*FUDMA Journal of Sciences (FJS)*
*ISSN online: 2616-1370*
*ISSN print: 2645 - 2944*
*Vol. 9 No. 12, December (Special Issue), 2025, pp 837 – 842*
*DOI: https://doi.org/10.33003/fjs-2025-0912-4356*

# BAYESIAN-OPTIMIZED ENSEMBLE SUPPORT VECTOR MACHINE MODEL FOR PHISHING EMAIL DETECTION

***Aji L. Igba, Ismaila Idris, Sikiru. O. Subairu, Moses. D. Noel and Sulemam Ahmed**

Department of Cyber Security Science, Federal University of Technology, Minna, Niger State, Nigeria.

*Corresponding authors' email: movdanzelaji@gmail.com

**ABSTRACT**

With the rapid growth of email use, phishing and malware attacks have become more frequent and sophisticated, often slipping past traditional defenses such as blacklists and rule-based filters. Existing detection models, including SVM, XGBoost, and CNN, have improved accuracy but still depend heavily on manually crafted features and struggle to adapt to new or evolving attack patterns. This challenge creates the need for a more flexible and intelligent detection approach capable of learning and adapting to emerging email threats. This study aims to develop an ensemble phishing email detection model combining SVM and XGBoost, optimize it using Bayesian tuning, and evaluate its performance through accuracy, precision, recall, F1-score, and ROC-AUC metrics. This study used an ensemble approach that combines SVM and XGBoost to detect phishing emails. Various SVM models, including Baseline, Grid Search, SGD, and Bayesian-optimized versions, were developed and tested. An optimized Bayesian model was developed to improve accuracy, with performance evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. A well-known Kaggle phishing dataset was used for fair comparison. After cleaning and reducing 10,000 emails with 1,250 features to 9,872 emails and 500 cleaned features, the Baseline SVM reached 0.9287 accuracy, Grid Search SVM improved to 0.96, and SGD SVM slightly dropped to 0.92. The Bayesian SVM performed best at 0.9667, showing greater stability and generalization. The Bayesian-optimized Hybrid Ensemble SVM–XGBoost achieved 0.992 accuracy and 0.9992 ROC-AUC, confirming its strong reliability and effectiveness in phishing detection. Stacking substantially enhanced model stability, generalization, and real-time reliability for phishing detection.

**Keywords**: Phishing Detection, Ensemble Learning, Bayesian Optimization

## INTRODUCTION

Phishing continues to pose a major cybersecurity challenge, exploiting social engineering and deceptive website designs to obtain users' credentials, financial data, and other sensitive information. Despite advancements in email filtering systems, attackers constantly modify tactics, making static defenses such as blacklists and rule-based filters inadequate, as reviewed in the research work of Ramesh & Hafeez (2024). Phishing emails are a dangerous subset of spam, aimed at stealing sensitive information. While phishing detection builds on spam filtering, it requires additional checks like URL analysis, sender verification, and user behavior monitoring to effectively protect against security threats, this is credited to the research work of Tusher et al. (2024).

As the payload or content is delivered through that medium to infect systems or steal data. Malware is a growing cyber threat that can steal data, damage systems, or take control of devices. It often spreads through phishing emails, unsafe downloads, or infected USBs, tricking users into clicking links or opening files. Once inside, it can bypass security measures to steal information, lock files, or hijack systems, and can further spread via connected devices, compromised accounts, or websites. Staying safe requires updated software, strong passwords, and cautious online behavior, as asserted in the study Rashid et al., (2025). Recent research has therefore focused on data-driven and intelligent learning approaches capable of adapting to evolving phishing techniques, hence the study Kiseki et al., (2024), analyzed AI-based phishing detection methods, identifying supervised machine learning, particularly the Gradient Boosting Classifier, as highly effective.

