# A Hybrid Machine Learning and Reinforcement Learning Approach for Adaptive Phishing Email Detection

1st Mrs. Kodali Lakshmi Padmavathi
Department of Artificial Intelligence and Data Science
Lakireddy Balireddy College Of Engineering
Mylavaram, Andhra Pradesh, India
lakshmik@lbrce.ac.in

2nd Bhedam Pavan Kumar
Department of Artificial Intelligence and Data Science
Lakireddy Balireddy College Of Engineering
Mylavaram, Andhra Pradesh, India
pavankumarbhedam@gmail.com

3rd Dangatla Bhanu Sri
Department of Artificial Intelligence and Data Science
Lakireddy Balireddy College Of Engineering
Mylavaram, Andhra Pradesh, India
dangatlabhanusri@gmail.com

4th Kathula Jaithra
Department of Artificial Intelligence and Data Science
Lakireddy Balireddy College Of Engineering
Mylavaram, Andhra Pradesh, India
jaithrakathula3@gmail.com

***Abstract*— Phishing emails are the fake or fraudulent emails that are designed to trick people into giving personal information which may includes passwords or credit card details or any other sensitive data .The goal of our project is to create a system that can automatically identify these fake emails and filter them out, and making email communication more safer .The system will analyse the content and patterns in emails to differentiate between the legitimate and the fraudulent emails. This project aims to develop a robust, automated system for identifying and filtering phishing emails, enhancing the security of email communication. Our approach combines machine learning techniques, specifically Random Forest classification, and reinforcement learning via Deep Q-Networks (DQN), to analyse the metadata, textual content, and embedded URLs of emails, discerning legitimate messages from phishing attempts. Through a feedback-driven mechanism, the model can learn from incorrect predictions and adapt over time, providing an evolving and proactive defence against phishing threats. This system holds significant potential for deployment in both corporate and personal email environments, contributing to safer email practices across diverse use cases.**

***Index Terms*—Phishing emails, Fraudulent emails, Sensitive data, Email security, Data analysis, Random Forest, Reinforcement Learning, Deep Q-Network (DQN), Feedback-driven learning, Metadata analysis, Textual content analysis, Adaptability**

## I. INTRODUCTION

### *A. Background Information*

Phishing emails have become one of the most common forms of cyberattacks, targeting individuals and organizations with the intent of stealing sensitive information such as usernames, passwords, and financial data. As digital communication continues to expand, phishing attacks have grown more sophisticated, exploiting unsuspecting email users by mimicking trusted sources. This therefore calls for the need for dependable, automatic systems that can accurately identify and prevent phishing attacks with precision, especially in a scenario where large volumes of email traffic make manual detection infeasible.

Traditional approaches to detecting phishing normally employ rule-based systems or blacklists. Yet, both fail to maintain pace with the ever-changing tactics of phishers. Usually, the lists need updating, a task that proves to be very resource intensive and rather inefficient. Lately, techniques of machine learning seem promising in automatically defining phishing characters in models, based on patterns in the data. Although effective, ML methods alone are limited by their static nature and make it hard to update the systems in real-time in accordance with new phishing strategies.

### *B. Research Problem:*

While machine learning has demonstrated potential in the detection of phishing, its static nature hinders adaptability to novel phishing strategies and rapidly evolving attack patterns. Traditional rule-based systems and blacklists are resource-intensive and inefficient, which further complicates the challenge of maintaining effective detection systems.

There is a need for a system that achieves high accuracy in detecting phishing emails but at the same time dynamically adapts to emerging threats in real-time. Existing ML techniques are limited in addressing this adaptability, thus a hybrid approach is required which combines static classification with dynamic learning mechanisms to handle evolving phishing tactics.

### *C. Significance of the Research*

This paper describes a novel hybrid phishing detection model, bridging traditional machine learning and adaptive reinforcement learning, thereby providing a comprehensive approach to solving the challenges of phishing attacks. It uses a Random Forest classifier for an accurate and robust primary

detection while integrating a DQN that enables adaptability through feedback-driven reinforcement learning.

The primary contributions of this research include:

1. A hybrid detection model that combines the strengths of Random Forest and DQN, offering both accuracy and adaptability.
2. A feedback mechanism that enables continuous improvement by learning from incorrect predictions and identifying new phishing patterns.
3. Evaluation of the model across diverse datasets, demonstrating its capability to maintain high accuracy while adapting to novel phishing attempts.

