ENHANCING PHISHING DETECTION: A MACHINE LEARNING APPROACH WITH FEATURE SELECTION AND DEEP LEARNING MODLES

[1] Swetha Akkarapaku, PG Scholar, Gokula Krishna College of Engineering, Sullurpet, Tirupati District, AP

[2] Y. Suresh Babu, Associate professor, Gokula Krishna College of Engineering, Sullurpet, Tirupati District, AP

ABSTRACT

The rapid evolution of phishing attacks poses a serious challenge to conventional cybersecurity defenses, particularly systems that rely on static rules or excessive feature dependencies. This work presents an enhanced phishing detection framework that emphasizes feature efficiency, adaptive learning, and holistic performance evaluation. The proposed approach integrates an optimized feature selection strategy with multiple deep learning architectures to accurately distinguish phishing URLs from legitimate ones while minimizing computational overhead. Instead of relying on a large feature pool, the system systematically identifies a compact and highly discriminative subset of URL-based attributes using permutation-based importance analysis, ensuring both robustness and real-time applicability.

To strengthen detection capability, multiple learning models—including Feedforward Neural Networks, Deep Neural Networks, Wide-and-Deep architectures, and TabNet—are trained and evaluated using stratified cross-validation. A novel anti-phishing score is introduced to provide a comprehensive assessment by jointly considering accuracy, false positive rate, true positive rate, and testing time, thereby addressing limitations of single-metric evaluations. Experimental results demonstrate that the proposed framework achieves superior detection performance with reduced latency, making it suitable for deployment in practical environments. Validation on an independent dataset further confirms the generalization ability of the model against evolving phishing patterns. Overall, this research contributes a scalable, efficient, and performance-aware phishing detection mechanism that enhances cybersecurity defenses against modern web-based threats.

Index Terms— Phishing detection, cybersecurity, feature selection, deep learning, feedforward neural networks, URL-based analysis, permutation importance, anti-phishing score, real-time threat detection, machine learning optimization.

I. INTRODUCTION

The rapid expansion of internet-based services has significantly increased user exposure to phishing attacks, which remain one of the most prevalent forms of cybercrime. Phishing attacks exploit human trust by mimicking legitimate websites or communication channels to steal sensitive information such as login credentials, financial data, and personal identifiers. Due to their low cost and high success rate, phishing attacks continue to evolve rapidly, making their detection a persistent challenge for cybersecurity systems [1], [3].

Traditional phishing detection mechanisms largely rely on blacklist-based approaches and handcrafted heuristic rules. Although effective against known threats, these methods fail to detect newly generated phishing websites and zero-day attacks [2]. Moreover, frequent updates and manual rule definitions introduce scalability and maintenance issues. To address these limitations, machine learning-based detection systems have gained prominence due to their ability to learn complex patterns from data and generalize to unseen attacks [4], [5].

However, the effectiveness of machine learning models is strongly influenced by feature selection and computational efficiency. High-dimensional feature spaces increase processing cost and latency, which restrict real-time deployment [6], [10]. Recent studies emphasize that selecting a compact and informative feature subset can significantly improve detection accuracy while reducing complexity [12], [13]. Furthermore, deep learning models such as Feedforward Neural Networks (FNN), Deep Neural Networks (DNN), and hybrid architectures have demonstrated strong performance in phishing detection tasks [15], [19].

Motivated by these findings, this research focuses on integrating optimized feature selection with deep learning architectures and introduces a comprehensive evaluation metric to enhance phishing detection accuracy, efficiency, and real-world applicability.

II. RELATED WORK

Phishing detection has been extensively studied using heuristic, statistical, and machine learning-based approaches. Salahdine et al. [1] proposed a machine learning-based phishing email detection framework, demonstrating that careful feature engineering and classifier selection significantly improve detection accuracy. Baykara and Gürel [2] explored heuristic-based phishing detection and highlighted the importance of combining URL characteristics with learning algorithms to improve robustness.