However, phishing attackers employ various technical tactics to deceive users and circumvent security systems. These tactics exploit human trust and digital communication channels to spread malicious content or steal sensitive information. Highlight generally that phishing techniques fall into four main categories: link-based, attachment-based, impersonation-based, and hybrid/social-engineering methods. To effectively detect and defend against them, organizations need layered and flexible security measures that combine signature analysis, anomaly detection, and continuous user awareness, in agreement with the study by Abdillah et al. (2022), affirmed that phishing emails frequently impersonate legitimate companies, such as banks, cloud platforms, or email providers, to trick recipients into divulging personal information. According to Butt et al. (2023), these emails can appear so authentic that they are difficult to identify. Recent work uses email features to train models for phishing detection, treating it as a classification task. Particularly when used on carefully prepared, labeled datasets, methods such as SVM, Naive Bayes, and LSTM have proven successful in distinguishing between authentic and fraudulent service emails. Furthermore, the work of Geest et al. (2024), viewed phishing attacks as a serious cybersecurity threat, necessitating more intelligent and flexible defenses, by combining multiple models to improve accuracy, speed, and resilience, this study demonstrates the viability of a hybrid framework for automated phishing detection. The method exhibits significant promise for dependable, real-time phishing protection with a 97.44% accuracy rate and robust resistance to evasion

According to the study by Birthriya et al., (2025), phishing websites are an increasing threat that traditional detection methods often fail to handle. Their research work combines XGBoost with the Bat Algorithm, achieving 94.27% accuracy and demonstrating a reliable and efficient phishing website detection approach. Machine learning has become a key tool in cyber defense, especially for phishing email detection.

Supervised learning using labeled data with algorithms such as Random Forest, SVM, and deep learning accurately identifies known threats but needs frequent retraining to adapt to new ones. Unsupervised learning, by contrast, detects emerging threats by clustering behaviors and identifying anomalies, making it useful for event logs and unstructured data analysis. Combining both methods provides a stronger defense, where supervised learning ensures precision and unsupervised learning adds adaptability and early detection, creating an intelligent, evolving cybersecurity system.

Similarly, the study by Anirudh et al., (2024), has highlighted phishing emails as a major cybersecurity concern that uses deception to steal sensitive data. Their ensemble classification model analyzes email content and structure, adapting to new attack tactics using real-time data while emphasizing user awareness and organizational vigilance. The study by Ntayagabiri et al., (2025), applied machine learning and ensemble methods to detect abnormal behaviors in communication networks. By combining optimized LightGBM and XGBoost models, they enhance accuracy and efficiency in phishing detection, effectively handling large, imbalanced data for real-time, reliable threat identification.

As shown in the study by Sadaram et al., (2023), ensemble learning combines supervised and unsupervised methods to improve phishing detection. Supervised models identify known threats, while unsupervised models detect new or unusual patterns. Together, they create a flexible, adaptive, and precise defense framework that strengthens intrusion and anomaly detection systems.

Chinta et al., (2025), describe phishing as one of the most damaging cyberattacks spread through fake emails and websites. They emphasize deep learning (DL) and machine learning (ML), particularly boosting techniques like XGBoost, which improve accuracy by focusing on previous errors. When integrated with deep learning models such as BERT-LSTM, these approaches achieve over 99% accuracy and F1-scores, showing strong generalization and minimal overfitting compared to conventional models like Naïve Bayes and SVM.

In another novel, the research work of Ibrahim R. B., (2023), addressed the rising issue of phishing emails used to steal confidential information. Their study combines ensemble and machine learning to improve accuracy, using a hybrid filter-wrapper feature selection and a bagging ensemble of Decision Tree (CART), Naïve Bayes, and Logistic Regression. With SMOTE for data balancing, the Decision Tree bagging ensemble achieved 98.13% accuracy, offering a fast, reliable, and less overfitted real-time detection model. Furthermore, this study, Saravana Kumar (2022), introduces an Adaptive Ensemble Learning Framework that merges language understanding, behavior analysis, and deep learning for more accurate phishing detection. Achieving over 96% accuracy, it adapts to new attack patterns and demonstrates strong potential for large-scale, intelligent threat defense.

