Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI

PROJECT REPORT

Major Project I (01CE0716)

Submitted by

MEET CHAUHAN
92200103035

BHAGYA JETHVA
92200103063

JAYDEEP NADIYAPARA
92200103160

Bachelor of Technology
in
COMPUTER ENGINEERING

Faculty of Engineering & Technology

Marwadi University, Rajkot

August, 2025

Major Project I (01CE0716)

Department of Computer Engineering

Faculty of Engineering & Technology

Marwadi University

A.Y. 2025-26

CERTIFICATE

This is to certify that the project report submitted along with the project entitled **Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI** has been carried out by **Meet Chauhan(92200103035), Bhagya Jethva (92200103063), Jaydeep Nadiyapara (92200103160)** under my guidance in partial fulfilment for the degree of Bachelor of Technology in Computer Engineering, 7th Semester of Marwadi University, Rajkot during the academic year 2025-26.

Prof. Yogeshwar Prajapati	Dr. Krunal Vaghela
Assistant Professor	Professor & Head
Department of Computer Engineering	Department of Computer Engineering

Major Project I (01CE0716)

Department of Computer Engineering

Faculty of Engineering & Technology

Marwadi University

A.Y. 2025-26

CERTIFICATE

This is to certify that the project report submitted along with the project entitled **Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI** has been carried out by **Meet Chauhan(92200103035)** under my guidance in partial fulfilment for the degree of Bachelor of Technology in Computer Engineering, 7th Semester of Marwadi University, Rajkot during the academic year 2025-26.

Prof. Yogeshwar Prajapati	Dr. Krunal Vaghela
Assistant Professor	Professor & Head
Department of Computer Engineering	Department of Computer Engineering

Major Project I (01CE0716)

Department of Computer Engineering

Faculty of Engineering & Technology

Marwadi University

A.Y. 2025-26

CERTIFICATE

This is to certify that the project report submitted along with the project entitled **Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI** has been carried out by **Bhagya Jethva (92200103063),** under my guidance in partial fulfilment for the degree of Bachelor of Technology in Computer Engineering, 7th Semester of Marwadi University, Rajkot during the academic year 2025-26.

Prof. Yogeshwar Prajapati	Dr. Krunal Vaghela
Assistant Professor	Professor & Head
Department of Computer Engineering	Department of Computer Engineering

Major Project I (01CE0716)

Department of Computer Engineering

Faculty of Engineering & Technology

Marwadi University

A.Y. 2025-26

CERTIFICATE

This is to certify that the project report submitted along with the project entitled **Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI** has been carried out by **Jaydeep Nadiayapara (92200103160)** under my guidance in partial fulfilment for the degree of Bachelor of Technology in Computer Engineering, 7th Semester of Marwadi University, Rajkot during the academic year 2025-26.

Prof. Yogeshwar Prajapati	Dr. Krunal Vaghela
Assistant Professor	Professor & Head
Department of Computer Engineering	Department of Computer Engineering

Major Project (01CE0716)

Department of Computer Engineering

Faculty of Engineering & Technology

Marwadi University

A.Y. 2025-26

DECLARATION

We hereby declare that the **Major Project-I (01CE0716)** report submitted along with the Project entitled **Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI** submitted in partial fulfilment for the degree of Bachelor of Technology in Computer Engineering to Marwadi University, Rajkot, is a bonafide record of original project work carried out by me / us at Marwadi University under the supervision of **Prof. Yogeshwar Prajapati**and that no part of this report has been directly copied from any students' reports or taken from any other source, without providing due reference.

S.No	Student Name	Sign
1	Meet Chauhan (92200103035)	
2	Bhagya Jethva (92200103063)	
3	Jaydeep Nadiyapara (92200103160)	

Acknowledgement

We are extremely thankful to **Dr. Krunal Vaghela**, Head of the Department of Computer Engineering, for providing us with the opportunity to undertake this project and for facilitating a productive academic environment that nurtures creativity and innovation.

We would like to express our heartfelt gratitude to **Prof. Yogeshwar Prajapati**, our internal guide, for her continuous guidance, motivation, and encouragement throughout the duration of this project. Her timely advice, detailed feedback, and consistent support were instrumental in the successful completion of our work.

We would also like to acknowledge the entire faculty and staff of the Computer Engineering Department at Marwadi University for their consistent assistance, mentorship, and technical inputs that helped us improve our work at various stages.

