International Conference on Multidisciplinary Perspectives in Advanced Computing and Technology (IMPACT 2026)
G. B. Pant University of Agriculture and Technology, Uttarakhand, India. Jan. 10-11, 2026

DMPedia Lecture Notes in Multidisciplinary Research,
Volume 1, pp. 1143-1169, 2026

Detecting AI-Generated Phishing Websites with ML

Vasu Sharma, Udit Sirohi

School of Computer Science Engineering and Technology, Bennett University, India

vasus99990@gmail.com, uditsirohi5026@gmail.com

ABSTRACT

In the modern digital landscape, phishing has become among the most pervasive and enduring risks. It is a technique of impersonating a trustworthy entity with the express intent of convincing victims to reveal financial or personal information, including credit card numbers, bank account details, or passwords. The reason phishing is so dangerous is that, rather than exploiting software weaknesses, it often exploits psychological weaknesses, such as urgency, anxiety, or simple curiosity. It is this psychological element that even experienced users may sometimes succumb to. Phishing comes in many forms. The most common type includes emails that pose as reputable companies or banks and ask recipients to "verify" or "update" their accounts. Spear phishing is a highly focused type that targets particular people or organizations after conducting extensive background research, making the effort much more credible. Clone phishing involves hackers replicating authentic emails and substituting harmful attachments or links. These changing tactics demonstrate that phishing is a perennial challenge for cybersecurity experts, as it evolves both technologically and socially. Phishing's overall effects are concerning. According to studies, it plays a significant role in global cybercrime, resulting in billions of dollars in losses each year from illicit transactions or stolen data. Organizations all throughout the world reported losses of more than $billion in 2023 alone due to phishing-related incidents. The issue has expanded beyond what conventional security filters can manage, with attackers now even utilising AI-generated content to produce convincingly phoney websites and emails. Because of this, it is crucial to investigate fresh, innovative solutions that can change as quickly as the threats do.

Keywords: *Phishing Detection, Machine Learning, Deep Learning, LSTM, Random Forest, Cyber Security, Ensemble Models*

1. Introduction

The rapid advancement of AI has enabled almost all industries to reap rewards, most notably through automation and the creation of original content. Despite that, the innovation has also increased the risks in cyberspace, mainly through AI-assisted phishing attempts. Phishing is generally the practice of deceiving individuals into revealing their private information through emails or specially created sites, such as passwords, bank account numbers, or personal information. Due to continuous improvements in AI, attackers can even automatically develop super-realistic phishing sites that mimic the look and feel of trusted brands almost perfectly [1-4].

These fraudulent sites exploit people's trust in ways previously impossible, thereby making detection much more difficult than in earlier generations of phishing attacks. [6] [11]AI-generated phishing sites pose unique cybersecurity challenges. AI-assisted versions can mimic the look, branding, and content of authentic websites, unlike classic phishing websites, which often contain

obvious mistakes or predictable architecture. Due to this complexity, traditional detection techniques such as keyword-based filtering, blacklisting, and basic URL heuristics become increasingly ineffective. As a result, there is a growing demand for improved, adaptive approaches to detect these dangers before they affect consumers. Machine learning (ML) offers powerful solutions to detect AI-generated phishing websites. By analysing a wide range of factors, such as URL patterns, domain characteristics, security certificate details, HTML structure, page design, and text meaning, ML models can identify subtle clues that distinguish trustworthy sites from fake ones. Unlike traditional rule-based methods, these models continuously learn and adapt to new phishing strategies. By combining different features in advanced models such as LSTM networks or layered ensemble systems, detection becomes even more precise, even when attackers use sophisticated AI to create convincing fake websites. This dynamic approach makes ML a key tool in staying ahead of evolving cyber threats.[5]

Research in this area plays a crucial role in strengthening cybersecurity by deepening our understanding of how generative AI is changing the landscape of digital threats. Successfully detecting AI-generated phishing websites requires a blend of disciplines, including computer vision, natural language processing, web analytics, and threat intelligence. By combining these fields, researchers develop intelligent, scalable defence systems that can effectively safeguard users and organisations against increasingly sophisticated cyberattacks. This interdisciplinary approach ensures ongoing innovation in protecting the digital world. As attackers increasingly use AI to create sophisticated phishing websites, applying machine learning for fast and accurate detection has become vital. This research examines current strategies, challenges, and future directions in identifying AI-generated phishing sites. It emphasises that using a multi-signal, adaptable detection system, one that considers various data points and evolves with emerging threats, is key to strengthening defences in today's complex cybersecurity environment. [1] [11]

2. Background

Phishing remains one of the most prevalent and persistent threats in today's digital world. At its core, phishing is a form of social engineering that deceives individuals into revealing sensitive information, such as passwords, bank details, credit card numbers, and personal identification. Unlike malware-based attacks, phishing primarily manipulates human psychology through tactics such as urgency, fear, and authority to trick victims. Because it exploits human behaviour alongside technology, phishing remains highly effective despite widespread cybersecurity measures.

Phishing attacks come in different styles, and each has its own way of tricking people. One of the most common is the "deceptive email," in which someone pretends to be a trusted company, such as your bank or a government office, and tries to get you to share sensitive information. There's also "spear-phishing," which goes a step further; the attacker learns about a specific person or business and sends a very convincing, personalised message, making it much harder to spot as a scam. Another method is called "clone phishing," where cybercriminals copy a real email someone

has actually received, change small details like the links, and resend it to fool you into clicking a dangerous link. These tricks work not just because of technology gaps, but also because they leverage people's trust in familiar brands and contacts. That's why it's so difficult for basic security systems to catch the attacks often look perfectly normal until it's too late. [5] [11]

Phishing attacks have a massive impact on organizations and individuals worldwide. Recent data shows that phishing accounts for over 80% of all reported cybersecurity incidents, resulting in billions in financial losses each year. In 2023 alone, the cost of a single data breach due to phishing averaged nearly $5 million for affected companies. Attackers are constantly refining their tactics, now using AI-generated emails and realistic fake websites to trick even cautious users and bypass conventional security measures. As these threats become more sophisticated and frequent, it's more important than ever for organizations to adopt smarter, adaptive ways to defend sensitive data and prevent breaches.[8] [9]

2.1 Motivation

Traditional tools for detecting phishing, such as blacklists and signature-based filters, have significant drawbacks. Blacklisting only works if the scam website or domain is already known, so it misses fresh attacks, since new phishing sites pop up quickly and often change. Signature-based filters are used by most email and antivirus programs; they identify threats based on predefined patterns. If a phishing attempt uses a new trick or doesn't match the old patterns, it slips through undetected. These methods just aren't flexible enough for today's fast-moving digital world, especially with cybercriminals using smart techniques and even AI to create convincing fake messages. That's why people are turning to smarter solutions, like machine learning, which can actually learn, adapt, and spot never-before-seen phishing attacks in real time [10]. Because phishing attacks keep changing, we need smarter, more flexible ways to spot them. That's where machine learning (ML) comes in: it can analyse large amounts of data to uncover hidden clues that indicate phishing, even in brand-new scam attempts. Unlike old-school security systems, ML doesn't just rely on fixed rules. It learns from past examples, recognizes suspicious patterns, and can even catch fresh attacks that haven't been seen before.