In a more fundamental study, which has formed the basis for a novel research work, the study paper of Fares et al., (2024), asserted that traditional ML models, such as Decision Tree, Random Forest, Naïve Bayes, and SVM, perform well but face challenges like overfitting, redundant features, and parameter sensitivity. To address this, the study developed a Bayesian-optimized hybrid SVM–XGBoost model. Bayesian optimization fine-tunes SVM hyperparameters, and the optimized classifiers are integrated into a stacking ensemble. Evaluated on phishing email datasets, the model outperformed baseline methods, showing better generalization, scalability, and accuracy for real-time cybersecurity applications.

## MATERIALS AND METHODS

This section outlines the research methods used to enhance the robustness of machine learning in phishing email detection. After setting up the environment, the process included dataset selection, preprocessing, model training, test classification, and performance evaluation. The overall workflow is shown in Figure 1.

### Dataset Description

The dataset, sourced from a publicly available Kaggle phishing email benchmark used in prior research by Fares et al. (2024), contains 10,000 email samples and 1,250 features, including both phishing and legitimate messages. It was split into 80% for training and 20% for testing to ensure fair comparison and model reliability. Chosen for its balance, quality, and relevance, the dataset provided a strong foundation for model training and evaluation.

The dataset consists of two key components:

i. ***Email_Text*:** The email's main content, capturing wording, tone, and structure crucial for identifying deceptive language in phishing emails.
ii. ***Email_Type*:** The label indicating whether an email is phishing or legitimate, guiding the model's learning process.

### Data Preprocessing

Preprocessing improved dataset quality and consistency by removing redundant attributes and handling missing values. Key steps included:

i. Noise removal (extra spaces, punctuation, repeated tokens)
ii. HTML and URL cleaning
iii. Tokenization and TF–IDF vectorization to convert text into numeric features
iv. Dimensionality reduction, refining 1,250 features to 500 using correlation-based selection

These steps produced a compact, discriminative feature set that minimized overfitting and reduced computation time.

### Model Framework

The proposed hybrid model combines two powerful algorithms, Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), within a stacking ensemble framework. SVM effectively handles high-dimensional data and defines strong decision boundaries, while XGBoost provides high accuracy and speed in modeling complex, nonlinear patterns. Bayesian optimization is applied to fine-tune parameters, minimize errors, and boost overall performance. This hybrid stacking ensemble delivers a more accurate, efficient, and reliable phishing email classification system. Figure 1, below is the schematic schematic representation for this research model.