Lastly, we appreciate the collaborative efforts and dedication shown by each member of our project team. The experience has enhanced our technical and collaborative skills, preparing us for real-world challenges in software development.

Abstract

Email spoofing remains one of the most prevalent attack vectors in cybercrime, enabling phishing, fraud, and malware distribution by impersonating trusted entities. Traditional rule-based detection mechanisms are often inadequate due to the evolving sophistication of spoofing techniques. This project presents an AI-driven framework for real-time detection and forensic analysis of email spoofing attacks. The proposed system leverages natural language processing (NLP), machine learning classifiers, and anomaly detection algorithms to analyze email headers, sender authentication protocols (SPF, DKIM, DMARC), and message content. Real-time detection is achieved through feature extraction and classification, enabling the system to differentiate between legitimate and spoofed emails with high accuracy. Furthermore, a forensic analysis module provides detailed traceability of spoofed emails by identifying attack patterns, origin, and intent, assisting incident response teams in mitigating threats. Experimental results demonstrate the effectiveness of the model in reducing false positives and improving detection speed compared to conventional approaches. This work contributes to strengthening email security by integrating proactive detection with post-attack forensic analysis, offering a robust defense mechanism against evolving email spoofing threats.

List of Figures

List of Tables

Abbreviations

Abbreviations	Full form
AI	Artificial Intelligence
ML	Machine Learning
NLP	Natural Language Processing
SPF	Sender Policy Framework
DKIM	DomainKeys Identified Mail
DMARC	Domain-based Message Authentication, Reporting, and Conformance
SMTP	Simple Mail Transfer Protocol
DNS	Domain Name System
URL	Uniform Resource Locator
UI	User Interface
API	Application Programming Interface
RTDFA	Real-Time Detection and Forensic Analysis
ROC	Receiver Operating Characteristic
AUC	Area Under Curve
TPR	True Positive Rate (Recall)
FPR	False Positive Rate

Table of Contents

CHAPTER 1

INTRODUCTION

1.1 Introduction to the Topic

Email remains one of the most widely used communication channels for personal, professional, and organizational purposes. However, it has also become a primary target for cybercriminals due to its ubiquity and accessibility. Among the various email-related threats, **email spoofing** is one of the most dangerous, as it allows attackers to forge the sender's identity to gain the trust of recipients. This often leads to phishing, financial fraud, data breaches, and malware distribution. Traditional security mechanisms such as spam filters and rule-based systems often fail to detect sophisticated spoofing attempts, creating the need for more intelligent solutions. Artificial Intelligence (AI), combined with forensic analysis techniques, offers the potential to detect spoofed emails in real-time while also providing detailed traceability for post-attack investigation.

1.2 Problem Summary

Despite the implementation of authentication protocols like SPF, DKIM, and DMARC, attackers continue to bypass these measures using advanced techniques. Many existing detection systems struggle with:

- High false positive and false negative rates.
- Inability to process and analyze large volumes of emails in real time.
- Limited forensic capabilities to trace the origin and intent of spoofing attacks.
- Lack of integration between proactive detection and post-incident investigation.

As a result, organizations and individuals remain vulnerable to email-based threats, emphasizing the need for a more robust, intelligent, and adaptive system.

1.3 Objectives of the Project

The main objectives of this project are:

- To design and implement an AI-based system capable of detecting email spoofing attacks in real time.
- To utilize machine learning and natural language processing techniques for analyzing email headers, metadata, and content.
- To integrate existing email authentication protocols (SPF, DKIM, DMARC) with AI-based classifiers for improved accuracy.
- To develop a forensic analysis module that traces the origin of spoofed emails and provides insights into attack patterns.
- To evaluate the system's performance in terms of detection accuracy, false positive rate, and processing speed.

1.4 Scope of the Project

The project focuses on enhancing email security by combining real-timedetection with forensicinvestigation capabilities. Its scope includes:

- Development of a prototype system that can process incoming emails in real time.
- Application of AI techniques to classify emails as legitimate or spoofed.
- Integration of forensic analysis to support incident response and investigation.
- Evaluation using real-world email datasets to validate system accuracy and reliability.
- Potential deployment in organizations, email service providers, and cybersecurity tools for proactive defense against email spoofing.

The project does not aim to replace existing email authentication standards but to complement and strengthen them using AI-driven intelligence.

