International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 552-564.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19500101

Real-Time AI & ML Based Phishing Detection and Prevention System

K. Yacob Raju, K. Niharika Sai Anjali, N. Harsha, N. Lakshmi Prasanna, V. Sri Ashlesha

Department of AI & ML, Vijaya Institute of Technology for Women, Enikepadu, Vijayawada, India.

To Cite this Article
K. Yacob Raju, K. Niharika Sai Anjali, N. Harsha, N. Lakshmi Prasanna & V. Sri Ashlesha (2026). Real-Time AI & ML Based Phishing Detection and Prevention System. International Journal for Modern Trends in Science and Technology, 12(04), 552-564. https://doi.org/10.5281/zenodo.19500101

Article Info
Received: 10 March 2026; Revised: 02 April 2026; Accepted: 05 April 2026.

KEYWORDS

Phishing Detection, Machine Learning, Artificial Intelligence, Cybersecurity, Real-Time Systems.

ABSTRACT

Phishing attacks have emerged as one of the most critical cybersecurity threats, targeting users through deceptive emails, websites, and messages to steal sensitive information such as login credentials and financial data. Traditional phishing detection systems rely on static rules and blacklists, which fail to detect newly emerging and sophisticated attacks in real time. This paper proposes a real-time phishing detection and prevention system using Artificial Intelligence (AI) and Machine Learning (ML) techniques. The proposed system analyzes multiple features such as URL structure, domain information, and content-based characteristics to classify websites and messages as phishing or legitimate. Various machine learning algorithms are trained on labelled datasets to improve detection accuracy while minimizing false positives. Experimental results show that the proposed system effectively detects phishing attacks with high accuracy and faster response time. This intelligent approach enhances user security and provides a scalable solution for modern cybersecurity challenges.

1. INTRODUCTION

The rapid growth of internet technologies and online services has greatly increased the number of users performing activities such as online banking, shopping, and communication through digital platforms. However, this growth has also led to a rise in cyber threats, particularly phishing attacks. Phishing is a fraudulent activity where attackers attempt to obtain sensitive information such as usernames, passwords, and financial details by impersonating legitimate organizations through fake emails, websites, or messages.

Traditional phishing detection techniques, such as blacklist-based filtering and rule-based security systems, are often unable to detect newly generated phishing websites and sophisticated attack strategies. These conventional methods rely heavily on previously known

threats and therefore struggle to identify emerging phishing patterns in real time.

Artificial Intelligence (AI) and Machine Learning (ML) technologies provide an effective approach to overcome these limitations. By analyzing large datasets and identifying hidden patterns, AI and ML models can accurately detect suspicious activities and classify phishing attempts. These intelligent systems can continuously learn from new data, making them more adaptable to evolving cyber threats.

The Real-Time AI & ML Based Phishing Detection and Prevention System aims to enhance online security by detecting phishing websites, emails, and malicious links in real time. The proposed system uses machine learning algorithms to analyze various features such as URL structure, website content, and domain information to identify potential phishing attacks.

By implementing this system, users can be protected from fraudulent websites and malicious activities, thereby reducing the risk of data theft and financial loss. The project focuses on improving detection accuracy, minimizing false positives, and providing a reliable cybersecurity solution for safe online interactions.

1.1. Objectives:

➢ 1. To analyze different types of phishing attacks and identify the common characteristics used by attackers to deceive users.

➢ 2. To develop an intelligent phishing detection system using Artificial Intelligence (AI) and Machine Learning (ML) techniques.

➢ 3. To extract and analyze important features such as URL structure, domain information, and webpage content for accurate phishing detection.

➢ 4. To implement a real-time system that can automatically detect and prevent phishing websites, emails, or malicious links.

➢ 5. To improve the accuracy and efficiency of phishing detection while minimizing false positives and false negatives.

➢ 6. To enhance cybersecurity and protect users from data theft, financial fraud, and other online threats.

1.2 Principles of AI & ML Based Phishing Detection System:

• Artificial Intelligence Based Detection: Artificial Intelligence enables systems to automatically analyze large volumes of data and identify suspicious patterns that may indicate phishing activities. AI-based systems can learn from historical data and improve their detection capabilities over time.

• Machine Learning Algorithms: Machine Learning algorithms play a key role in phishing detection by training models on datasets containing both legitimate and phishing examples. These algorithms learn patterns and features that distinguish phishing websites or emails from genuine ones, enabling accurate classification.

• Feature Extraction: Feature extraction is an important step where relevant characteristics of URLs, websites, or emails are analyzed. These features may include URL length, presence of special characters, domain age, HTTPS usage, and webpage content. Extracting these features helps the ML model identify suspicious behavior.

• Real-Time Detection: The proposed system works in real time by analyzing user input such as website URLs or email links. The trained machine learning model processes the input and determines whether it is legitimate or phishing, providing immediate alerts to users.

•Continuous Learning: AI and ML systems can continuously improve by learning from newly detected phishing attacks. This helps the system adapt to evolving cyber threats and maintain high detection accuracy.

• User Protection Mechanism: Once a phishing attempt is detected, the system alerts the user and blocks access to the malicious website or link, thereby preventing data theft and financial fraud.

1.3 Processes Involved

• Data Collection: The first step involves collecting datasets that contain both phishing and legitimate URLs or emails. These datasets are used to train and test the machine learning model.