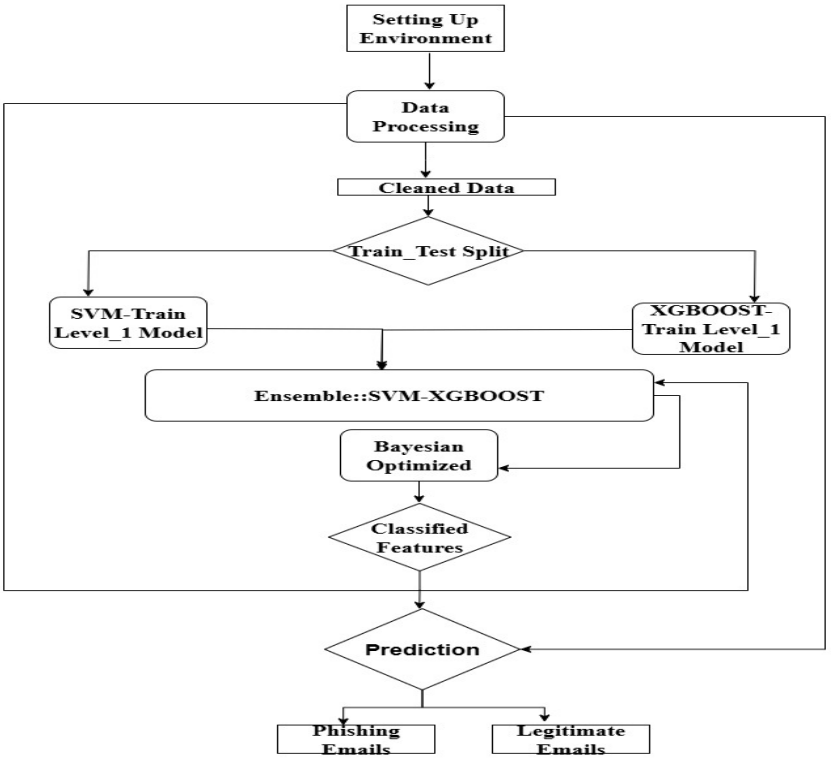


Figure 1: Research Process Framework

i. SVM: Acts as the base learner for high-dimensional feature classification, with kernel type, penalty factor (C), and gamma parameters optimized through Bayesian search.
ii. XGBoost: Delivers efficient gradient boosting with built-in regularization to reduce variance and bias.
iii. Meta-learner: Combines outputs from both models using a Bayesian optimizer to generate the final predictions.

**Bayesian Optimization Process**

The approach of Bayesian optimization efficiently tunes model hyperparameters by balancing exploration of new search areas with exploitation of known high-performing regions. It models the objective function as a Gaussian Process to manage uncertainty and predict the best parameter settings. By maximizing expected improvement, it targets configurations with the greatest potential gain. Unlike manual or grid search methods, it finds near-optimal results faster and with fewer evaluations, making it a smarter and more efficient tuning approach. The algorithmic representation of by the process named Bayesian SVM Algorithm as shown below.

**Algorithm – Bayesian SVM for Phishing Email Detection**

1. Input dataset

$$D = \left\{\left(x_i, y_i\right)\right\}_{i=1}^{N},\ y_i \in \{0,1\}$$

where $x_i$ is the raw email text $y_i$ is the class label (0 = legitmate, 1 = phishing).

2. Data processing
   → Text cleaning: remove stopwords, puntuation, numbers, and convert text to lowercase
   → Tokenization:split email text into words/tokens
   → Vectorization( TF − IDF):

$$TF(t,d) = \frac{f_{t,d}}{\sum f_{u,d}},\ IDF(t) = \log\left(\frac{N}{1+n_t}\right),\ v_d = TF.IDF$$

This produces feature vectors for each email.

3 Data split
Divide dataset into train and test sets:

$$D \rightarrow D_{train}, D_{test\ (e.g.\ 70/30)}$$

4. Bayesian Prior Selection
Place prior on model parameters

$$p(w) \sim N(0, \delta^2 I),\ p(b) \sim N(0, \delta^2)$$

5. Likelihood Function (Bayesian SVM)
*Use hinge − loss likelihood in a probability setting:

$$p(y_i | x_i, w, b) \alpha \exp(-C.\ \max(0, 1 - y, (w^T x_i + b)))$$

6. Preterio Inference
Drive Posterior distribution:

$$P(w,b|D) \alpha P(w) p(b) \prod_{i=1}^{N} p(y_i | x_i, w, b)$$

Appoxmate via MCMC or Variationl Bayes.

7. Prediction
For new email x:

**Stacking an Architecture, SVM, and XGBoost**

This research integrates a stacking ensemble architecture that combines Support Vector Machine (SVM) and XGBoost to establish a comprehensive and efficient phishing email detection framework. The stacking design enhances model performance by exploiting the complementary strengths of both algorithms.