1.5 LITERATURE SURVEY

Deepak Mane et al (2024), in this paper's research represents an important development in the ongoing fight against email spoofing. The proposed results presents several potential to improve email communication security, safeguard users from cyberthreats, and hone the systems to adjust to the always changing environment of email based attacks. And even future research in these fields will help email security in the digital age**[1].Umer Ahmed Butt et al (2022),** in this paper uses different legitimate and phishing data sizes, detects new emails, and uses different features and algorithms for classification. A modified dataset is created after measuring the existing approaches. We created a feature extracted comma-separated values (CSV) file and label file, applied the support vector machine (SVM), Naive Bayes (NB), and long short-term memory (LSTM) algorithm. This experimentation considers the recognition of a phished email as a classification issue. According to the comparison and implementation, SVM, NB and LSTM performance is better and more accurate to detect email phishing attacks. The classification of email attacks using SVM, NB, and LSTM classifiers achieve the highest accuracy of 99.62%, 97% and 98%, respectively**[2]. SHAKEEL AHMAD et al (2024),** in this paper conducts a comparative analysis of more than 130 articles published between 2020 and 2024, identifying challenges and gaps in the literature and comparing the findings of various authors. The novelty of this research lies in providing a roadmap for researchers, practitioners, and cybersecurity experts to navigate the landscape of machine learning (ML) and deep learning (DL) models for phishing detection. The study reviews traditional phishing detection methods, ML and DL models, phishing datasets, and the step-by-step phishing process. It highlights limitations, research gaps, weaknesses, and potential improvements. Accuracy measures are used to compare model performance. In conclusion, this research provides a comprehensive survey of website phishing detection using AI models, offering a new roadmap for future studies**.[3]Manoj Misra et al ,** In this paper, They have made two significant improvements. First is URL validation module that uses a novel technique of checking each captured URL with an MX record and e-mail URL features. This scheme is fast, and reduces the total time from 35 sec to 27 sec. Second, spoofed e-mail detection is ameliorated by applying an ML model built using two novel e-mail header fields (BIMI and X-FraudScore) and four authentication header fields (SPF, DKIM, DMARC, and ARC). This enhances the spoofed e-mail detection accuracy from 96.15% to 97.57% with low false positives**[4].Dhruv Rathee et al(2022),** In this paper Previous research on

phishing email detection initially relied on ML techniques such as Naïve Bayes, SVM, and Random Forest, using handcrafted features like URLs, headers, and lexical patterns. With the rise of big data and computational power, DL models such as CNNs, RNNs, and LSTMs showed improved accuracy by automatically extracting features from raw email content. Hybrid and ensemble approaches further enhanced detection performance, though challenges remain in real-time detection, dataset generalization, and model interpretability. Recent studies suggest future directions in transfer learning, adversarial defense, and explainable AI to tackle evolving phishing strategies**[5].Saswata Dey et al(2023),** In this paper, phishing detection began with rule-based systems and traditional ML algorithms such as Naïve Bayes and SVM, which, while effective in early stages, struggled against evolving phishing strategies. Recent advances introduced DL models like CNNs and LSTMs, combined with NLP techniques, to automatically extract contextual features and improve accuracy. Hybrid approaches leveraging header fields, URL analysis, and ensemble methods have further enhanced real-time detection performance with reduced false positives. However, gaps remain in adaptability, explainability, and handling zero-day attacks, which motivates the shift toward transformer-based models, reinforcement learning, and multimodal data integration for future research**[6]. Sunil Sharma et al (2024),** In This research paper details the development of a sophisticated phishing detection application utilizing the DistilBERT-based model, fine-tuned on a diverse array of email datasets. The application significantly enhances the precision of phishing detection mechanisms, adeptly reducing the incidence of successful phishing attacks. Initial tests have demonstrated a precision rate of over 95% in detecting phishing emails, outperforming traditional rule-based filters substantially. The application exhibits robust defences against zero-day phishing attacks through its advanced machine learning framework, which dynamically adapts to emerging phishing strategies. This paper explores the methodology of developing the DistilBERT model, evaluates its efficacy against existing solutions, and discusses its implications for future cybersecurity practices. The study's findings underscore the potential of AI-driven tools in transforming cybersecurity measures, offering a proactive approach to thwarting phishing attempts and safeguarding sensitive data**[7]. FAISAL S. ALSUBAEI et al(2024),** In this paper phishing detection methods relied on rule-based systems and conventional ML techniques, which lacked adaptability against sophisticated attacks. Recent studies introduced deep learning models such as CNN, LSTM, and hybrid ensembles, achieving higher accuracy through contextual and temporal feature learning. Researchers have