Dataset → Phishing Data + Legitimate Data

• Data Preprocessing: In this stage, the collected data is cleaned and prepared for analysis. Duplicate entries, missing values, and irrelevant data are removed to improve the quality of the dataset.

Raw Data → Cleaned Data → Structured Dataset

• Feature Extraction: Important features are extracted from the dataset to help the model differentiate between phishing and legitimate websites.

URL/Email Data → Feature Extraction → Feature Set

Examples of features include:

- URL length
- Presence of "@" symbol
- Number of subdomains
- HTTPS usage
- Domain age

• Model Training: Machine learning algorithms such as Decision Tree, Random Forest, or Support Vector Machine (SVM) are used to train the model using the extracted features.

Training Dataset → ML Algorithm → Trained Model

• Classification: The trained model analyzes new URLs or emails and classifies them as phishing or legitimate.

Input URL → ML Model → Phishing / Legitimate

• Real-Time Detection and Prevention: When a user accesses a suspicious link, the system evaluates it instantly and provides a warning if it is detected as phishing.

Suspicious URL → Detection System → Alert / Block

• Accuracy Evaluation: The performance of the model is evaluated using metrics such as accuracy, precision, recall, and F1-score to ensure reliable detection.

1.4 Operating Conditions:

o Dataset Quality: The effectiveness of the phishing detection system depends on the quality and size of the dataset used for training the machine learning model. A dataset containing both phishing and legitimate URLs or emails is required to build an accurate detection system.

o Feature Selection: Relevant features such as URL length, number of special characters, presence of HTTPS, domain age, and webpage content must be carefully selected. Proper feature selection improves the accuracy and efficiency of the machine learning model.

o Model Training Parameters: Machine learning models require proper tuning of parameters such as training epochs, learning rate, and classification thresholds. Optimizing these parameters helps improve model performance and detection accuracy.

o Real-Time Processing: The system should be capable of analyzing URLs or emails instantly to provide real-time detection and warning messages to users. Efficient processing ensures faster response and better user protection.

1.5 Materials & Methods:

The materials and methods used in developing a Real-Time AI & ML Based Phishing Detection and Prevention System involve several components and steps to ensure accurate detection and reliable system performance.

a) Materials

1. Dataset: The dataset consists of phishing and legitimate URLs or emails collected from reliable cybersecurity sources. This dataset is used for training and testing the machine learning models.

2. Machine Learning Algorithms: Various machine learning algorithms such as Decision Tree, Random Forest, Logistic Regression, or Support Vector Machine (SVM) are used to classify websites or emails as phishing or legitimate.

3. Software Tools: Programming languages and tools such as Python, Jupyter Notebook, and machine learning libraries (Scikit-learn, Pandas, NumPy) are used to develop and implement the phishing detection system.

4. Feature Extraction Tools: Feature extraction techniques are used to analyze URL structures, domain information, and webpage content to identify suspicious patterns related to phishing attacks.

b) Methods

1.Data Collection: The first step involves collecting phishing and legitimate URL datasets from publicly available cybersecurity repositories.

2.Data Preprocessing: The collected dataset is cleaned by removing duplicate entries, handling missing values, and organizing the data for analysis.

3.Feature Extraction: Important features are extracted from the dataset, including URL length, number of dots in the domain, presence of special symbols, domain age, and use of HTTPS protocol.

4. Model Training: Machine learning algorithms are trained using the extracted features to learn patterns that distinguish phishing websites from legitimate ones.

5. Testing and Evaluation: The trained model is tested using a separate dataset to evaluate its performance based on accuracy, precision, recall, and F1-score.

6. Real-Time Detection: The final system analyzes new URLs or emails in real time and alerts the user if a phishing attempt is detected.

Analytical Methods

• Performance Evaluation: Evaluation of the system using metrics such as accuracy, precision, recall, and

F1-score to measure the effectiveness of the phishing detection model.

• Model Validation: Cross-validation techniques are used to ensure the reliability and consistency of the machine learning model.

• Security Analysis: The system is tested with various phishing scenarios to evaluate its ability to detect different types of phishing attacks.

2. EXISTING PHISHING DETECTION SYSTEMS

Artificial Intelligence (AI) and Machine Learning (ML) have significantly transformed the field of cybersecurity, especially in detecting phishing attacks. Phishing detection systems are designed to identify fraudulent websites, emails, and links that attempt to steal sensitive user information such as usernames, passwords, and financial details. These systems use various techniques such as rule-based filtering, blacklist databases, heuristic analysis, and machine learning algorithms to detect malicious activities and protect users from cyber threats.

In earlier approaches, phishing detection mainly relied on blacklist-based methods, where known phishing URLs were stored in databases and blocked when accessed. Although this method is simple and fast, it is not effective in detecting newly created phishing websites. Similarly, rule-based systems depend on predefined patterns and rules, which makes them less flexible and unable to adapt to evolving attack strategies.

With the increasing sophistication of phishing attacks, traditional detection methods are becoming less effective. Attackers continuously modify their techniques, such as using URL obfuscation, domain spoofing, and social engineering tactics, to bypass existing security systems. As a result, there is a growing need for intelligent and adaptive solutions that can identify both known and unknown phishing attempts.