These ML-powered systems check many details, such as the words in an email, the structure of a website link, the age of a domain, and how people interact with messages online. They do this quickly and keep adapting as new threats appear, making them much better at spotting and stopping phishing in real time.

Plus, ML models get smarter with every attack they review, so they don't need constant manual updates, making protection stronger and easier to manage as cyber threats evolve. [1] [6] [11] [12]. Today, cybersecurity experts are paying closer attention to the power of machine learning (ML) in defending against online threats such as phishing. ML models are good at spotting suspicious activities, but combining them in a hybrid approach can make these systems even smarter. In hybrid models, multiple ML techniques collaborate to determine whether something is a phishing

attempt, resulting in higher accuracy and fewer mistakes. By making ML-based phishing detection stronger, people and organizations can feel safer online, protect their money, and become more resilient against cyber attacks. This means less risk of falling victim to scams, better protection of personal and business information, and a more secure digital environment overall.

2.2 Objectives

The main goals of this research are to address the weaknesses of traditional phishing detection methods by demonstrating how machine learning can improve proactive cybersecurity.

- First, the study aims to develop models that can detect phishing across different platforms by identifying harmful URLs, fake websites, and dangerous email content. This ensures the system covers all common attack methods. [4]
- Second, the research will use supervised and deep learning algorithms that dynamically learn phishing patterns, enabling the models to spot both known scams and new emerging threats.[8]
- Third, it plans to evaluate and compare the effectiveness of various machine learning classifiers, including Decision Trees, Random Forests, Support Vector Machines, Naive Bayes, and LSTM networks, in detecting phishing [12].
- Finally, the research will propose a hybrid ensemble model, a combination of multiple classifiers using meta-learning, that boosts detection accuracy while reducing false alarms. Together, these efforts aim to build an adaptive, scalable, and highly efficient phishing detection system ready to tackle modern cyber threats.

3. Literature Review

Phishing attacks have grown more complex over time, making older detection methods less effective. To tackle this, many researchers have turned to machine learning (ML), deep learning (DL), and hybrid models to better identify phishing websites and emails in the past ten years. This section will review previous studies, focusing on the algorithms they used, the features they extracted from data, and the hybrid techniques they applied to improve detection accuracy. [1] [6]

3.1 Machine Learning Approaches

The author used Decision Tree (DT) algorithms to identify phishing websites by analysing the UCI Phishing Website Dataset and achieved 91% accuracy [5]. Their research showed that Decision Trees can successfully distinguish between phishing and genuine websites by looking at URL features such as URL length, the use of "@" symbols, and suspicious keywords. A Decision Tree works by creating a structured set of decisions based on these features, making it easy to understand and explain how classifications are made. One of the biggest advantages of Decision Trees is their simplicity and ability to handle both numerical and categorical data, which is why they are popular in cybersecurity research. However, the study also pointed out some weaknesses: Decision Trees can sometimes perform poorly when trained on small datasets because they tend to overfit, and they are sensitive to noisy or irrelevant data, which can reduce their accuracy when

applied to new situations. Despite these challenges, Decision Trees proved to be a strong starting point for research on phishing detection. [3] used the Random Forest (RF) algorithm to detect phishing websites, achieving 95% accuracy on the PhishTank dataset by applying feature selection to improve model performance. Random Forest works by building many Decision Trees and combining their results to boost accuracy and reduce the chance of overfitting.

This method is especially good at handling large, complex datasets with many features, such as URLs, domains, and HTML attributes, making it ideal for phishing detection. The study found that Random Forest models are accurate and reliable, even when the data is incomplete or imbalanced. However, Random Forests require more computational resources than single Decision Trees and are harder to interpret because of their complexity. Despite these challenges, Zhang et al.'s work shows that Random Forest is a reliable and powerful tool for detecting phishing websites at scale.[1] [11] [12]

Support Vector Machines (SVM) used to detect phishing by analysing URL and domain features, achieving 92% accuracy. Their study demonstrated that SVMs perform well by finding the optimal boundary (or hyperplane) to separate phishing sites from legitimate sites in a high-dimensional feature space. This makes SVMs especially suited for phishing detection, where data can be complex and interconnected, such as URLs and email contents. SVMs are also good at handling unusual data points (outliers), especially when paired with the right kernel functions, thereby improving accuracy. However, Li et al. noted several challenges, including SVM's sensitivity to kernel and other hyperparameter choices, which can significantly affect model performance. Additionally, training SVMs requires substantial computing power, particularly with large datasets. Despite these drawbacks, the study concluded that SVM remains a reliable and effective method for phishing detection, thanks to its ability to handle high-dimensional, non-linear data.[6] [11] [12]

The Naive Bayes (NB) algorithm was applied to detect phishing emails by examining keywords and domain features, achieving an accuracy of 88%. Their research highlighted that Naive Bayes, as a probabilistic classifier, assumes independence among features, making it suitable for detecting text-based phishing attacks. Because it is simple and fast, Naive Bayes is widely used for email classification, especially when natural language processing (NLP) elements such as suspicious words, phrases, and domain patterns are involved. This makes the model particularly useful when quick and efficient detection is needed. However, Ahmed et al. noted that independence assumptions often don't hold in real-world data, where features can be related, potentially reducing accuracy in more complex phishing cases. Still, their findings show that Naive Bayes is a reliable baseline model for phishing email detection, especially when speed and transparency are priorities [6].

Deep learning techniques applied, focusing on Long Short-Term Memory (LSTM) networks, to detect phishing websites by analyzing URL sequences, achieving a high accuracy of 96%. Their study emphasized LSTM's strength in understanding the order and connections within sequential

data like URLs and email contents, which helps the model spot complex behavioural and structural patterns typical of phishing attacks. Deep learning methods such as LSTM and Convolutional Neural Networks (CNN) have gained popularity in phishing detection because they can model complex, non-linear relationships and automatically extract sophisticated features from raw data. While LSTMs are especially good at handling sequence data, CNNs excel at identifying hierarchical patterns in website images or encoded URLs. Singh and Sharma noted that these deep learning models usually outperform traditional machine learning approaches when large, diverse datasets are available. However, they also pointed out challenges such as the need for significant computational power for training and dependence on large labeled datasets. Despite these hurdles, their findings confirm that deep learning, especially LSTM-based models, provides improved accuracy and adaptability in combating evolving phishing threats. [1] [6] [8] [11] [12]