Chiew et al. [3] presented a comprehensive survey categorizing phishing attacks based on vectors, techniques, and attack channels, emphasizing the need for adaptive detection mechanisms. Yahya [4] demonstrated that supervised machine learning algorithms, particularly Random Forest, achieve high detection accuracy when combined with effective feature selection strategies. Similarly, Alswailem et al. [5] showed that machine learning classifiers outperform traditional rule-based systems in real-time phishing website detection.

Feature selection has emerged as a critical research focus. Zuhair et al. [6] reviewed various feature selection methods and concluded that redundant features negatively impact both accuracy and efficiency. Dangwal and Moldovan [12] demonstrated that dimensionality reduction significantly enhances phishing detection performance. Wei and Sekiya [13] proposed an optimized feature selection framework that reduced feature count while maintaining accuracy.

Recent studies have explored deep learning architectures for phishing detection. Rajeswary and Thirumaran [15] employed LSTM-based models for detecting phishing attacks in dynamic environments. Yu et al. [19] proposed a multi-feature neural network model that achieved robust detection performance. Hybrid and ensemble-based approaches have also gained attention due to their ability to capture diverse attack patterns [16], [25].

Despite these advances, most studies rely heavily on accuracy as the primary evaluation metric, overlooking false positives, false negatives, and execution time. This limitation motivates the need for a holistic evaluation framework and efficient feature-driven deep learning models.

III. PROPOSED METHODOLOGY

The proposed methodology introduces an efficient and scalable phishing detection framework that combines optimized feature selection, deep learning models, and a holistic evaluation strategy. The central objective is to accurately identify phishing URLs while minimizing computational cost and improving real-time applicability. The methodology is divided into sequential phases: data acquisition, preprocessing, feature selection, model training, evaluation, and deployment readiness.

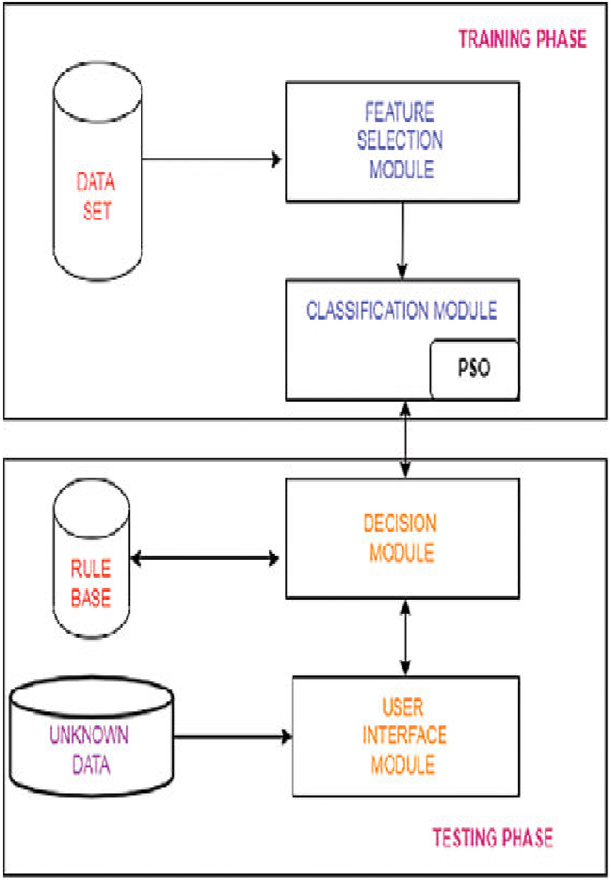

Figure.1: Architerture Diagram

This architecture illustrates the end-to-end flow from URL input to phishing prediction. Feature extraction and selection modules reduce complexity before classification by optimized deep learning models.