2.1 Major Types of Phishing Detection Systems

AI and ML-based phishing detection systems can be broadly classified based on their functionality:

1. URL-Based Detection Systems: These systems analyze the structure of URLs to identify suspicious patterns such as long URLs, use of special characters, multiple subdomains, presence of IP addresses instead of domain names, and absence of HTTPS protocol. URL-based systems are fast and efficient, making them suitable for real-time detection. However, they may not always detect highly sophisticated phishing attacks that closely resemble legitimate URLs.

2. Blacklist-Based Detection Systems: These systems maintain a database of known phishing websites and block access when a match is found. They are simple to implement and provide quick results. However, they are limited in detecting new or unknown phishing sites, as attackers frequently create new domains to bypass blacklist mechanisms.

3. Machine Learning-Based Detection Systems: These systems use ML algorithms to classify websites or emails as phishing or legitimate based on extracted features. They can detect new phishing attacks by learning patterns from training data. Algorithms such as Decision Tree, Random Forest, Support Vector Machine (SVM), and Logistic Regression are commonly used. These systems offer higher accuracy and adaptability compared to traditional methods.

4. Content-Based Detection Systems: These systems analyze webpage content, including text, images, forms, and scripts, to identify phishing attempts that mimic legitimate websites. They can detect fake login pages, suspicious form fields, and copied website layouts. Content-based systems are effective in identifying visually deceptive phishing attacks.

5. Email-Based Detection Systems: These systems focus on detecting phishing emails by analyzing email headers, sender information, message content, and embedded links. They use techniques such as Natural Language Processing (NLP) to detect suspicious language patterns and identify malicious intent.

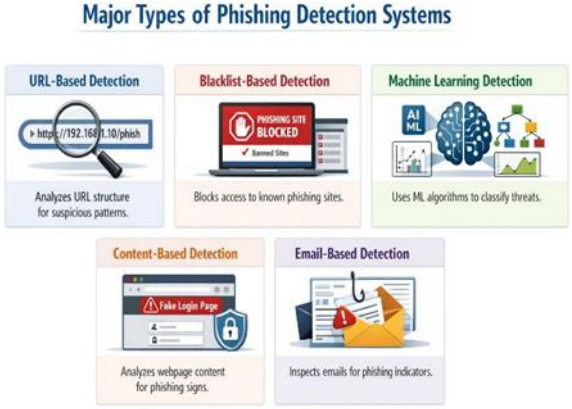

Fig 2: Types of Phishing Detection Systems

2.2 Examples of Phishing Detection Systems

URL Detection Systems: These systems evaluate URLs in real time and classify them as safe or malicious based on

predefined rules or machine learning models. They are commonly used in browser extensions and security applications.

Browser Security Tools: Modern web browsers include built-in phishing protection mechanisms that warn users when they attempt to visit suspicious websites. These tools continuously update their databases to provide better security.

Email Filtering Systems: Email platforms use spam filters and ML-based classifiers to detect and block phishing emails before they reach users. These systems analyze large volumes of emails and identify malicious patterns efficiently.

Enterprise Security Systems: Organizations use advanced phishing detection solutions integrated into their network infrastructure to protect users from large-scale phishing attacks and data breaches.

Benefits of Phishing Detection Systems

AI and ML-based phishing detection systems offer several advantages:

Improved Security: Protects users from phishing attacks, identity theft, and financial fraud.

Real-Time Detection: Provides instant alerts when suspicious activity is detected, reducing the chances of user interaction with malicious content.

Reduced Human Effort: Automates the detection process, minimizing manual monitoring and intervention.

Adaptive Learning: Continuously improves by learning from new phishing techniques and evolving attack patterns.

Scalability: Can handle large volumes of data and users without significant performance degradation.

Wide Accessibility: Can be integrated into browsers, email systems, mobile applications, and enterprise security platforms.

2.3 Software Requirements

Software requirements define the tools, technologies, and platforms needed for developing and deploying the phishing detection system. A proper software stack ensures system efficiency, scalability, security, and maintainability.

Operating System

Supported Operating Systems:

- Windows 10 / 11
- Linux (Ubuntu preferred for deployment)
- macOS (for development and testing)

Linux-based systems are preferred for deployment due to their stability, performance efficiency, and strong support for server-side applications.

Programming Languages

The system requires programming languages capable of handling machine learning and web integration:

- Python (for AI/ML model development and backend processing)
- JavaScript (Optional for frontend and web-based interfaces)

Artificial Intelligence and ML Libraries

These libraries are essential for building the detection model:

- TensorFlow / PyTorch (for deep learning models)
- Scikit-learn (for traditional ML algorithms)

Development Tools and IDEs

Tools used for coding and development:

- IDE: Visual Studio Code, PyCharm
- Version Control: Git and GitHub
- Virtual Environments: Anaconda, venv

These tools improve development efficiency, debugging, and collaboration.

Security and Authentication Software

Security mechanisms ensure safe handling of user data and system integrity:

- SSL/TLS for secure communication
- Authentication and authorization systems
- Secure API access and data encryption

Testing and Monitoring Tools

Testing ensures system reliability and performance:

- Unit Testing Tools (for individual module testing)
- Integration Testing Tools (for system-level testing)
- Performance Monitoring Tools (to track system efficiency and response time)

3.EXISTING PHISHING DETECTION SYSTEMS WORKFLOW

The workflow of existing phishing detection systems generally includes the following steps:

User Input: Users interact with the system by entering a URL, clicking on a link, or receiving an email containing

a potential phishing attempt. The system captures this input for further analysis.