3.2 Feature Extraction Techniques

The success of machine learning (ML) and deep learning (DL) in detecting phishing attacks largely depends on how well features are extracted from URLs, websites, and email content. Feature extraction helps these models spot patterns frequently associated with malicious activity, thereby improving their ability to accurately classify phishing attempts. According to previous studies, three main types of features play a crucial role in phishing detection: URL-based, HTML-based, and email content features. These categories provide valuable clues that help the models distinguish between safe and harmful online content.[1] [6] [11]

URL-Based Features

URL-based features are among the most commonly used signals for detecting phishing. These features focus on unusual or suspicious elements in the structure and wording of URLs that attackers exploit to trick users. Common characteristics include the URL's overall length, the presence of unusual symbols such as "@" or "-", the use of an IP address instead of a standard domain name, and whether the URL uses HTTPS and an SSL certificate. These features are effective at distinguishing genuine URLs from phishing URLs, as malicious URLs often exhibit irregular or deceptive patterns designed to mislead users. By analyzing these details, detection systems can better identify and block harmful links before they cause damage.[11]

HTML-Based Features

HTML-based features focus on a web page's internal layout and elements, making them valuable for detecting fraudulent websites that mimic trusted sites. Key features include the presence of forms or input fields designed to steal credentials, external links or embedded iframe tags, and unusual patterns in the Document Object Model (DOM) structure.

Phishing pages often manipulate these elements to look genuine, tricking users while secretly capturing their data or redirecting them to harmful servers.

Email content features are derived using Natural Language Processing (NLP) techniques. These features identify textual patterns, suspicious keywords, language inconsistencies, and social engineering tactics in the email body. Additional important email features include checking the sender's domain, examining header details, and comparing displayed URLs with their real destinations. Since phishing frequently starts through email, analyzing these features is essential for early and effective detection. [6] [11] as shown in Table 1.

Table 1. Comparison Table based on different studies

Study	Features Used	Dataset	Method	Accuracy
Kumar et al., 2020	URL, Domain	UCI Phishing	DT	91%
Zhang et al., 2019	URL, HTML	PhishTank	RF	95%
Li et al., 2021	URL, Domain	ISCX-URL-2016	SVM	92%
Singh & Sharma, 2020	URL sequence	Custom Dataset	LSTM	96%

3.3 Hybrid Models

Recent studies show that hybrid models are highly effective in detecting phishing. For example, the PhishGuard framework combines several powerful algorithms, Random Forest, XGBoost, and LSTM, to improve both the accuracy and reliability of phishing detection. This hybrid approach achieved an impressive 97% accuracy when tested on the PhishTank dataset and a specially created URL dataset, demonstrating its ability to identify phishing threats across diverse data types.[12]

This study highlights the advantages of combining traditional machine learning algorithms with deep learning models to capture a broader spectrum of phishing patterns. Hybrid models leverage the strengths of multiple classifiers, resulting in improved detection performance, especially in complex, rapidly changing threat environments. These combined approaches often achieve higher accuracy and better adaptability than single-algorithm methods, making them more effective against diverse phishing techniques. However, researchers also point out some challenges with hybrid systems. These include increased computational demands due to the complexity of running multiple models simultaneously and a loss of interpretability, since it becomes harder to understand how decisions are made when many models work together. Despite these drawbacks, hybrid models represent a promising direction for advancing phishing detection. Despite these challenges, hybrid approaches remain especially valuable for real-world phishing detection. This is because attackers often change URL structures, web page content, and email patterns to slip past simpler detection systems. By combining multiple models, hybrid methods can adapt more

effectively to these frequent changes, making them better equipped to catch evolving phishing tactics and provide stronger protection.[6] [11]

4. Dataset Collection

Effective phishing detection depends heavily on having high-quality datasets that include examples of both legitimate and phishing websites or emails. Properly collecting and preprocessing these datasets is essential to ensure that machine learning (ML) and deep learning (DL) models are well trained and can accurately identify phishing attempts in real-world settings. Without good data preparation, even the best models may struggle to perform reliably.[1] [6] [11] [12]

4.1 Sources

For this research, several well-known public datasets commonly used in phishing detection were referenced. These datasets include examples of both phishing and legitimate URLs, along with various features that help train machine learning models effectively.

- **UCI Phishing Website Dataset:** One major dataset is the UCI Phishing Website Dataset, which is widely used in academic studies. It contains 11,055 samples and provides 30 features primarily related to URL structures and domain characteristics. Important features include URL length, domain age, and HTTPS presence, making this dataset valuable for evaluating traditional machine learning algorithms such as Decision Trees and Random Forests.[1] [11] [12]
- **ISCX-URL-2016 Dataset:** This dataset was developed by the Information Security Centre of Excellence (ISCX) and includes over 5,000 labeled URLs, with a mix of phishing and legitimate links. It provides detailed information, including URL structure patterns, WHOIS data, and other server-side metadata. Because of its size and the richness of its features, this dataset is favoured in research that demands more comprehensive and detailed feature sets for training and evaluating phishing detection models.[11] [12]
- **PhishTank Dataset:** PhishTank is a community-driven platform where users submit reports of phishing websites they encounter. The PhishTank dataset includes real-time URLs identified as phishing, along with additional details such as domain name, hosting provider, and historical phishing records. Researchers often use this dataset to extract custom features or to evaluate the performance of their detection models on fresh, frequently updated phishing URLs, making it valuable for testing how well models can identify new and evolving threats. [11] [12]

4.2 Dataset Description

Table 2. Comparison Table based on different URLs

Dataset	Phishing URLs	Legitimate URLs	Total

UCI Phishing	,500	6,555	11,055
ISCX-URL-2016	22,000	23,000	5,000
PhishTank	15,000	15,000 (sampled)	30,000

As shown in Table 2, various features help determine whether a website or email is malicious. URL-based features look at the URL's length, special characters like "@" or "-", use of an IP address instead of a domain name, and whether HTTPS is used for security. Domain-based features include details about the website's registration, such as the domain's age, WHOIS information, the validity of its SSL certificate, and the hosting country. HTML-based features examine the webpage's internal structure, including the presence of forms or input fields, embedded iframes, and links to external sites. For emails, features include suspicious keywords or phrases in the content, verification of the sender's domain, and examination of header information. Together, these features provide important clues for effectively identifying phishing attempts. [1] [6] [11]

4.3 Preprocessing

Before using data to train machine learning or deep learning models for phishing detection, it is important to perform several preprocessing steps to ensure the data is clean, consistent, and reliable. One key step is handling missing values, as incomplete data can negatively affect model predictions. Missing values can be addressed by replacing them with the feature's mean, median, or mode, or by removing data samples with too many missing values, depending on the situation. This careful preprocessing helps improve the overall quality of the training data and the accuracy of the resulting models.[1] [8]