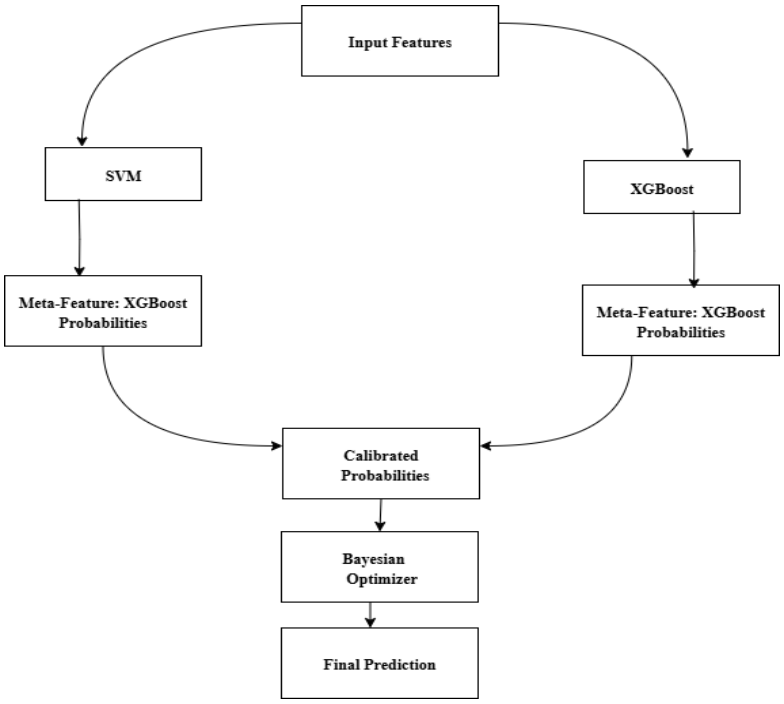


Figure 2: Stacking Architecture Ensemble

SVMs provide strong generalization on high-dimensional text data, making them effective at detecting phishing emails. It performs well with limited data and captures subtle textual cues of malicious intent, but is computationally intensive, sensitive to hyperparameters, and less scalable for real-time use. XGBoost offers fast, scalable learning, capturing complex non-linear patterns through gradient boosting and regularization to reduce overfitting, though it can be memory-heavy and less interpretable.

The proposed stacking framework combines SVM's generalization with XGBoost's efficiency through a meta-learner, while Bayesian optimization fine-tunes hyperparameters to balance bias and variance. This hybrid design enhances accuracy, adaptability, and scalability, making it a strong phishing detection model for real-world use (Figure 2).

**Evaluation Metrics**

Model performance was measured using standard classification metrics: Accuracy, Precision, Recall, F1-score, and ROC-AUC. These metrics provide insight into detection capability, generalization, and false-positive control. Accuracy indicates overall correctness; Precision reflects false alarm reduction; Recall measures successful detection of phishing; and ROC-AUC quantifies discriminative strength.

**RESULTS AND DISCUSSION**

The experiment was performed using Python-based machine learning libraries on a workstation with 16 GB RAM and an Intel Core i7 processor. All SVM variants and the hybrid model were trained under identical settings for fair comparison. The baseline SVM achieved 92.87% accuracy, serving as the benchmark. Grid Search tuning raised accuracy to 96.00%, highlighting the value of parameter optimization. The SGD SVM achieved 92.00%, slightly lower due to its sensitivity to learning rate and scaling. Bayesian Optimization further improved accuracy to 96.67%, enhancing generalization and stability. When integrated with XGBoost in a hybrid stacking ensemble, the model achieved 99.2% accuracy and a ROC-AUC of 0.9992, outperforming all individual classifiers.

**Proposed Hybrid Optimized Phishing Detection Model**

Table 1. presents the performance of the proposed hybrid phishing detection model, showing outstanding classification results. The model achieved about 99% precision and recall, accurately identifying both phishing and legitimate emails with minimal errors. With an F1-score of 99.1% and an overall accuracy of 99.2%, it demonstrates high reliability and consistency in distinguishing phishing emails from genuine ones.