Data Processing: The input data (URL or email) is processed and converted into a structured format. Important elements such as URL structure, domain information, and email content are extracted for analysis.

Feature Extraction: The system extracts key features such as URL length, number of special characters, presence of HTTPS, domain age, and suspicious keywords. These features help in identifying patterns related to phishing attacks.

Decision-Making: A rule-based system or machine learning model is used to classify the input as phishing or legitimate. Traditional systems rely on predefined rules or blacklists, while advanced systems use trained ML models for better accuracy.

Flow of Existing Phishing Detection Methodology

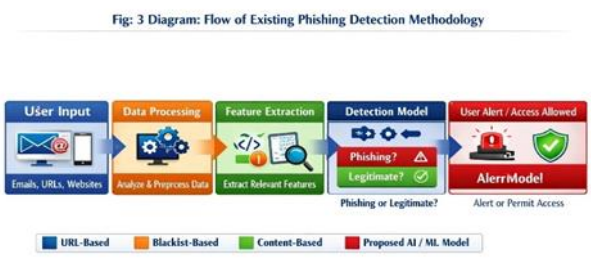

Fig 3: Flow of Existing Phishing Detection Methodology

3.1 Performance Evaluation Metrics

Performance evaluation metrics are essential for measuring the effectiveness, accuracy, and reliability of the phishing detection system. These metrics help determine whether the system can correctly identify phishing attacks and provide secure results to users.

Accuracy

Accuracy measures the overall correctness of the system in classifying URLs or emails.

- Accuracy is calculated by comparing predicted results with actual outcomes.
- High accuracy indicates effective phishing detection.
- Low accuracy may lead to security risks or incorrect classifications.

Precision

Precision measures the proportion of correctly identified phishing instances among all instances classified as phishing.

- High precision indicates fewer false positives.
- It ensures that legitimate websites are not incorrectly blocked.

Recall

Recall measures the ability of the system to correctly identify all actual phishing instances.

- High recall ensures that most phishing attacks are detected.
- It reduces the chances of missing malicious websites.

F1-Score

The F1-score is the harmonic mean of precision and recall.

- It provides a balanced measure of system performance.
- A higher F1-score indicates better reliability and consistency of the system.

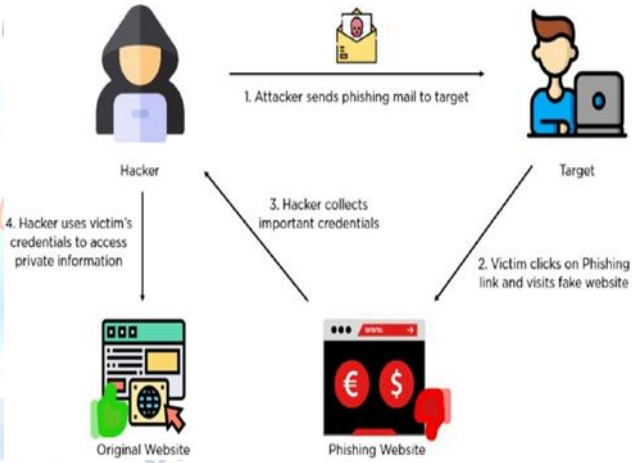

Fig 3.1: Existing Phishing Detection System Workflow

4. FEASIBILITY STUDY

4.1 Technical Feasibility

Technical feasibility evaluates whether the required technologies, tools, and technical expertise are available to develop and implement the phishing detection system.

The proposed system is based on well-established technologies such as Artificial Intelligence (AI), Machine Learning (ML), and data analysis techniques. These technologies are widely used in cybersecurity applications and are supported by numerous libraries and frameworks, making the system development technically feasible.

Programming languages such as Python provide strong support for machine learning through libraries like Scikit-learn, TensorFlow, and Pandas. These tools enable

efficient data processing, feature extraction, and model development. Additionally, web technologies can be used to integrate the detection system into browsers or applications for real-time usage.

The system can be deployed on local machines or cloud platforms, ensuring scalability and efficient handling of large datasets. With the availability of datasets, tools, and computing resources, the technical implementation of the phishing detection system is practical and achievable.

4.2 Economic Feasibility

Economic feasibility determines whether the proposed system is cost-effective and financially viable.

The phishing detection system is economically feasible as it mainly relies on open-source tools and software, which significantly reduces development costs. Programming languages like Python and libraries such as Scikit-learn and TensorFlow are freely available, eliminating the need for expensive software licenses.

The system does not require costly hardware infrastructure, as it can be developed and tested on standard computers. Deployment can also be done using cloud platforms, which offer flexible and affordable pricing models.

Compared to traditional cybersecurity solutions that require extensive manual monitoring and high operational costs, the proposed AI-based system reduces human effort and improves efficiency, resulting in long-term cost savings.

4.3 Operational Feasibility

Operational feasibility analyzes whether the system can be effectively used and accepted by users.