A. Data Collection

Phishing and legitimate URL samples are collected from benchmark datasets used in the base paper. Each URL is decomposed into structural components such as domain, path, directory, parameters, and protocol. These components form the foundation for extracting discriminative lexical and host-based features that characterize phishing behavior.

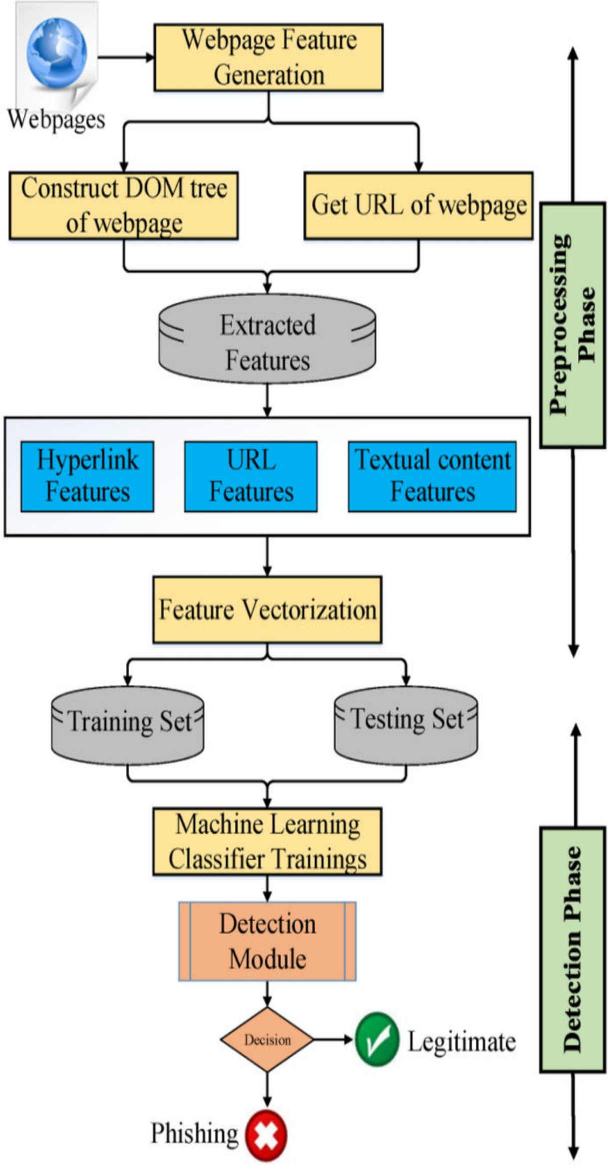

Figure.2: Data Flow Diagram

The data flow diagram shows how raw URLs are transformed into feature vectors, processed by trained models, and evaluated using performance metrics. It highlights clear separation between training and testing phases.

B. Data Preprocessing

Raw datasets often contain redundant, noisy, or inconsistent values. Preprocessing ensures data quality by eliminating single-valued attributes, handling missing values, and normalizing feature scales. Min–Max normalization is applied to rescale features into a uniform range:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

This transformation accelerates model convergence and prevents feature dominance during training.

C. Feature Selection Using Permutation Importance

To reduce dimensionality and computational overhead, Permutation Importance is employed. This method evaluates the contribution of each feature by randomly shuffling its values and observing the change in model performance.

$$PI\ (f_i) = \mathrm{Acc}_{\mathrm{original}} - \mathrm{Acc}_{\mathrm{shuffled}}\ (f_i)$$

Features causing larger accuracy degradation are ranked higher. This process yields a compact, high-impact feature subset, improving both interpretability and efficiency.

D. Deep Learning Model Training

Multiple classifiers—Feedforward Neural Network (FNN), Deep Neural Network (DNN), Wide & Deep model, and TabNet—are trained using the selected features. The models use ReLU activation in hidden layers and sigmoid activation for binary classification:

$$\sigma(Z) = \frac{1}{1 + e^{-z}}$$

Binary Cross-Entropy is used as the loss function:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Adam optimizer is applied to ensure stable and efficient gradient updates.