**Table 1: Proposed Model Performance**

| Class | Precision(%) | Recall(%) | F1-Score(%) | Support(%) |
|---|---|---|---|---|
| Legitimate (0) | 99.2 | 99.0 | 99.1 | 500 |
| Phishing (1) | 99.0 | 99.0 | 99.1 | 500 |
| Accuracy | 99.2 | | | 1000 |
| Macro Avg. | 99.1 | 99.1 | 99.1 | 1000 |
| Weight Avg. | 99.1 | 99.2 | 99.1 | 1000 |

**Confusion Matrix**

The confusion matrix shows that the hybrid ensemble model effectively distinguished phishing from legitimate emails, with TN = 4952, FP = 48, FN = 32, and TP = 4968. It made a few misclassifications, reflecting strong precision, recall, and overall accuracy. As shown in Figure 3.3, its low error rates highlight the model's reliability, efficiency, and suitability for real-world cybersecurity applications**.**

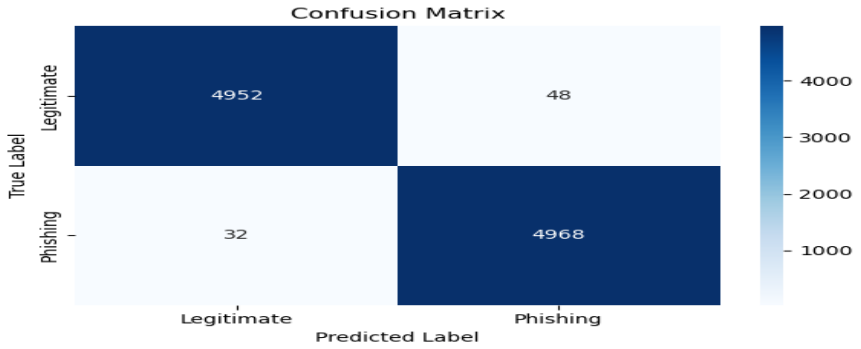


Figure 3: Confusion Matrix for Hybrid

**Comparative Analysis**

Table 2 shows that the proposed hybrid ensemble model outperformed all others. The baseline linear model achieved 92.9% accuracy, the Grid Search SVM reached 96.0%, and the Bayesian Optimized model reached 96.7%. The hybrid ensemble achieved the highest accuracy of 99.2%, a 6.3% improvement over the baseline, confirming its superior effectiveness in phishing detection.

**Table 2: Result Comparison with Baseline Models**

| Model | Accuracy (%) | Improvement over baseline (%) |
|---|---|---|
| Baseline (linear) | 92.9 | - |
| Grid Search SVM | 96.0 | 3.1 |
| Bayesian Optimized | 96.7 | 3.8 |
| Hybrid Ensemble (Ours) | 99.2 | 6.3 |

**Evaluation of Model Performance Metrics**

As shown in Table 3.4 (Performance Metrics), the phishing detection model achieved excellent accuracy, precision, and recall, effectively identifying phishing emails with minimal false alerts. Its high ROC-AUC confirms a reliable distinction between legitimate and malicious messages, making the model efficient and suitable for real-world use.

**Table 3*:* Model Performance Metrics**

| Metric | Score (%) | Interpretation |
|---|---|---|
| Accuracy | 99.2 | The model correctly identified almost all emails, whether phishing or legitimate. This shows a strong overall performance and reliability. |
| Precision (phishing | 99 | The model rarely flagged genuine emails as phishing, which means it keeps false alarms very low. This helps maintain smooth and uninterrupted communication. |
| Recall (phishing) | 99.2 | The model successfully detected nearly every phishing attempt, missing very few. This is vital for security since any missed phishing email could pose a serious threat. |
| ROC-AUC | 99.92 | The model shows an almost perfect ability to tell phishing emails apart from legitimate ones, even under different testing conditions. |

**Discussion**

The proposed phishing detection model showed outstanding performance in distinguishing phishing from legitimate emails. With a False Positive Rate of 0.96% and a False Negative Rate of 0.64%, it rarely misclassifies genuine messages or misses phishing attempts, ensuring both accuracy and security. Its high Recall (99.36%), Specificity (99.04%), ROC-AUC (0.9992), and Average Precision (0.9989) confirm near-perfect classification capability.