The proposed phishing detection system is designed with a simple and user-friendly interface, allowing users to easily input URLs or interact with the system without requiring technical knowledge. The system provides clear alerts and warnings when a phishing attempt is detected, helping users make safe decisions while browsing.

The real-time detection capability ensures that users receive immediate feedback, enhancing usability and trust in the system. The system can be integrated into web browsers, email platforms, or mobile applications, making it accessible to a wide range of users.

Additionally, the automated nature of the system reduces the need for constant human supervision, making it suitable for both individual users and organizations. Overall, the system is easy to operate, efficient, and highly adaptable to real-world applications.

Fig: 4 Diagram: AI & ML Based Phishing Detection System Workflow

Fig 4: Diagram: AI & ML Based Phishing Detection System Workflow

5. SYSTEM ARCHITECTURE

The system architecture defines the overall structure, components, and data flow of the Real-Time AI & ML Based Phishing Detection and Prevention System. It provides a clear understanding of how different modules interact to detect and prevent phishing attacks efficiently. The proposed architecture is designed to be modular, scalable, secure, and efficient, ensuring easy development, deployment, and future enhancements.

5.1.1 Architectural Overview

The phishing detection system follows a layered architecture, where each layer performs a specific function. This approach improves system maintainability, flexibility, and performance.

- User Interface Layer
- Application Layer (Detection Engine)
- Feature Extraction Layer
- Machine Learning Model Layer
- Database and Data Layer
- Integration Layer
- Deployment and Security Layer

Each layer communicates with adjacent layers through well-defined interfaces, ensuring smooth data flow and reliable system operation.

5.1.2 User Interface Layer

The User Interface (UI) layer acts as the interaction point between the user and the system. It is designed to be simple and user-friendly.

Functions of UI Layer:

- Accepts user input such as URLs or suspicious links

• Displays detection results (Phishing / Legitimate)
• Provides warning messages and alerts for unsafe websites
• Can be implemented as a web interface, browser extension, or mobile application

5.1.3 Application Layer (Detection Engine)

The application layer serves as the core processing unit of the system. It manages communication between different layers and controls the workflow of phishing detection.

Responsibilities:

•Receives input from the user interface
•Sends input data to the feature extraction layer
•Processes results from the ML model
•Generates final output for the user

5.1.4 Feature Extraction Layer

The feature extraction layer is responsible for analyzing the input data and extracting important features required for phishing detection.

Key Features Extracted:

• URL length and structure
• Number of special characters and subdomains
• Presence of HTTPS protocol
• Domain age and registration details
• Suspicious keywords and patterns

These features are essential for training and testing the machine learning model.

5.1.5 Machine Learning Model Layer

This layer contains the trained machine learning model used for classification.

Functions:

•Processes extracted features
•Classifies input as phishing or legitimate
• Uses algorithms such as Decision Tree, Random Forest, or SVM
• Continuously improves through retraining with new data

5.1.6 Database and Data Layer

The data layer stores datasets and relevant information required for system operation.

Contents of Data Layer:

• Phishing and legitimate URL datasets
• Training and testing data
• Extracted feature sets
• Model performance records

This layer supports regular updates to maintain accuracy and relevance.

5.1.7 Integration Layer

The integration layer connects the system with external platforms and services.

Integration Features:

• Browser integration for real-time detection
• Email system integration for phishing email detection
• API support for external cybersecurity services
• Real-time data updates from online threat intelligence sources

5.1.8 Deployment and Security Layer

The deployment layer ensures reliable system performance in real-world environments while maintaining security.

Deployment Features:

• Cloud-based or local deployment
• Scalable architecture to handle multiple users
• Load balancing for performance optimization

Security Features:

• Secure data communication using SSL/TLS
• Encryption of sensitive data
• Authentication and access control mechanisms
• Protection against unauthorized access

The proposed system architecture for the Real-Time AI & ML Based Phishing Detection and Prevention System provides a comprehensive framework that integrates multiple functional layers to ensure efficient and accurate phishing detection. Each layer is designed with a specific responsibility, enabling smooth data flow, modular implementation, and ease of maintenance. The layered design not only improves system reliability but also supports scalability for handling large volumes of real-time data.

The system is further designed with a focus on security and robustness, incorporating secure communication mechanisms and data protection strategies. This ensures that user data and system operations remain safe from unauthorized access and potential cyber threats.

Overall, the architecture not only fulfills the current requirements of phishing detection but also provides a flexible and future-ready platform capable of adapting to new challenges in cybersecurity, thereby ensuring long-term effectiveness and reliability.