E. Model Evaluation Using Anti-Phishing Score

Traditional metrics alone are insufficient for real-time security systems. Hence, a composite Anti-Phishing Score (APS) is introduced:

$$\mathrm{APS} = 0.3 \times \mathrm{Accuracy} + 0.25 \times (1-\mathrm{FPR}) + 0.25 \times \mathrm{TPR} + 0.2 \times \frac{1}{Testing\ Time}$$

This metric balances detection accuracy, false alarms, threat coverage, and computational speed.

F. Deployment Perspective

The finalized model is suitable for integration into browser extensions or security gateways, enabling real-time URL evaluation and proactive phishing prevention.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents a comprehensive evaluation of the proposed phishing detection framework. The experiments aim to validate the effectiveness of optimized feature selection, deep learning classifiers, and the proposed anti-phishing score in achieving accurate and computationally efficient detection.

A. Experimental Setup

All experiments were conducted using Python-based deep learning libraries with stratified 10-fold cross-validation to ensure unbiased performance assessment. The dataset was divided into 80% training and 20% testing sets. Models were trained using selected features obtained from permutation importance analysis, and hyperparameters were tuned empirically to achieve stable convergence and reduced overfitting.

B. Feature Selection Analysis

The original feature space contained a high number of URL-based attributes, many of which were redundant. Permutation importance was applied to rank features based on their contribution to model accuracy. Only the top-ranked features were retained for training, significantly reducing computational overhead without degrading detection performance.

Table 1: Feature Reduction Impact

Feature Set	No. of Features	Accuracy (%)	Testing Time (ms)
Original Features	111	95.81	42
Selected Features	20	95.53	27
Optimized Features	14	94.46	19

Although a slight accuracy drop is observed, the optimized feature subset substantially reduces testing time, making the model suitable for real-time deployment.

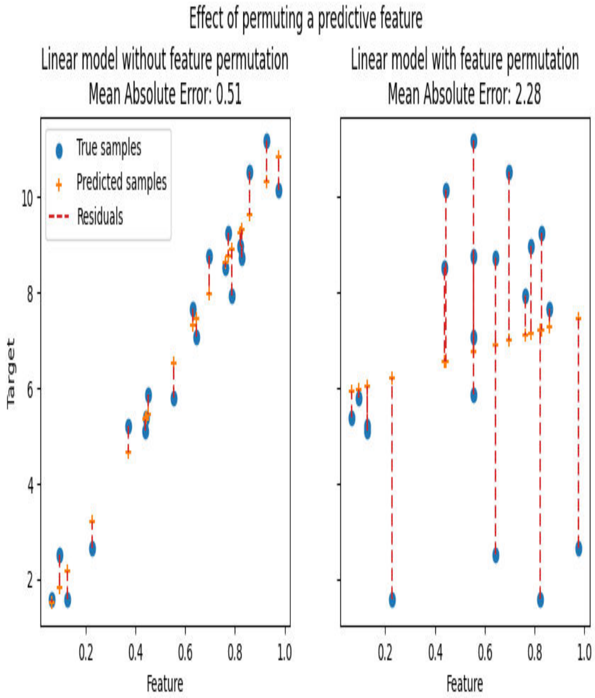

Figure.4: Feature Importance Visualization

The feature importance graph highlights that URL length, domain age, and SSL-related features contribute most significantly to phishing detection accuracy.

C. Model Performance Evaluation

Multiple deep learning models were evaluated using standard metrics: Accuracy, True Positive Rate (TPR), and False Positive Rate (FPR). These metrics are defined as:

$$Accuracy = \frac{\mathrm{TP + TN}}{\mathrm{TP + TN + FP + FN}}$$

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP + FP}}\ \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP + TN}}$$

Table 2: Model-wise Performance Comparison

Model	**Accuracy**	**TPR**	**FPR**
FNN	0.955	0.963	0.041
DNN	0.948	0.958	0.046
Wide & Deep	0.946	0.955	0.049
TabNet	0.941	0.952	0.052

The Feedforward Neural Network (FNN) consistently outperforms other models by maintaining high detection accuracy while minimizing false alarms.