Compared to earlier models like SVM, Naïve Bayes, and Decision Trees (90–95% accuracy), the Bayesian-optimized ensemble improved SVM performance by 6.3%, highlighting the value of optimization in boosting accuracy. The model's near-zero false negatives ensure robust protection against phishing threats, while the low false positives maintain smooth, reliable communication, making it both secure and user-friendly.

**Contribution to Knowledge**

This research makes several important contributions to the field of phishing detection. Firstly, it shows how Bayesian optimization improved machine learning models for peak performance. The findings show that optimized ensemble learning outperforms single classifiers by leveraging their collective strengths. The study also proves that a hybrid model can achieve near-perfect accuracy with minimal error. Crucially, this approach solves a major industry hurdle, reducing false negatives without hurting precision, ensuring tight security remains user-friendly. Ultimately, this framework offers a reliable, high-precision solution for modern cybersecurity.

**CONCLUSION**

This study demonstrated that hyperparameter optimization greatly enhances phishing email detection. Using Bayesian

optimization with an RBF kernel, the baseline SVM model's accuracy increased from 0.9287 to 0.9667, emphasizing the importance of kernel selection and tuning. The optimized hybrid stacking ensemble of SVM and XGBoost achieved 99.2% accuracy, 0.994 precision and recall, a ROC-AUC of 0.9992, and minimal misclassification (False Positive Rate 0.096, False Negative Rate 0.064). High True Positive (0.9936) and True Negative (0.9904) rates confirm its reliability and suitability for real-world phishing detection. These results highlight the effectiveness of ensemble learning combined with Bayesian optimization. Future improvements should focus on explainable AI (XAI) to make model decisions transparent and build user trust, as well as real-time adaptability to counter evolving phishing tactics. Scalability is also essential to efficiently handle large email streams across enterprise and cloud platforms. By integrating accuracy, explainability, adaptability, and scalability, this research demonstrates a practical, resilient, and trustworthy solution for real-time phishing detection.

## REFERENCES


Abdillah, R., Shukur, Z., Mohd, M., & Murah, T. M. Z. (2022). Phishing Classification Techniques: A Systematic Literature Review. In *IEEE Access* (Vol. 10, pp. 41574–41591). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3166474

Alam, S., Jameel, A., Parveen, Z., & Alnfrawy, E. (2025). *Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000. SHRED: An Ensemble-Based Machine Learning Model to Sift Email Messages for Real-Time Spam Detection*. https://doi.org/10.1109/ACCESS.2025.DOI

Anirudh, S., Radha Nishant, P., Baitha, S., & Dinesh Kumar, K. (2024). An Ensemble Classification Model for Phishing Mail Detection. *Procedia Computer Science*, *233*, 970–978. https://doi.org/10.1016/j.procs.2024.03.286

Birthriya, S. K., Ahlawat, P., & Jain, A. K. (2025). Phishing Website Detection with XGBoost and Adaptive Hyperparameter Optimization using the Bat Algorithm. *Procedia Computer Science*, *258*, 1774–1782. https://doi.org/10.1016/j.procs.2025.04.429

Butt, U. A., Amin, R., Aldabbas, H., Mohan, S., Alouffi, B., & Ahmadian, A. (2023). Cloud-based email phishing attack using machine and deep learning algorithm. *Complex and Intelligent Systems*, *9*(3), 3043–3070. https://doi.org/10.1007/s40747-022-00760-3