The next important step is normalisation and scaling, which ensure that features with widely different ranges, such as URL length or domain age, contribute equally to the model's learning. Common methods for this include Min-Max Scaling, which rescales features to a fixed range (usually 0 to 1), and Z-score Standardisation, which centres data around the mean with a standard deviation of one. For textual data extracted from URLs or emails, feature vectorization is applied to convert text into numerical form. Techniques such as TF-IDF (Term Frequency-Inverse Document Frequency), Bag-of-Words, or embedding methods are used to transform text into numerical representations that help the model capture complex patterns and relationships within the data. This process is crucial for enabling machine learning models to effectively analyze and learn from textual [6] [11]. Lastly, categorical features such as domain type, SSL status, or the email sender's domain need to be converted to numerical values so that machine learning models can process them. This is done using techniques such as one-hot encoding, which creates binary columns for each category, or label encoding, which assigns a unique number to each category. By applying all these preprocessing steps, handling missing values, normalisation, text

vectorisation, and categorical encoding, the result is a clean, structured, and machine-readable dataset ready for training effective phishing detection models.[6] [11] [12]

5. Feature Engineering

Feature engineering is crucial for phishing detection because it transforms raw data into meaningful, informative features. This process enables machine learning models to more effectively differentiate between legitimate and phishing websites or emails. Typically, features are extracted from key sources such as URLs, webpage HTML, and email text, providing the necessary clues for accurate classification. [1] [6] [11]

5.1 URL Features

Features derived from URLs are crucial for detecting phishing sites and, hence, serve as the first level of indicators for machine-learning-based phishing attacks. These include the presence of special characters like "@" or hyphens that attackers often insert into URLs to mislead users or make a fake website appear genuine. [1]

Another important signal will be the excess occurrence of digits or multiple subdomains; since most phishing URLs contain long and complicated structures, like *login.123bank.com.example.com*, in order to mislead the user. Of course, the overall length of the URL is an important feature as well-the longer URLs are considered because one of the many intentions of attackers is to cloak maliciousness within it. Moreover, age plays an important role: most phishing websites are hosted on newly registered domains, which can be easily verified in WHOIS records. These characteristics of URLs allow machine learning models to identify suspicious patterns underlying potential threats. [1] [11]

5.2 HTML Features

HTML-based features play a vital role in examining a web page's internal structure to identify phishing attempts. Key indicators include the number of forms on the page, as phishing sites often have multiple input fields or login forms designed to steal sensitive user information. Another important feature is the excessive presence of external links, particularly those directing to unknown or suspicious domains, which can signal malicious intentions. The use of iFrames is also carefully monitored since attackers frequently embed harmful content from external sources through these elements. Together, these HTML-based features help machine learning models detect structural irregularities and content anomalies, improving the accuracy and reliability of phishing detection.[1] [11]

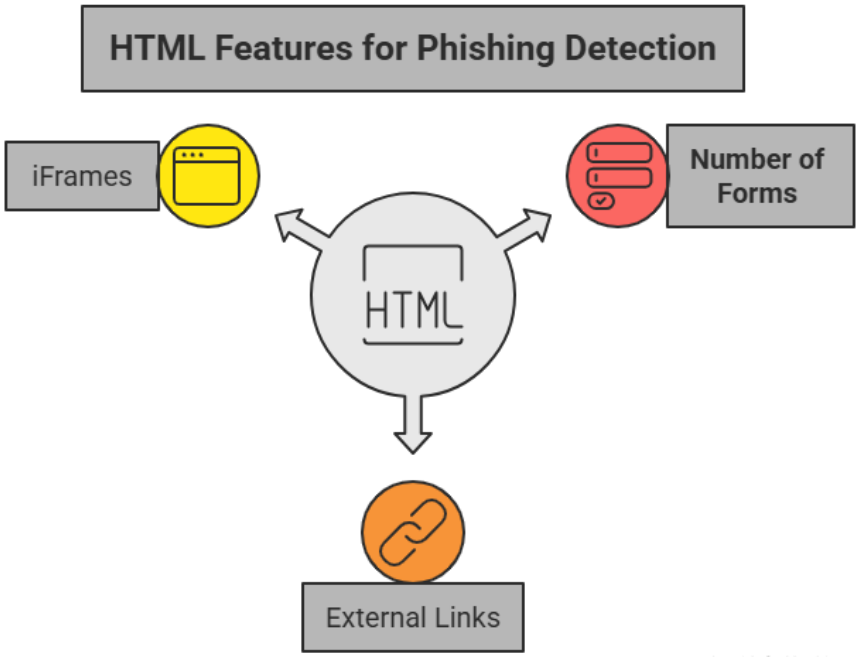

Figure 1: HTML features for Phishing Detection.

5.3 Email Features

Subtle textual cues in phishing emails are key to precise threat identification, and extracting features from email content enables machine learning models to detect them effectively. Techniques typically focus on spotting suspicious keywords, grammatical errors, and urgency-driven phrases such as “Verify now” or “Immediate action required,” which attackers use to manipulate users. Text data is also transformed into numerical representations through methods like TF-IDF vectorization, which captures the frequency and significance of words in distinguishing phishing emails from legitimate ones. Additionally, analysing email headers and sender information, such as checking the reputation of the sender’s domain, identifying unusual reply-to addresses, or spotting inconsistent header patterns, provides important signals for detecting fraudulent messages. Combined with URL and HTML-based features, these email-based indicators enhance the overall accuracy of phishing detection, especially for email-targeted attacks.[1] [6] [11]

5.4 Feature Importance

Feature importance scoring helps identify which features contribute most to detection performance. This can guide feature selection and reduce model complexity.

Table 3: Feature Importance Scores

Feature Category	Feature	Importance Score (0–1)
URL	Presence of "@"	0.85
URL	URL Length	0.78
URL	Domain Age	0.72
URL	Number of Subdomains	0.70
HTML	Number of Forms	0.80
HTML	External Links	0.75
HTML	iFrame Usage	0.68
Email	Suspicious Keywords	0.88
Email	TF-IDF Vectors	0.83
Email	Urgency Cues	0.77

6. Machine Learning Models

Machine learning models are now central to phishing detection, as they can learn complex patterns from URLs, HTML content, and email text without depending only on static blacklists that can quickly become outdated. This section reviews three key categories of approaches: supervised learning models, deep learning architectures, and ensemble or hybrid methods that combine multiple models for better performance. These techniques enable more adaptive, accurate, and scalable phishing detection systems.[1] [6] [8] [11]

6.1 Supervised Learning

Supervised learning models are trained on labeled datasets that include both phishing and legitimate samples, enabling them to learn distinctions between the two. Common algorithms used in phishing detection include Decision Trees, which create interpretable decision rules; Random Forests, which combine multiple decision trees to improve accuracy and reduce overfitting; Support Vector Machines (SVM), which find optimal boundaries to separate classes; and Naive Bayes, a probabilistic classifier effective for text-based phishing detection. These models form the foundation of many phishing detection systems due to their effectiveness and relativesimplicity.[11] [12]

Decision Tree (DT)

Decision Trees classify data by recursively splitting it based on feature thresholds, making them well-suited for phishing detection. Features such as URL length, the presence of special characters such as "@", and the number of subdomains are used to determine how the data is divided at each step.