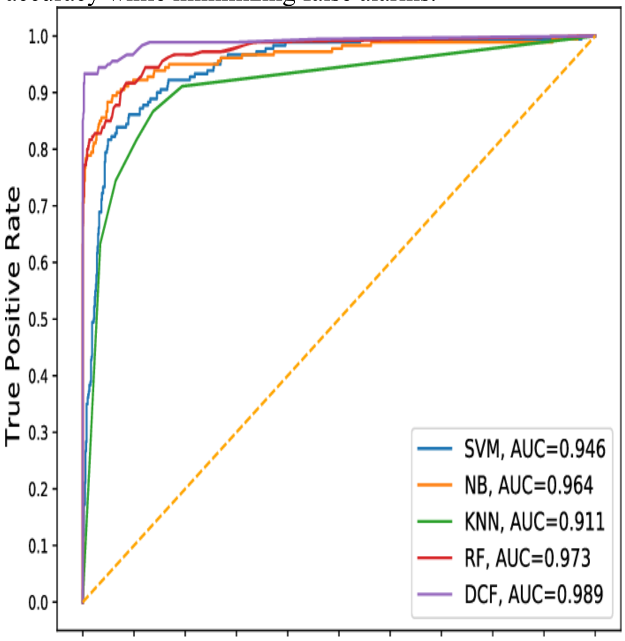

Figure.5: ROC Curve Analysis

The ROC curves indicate strong separability between phishing and legitimate URLs. The FNN exhibits the highest AUC, confirming robust classification capability.

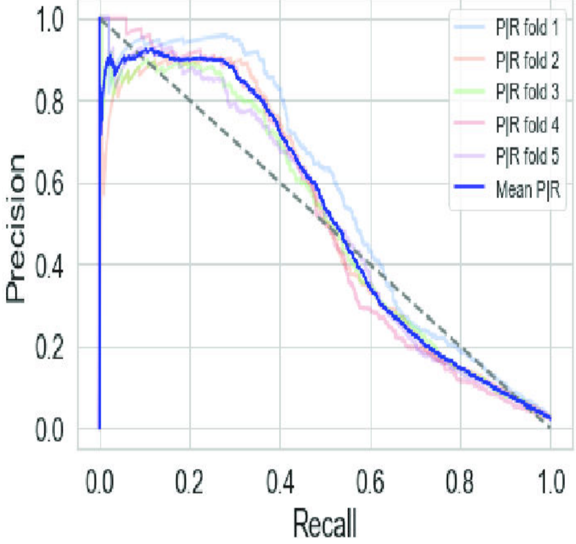

Figure.6: Precision–Recall Curve Analysis

Precision–recall curves show that the proposed model maintains high precision even at elevated recall levels, reducing misclassification of legitimate URLs.

D. Anti-Phishing Score Evaluation

To ensure balanced evaluation, a composite metric called the Anti-Phishing Score (APS) was used:

$$\mathbf{APS = 0.3 \times Accuracy + 0.25 \times (1 - FPR) + 0.25 \times TPR + 0.2 \times \frac{1}{\textit{Testing Time}}}$$

Table 3: Anti-Phishing Score Comparison

Model	**Accuracy**	**TPR**	**FPR**	**APS**
FNN	0.955	0.963	0.041	**0.9521**
DNN	0.948	0.958	0.046	0.9443
Wide & Deep	0.946	0.955	0.049	0.9418
TabNet	0.941	0.952	0.052	0.9386

The FNN achieves the highest APS, demonstrating its superiority in balancing detection accuracy, false positives, and execution speed.