Chinta, P. C. R., Moore, C. S., Karaka, L. M., Sakuru, M., Bodepudi, V., & Maka, S. R. (2025). Building an Intelligent Phishing Email Detection System Using Machine Learning and Feature Engineering. *European Journal of Applied Science, Engineering and Technology*, *3*(2), 41–54. https://doi.org/10.59324/ejaset.2025.3(2).04

Fares, H., Kilani, J., Fagroud, F. E., Toumi, H., Lakrami, F., Baddi, Y., & Aknin, N. (2024). Machine Learning Approach for Email Phishing Detection. *Procedia Computer Science*, *251*, 746–751. https://doi.org/10.1016/j.procs.2024.11.179

Ibrahim R. B, M. S. A. I. M. U. (2023). *Development of an Ensemble Classification Model Based On Hybrid Filter-Wrapper Feature Selection for Email Phshing Detection*.

Iqbal, A., Younas, M., Iftikhar, S., Fatima, F., & Saleem, R. (2025). Spam detection using hybrid model on fusion of spammer behavior and linguistics features. *Egyptian Informatics Journal*, *29*. https://doi.org/10.1016/j.eij.2024.100605

Kalabarige, L. R., Rao, R. S., Abraham, A., & Gabralla, L. A. (2022). Multilayer Stacked Ensemble Learning Model to Detect Phishing Websites. *IEEE Access*, *10*, 79543–79552. https://doi.org/10.1109/ACCESS.2022.3194672

Kiseki, D. W., Havyarimana, V., Zabagunda, D. L., Wail, W. I., & Niyonsaba, T. (2024). Artificial Intelligence in Cybersecurity to Detect Phishing. *Journal of Computer and Communications*, *12*(12), 91–115. https://doi.org/10.4236/jcc.2024.1212007

Ntayagabiri, J. P., Bentaleb, Y., Ndikumagenge, J., & El Makhtoum, H. (2025). OMIC: A Bagging-Based Ensemble Learning Framework for Large-Scale IoT Intrusion Detection. *Journal of Future Artificial Intelligence and Technologies*, *1*(4), 401–416. https://doi.org/10.62411/faith.3048-3719-63

Qiqieh, I., Alzubi, O., Alzubi, J., Sreedhar, K. C., & Al-Zoubi, A. M. (2024). An intelligent cyber threat detection: A swarm-optimized machine learning approach. *Alexandria Engineering Journal*. https://doi.org/10.1016/j.aej.2024.12.039

Ramesh, K., & Hafeez, K. (2024). *Phishing Detection and Mitigation: A Cybersecurity and Machine Learning Approach MSc Research Project MSc Cyber Security*.

Rashid, M. U., Qureshi, S., Abid, A., Alqahtany, S. S., Alqazzaz, A., ul Hassan, M., Al Reshan, M. S., & Shaikh, A. (2025). Hybrid Android Malware Detection and Classification Using Deep Neural Networks. *International Journal of Computational Intelligence Systems*, *18*(1). https://doi.org/10.1007/s44196-025-00783-x

Sankaine, L., Ndia, J. G., & Kaburu, D. (2025). An English-Swahili Email Spam Detection Model for Improved Accuracy Using Convolutional Neural Networks. *Mesopotamian Journal of CyberSecurity*, *5*(2), 590–605. https://doi.org/10.58496/MJCS/2025/036

Saravana Kumar, S. (2022). *Adaptive Ensemble Learning Framework for Evolving Social Engineering Threats*. www.ijesat.com

Tusher, E. H., Ismail, M. A., Rahman, M. A., Alenezi, A. H., & Uddin, M. (2024). Email Spam: A Comprehensive Review of Optimize Detection Methods, Challenges, and Open Research Problems. *IEEE Access*, *12*, 143627–143657. https://doi.org/10.1109/ACCESS.2024.3467996

van Geest, R. J., Cascavilla, G., Hulstijn, J., & Zannone, N. (2024). The applicability of a hybrid framework for automated phishing detection. *Computers and Security*, *139*. https://doi.org/10.1016/j.cose.2024.103736