By treating phishing detection as a sequential decision-making problem, the DQN component ensures that the system learns optimal detection policies and iteratively refines its criteria over time. This adaptability is critical to handling the diverse and rapidly changing nature of phishing techniques, thereby significantly reducing the risk of misclassification.

This work contributes to the development of intelligent, self-improving cybersecurity systems that address current limitations and proactively face future challenges, which is in turn a contribution to advancing the phishing detection field toward more resilient and autonomous threat detection solutions.

## II. LITERATURE SURVEY

### *A. Overview of relevant literature :*

Lingampally Shalini, Sunil Kumar S. Manvi, Naveen Chandra Gowda, Manasa K N [1],This paper addresses phishing email detection, a critical cybersecurity challenge that can cause significant financial and data losses. It examines the use of machine learning (ML) and deep learning (DL) algorithms to detect and classify phishing emails based on textual and metadata features. The study splits the dataset into training and testing sets with 70-30% and 80-20% ratios. It compares the performance of variousprocess: first, a sample expansion stage using KNN and K-Means algorithms to augment the dataset, and second, a testing stage where the LSTM model is trained on preprocessed data, including word segmentation and vector generation. The model's performance is assessed using metrics such as accuracy, highlighting a 95% detection rate. The findings demonstrate the effectiveness of LSTM in handling large-scale email data, emphasizing the model's superiority in adapting to the evolving nature of phishing attacks and improving detection accuracy over traditional methods

Melad Mohamed Al-Daeef, Nurlida Basir, Madihah Mohd Saudi [2], The research paper focuses on identifying the most effective detection features, particularly those extracted from the email body, such as Keywords and URLs. The authors propose a novel feature selection framework based on sender information, email content, and recipient interaction to accurately distinguish phishing emails from legitimate ones. Through an analysis of over 8,000 emails, the study finds that URL-based features are more reliable than Keywords, showing a stronger correlation with phishing activity. Additionally, the research emphasizes the importance of sender and content-based features, highlighting the crucial role of URL analysis in phishing detection. The paper concludes that optimizing feature selection using these criteria significantly improves detection model performance, providing a more effective approach to combating phishing threats in practical scenarios.

Rohit Valecha,Pranali Mandaokar, H. Raghav Rao [3], This paper addresses phishing email detection, a significant cybersecurity challenge that threatens both financial security and data integrity. It explores the use of machine learning algorithms to detect phishing emails by analyzing persuasion cues within the email content. The study employs various ML algorithms, including SVM, Random Forest, Naive Bayes, Decision Trees, and Logistic Regression, to classify emails based on these cues. The dataset is split into training and testing sets to evaluate the models' effectiveness. The performance of each algorithm is assessed using metrics such as accuracy, precision, recall, and F1 score. The results highlight the critical role of persuasion cues in enhancing phishing detection and demonstrate the superior performance of Random Forest and SVM models in accurately identifying phishing attempts.

Ayman El Aassal, Shahryar Baki, Avisha Das, Rakesh M. Verma[4], This paper addresses the benchmarking and evaluation of phishing detection methods, a crucial aspect of cybersecurity aimed at preventing significant financial and data breaches. It introduces the "PhishBench" framework, designed to systematically compare and assess various phishing detection techniques across different datasets. The study evaluates the effectiveness of machine learning classifiers, including supervised and deep learning models, in detecting phishing attacks based on URL, website, and email features. The paper also examines the challenges posed by imbalanced datasets, which are common in phishing detection, and highlights the importance of feature selection in enhancing detection accuracy. The results emphasize the need for innovative approaches to improve phishing detection systems' robustness against evolving threats

Eder Souza Gualberto, Rafael Timoteo De Sousa, Thiago Pereira De Brito Vieira, João Paulo Carvalho Lustosa Da Costa, Cláudio Gottschalg Duque[5] ,This paper introduces an advanced phishing detection approach that leverages text-based features through a multistage methodology. By extracting and refining key features from web content—such as URLs, HTML, and metadata—the authors optimize machine learning models to effectively distinguish between legitimate and phishing sites. This multi-stage process enhances detection accuracy by reducing false positives, and the methodology is validated on large datasets, demonstrating high precision and recall in identifying phishing attempts. The research offers a comprehensive framework that combines feature engineering with machine learning to address a critical cybersecurity challenge.