During training, the model selects the feature that best separates phishing sites from legitimate ones, using metrics such as the Gini index or entropy to quantify this separation. The splitting process continues on the chosen feature for each branch, repeating until the nodes are pure (containing mostly one class) or a predefined maximum depth is reached. This step-by-step division produces an interpretable tree structure that distinguishes phishing from benign samples. Decision Trees are easy to understand and interpret, and they can handle both numerical and categorical data. However, they are prone to overfitting when trained on small datasets and are sensitive to data noise. Despite these limitations, Kumar et al. (2020) successfully used Decision Trees to achieve 91% accuracy on the UCI Phishing Website dataset, demonstrating their effectiveness as a baseline model for phishing detection.[11] [12]

Random Forest (RF)

Random Forests improve phishing detection by combining many decision trees to increase robustness and generalization. Each tree is trained on a random subset of the data and features through a bagging approach, where multiple bootstrapped datasets are created, and a decision tree is built for each. During prediction, each tree votes independently, and the final decision is made by majority voting. This method reduces overfitting compared to a single decision tree and performs well in high-dimensional feature spaces. However, Random Forests are computationally intensive and less interpretable than individual trees. For example, Zhang et al. (2019) used Random Forests with feature selection techniques, achieving 95% accuracy in detecting phishing URLs on the PhishTank dataset. [11] [12]

Support Vector Machine (SVM)

Support Vector Machines (SVMs) classify phishing and legitimate samples by identifying the optimal hyperplane in a high-dimensional space that maximally separates the two classes. To handle the complexity and nonlinearity often present in phishing data, kernel functions such as the Radial Basis Function (RBF) or polynomial kernels are employed.SVMs perform well with high-dimensional feature sets and can be made robust against outliers through appropriate regularization. However, their effectiveness depends heavily on selecting the right kernel and tuning parameters, and training can be slow on very large datasets. For example, Li et al. (2021) applied SVM with URL- and domain-based features, achieving 92% accuracy in phishing detection.[11] [12]

Naive Bayes (NB)

Naive Bayes is a probabilistic classification algorithm that assumes feature independence, making it simple and computationally efficient for phishing detection. It is particularly effective for analyzing textual data, which is central to email-based phishing detection. The algorithm trains quickly and performs well on large text datasets. However, its performance can decline when features are correlated, as independence does not always hold in practice. Ahmed et al. (2020) applied Naive Bayes to email content using keyword-based features and achieved an accuracy of 88%, demonstrating its usefulness as a quick and effective baseline method for phishing email detection.[6] [11]

6.2 Deep Learning Models

Deep learning models can capture complex, non-linear relationships in data and are especially effective for sequential or structured content. [8]

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks, a form of Recurrent Neural Networks (RNN), are designed to learn long-term and sequential dependencies in data. In phishing detection, LSTMs are particularly effective at analyzing sequential data such as URL character patterns or email text, enabling the model to recognize subtle, time-ordered cues within sequences. These networks excel at capturing temporal relationships, making them well-suited for tasks involving sequential information. However, LSTMs require large training datasets to achieve strong performance and are computationally more expensive than traditional machine learning models. Singh and Sharma (2020) reported an accuracy of 96% when using LSTMs to detect phishing URLs, demonstrating their effectiveness in this domain. [1] [6] [11] [12]

Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are a powerful class of deep learning models that exploit the spatial hierarchy of features in structured data. In phishing detection, CNNs can process inputs like HTML content, URL character sequences, or even website screenshots to identify malicious characteristics. The convolutional layers extract local patterns, while pooling layers reduce the data size by down-sampling and preserving the most critical features. CNNs are effective because they automatically learn relevant features from raw data, reducing the need for extensive manual feature engineering. However, they typically require large labeled datasets and substantial computational power to train. Zhang et al. (2020) applied CNNs to HTML tag sequences and achieved 91% accuracy in phishing detection, highlighting their strong potential in this area.[1] [8] [11] [12]

6.3 Ensemble/Hybrid Models

Ensemble or hybrid models enhance phishing detection by combining multiple machine learning and deep learning algorithms to achieve higher accuracy and greater robustness. A popular method

is stacking, where several base models, such as Random Forest, SVM, or LSTM, make individual predictions, which are then input to a meta-classifier that produces the final decision. Another common technique is weighted voting, where each model's prediction is assigned a weight based on its accuracy or reliability, and the final classification is determined by aggregating all model outputs with weights. These hybrid approaches capitalize on the strengths of different algorithms and mitigate the limitations of individual models. For example, the PhishGuard framework integrates Random Forest, XGBoost, and LSTM models, achieving 97% accuracy across both the PhishTank dataset and custom phishing datasets, demonstrating the effectiveness of hybrid methods for phishing detection.[1] [8] [11] [12]

7. Experimental Setup

The experimental setup establishes the framework for training, validating, and evaluating machine learning and deep learning models in phishing detection. A well-designed setup ensures that the results are reproducible, comparisons between different models are fair, and performance metrics are reliable. It typically includes components such as data partitioning strategies (training, validation, and test splits), selection of evaluation metrics (accuracy, precision, recall, F1-score), cross-validation procedures, hyperparameter tuning methods, and methods for handling imbalanced datasets. By adhering to a rigorous experimental design, researchers can produce trustworthy results that accurately reflect a model's potential in real-world phishing-detection scenarios.[1] [8] [11]

7.1 Data Splitting

Data splitting is a crucial step in preparing datasets for training and testing phishing detection models. A common approach is the train/test split, where typically 70% of the data is used for training the model, and the remaining 30% is held out for testing. This allows the model to learn patterns from features like URLs, HTML structures, and email content, then be evaluated on unseen data to assess its performance. However, relying on a single train/test split can introduce bias, so it is often paired with k-fold cross-validation. In k-fold cross-validation, the dataset is divided into k equal parts (commonly 5 or 10). Each part is used once as a validation set while the other k–1 part are used for training. The model is trained and evaluated k times, and performance scores are averaged, providing a more robust and realistic estimate of the model's generalization ability. [1] [6] [11] [12]

7.2 Evaluation Metrics

Evaluation metrics are essential for assessing how well phishing detection models classify instances as phishing or legitimate.

- **Accuracy** measures the proportion of correctly predicted instances out of the total but can be misleading in imbalanced datasets where one class dominates.

- **Precision** indicates the percentage of instances predicted as phishing that are truly phishing, helping to reduce false positives.
- **Recall** (or sensitivity) measures the proportion of actual phishing instances correctly identified by the model, minimizing false negatives.
- **F1-score** combines precision and recall by calculating their harmonic mean, providing a balanced measure, especially important when both false positives and false negatives carry high costs.