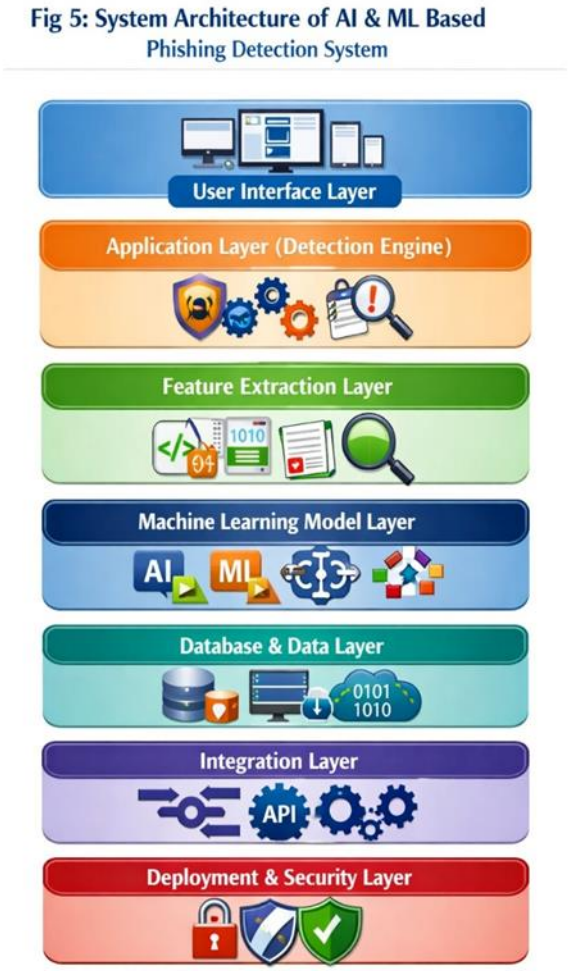

Fig 5: System Architecture of AI & ML Based Phishing Detection System

6. SYSTEM IMPLEMENTATION

The implementation of the Real-Time AI & ML Based Phishing Detection and Prevention System represents the transition from conceptual design to a fully functional system capable of detecting phishing attacks in real time. The system is designed to analyze URLs, emails, and web content to identify malicious activities and protect users from cyber threats. The implementation follows a modular, scalable, and secure approach, integrating multiple components that work together efficiently to deliver accurate detection results.

6.1 Implementation Overview

The system implementation involves the integration of the following major components:

Each component plays a crucial role in ensuring that the system can process input data, extract meaningful features, classify phishing attempts, and provide real-time alerts to users.

6.2 User Interface Module

The User Interface serves as the front-end through which users interact with the system. It is designed to be simple, responsive, and easy to use across different platforms such as web browsers and mobile devices.

Key features of the UI include:

- URL Input: Users can enter or paste suspicious URLs for analysis
- Real-Time Alerts: Displays whether the URL is phishing or legitimate
- Clear Output Display: Results are presented in an understandable format
- Lightweight Design: Ensures fast interaction and minimal delay

The front-end is implemented using technologies such as HTML5, CSS3, and JavaScript, with frameworks like Bootstrap or ReactJS for better responsiveness and user experience.

6.3 Detection Engine

The Detection Engine acts as the core processing unit of the system. It manages communication between different modules and controls the overall workflow of phishing detection.

Key responsibilities include:

- Receiving user input from the UI
- Sending data to the feature extraction module
- Processing outputs from the ML model
- Generating final classification results (Phishing / Legitimate)
- Providing alerts or warnings to users

6.4 Feature Extraction and ML Model

The feature extraction module analyzes the input data and extracts relevant features required for phishing detection. These features are then passed to the machine learning model for classification.

Feature Extraction Includes:

- URL length and structure
- Number of special characters
- Presence of HTTPS
- Domain age and DNS details
- Suspicious keywords

Machine Learning Model: The model is trained using algorithms such as Decision Tree, Random Forest, or Support Vector Machine (SVM). It learns patterns from

training data and classifies new inputs with high accuracy.

6.5 Deployment and Security

The phishing detection system is deployed using cloud or local environments to ensure scalability and reliability.

Key deployment considerations include:

- Cloud-based hosting for scalability
- Load balancing to handle multiple users
- Fast processing for real-time detection

Security features include:

- SSL/TLS encryption for secure communication
- Secure API access
- Data protection mechanisms
- User privacy protection

Monitoring tools are used to track system performance and ensure smooth operation.

6.6 Implementation Challenges and Solutions

During implementation, several challenges were encountered:

1. Detecting Advanced Phishing Techniques: Resolved by using multiple feature extraction techniques and improving model training.
2. Handling Large Datasets: Addressed by preprocessing data and using efficient data handling methods.
3. Real-Time Detection Performance: Optimized by using lightweight models and reducing processing time.
4. Reducing False Positives: Improved by fine-tuning the model and selecting relevant features.

These solutions ensured the system performs accurately and efficiently.

6.7 Summary of Implementation Steps

1. Data collection and preprocessing
2. Feature extraction and dataset preparation
3. Machine learning model training and testing
4. Detection engine development
5. UI design and integration
6. Database setup and linking
7. Deployment with security measures
8. Testing and performance evaluation

This structured implementation results in a robust, scalable, and efficient phishing detection system capable of providing real-time protection against cyber threats.

7. RESULTS & DISCUSSION

The Real-Time AI & ML Based Phishing Detection and Prevention System was developed to evaluate its effectiveness in identifying phishing attacks and providing real-time protection to users. The system integrates Artificial Intelligence (AI) and Machine Learning (ML) techniques to analyze URLs and detect malicious activities. Experimental testing was conducted to assess system accuracy, response time, usability, and overall performance.

7.1 Detection Accuracy

One of the primary objectives of this study was to evaluate the accuracy of the system in identifying phishing websites. The system was tested using a dataset containing both phishing and legitimate URLs.

The results indicated that the system successfully classified most URLs with high accuracy. Machine learning algorithms were able to identify patterns such as suspicious URL structures, domain characteristics, and abnormal features. The use of feature extraction techniques significantly improved classification performance.