Yong Fang, Cheng Zhang, Cheng Huang, Liang Liu,Yue Yang [6],The study introduces an improved Recurrent Convolutional Neural Network (RCNN) model designed to enhance phishing email detection by leveraging multilevel vectors and an attention mechanism. The model combines convolutional layers to capture local features and recurrent layers to capture sequential information from email content. Additionally, the attention mechanism helps the model focus on important parts of the email, improving classification accuracy. The study compares the performance of the proposed RCNN model against other models such as CNN, RNN, and LSTM. Performance is evaluated using metrics like accuracy, precision, recall, and F1 score. The results demonstrate that the improved RCNN model with multilevel vectors and attention mechanism significantly outperforms traditional models, providing a robust solution for detecting phishing emails.

Nguyet Quang Do, Ali Selamat, Ondrej Krejcar, Enrique Herrera-Viedma[7], The study presents a comprehensive review of deep learning (DL) techniques applied to phishing detection, offering a detailed taxonomy of these methods. It examines various DL algorithms, categorizing them based on their approaches and effectiveness in identifying phishing attempts. The paper highlights the major challenges faced by existing DL models, such as the need for manual parameter tuning, long training times, and difficulties in detecting unknown phishing attacks. Additionally, the study provides an empirical analysis of these models, evaluating their performance using metrics like accuracy, precision, recall, and F1 score. This analysis helps to identify common issues and areas where improvements are needed. Furthermore, the paper discusses the future directions for research in this field, suggesting that advancements in DL techniques could lead to more accurate and efficient phishing detection systems. The study emphasizes the importance of addressing current challenges to enhance the robustness of phishing detection models.

### *B. Key theories and Concepts:*

Literature on the subject of phishing detection reveals an abundance of approaches based on the techniques of machine learning (ML) and deep learning (DL). In this, Lingampally Shalini et al. [1] explained how ML and DL algorithms will analyse the textual and metadata features to detect phishing e-mail through the KNN, K-Means technique of data augmentation, along with the LSTM classification where 95% of them can be detected. Similarly, Melad Mohamed Al-Daeef et al. [2], highlights the feature selection importance and, in particular, URL-based features as being highly reliable for phishing activity indicators. Rohit Valecha et al. [3], study the role of persuasion cues in phishing emails and identify Random Forest and SVM as the best ML models for classification. Eder Souza Gualberto et al. [5] introduced a multistage methodology that uses refined text-based features, such as URLs and HTML metadata, to enhance detection accuracy and reduce false positives. Advanced neural network architectures also play a significant role in the literature. Yong Fang et al. [6] proposed an improved RCNN model combining convolutional layers for local feature extraction, recurrent layers for sequential data, and an attention mechanism to focus on important email parts, outperforming traditional models like CNN, RNN, and LSTM. For this problem, an ample taxonomy by Nguyet Quang Do et al.[7] is reviewed for detailing the DL technique for phishing detection, comparing its advantages and disadvantages by manually fine-tuning its parameter. Moreover, Ayman El Aassal et al.[4], discuss the "PhishBench" framework for systematically benchmarking methods for detecting phishing attacks which deals with the challenges presented as imbalanced datasets and places more focus on effective selection of features.

### *C. Controversies in the Literature or Gaps:*

Despite advancements, several controversies persist in phishing detection research. One key debate revolves around feature selection. While Melad Mohamed Al-Daeef et al. [2], argue that URL-based features are the most reliable, Lingampally Shalini et al. [1] and Eder Souza Gualberto et al. [5] highlight the importance of textual and metadata features, indicating that feature effectiveness may depend on the dataset and phishing attack characteristics.

Another area of debate is the performance trade-offs between traditional models and advanced neural networks. Yong Fang et al. [6] claim superiority of RCNN models with attention mechanisms, while Nguyet Quang Do et al. [7] warn that deep learning models often suffer from problems such as manual parameter tuning and long training times, thus raising doubts about their feasibility compared to simpler ML methods. Another topic of debate includes dataset augmentation. Lingampally Shalini et al. [1] support augmenting with KNN and K-Means, whereas Ayman El Aassal et al. [4], mention that the risk of using imbalanced datasets is that these might influence the evaluation process.