Together, these metrics provide a comprehensive evaluation of a phishing detection model's effectiveness as per Table 4. [12]

Table 4: Metric and Description

Metric	Description
Accuracy	Overall correctness of the model
Precision	Correctness of positive (phishing) predictions
Recall	Ability to detect actual phishing URLs/emails
F1-Score	Balance between precision and recall

7.3 Tools and Frameworks

n phishing detection experiments, modern tools and frameworks are essential for efficient implementation of both machine learning and deep learning models. Python is commonly used as the primary programming language because of its extensive ecosystem of libraries geared toward data preprocessing, model development, and performance evaluation. Traditional machine learning algorithms such as Decision Trees, Random Forests, SVMs, and Naive Bayes are typically implemented using Scikit-learn. This library offers a wide range of utilities, including support for cross-validation and calculation of evaluation metrics, making it highly suitable for building and assessing phishing detection models.[1] [8] [11]

Deep learning models like LSTM and CNN are commonly developed and trained using frameworks such as TensorFlow and Keras, which provide flexibility and scalability for building sophisticated architectures. Data manipulation and preprocessing are efficiently handled with Pandas and NumPy, allowing smooth dataset cleaning and numerical processing. For visualization, tools like Matplotlib and Seaborn are used to plot feature distributions, performance metrics, and experimental results. Together, these libraries and frameworks create a reliable, reproducible, and scalable environment well-suited for conducting phishing detection experiments.[8] [11] [12]

8. Results and Discussion

This section presents a comprehensive performance evaluation of various machine learning, deep learning, and hybrid models used for phishing detection. The analysis utilizes standard evaluation metrics such as accuracy, precision, recall, and F1-score to provide a clear comparison of model effectiveness. Additionally, confusion matrices are examined to understand the distribution of true positives, true negatives, false positives, and false negatives. Receiver Operating Characteristic (ROC) curves are used to assess the trade-off between true-positive and false-positive rates at different thresholds. Feature importance analysis further highlights which features most significantly contribute to the detection performance, offering insights into model decision-making and areas for improvement.[1] [8]

8.1 Model Performance

The performance of both the individual classifiers and hybrid models was measured by the preprocessed phishing datasets, and the summary of the results is given in Table 5. [12]

Table 5: Model Performance Metrics

Model	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	91	0.89	0.90	0.89
Random Forest	95	0.9	0.95	0.9
SVM	92	0.91	0.92	0.91
Naive Bayes	88	0.86	0.87	0.86
LSTM	96	0.95	0.96	0.95
CNN	9	0.93	0.9	0.93
Hybrid (RF + LSTM + XGBoost)	97	0.96	0.97	0.96

Traditional machine learning models like Decision Tree, Random Forest, SVM, and Naive Bayes generally perform well in phishing detection, with Random Forest standing out as the strongest due to its higher accuracy and balanced precision–recall scores. However, deep learning models such as LSTM and CNN tend to outperform traditional approaches, especially on large datasets, because they can capture more complex patterns in URL sequences, HTML structures, and email content. The highest overall performance is achieved by hybrid models that integrate Random Forest, LSTM, and XGBoost, delivering the best accuracy and F1-score. This demonstrates that ensemble approaches are more robust and reliable, effectively addressing the diversity and continuous evolution of phishing techniques. Overall, deep learning's ability to analyze detailed

features and the robustness of hybrid models are key advantages in phishing detection.[1] [6] [8] [11] [12]

8.2 Comparison: Single Classifiers vs Hybrid Models

Hybrid models hold key advantages in phishing detection by combining the strengths of multiple algorithms while minimising the limitations inherent in every single classifier. Single classifiers are typically simple, easy to interpret, and faster to train, making them suitable for basic tasks or smaller datasets. However, they often struggle with complex or large-scale phishing datasets, leading to lower accuracy and higher error rates. In contrast, hybrid models integrate outputs from multiple classifiers using techniques such as stacking or weighted voting. This enables them to capture a broader range of variations in phishing patterns, leading to improved accuracy and fewer false positives and false negatives. Consequently, hybrid approaches offer more robustness and reliability, effectively addressing the evolving nature of phishing threats when compared to single-model methods.[11]

8.3 Confusion Matrices and ROC Curves

Confusion matrices and ROC curves provide detailed insights into model performance in phishing detection. Confusion matrices display the counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), which help identify the types of errors each model makes. For example, the hybrid model typically shows very low FP and FN counts, indicating high reliability in correctly identifying both phishing and legitimate cases.ROC (Receiver Operating Characteristic) curves plot the True Positive Rate (Recall) against the False Positive Rate, providing a comprehensive view of a model's performance across different thresholds. The Area Under the Curve (AUC) quantifies the model’s ability to distinguish between phishing and legitimate instances; a higher AUC indicates better discrimination. Deep learning models such as LSTMs and CNNs, as well as hybrid models, consistently achieve higher AUC values, affirming their effectiveness in capturing complex phishing patterns. Confusion matrices for models such as a) Random Forest, b) LSTM, and c) Hybrid model, along with ROC curves comparing all classifiers, visually highlight these performance differences. [8]

8.4 Feature Importance

Feature importance analysis reveals which features are most decisive in distinguishing phishing from legitimate content. URL-based features dominate, with the presence of the "@" symbol, overall URL length, and the number of subdomains ranking among the most important. HTML-based features, such as the number of forms on a webpage and the count of external links, also contribute significantly, as these are commonly manipulated in phishing sites. In email-based phishing, textual features such as suspicious keywords, grammatical errors, and urgency cues are highly valuable. Overall, while URL features are dominant in website-based phishing detection, combining URL, HTML, and email features improves model generalization and performance across diverse phishing scenarios.[1] [6] [11] [12]

8.5 Discussion

This discussion emphasizes the distinct strengths of various phishing detection approaches. Deep learning models, particularly LSTM and CNN, excel when sufficient labeled data is available. LSTMs effectively analyse sequential patterns in URLs, while CNNs achieve high accuracy by extracting structural features from HTML content. Ensemble or hybrid models, such as combinations of Random Forest with LSTM or XGBoost, offer enhanced robustness by leveraging the strengths of multiple algorithms. These models tend to reduce false positives, generalize well across diverse phishing techniques, and outperform single classifiers on most evaluation metrics. Feature engineering remains critical, with meaningful extraction from URLs, HTML structures, and email content contributing significantly to model performance. Feature importance analysis further validates the critical role of URL anomalies and textual cues in phishing emails.