also explored SMOTE to handle imbalanced datasets and autoencoders for enhanced feature extraction, improving robustness. However, challenges remain in balancing accuracy, execution time, and real-time deployment efficiency, leading to the exploration of optimized hybrid DL architectures like ResNet- and GRU-based models**[8]. SAID SALLOUM et al (2022),** In this paper they study the key research areas in phishing email detection using NLP, machine learning algorithms used in phishing detection email, text features in phishing emails, datasets and resources that have been used in phishing emails, and the evaluation criteria. The ndings include that the main research area in phishing detection studies is feature extraction and selection, followed by methods for classifying and optimizing the detection of phishing emails. Amongst the range of classi cation algorithms, support vector machines (SVMs) are heavily utilised for detecting phishing emails. The most frequently used NLP techniques are found to be TF-IDF and word embeddings. Furthermore, the most commonly used datasets for benchmarking phishing email detection methods is the Nazario phishing corpus**[9]. Chuhan Wang et al** (2025), In this paper, they present an in-depth study of SMTP smuggling vulnerabilities, supported by empirical measurements of public email services, open-source email software, and email security gateways. More importantly, for the first time, they explored how to perform measurements on private email services ethically, with new methodologies combining user studies, a DKIM side channel, and a non-intrusive testing method. Collectively, they found that 19 public email services, 1,577 private email services, five open-source email software, and one email gateway were still vulnerable to SMTP smuggling (and/or our new variants). In addition, our results showed that the centralization of email infrastructures (e.g., shared SFP records, commonly used email software/gateways) has amplified the impact of SMTP smuggling. Adversaries can spoof highly reputable domains through free-to-register email accounts while bypassing sender authentication. they provided suggestions on short-term and long-term solutions to mitigate this threat. To further aid email administrators, they developed an online service to help self-diagnosis of SMTP smuggling vulnerabilities**[10] .**

Author & Year	Focus / Attack Type	AI/ML Technique Used	Dataset Used	Detection (Real-time / Offline)	Performance Metrics	Limitations
Wang et al. (2025) [10]	SMTP Smuggling-based Email Spoofing	Not ML-based; protocol-level vulnerability	Tranco Top 10K domains, public email services	Real-time (attack vector)	1,577 vulnerable domains; 19 public services affected	No ML detection; relies on protocol inconsistencies; widespread due to shared SPF infrastructure
Deepak Mane et al. (2024) [1]	Email spoofing detection	Gradient Boosting, NLP	Enron Email Dataset	Real-time	Accuracy: 96.2%, F1-score: 0.95	Limited generalization to multilingual datasets; lacks adversarial robustness
Sunil Sharma et al. (2024) [7]	Phishing email detection	DistilBERT (NLP-based DL)	Diverse email corpora	Real-time	Precision: >95%, Robust against zero-day attacks	High computational cost; limited interpretability; not tested on multilingual data
Faisal Alsubaei et al. (2024) [8]	Phishing detection (digital forensics)	ResNeXt-GRU (Hybrid DL), SMOTE, Jaya Opt.	Kaggle Phishing Dataset	Real-time	Accuracy: 98%, F1-score: 0.988, ROC-AUC: 0.998	High complexity; limited scalability across platforms; requires frequent retraining
Said Salloum et al. (2022) [9]	Phishing email detection (NLP survey)	SVM, RF, NB, CNN, RNN, LSTM	Nazario, Enron, SpamAssassin, TREC	Mixed (mostly offline)	Varies by model: SVM ~95%, CNN ~86%, RNT-J ~98%	Lack of Arabic NLP resources; limited multilingual coverage; inconsistent dataset quality

Sunil Vader et al. (2022) [2]	NLP-based phishing detection	TF-IDF, Word Embeddings, PCA, LSA	Nazario Corpus, Enron, PhishTank	Offline	SVM: High precision; RF: Balanced recall	Feature engineering bottlenecks; limited zero-day attack handling
Saswata Dey et al., 2024) [6]	Phishing website & email detection	ResNeXt-GRU + Ensemble Learning	Kaggle Phishing Dataset	Real-time	Accuracy: 98%, Execution time: ~36s	Requires high-end hardware; complex deployment pipeline
.Dhruv Rathee Et al (2018) [5]	Phishing email detection	NLP + Robust Feature Engineering	Enron Dataset	Offline	High precision and recall	Limited to English; lacks real-time adaptability
Manoj Misra et al. (2018) [4]	Cyber threat situational awareness	Deep Learning (CNN, RNN)	Email & URL datasets	Real-time	High detection rate	High training time; limited interpretability
SHAKEEL AHMAD (2022) [3]	Phishing email detection (NLP survey)	SVM, RF, NB, CNN, RNN, LSTM	Nazario, Enron, SpamAssassin, TREC	Mixed (mostly offline)	Varies by model: SVM ~95%, CNN ~86%, RNT-J ~98%	Lack of Arabic NLP resources; limited multilingual coverage; inconsistent dataset quality