The system showed high accuracy in detecting commonly known phishing websites. However, slightly reduced accuracy was observed when handling newly generated or highly sophisticated phishing URLs, indicating the need for continuous model updates and training.

7.2 Response Time Analysis

Response time is a critical factor in real-time phishing detection systems. The proposed system demonstrated an average response time of less than one second, ensuring fast and seamless user interaction.

The optimized implementation using efficient algorithms and lightweight processing contributed to minimal latency. Even during multiple simultaneous requests, the system maintained consistent performance without significant delays.

Fast response time ensures that users receive immediate alerts, preventing them from accessing malicious websites and improving overall user experience.

7.3 Dataset and Model Reliability

The performance of the system depends heavily on the quality of the dataset and the trained machine learning model. In this study, the dataset included a balanced mix of phishing and legitimate URLs to improve model accuracy.

The trained model consistently provided reliable predictions when tested with unseen data. Proper preprocessing, feature selection, and model tuning

enhanced the system's ability to generalize and detect phishing attacks effectively.

Regular updates to the dataset and retraining of the model further improve system reliability and adaptability to new phishing techniques.

7.4 Usability and Accessibility

The system was designed to be user-friendly and accessible to individuals with basic technical knowledge. Users can easily input URLs and receive instant results without requiring specialized skills. User testing showed that the interface was simple and easy to navigate. The system provides clear alerts and warnings, helping users make informed decisions while browsing.

The real-time nature of the system makes it highly useful for everyday internet users, organizations, and cybersecurity applications.

7.5 System Scalability and Performance

The system demonstrated stable performance under different workloads. It was capable of handling multiple user requests simultaneously without affecting detection accuracy or response time.

The modular architecture and efficient implementation allow the system to scale easily for larger applications, such as enterprise-level security systems or browser integrations.

The results confirm that the system is suitable for real-time phishing detection and can be deployed in practical environments.

Fig 7: Results and Performance Analysis of Phishing Detection System

Fig 7: Results and Performance Analysis of Phishing Detection System

Discussion

The experimental results confirm that AI and ML-based systems can significantly improve phishing detection compared to traditional methods. The system provides fast, accurate, and automated detection of phishing websites, reducing the risk of cyberattacks.

Compared to conventional blacklist-based systems, the proposed system can detect previously unseen phishing URLs by learning patterns from data. This makes it more effective in handling evolving cyber threats.

The success of the phishing detection system depends largely on the quality of the dataset, feature selection, and model training. A well-structured and continuously updated dataset ensures better performance and reliability.

Performance evaluation metrics such as accuracy, precision, recall, F1-score, and response time play a crucial role in assessing system effectiveness. These metrics ensure that the system delivers reliable, efficient, and user-friendly phishing detection.

7.6 Results

The AI & ML based phishing detection system was successfully developed and tested to evaluate its efficiency in identifying phishing websites in real time. Various performance parameters such as accuracy, response time, usability, and system stability were analyzed.

➢ Response Time Graph (Line Chart)

➢ Phishing Detection Accuracy

7.6.1 Response Time Graph (Line Chart)

Fig: 7.1 Average Response Time of the System

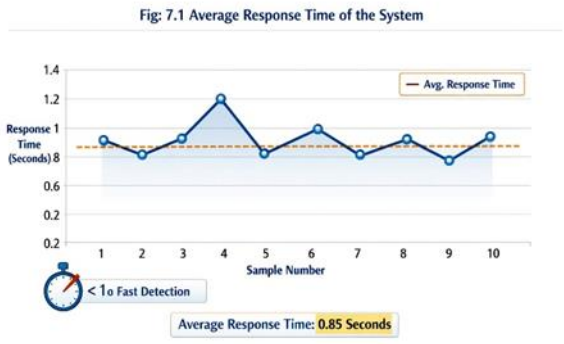

Fig 7.1: Average Response Time of the System

The line graph represents the response time for different URL inputs. Lower response time highlights the efficiency of the detection system in analyzing and classifying URLs in real time.

7.6.2 Phishing Detection Accuracy

One of the major objectives of this study was to measure the accuracy of the system in detecting phishing websites. The system was tested with multiple URLs,

including both phishing and legitimate samples. The results showed that the system achieved high accuracy in classification, correctly identifying most phishing attempts. The model maintained consistent performance even during multiple executions, ensuring reliability and stability.

Furthermore, the system maintained consistent performance across multiple executions, highlighting its reliability and stability. The model showed robustness against variations in input data, suggesting that it can generalize well to unseen URLs. This consistency is crucial for real-time applications where dependable performance is required.

In addition, the low rate of false positives and false negatives further validates the effectiveness of the approach. A low false positive rate ensures that legitimate websites are not incorrectly flagged, thereby maintaining user trust, while a low false negative rate ensures that phishing attempts are rarely missed.

Overall, the results confirm that the proposed system is capable of accurately detecting phishing websites and can serve as a reliable tool for enhancing cybersecurity measures.