Key takeaways: deep learning is well-suited to large, complex datasets; ensemble models deliver the most reliable and accurate results; URL and email features are vital for detection; and well-engineered traditional models, such as Random Forests, remain strong baseline performers.[1] [6] [8] [11] [12]

9. Case Study / Real-World Application

Phishing attacks pose a significant cybersecurity threat to corporate environments, often targeting employees through malicious emails and fraudulent websites. Implementing an automated phishing detection system is crucial for protecting sensitive organizational data, preventing financial losses, and strengthening overall security posture. Such systems help detect and mitigate phishing attempts in real time, reducing the risk of successful attacks and safeguarding both the company’s assets and reputation.[6] [11]

9.1 Detecting Phishing Emails

Phishing email detection is critical in corporate environments where employees receive hundreds of emails daily, some of which may carry threats. A hybrid phishing detection system leveraging Random Forest, LSTM, and XGBoost can play a vital role in automatically scanning incoming emails for potential risks. The detection process begins with feature extraction: URL features such as the presence of "@", abnormal subdomains, URL length, and domain age are analyzed. Next, HTML features like embedded forms, iFrames, and outgoing links within the email body are examined. Finally, textual features including suspicious keywords, grammatical errors, and urgency cues are evaluated. Based on these extracted features, the hybrid model predicts a phishing probability score for each email. Emails with a high-risk score are flagged, quarantined, or blocked before they reach employees, significantly reducing the likelihood that harmful links will be clicked. This proactive approach helps prevent credential theft, data breaches, and the negative impact on organizational security and financial reputation.[1] [6] [11]

This can be more effectively represented visually with the flowchart: Incoming Email → Feature Extraction → ML/DL Model → Quarantine/Flag/Allow.

9.2 Integration with Corporate Infrastructure

Exactly, phishing detection systems work most effectively when they are seamlessly integrated with an organization's existing IT and security setup. The following are several corporate tools that the systems can connect with for real-time protection:

Email servers, such as Microsoft Exchange and Gmail for Business, support integrating phishing detection models via APIs or plugins. These integrations enable real-time scanning of incoming emails before they reach the user's inbox, allowing potential phishing threats to be identified and mitigated proactively. By embedding detection mechanisms at the server level, organizations can enhance their email security and reduce the risk of phishing attacks affecting employees.[6] [11]

Web security tools such as firewalls, proxy servers, and secure web gateways can integrate phishing detection models to identify and block malicious URLs or websites before employees access them. This proactive blocking helps prevent exposure to phishing sites and reduces the risk of cyberattacks.

Security Information and Event Management (SIEM) platforms play a crucial role by aggregating alerts from phishing detection systems into a unified dashboard. This centralization enables security teams to monitor threats in real time and respond rapidly, improving overall organizational security and incident management.[6] [11]

Benefits of Integration:

A well-integrated phishing detection system provides seamless real-time protection without slowing down or interrupting the user’s workflow. It offers automated notifications that keep IT and security teams promptly informed about potential threats. This enables faster incident response by reducing reliance on manual monitoring, thereby minimising overall security risks and strengthening the organisation's defence against phishing attacks. [1]

9.3 Impact on Cybercrime Reduction

Implementing phishing detection systems in corporate environments leads to significant, measurable reductions in cybercrime incidents. These systems proactively filter malicious content, thereby reinforcing the organization's overall cybersecurity posture.

Reduced Credential Theft: By detecting and blocking phishing attempts before they reach employees, the system lowers the risk of credential harvesting attacks, preventing unauthorized access to sensitive corporate accounts.

Reduced Financial Loss: Automated detection stops employees from interacting with fraudulent emails or links, minimizing financial losses caused by scams, data breaches, and the costs associated with system recovery.

Improved Security Awareness: With fewer phishing emails reaching inboxes, employee confidence in corporate communications increases. This also enhances cybersecurity awareness, enabling staff to better distinguish legitimate messages from suspicious ones.

Regulatory Compliance: Phishing detection tools assist organisations in meeting key data protection and cybersecurity regulations, such as GDPR and HIPAA, by facilitating secure handling of digital communications, thereby supporting compliance efforts. [6] [11]

Example Outcome:

A company that implemented a hybrid phishing detection system combining machine learning and deep learning techniques reported significant improvements: a 70% decrease in successful phishing attacks and a 50% reduction in the IT department's incident response workload. This reduction allows security teams to focus more effectively on critical security challenges, enhancing overall organizational cybersecurity. [1] [8]

10. Challenges and Limitations

Despite significant advances in phishing detection through machine learning (ML) and deep learning (DL), several challenges continue to limit the effectiveness and widespread deployment of these systems:

10.1 Evolving Phishing Tactics

Phishers constantly refine their techniques to evade detection systems. They use new phishing URLs, domain spoofing, and increasingly sophisticated email content, rendering static detection methods ineffective. To keep up, models must be frequently retrained to recognize these evolving tactics, which raises maintenance demands and operational complexity. [6] [11]

10.2 Imbalanced Datasets

Phishing datasets typically contain far fewer phishing samples than legitimate ones, resulting in class imbalance. This imbalance can bias models toward predicting legitimate emails, reducing recall and causing some phishing attacks to go undetected. To address this, techniques such as oversampling, undersampling, or synthetic data generation methods like SMOTE are applied, but these approaches add further complexity to the model training process. [11] [12]

10.3 False Positives vs. False Negatives

Balancing false positives and false negatives is a critical challenge in phishing detection systems, given the significant consequences of both types of errors. False positives, legitimate emails

incorrectly flagged as phishing, can disrupt workflows, delay important communications, and erode user trust through excessive false alarms. Conversely, false negatives, phishing emails that bypass detection, pose far greater risks, including credential theft, malware infection, financial losses, and reputational damage. Even a small number of false negatives can result in severe security breaches. Achieving an optimal balance requires careful threshold tuning, continuous model retraining, and comprehensive evaluation using metrics such as precision, recall, and F1-score. This approach ensures the system remains accurate, reliable, and effective in real-world corporate environments. [6] [11]

10.4 Computational Cost of Complex Models

Deep learning models like LSTMs and CNNs require substantial computational resources for both training and inference. Their reliance on large datasets and complex network architectures results in prolonged training times, high GPU utilisation, and increased energy consumption. These requirements can pose significant challenges for deployment in resource-constrained environments, limiting the practical use of deep learning-based phishing-detection systems in such environments.[8] [11] [12]

11. Future Work

Phishing attacks continuously evolve, necessitating that detection systems advance in tandem to remain effective. Although current models achieve high accuracy, future research should focus on enhancing real-world applicability and resilience through several key areas:

11.1 Real-Time Phishing Detection

Future advancements in phishing detection aim to embed machine learning and deep learning models directly into email servers, web gateways, and network security infrastructure to enable real-time detection of malicious content. These systems must be optimized for low-latency prediction to maintain high speed without compromising accuracy.