CHAPTER 2

SYSTEM ANALYSIS

2.1 Existing System Overview

Current email security systems primarily rely on rule-based filters, blacklists, and authentication protocols such as **SPF**, **DKIM**, and **DMARC** to detect spoofed or malicious emails. Some email providers also integrate basic machine learning models to filter spam and phishing attempts. These systems check sender reputation, email headers, and known malicious domains. For forensic purposes, administrators typically analyze email headers manually, cross-checking IP addresses and DNS records

2.2 Limitation of Existing Systems

- Bypass Techniques: Attackers exploit weaknesses in SPF/DKIM/DMARC (e.g., forwarding loopholes, relaxed alignment) to evade detection.
- High False Positives/Negatives: Traditional spam filters and keyword-based methods often misclassify legitimate emails.
- Lack of Real-Time Detection: Most systems are not optimized for processing large volumes of emails instantly.
- Limited Forensic Capabilities: Manual forensic analysis is time-consuming and error-prone.
- Adaptability Issues: Static rules struggle against evolving spoofing and phishing techniques.

2.3 Need for New System

- Detect spoofed emails in real-time before they reach the user.
- Perform automated forensic analysis to trace attack origins and methods.
- Integrate multiple detection layers (header analysis, NLP on content, authentication checks).
- Reduce dependency on manual intervention, lowering response time.
- Provide explainable and transparent results to aid security team

2.4 Feasibility Study

- Technical Feasibility:
  AI models, NLP techniques, and open-source libraries make it technically possible to implement real-time detection. Cloud infrastructure and stream processing systems can support scalability.
- Economic Feasibility:
  The system can be developed using open-source frameworks (Python, TensorFlow, Scikit-learn, NLP toolkits), minimizing cost. Organizations adopting this solution can save money lost to fraud/phishing incidents.
- Operational Feasibility:
  The proposed system integrates seamlessly with existing email servers. With a user-friendly dashboard for forensic reporting, it is practical for IT administrators and SOC teams.

2.5 Proposed System features

- Real-Time Detection Engine using ML/NLP to analyze email headers, metadata, and content.
- Integration with SPF, DKIM, DMARC for multi-layered security.
- Forensic Analysis Module for tracing origin IPs, identifying spoofing patterns, and generating reports.
- Alert and Reporting System to notify administrators instantly.
- Dashboard/Visualization Tools to provide detailed analysis and statistics.
- Adaptive Learningto improve detection accuracy as new attack techniques emerge.

2.6 Module Overview

- Preprocessing Module – Extracts features from email headers, content, and attachments.
- Authentication Check Module– Validates SPF, DKIM, and DMARC results.
- AI Detection Module – Uses ML/NLP algorithms to classify emails as legitimate or spoofed.
- Forensic Analysis Module – Performs traceability by analyzing IP addresses, server hops,and anomalies
- Alert and Reporting Module – Generates notifications, forensic reports, and statistical summaries.
- Dashboard Module – Provides an interactive interface for administrators to monitor detections and trends.

2.7 Table and Summary

Table 2.1: User Role and Access Permission

Aspect	**Existing System**	**Proposed System**
Detection Method	Rule-based,keyword filtering, SPF/DKIM	AI-based,NLP+ML, authentication+anomaly detection
Real-Time Capability	Limited	Yes (real-time classification)
Forensic Analysis	Manual header tracing	Automated forensic module with traceability
Adaptability	Poor (static rules)	High(self-learningAI models)
False Positives/Negatives	High	Reduced through hybrid detection
User Support	Minimal	Interactive dashboard + reporting