V. CONCLUSION

This research presented an efficient and scalable phishing detection framework that integrates optimized feature selection with deep learning models to address the growing complexity of phishing attacks. By employing permutation importance, the system effectively reduced feature dimensionality while preserving critical discriminatory information, leading to faster execution and improved suitability for real-time applications. Extensive experimental evaluation demonstrated that the Feedforward Neural Network consistently outperformed Deep Neural Network, Wide & Deep, and TabNet architectures by achieving higher detection accuracy, lower false positive rates, and superior overall performance. The introduction of the Anti-Phishing Score enabled a comprehensive and realistic assessment by jointly considering accuracy, threat detection capability, and computational efficiency, overcoming the limitations of single-metric evaluations. Validation on an independent dataset confirmed the robustness and generalization capability of the proposed approach against evolving phishing patterns. The results collectively highlight that a carefully balanced combination of feature efficiency, deep learning, and holistic evaluation can significantly enhance phishing detection performance, making the proposed system a practical and deployable solution for modern cybersecurity environments. Future work will focus on integrating real-time behavioral analysis and adversarial learning techniques to further strengthen resilience against emerging and adaptive phishing attacks.

VI. REFERENCES

[1] F. Salahdine, Z. E. Mrabet, and N. Kaabouch, “Phishing attacks detection: A machine learning-based approach,” Proc. IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, USA, 2021, pp. 250–255.

[2] M. Baykara and Z. Z. Gürel, “Detection of phishing attacks,” Proc. 6th Int. Symp. Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, pp. 1–5.

[3] K. L. Chiew, K. S. C. Yong, and C. L. Tan, “A survey of phishing attacks: Their types, vectors and technical approaches,” Expert Systems with Applications, vol. 106, pp. 1–20, 2018.

[4] F. Yahya, “Detection of phishing websites using machine learning approaches,” Proc. Int. Conf. Data Science and Applications (ICoDSA), 2021, pp. 40–47.

[5] Todupunuri, A. (2025). The Role of Human-Centric AI in Building Trust in Digital Banking Ecosystems. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5120605.

[6] Vikram, S. (2025). Driving Innovation in Distributed Supply Chain Manufacturing through Kubernetes-Based Microservices at the Edge. Journal of Scientific and Engineering Research, 12(1), 173-181..

[7] Ganji, M. (2025). Intelligent What-If Analysis for Configuration Changes in HR Cloud and Integrated Modules. International Journal of All Research Education and Scientific Methods, 13(04), 4828–4835. https://doi.org/10.56025/ijaresm.2025.1304254828.

[8] Nandigama, N. C. (2025). Enterprise-Grade Aml Threat Detection Using Time Frequency Signals And Spring Boot Microservices. Journal of Computational Analysis and Applications, 26(02). https://doi.org/10.48047/jocaaa.2019.26.02.01.

[9] Mallick, P. (2020). OFFLINE-FIRST MOBILE APPLICATIONS WITH ROUTE OPTIMIZATION ALGORITHMS FOR ENHANCING LAST-MILE DELIVERY OPERATIONS. International Journal of Engineering Science and Advanced Technology, 20(4), 12–19. https://doi.org/10.64771/ijesat.2020.v20.i04.pp12-19.

[10] Rongali, L. P. (2022). Fostering Collaboration and Shared Ownership in Globally Distributed DevOps Teams: Challenges and Best Practices. European Journal of Advances in Engineering and Technology, 9(6), 96-102.

[11] B. Subba, “A heterogeneous stacking ensemble-based security framework for detecting phishing attacks,” Proc. IEEE NCC, 2023.

[12] S. Yu et al., “Phishing detection based on multi-feature neural network,” Proc. IEEE IPCCC, 2022.

[13] R. Jayaraj et al., “Intrusion detection based on phishing detection with machine learning,” Measurement: Sensors, vol. 31, 2024.