Adaptability in detection systems to changing phishing strategies remains a significant issue. Relatedly, in this regard are papers by Rohit Valecha et al. [3], while Nguyet Quang Do et al. [7] focus a lot on changing attack patterns with dynamic counter mechanisms. Other similar solutions still provide no one general solution over the other: attention mechanism and feature optimization. Moreover, evaluation metrics have inconsistencies in their focus, from accuracy-based evaluation (Lingampally Shalini et al. [1],Eder Souza Gualberto et al. [5]) to precision, recall, and F1 score-based evaluation (Rohit Valecha et al. [3], Yong Fang et al. [6]), which makes direct comparison difficult and obscures a clear understanding of model performance.

## III. PROPOSED MODEL

The phishing email detection model we propose identifies phishing attempts by analysing both text and metadata features in emails. The workflow is structured as follows:

*Dataset Information:*

The dataset for this phishing email detection model was constructed by merging eight different datasets, as outlined in the paper "Several Curated Datasets and Feature Analysis for Phishing Email Detection with Machine Learning[8] This combined dataset consists of over 100,000 email samples, covering a wide range of both phishing and legitimate emails. Each sample includes various enriched features such as email metadata (e.g., sender and receiver details), textual content (subject and body), and URL characteristics, providing a multidimensional view of the data. By integrating these diverse features, the dataset enables the model to identify complex patterns associated with phishing attacks, supporting a more robust training process and enhancing its overall detection accuracy. This extensive and well-rounded dataset serves as a strong foundation for the model, improving its adaptability and reliability in identifying phishing emails

### *A. Data Preprocessing*

To prepare the email text, we clean the subject and body, removing noise elements such as punctuation, stopwords, and case variations. The text is converted to lowercase, and lemmatization is applied to reduce words to their root forms, ensuring focus on essential meanings. Metadata features (e.g., sender, receiver, presence of URLs) are also extracted for further analysis.

### *B. Text Vectorization with TF-IDF:*

After preprocessing, the text is converted into numerical vectors using TF-IDF (Term Frequency-Inverse Document Frequency). This transformation emphasizes critical words indicative of phishing, such as uncommon or suspicious terms

### *C. Metadata Feature Extraction*

Metadata features, including the presence of URLs, sender details, and email body and subject, are crucial for phishing detection as phishing emails often exhibit abnormal patterns. For instance, emails from unrecognized domains or emails with unusual sending times can be indicators of phishing.

### *D. Exploratory Data Analysis (EDA):*

EDA highlights patterns and insights between phishing and legitimate emails. Key analyses include label distribution to assess class balance, sender domain frequency, subject line language, and the presence of URLs. Visualizations such as word clouds, distribution plots, and correlation heatmaps offer insights into feature selection, guiding the subsequent steps in model training.

### *E. Feature Engineering and Model Training (Random Forest):*

Using the insights from EDA, we engineer key features, including textual details (e.g., word frequency, sentiment analysis) and URL characteristics (e.g., length, presence of suspicious keywords). We then apply a Random Forest classifier, chosen for its robustness and interpretability, as the initial model.

### *F. Random Forest as Initial Policy:*

The trained Random Forest model serves as the initial policy for the Deep Q-Network (DQN) agent. It provides a baseline classification of emails as phishing or legitimate, leveraging its knowledge to enhance the DQN agent's learning efficiency.

### *G. DQN Agent for Continuous Learning:*

To improve adaptability, we integrate a Deep Q-Network (DQN), an RL approach that enables the model to refine its phishing classification policy iteratively. In this setup, the DQN agent interacts with the environment (an email classification task), making decisions and receiving feedback based on the correctness of each classification. Positive rewards reinforce correct phishing classifications, while penalties discourage false positives, allowing the model to adapt to new patterns in phishing attacks.

### *H. Integration of DQN and Random Forest:*

To build more efficient and accurate Phishing email detection we combined the power of reinforcement learning and supervised learning. The DQN agent uses the Random Forest model to calculate the initial observation for the environment.

This observation represents the phishing probability, which helps the agent make its first classification decision. The agent then refines its decision-making process by continuously interacting with the environment, receiving rewards or penalties based on the accuracy of its classifications. As it learns, the DQN model gradually improves its performance, potentially outperforming the Random Forest model by adapting to new phishing techniques.