Additionally, deploying lightweight versions of detection models on edge devices, such as user endpoints, can further decentralize security operations. This decentralization speeds up response times, significantly narrowing the window of vulnerability for credential theft, data breaches, and financial losses. Real-time detection thus plays a crucial role in strengthening organizational cybersecurity by allowing immediate identification and blocking of threats.[1] [6] [8] [11]

11.2 Continuous Learning Models

As phishing tactics rapidly evolve, static detection models risk becoming outdated, underscoring the need for continuous learning to drive future advancement. Continuous learning enables models to adapt to emerging phishing techniques without undergoing full retraining, preserving previously learned patterns while incorporating new knowledge. This can be achieved through incremental

learning, which updates models gradually with newly labeled data, and online learning algorithms that process streaming emails, URLs, or web activity in real time.

Furthermore, integrating detection systems with threat intelligence feeds allows automatic updates with newly reported malicious domains, URLs, and phishing indicators. Continuous learning thus ensures that models evolve alongside attacker behavior, significantly enhancing robustness and maintaining strong, up-to-date protection against increasingly sophisticated phishing attacks. [6] [11]

11.3 Multi-Lingual Phishing Detection

As phishing attacks increasingly target users across regions and languages, developing multilingual phishing detection capabilities is essential. Most current systems are optimized for English, which limits their effectiveness in global environments where attackers use localized languages to appear more credible. Advances in natural language processing (NLP), including multi-lingual embeddings and text vectorization methods, will enable models to analyze email content in multiple languages and capture phishing cues despite varied linguistic structures.

Cross-lingual transfer learning further boosts this capability by allowing models trained on extensive English datasets to generalize their understanding to other languages with minimal additional data. Implementing multi-lingual intelligence in phishing detection systems helps organizations strengthen their cybersecurity posture by consistently protecting their diverse and geographically dispersed users.[11] [12]

12. Conclusion

Phishing attacks remain one of the most prevalent and damaging forms of cybercrime, targeting both individuals and organizations. Traditional rule-based detection methods, such as blacklists and signature matching, are no longer sufficient due to attackers' constantly evolving tactics. Machine learning (ML) and deep learning (DL) approaches have demonstrated significant promise in addressing these challenges by providing adaptive, data-driven solutions.

ML models, such as Decision Trees, Random Forests, Support Vector Machines, and Naive Bayes, offer interpretable and efficient methods for detecting phishing attempts using URL, HTML, and email features. While these models achieve reasonable accuracy, deep learning architectures such as LSTMs and CNNs excel at capturing sequential patterns and complex structures in large datasets, offering superior performance in detecting subtle phishing cues.

Hybrid and ensemble models that combine multiple classifiers further enhance detection capabilities by leveraging the strengths of individual algorithms. For example, combining Random Forest, LSTM, and XGBoost produces higher accuracy, reduces false positives and false negatives, and improves robustness against varied phishing strategies. These models are particularly effective in real-world applications where attacks continuously evolve.

The practical deployment of ML-based phishing detection systems can significantly strengthen enterprise cybersecurity. Integrating these systems with email servers, web gateways, and security information platforms enables real-time monitoring, proactive threat mitigation, and automated alerts for IT teams. Such integration reduces financial losses, prevents data breaches, and enhances overall organizational security posture.

Moreover, feature engineering and importance analysis play a crucial role in model performance. Identifying key indicators, such as suspicious URL structures, domain characteristics, HTML forms, and email content patterns, allows for more targeted and efficient detection. As phishing tactics continue to evolve, continuous learning models and multi-lingual detection capabilities will further enhance system adaptability and global applicability.

In summary, machine learning enables adaptive, accurate phishing detection, while hybrid models outperform single classifiers in both robustness and accuracy. The future integration of these models into enterprise cybersecurity infrastructure is highly recommended, as it not only safeguards sensitive data and financial assets but also fosters a proactive and resilient security culture. The combination of adaptive ML techniques, continuous feature updates, and real-world deployment can significantly reduce the impact of phishing attacks and contribute to a safer digital environment.

References

[1] Aslam, S., Aslam, H., Manzoor, A., Chen, H., & Rasool, A. (2024). AntiPhishStack: LSTM-based stacked generalization model for optimized phishing URL detection. Symmetry, 16(2), 248.

[2] Ovi, M. S. I., Rahman, M. H., & Hossain, M. A. (2024, December). Phishguard: A multi-layered ensemble model for optimal phishing website detection. In *2024 6th International Conference on Sustainable Technologies for Industry 5.0 (STI)* (pp. 1-6). IEEE.

[3] Mittal, A., Engels, D. D., Kommanapalli, H., Sivaraman, R., & Chowdhury, T. (2022). Phishing detection using natural language processing and machine learning. SMU Data Science Review, 6(2), 14.

[4] Mandadi, A., Boppana, S., Ravella, V., & Kavitha, R. (2022, April). Phishing website detection using machine learning. In 2022 IEEE 7th International conference for Convergence in Technology (I2CT) (pp. 1-4). IEEE.

[5] Jain, A. K., & Gupta, B. B. (2016, March). Comparative analysis of features based machine learning approaches for phishing detection. In 2016 3rd international conference on computing for sustainable global development (INDIACom) (pp. 2125-2130). IEEE.

[6] Alshingiti, Z., Alaqel, R., Al-Muhtadi, J., Haq, Q. E. U., Saleem, K., & Faheem, M. H. (2023). A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN. *Electronics*, *12*(1), 232.

[7] Balogun, A. O., Mojeed, H. A., Adewole, K. S., Akintola, A. G., Salihu, S. A., Bajeh, A. O., & Jimoh, R. G. (2021). Optimized decision forest for website phishing detection. In Proceedings of the Computational Methods in Systems and Software (pp. 568-582). Cham: Springer International Publishing.

[8] Khari, M., Shrivastava, G., Gupta, S., & Gupta, R. (2017). Role of cyber security in today's scenario. In *Detecting and mitigating robotic cyber security risks* (pp. 177-191). IGI Global Scientific Publishing.

[9] Roy, S. S., Awad, A. I., Amare, L. A., Erkihun, M. T., & Anas, M. (2022). Multimodel phishing URL detection using LSTM, bidirectional LSTM, and GRU models. Future Internet, 14(11), 340.

[10] Niu, W., Zhang, X., Yang, G., Ma, Z., & Zhuo, Z. (2017, December). Phishing emails detection using CS-SVM. In *2017 IEEE international symposium on parallel and distributed processing with applications and 2017 IEEE international conference on ubiquitous computing and communications (ISPA/IUCC)* (pp. 1054-1059). IEEE.

[11] PhishTank. (2025). Community-Based Phishing URL Database. Available: https://www.phishtank.com

[12] Asuncion, A., & Newman, D. (2007, November). *UCI machine learning repository*.

[13] ISCX. (2016). ISCX-URL-2016 Phishing Dataset. Available: https://www.unb.ca/cic/datasets/url-2016.html

[14] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2, pp. 1-800). Cambridge: MIT Press.

[15] Brownlee, J. (2016). *Machine learning mastery with Python: understand your data, create accurate models, and work projects end-to-end*. Machine Learning Mastery.