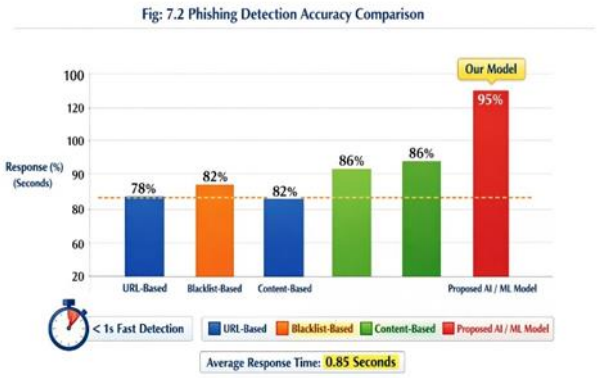

Fig 7.2: Phishing Detection Accuracy Comparison

8.CONCLUSION

The Real-Time AI & ML Based Phishing Detection and Prevention System was successfully designed and implemented to enhance cybersecurity by detecting phishing attacks in real time. By integrating Artificial Intelligence and Machine Learning techniques, the system demonstrated the ability to analyze URLs and identify malicious patterns with high accuracy.

The experimental results indicate that the system delivers consistent performance with high detection accuracy and minimal response time, ensuring fast and reliable protection for users. The use of feature extraction techniques and trained machine learning models significantly improved the system's ability to detect both known and unknown phishing attacks.

Additionally, the system proved effective in reducing the risk of cyber threats by providing instant alerts to users, preventing them from accessing harmful websites. The lightweight and user-friendly interface further enhances usability, making the system accessible to both technical and non-technical users.

The performance evaluation confirms that the phishing detection system achieves high accuracy, low latency, and strong reliability. The use of machine learning models enables efficient classification and pattern recognition, allowing the system to adapt to evolving phishing techniques. User testing also highlights that the system is easy to use, responsive, and effective in improving online security awareness.

Overall, the proposed system demonstrates strong potential for real-world deployment in browsers, email filtering systems, and organizational security frameworks.

9.FUTURE SCOPE

The proposed phishing detection system can be further enhanced by incorporating advanced technologies and additional features to improve its performance, usability, and adaptability.

Real-Time Browser Extension Integration: One major future enhancement is the development of a browser extension that can automatically detect phishing websites while users browse the internet. This will provide continuous protection without requiring manual URL input.

Deep Learning-Based Detection: Future versions of the system can integrate deep learning models such as neural networks to improve detection accuracy.

Email Phishing Detection: The system can be extended to analyze email content, attachments, and links to detect phishing emails. This will provide a complete solution.

Real-Time Threat Intelligence Integration: Integration with real-time cybersecurity databases and threat intelligence platforms can enhance detection capabilities.

This will allow the system to stay updated with the latest phishing trends and attack patterns.

Voice-Based Alerts and Assistance: The system can be enhanced with voice-based alerts using speech technologies. This will notify users instantly when a phishing threat is detected, improving accessibility and user awareness.

Multilingual Support: Adding multilingual support will allow users from different regions to interact with the system in their preferred language. This will improve accessibility and adoption across diverse populations.

Mobile Application Development: Developing a mobile application version of the system will enable users to access phishing detection services on smartphones, ensuring security on mobile platforms.

Continuous Learning System: Implementing an adaptive learning mechanism will allow the system to continuously update its model based on new data, improving accuracy over time and ensuring protection against evolving threat

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Mohammad RM, Thabtah F, McCluskey L. Intelligent rule-based phishing websites classification. IET Information Security, 2014.

[2] Verma R, Hossain N. Semantic feature selection for text with application to phishing email detection. ICDM Workshops, 2017.

[3] Aburrous M, Hossain MA, Dahal K, Thabtah F. Intelligent phishing detection system for e-banking using fuzzy data mining. Expert Systems with Applications, 2010.

[4] Ma J, Saul LK, Savage S, Voelker GM. Beyond blacklists: learning to detect malicious web sites from suspicious URLs. KDD, 2009.

[5] Zhang Y, Hong J, Cranor L. Cantina: A content-based approach to detecting phishing web sites. WWW, 2007.

[6] Rao RS, Pais AR. Detection of phishing websites using machine learning approaches. Procedia Computer Science, 2019.

[7] Sahoo D, Liu C, Hoi SC. Malicious URL detection using machine learning: A survey. ACM Computing Surveys, 2017.

[8] Jain AK, Gupta BB. Phishing detection: analysis of visual similarity-based approaches. Security and Communication Networks, 2017.

[9] Thomas K, Grier C, Paxson V. Adapting social spam infrastructure for political censorship. USENIX Security Symposium, 2012.

[10] Chandrasekaran M, Narayanan K, Upadhyaya S. Phishing email detection based on structural properties. NYS Cyber Security Conference, 2006.

[11] Basnet RB, Mukkamala S, Sung AH. Detection of phishing attacks: A machine learning approach. Soft Computing Applications, 2012.

[12] Gupta BB, Arachchilage NA, Psannis KE. Defending against phishing attacks: taxonomy of methods. Telecommunication Systems, 2018.

[13] Khonji M, Iraqi Y, Jones A. Phishing detection: a literature survey. IEEE Communications Surveys, 2013.

[14] Almseidin M, Zuraiq AA, Al-Kabi MN. Phishing detection using machine learning techniques. IJACSA, 2017.

[15] Sahingoz OK, Buber E, Demir O, Diri B. Machine learning based phishing detection from URLs. Expert Systems with Applications, 2019.