*I Model Flow Chart:*

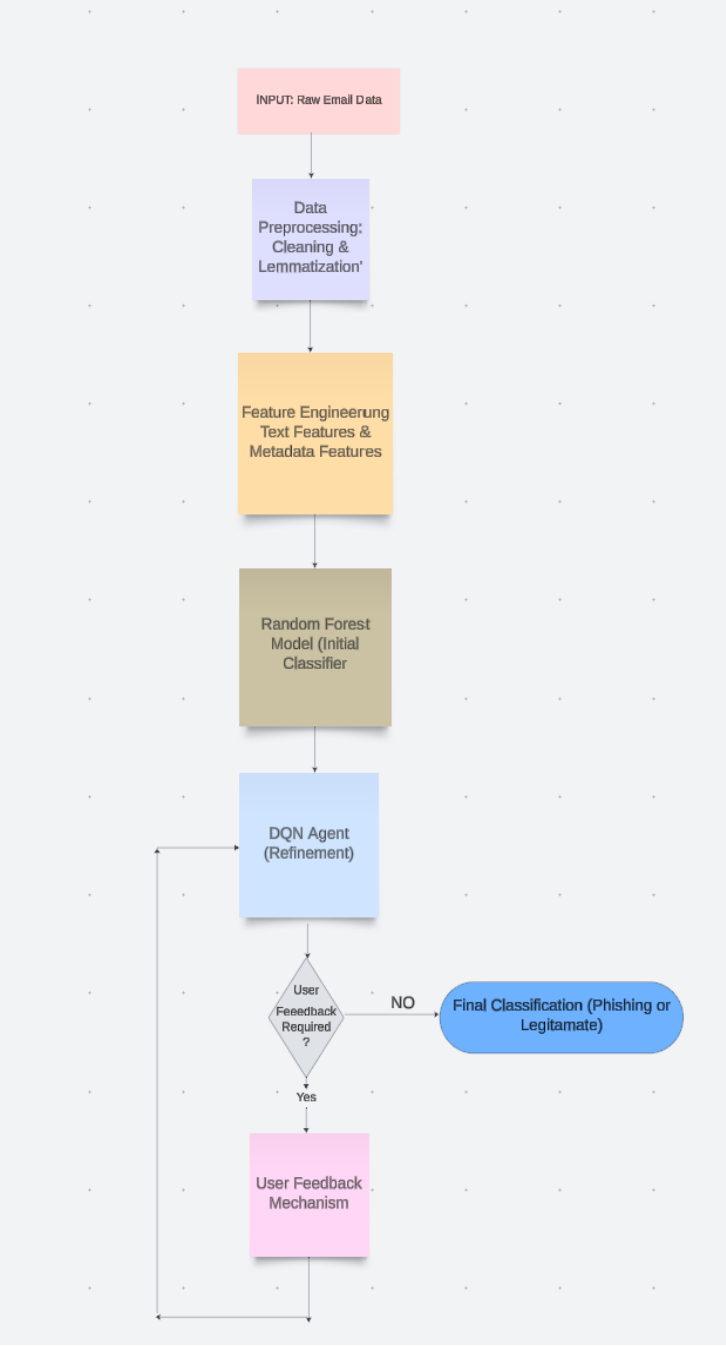


*Figure 3.1: Flowchart of the proposed phishing email detection model, illustrating the sequence from data preprocessing to classification*

*J. Model Architecture Diagram*

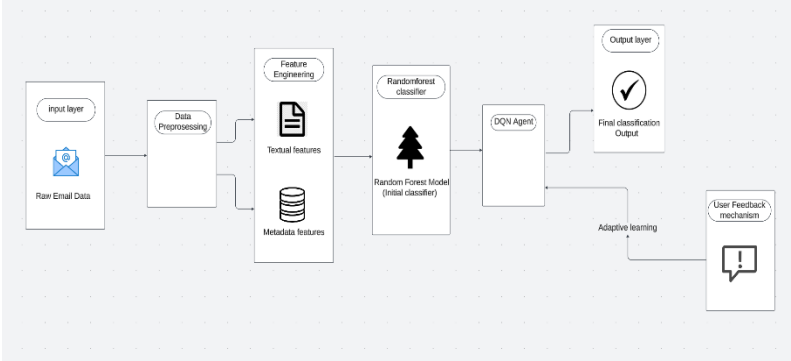


*Figure 3.2 : Architecture of the hybrid phishing detection model, illustrating the integration of Random Forest and Deep Q-Network for initial classification and iterative learning.*

