2019). Deep learning for phishing detection.

[2] Patil, S., & Patil, P. (2018). Comparative study of phishing detection using machine learning.

[3] Zhang, Y., et al. (2020). Evaluation metrics for phishing detection models.

[4] Brown, S., et al. (2021). The role of NLP in cybersecurity threat detection.

[5] Liu, Y., & Wang, J. (2022). Behavioral analysis in phishing attack prevention.

[6] J. Shad & S. Sharma (2018). A Novel Machine Learning Approach to Detect Phishing Websites.

[7] Rishikesh Mahajan, Irfan Siddavatam (2018). Phishing Website Detection using Machine Learning Algorithms.

[8] Verma, R., & Das, R. (2021). Machine Learning Approaches for Phishing Website Detection.