*K. Evaluation metrics:*

To evaluate our phishing detection model, we used accuracy, precision, recall, and F1 score as performance metrics:

1.**Accuracy**: Measures the percentage of correct predictions.

$$\text{Accuracy} = \frac{\text{True Positives+True Negatives}}{\text{Total Predictions}}$$

2.**Precision**: Indicates the proportion of true positive phishing detections among all predicted positives.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}}$$

3.**Recall** (Sensitivity): Shows the proportion of actual phishing emails correctly identified by the model.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives +False Negatives}}$$

**F1 Score**: The harmonic mean of precision and recall, providing a balanced metric for model performance.

$$\text{F1score} = \frac{\text{2*Precision * Recall}}{\text{Precision+Recall}}$$

## IV.RESULTS

*A. Presentation of Findings*

Our phishing detection model attained remarkable accuracy using a hybrid approach combining the Random Forest classification with the Deep Q-Network. The Random Forest classifier reached an accuracy of 98% in independent classification between phishing and legitimate emails, using textual and metadata features. This is comparable or better than the traditional machine learning approaches like Naive Bayes and Support Vector Machines, which achieve an accuracy of 90%-95% in similar scenarios. When DQN was incorporated into the system, the model achieved an accuracy of 99%, thus showing its adaptability and ability to learn new phishing strategies dynamically.

*B. Data Analysis and Interpretation*

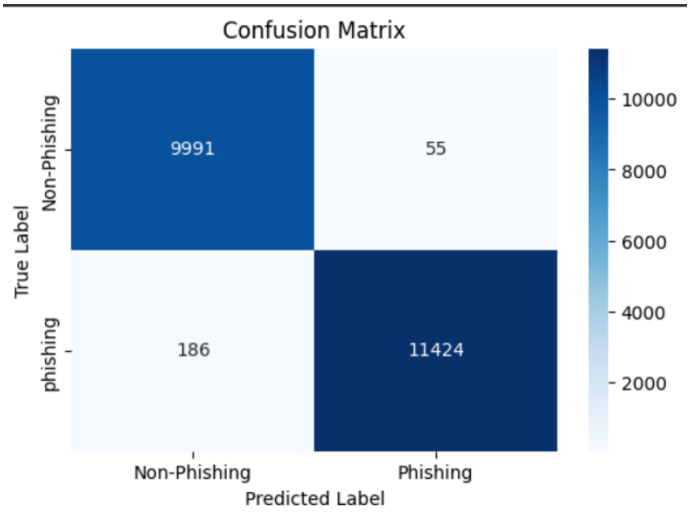


*Figure 4.1: Confusion Matrix illustrating the classification performance of the phishing detection model*

*TABLE I*
*PERFORMANCE EVALUATION TABLE for Random forest*

| | precision | Recall | f1-score | Support |
|---|---|---|---|---|
| 0 | 0.98 | 0.99 | 0.99 | 10046 |
| 1 | 1.00 | 0.98 | 0.99 | 11610 |
| Accuracy | | | 0.99 | 21656 |
| macro avg | 0.99 | 0.99 | 0.99 | 21656 |
| weighted avg | 0.99 | 0.99 | 0.99 | 21656 |

Table 1: reveal that the model exhibits balanced and high performance in detecting phishing and non-phishing emails.

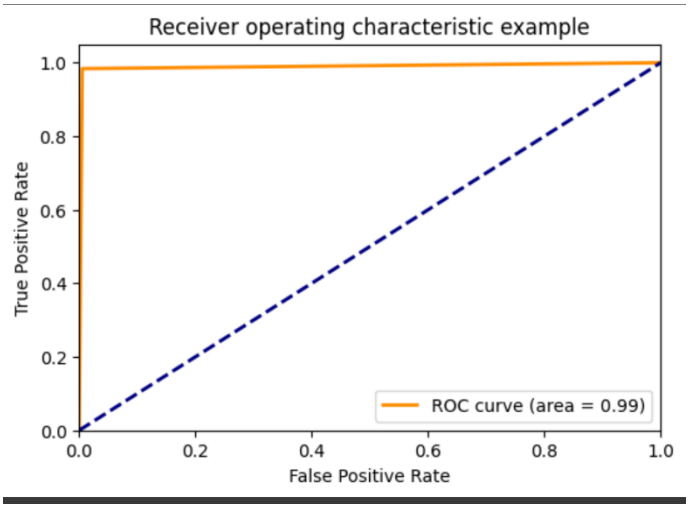


*Figure 4.2: ROC Curve demonstrating the true positive rate vs. false positive rate for the phishing detection model*

*C. Support for the Research Question or Hypothesis:*

Adaptability of the hybrid approach to new phishing tactics, with real-time feedback from DQN, addresses the research question of building robust phishing detection systems. This method offers a proactive, self-learning solution that is superior to static models, which typically require multiple retraining's.

## V.DISCUSSION:

*A. Interpretation of Results:*

The proposed phishing email detection model efficiently incorporates both machine learning and reinforcement learning to detect phishing emails with high accuracy and adaptability. The AutoPhishGuard interface, as presented in Figure 5.1, allows users to easily input email details and immediately obtain real-time predictions whether an email is phishing or legitimate. If an email is labelled as safe, the system provides a sense of relief with a "Safe Email" badge, as in Figure 5.2, and thus ensures confidence in its classification. In contrast, emails that are phishing-related display an explicit warning message, as seen in Figure 5.3, and thus ensure instantaneous threat detection. Moreover, the feedback mechanism, illustrated in Figure 5.4, further improves the system's effectiveness by refining predictions through user-provided corrections, thereby ensuring continuous enhancement in combating new phishing schemes.

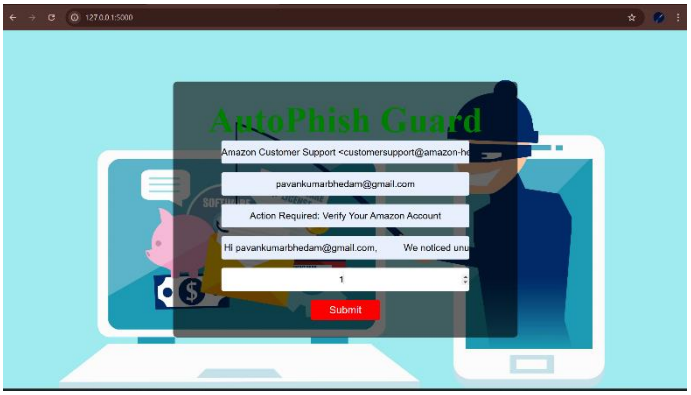


***Figure 5.1**: The AutoPhishGuard interface enables users to input email details and receive phishing or legitimate predictions.*


***Figure 5.2**: A "Safe Email" message reassures users when an email is determined to be legitimate.*


***Figure 5.3**: A phishing alert notifies users of potential threats for emails classified as phishing.*

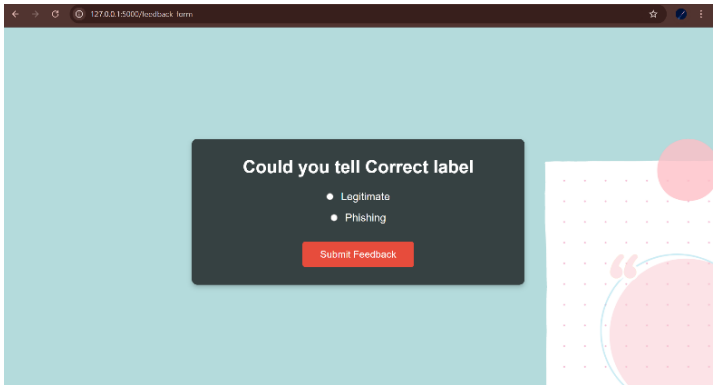


***Figure 5.4**: The feedback mechanism allows users to provide input for incorrect predictions, aiding in model refinement.*

*B.Comparison with Existing Literature:*

In comparison with past works, which mostly relied on static models such as Naive Bayes, Support Vector Machines, or standalone Random Forests with accuracy around 90% to 95%, our model outperformed these models with a significantly higher accuracy of 99%. Unlike these conventional techniques, reinforcement learning facilitates how our model learns in a dynamic way by adapting to evolving phishing techniques without the need to frequently retrain. This adaptability does make our model different from static classifiers, thus an advantage to a real-time phishing detection system.

*C. Implications and Limitations of the Study:*

The results of this study have great implications in real-world email security systems, providing a proactive and adaptive solution for the detection of phishing attacks. Also, with a mechanism of user feedback, it will keep on improving itself and become future-proof against new strategies of phishing. Some of the limitations include reliance on user feedback, which can introduce delay in model adaptation because of sparse or inaccurate feedback. Furthermore, the computational complexity of reinforcement learning models would prevent scalability in very large systems or in real-time environments. The issues are expected to be resolved through further research, and these factors could make the model more applicable and efficient.

## VI.CONCLUSION:

The proposed phishing email detection model demonstrates a robust and systematic approach to identifying phishing emails

by combining a Random Forest classifier with reinforcement learning through a Deep Q-Network (DQN). The hybrid model achieves high accuracy, outperforming traditional machine learning methods and adapting dynamically to new phishing techniques. The integration of the user feedback mechanism enhances the ability of the system to refine its predictions, which provides a proactive and reliable solution to real-world challenges in email security. With the user-friendly interface, further simplifying interaction, it makes the model accessible and effective for non-technical users. This study showcases the practicality of using hybrid models in addressing dynamic cybersecurity threats.

## VII.FUTURE SCOPE:

This phishing detection model can be further developed to improve the feedback system for better adaptability, integrate with real-time email platforms to mitigate threats instantly, and enhance multilingual support. Incorporating advanced features like behavioural analysis, testing on diverse datasets, and improving robustness will strengthen the model. Additionally, leveraging this model to analyse phishing trends can help develop proactive security measures.

## VIII REFERENCES


[1] L. Shalini, S. S. Manvi, N. C. Gowda and K. N. Manasa, "Detection of Phishing Emails using Machine Learning and Deep Learning," 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2022, pp. 1237-1243, doi: 10.1109/ICCES54183.2022.9835846.

[2] M. M. Al-Daeef, N. Basir and M. M. Saudi, "A Method to Measure the Efficiency of Phishing Emails Detection Features," 2014 International Conference on Information Science & Applications (ICISA), Seoul, Korea (South), 2019, pp. 1-5, doi: 10.1109/ICISA.2014.6847332.

[3] R. Valecha, P. Mandaokar and H. R. Rao, "Phishing Email Detection Using Persuasion Cues," in IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 2, pp. 747-756, 1 March-April 2022, doi: 10.1109/TDSC.2021.3118931.

[4] A. El Aassal, S. Baki, A. Das and R. M. Verma, "An In-Depth Benchmarking and Evaluation of Phishing Detection Research for Security Needs," in IEEE Access, vol. 8, pp. 22170-22192, 2020, doi: 10.1109/ACCESS.2020.2969780

[5] E. S. Gualberto, R. T. De Sousa, T. P. De Brito Vieira, J. P. C. L. Da Costa and C. G. Duque, "The Answer is in the Text: Multi-Stage Methods for Phishing Detection Based on Feature Engineering," in IEEE Access, vol. 8, pp. 223529-223547, 2020, doi: 10.1109/ACCESS.2020.3043396.

[6] Y. Fang, C. Zhang, C. Huang, L. Liu and Y. Yang, "Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism," in IEEE Access, vol. 7, pp. 56329-56340, 2019, doi: 10.1109/ACCESS.2019.2913705.

[7] N. Q. Do, A. Selamat, O. Krejcar, E. Herrera-Viedma and H. Fujita, "Deep Learning for Phishing Detection:

Taxonomy, Current Challenges and Future Directions," in IEEE Access, vol. 10, pp. 36429-36463, 2022, doi: 10.1109/ACCESS.2022.3151903.

[8] A. I. Champa, M. F. Rabbi and M. F. Zibran, "Curated Datasets and Feature Analysis for Phishing Email Detection with Machine Learning," 2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI), Mt Pleasant, MI, USA, 2024, pp. 1-7, doi: 10.1109/ICMI60790.2024.10585821.

[9] Gunikhan Sonowal, "A Model for Detecting Sounds-alike Phishing Email Contents for Persons with Visual Impairments", International Conference on Control Communication and Computing (ICCC), 2020 IEEE.

[10] Serfettin Senurk, Elif Yerli, Ibrahim Sogukpinar, "Email Phishing De- tection Prevention by Using Data Mining Techniques", International Conference on Computing Technologies (ICCT), 2017 IEEE.