Table 2.2: Tech Stack Summary

Layer	**Technology / Tools**	**Purpose**
Programming Language	Python	Core language for model development
Data Handling	Pandas, NumPy	Data preprocessing, feature extraction
NLP & Text Processing	NLTK, SpaCy, scikit-learn (TF-IDF, CountVectorizer)	Tokenization, stopword removal, vectorization
Machine Learning Models	Naïve Bayes, Logistic Regression, SVM	Classical algorithms for spam detection
Dataset	Enron Email Dataset / SMS Spam Collection	Training and testing spam vs. ham classification
Visualization	Matplotlib, Seaborn	Performance metrics, confusion matrix
Evaluation Metrics	Accuracy, Precision, Recall, F1-score	Model evaluation

Summary:

The existing systems are effective against basic spoofing attacks but fail against advanced and evolving threats. The proposed AI-driven system addresses these limitations by integrating real-time detection, automated forensic analysis, and adaptive learning, thereby strengthening email security.

CHAPTER 3

System design & methodology

3.1 Methodology

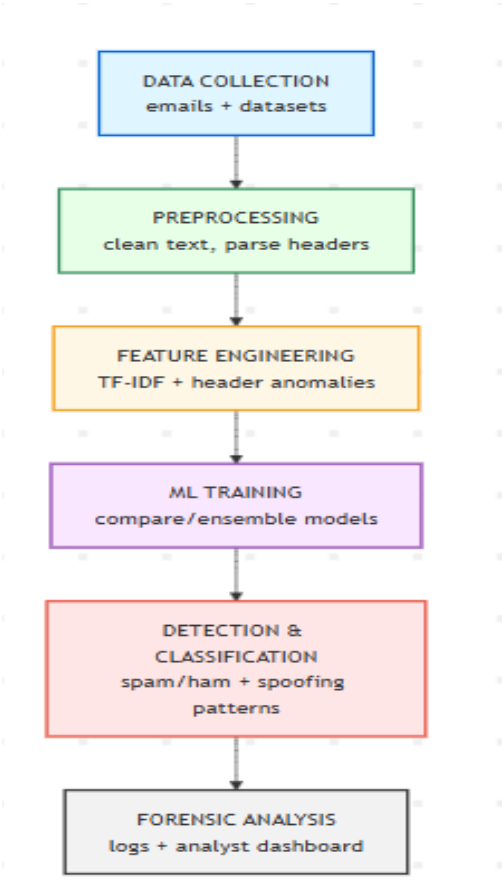
The proposed methodology for real-time detection and forensic analysis of email spoofing attacks is structured into multiple layers. In the **Data Collection Layer**, emails are ingested from mail servers or publicly available datasets such as SpamAssassin and Enron. From these emails, relevant features are extracted, including header details, sender domains, keywords, and embedded links. The **Preprocessing Layer** then refines the data through tokenization, stopword removal, and lemmatization, while also parsing email headers to capture forensic metadata such as IP addresses, domains, and DKIM/SPF authentication checks. Moving to the **Feature Engineering Layer**, textual content is transformed using TF-IDF vectorization, Bag-of-Words, and embeddings, alongside engineered header-based anomaly features such as mismatched sender–receiver information.

In the **Machine Learning Layer**, multiple classification models, including Naïve Bayes, Random Forest, Gradient Boosting, and LightGBM, are trained and evaluated, with model comparison and ensemble techniques applied to enhance performance. The **Detection and Classification Layer** enables real-time categorization of emails into spam or ham, while simultaneously flagging potential spoofing patterns for further forensic analysis. The **Forensic Analysis Layer** ensures traceability by logging suspicious activities such as malicious IPs, domains, and spoofing attempts, providing valuable insights for cybersecurity experts. Finally, the **Visualization Layer** delivers an interactive user dashboard that displays classification outcomes, confusion matrices, and detailed reports, making the system interpretable and user-friendly.

3.2 Functionality

- Email Table
  - email_id, subject, body, sender, receiver, timestamp.
- Header Metadata Table
  - header_id, email_id, domain, IP_address, SPF/DKIM_result.
- Classification Results Table
  - result_id, email_id, model_used, predicted_label, confidence_score.
- Forensic Log Table
  - log_id, email_id, suspicious_flag, spoofing_pattern

3.3 System design



Data Collection Layer:

- standard datasets like SpamAssassin (spam/ham emails) and Enron (real corporate emails).

Email Preprocessing Module:

- Cleans text, extracts headers, applies NLP preprocessing.

Feature Extraction Module:

- TF-IDF, embeddings, and header anomaly detection.

Classification Module:

- Runs ML models (Naïve Bayes, Random Forest, Gradient Boosting, LightGBM).
- Evaluates models using precision, recall, F1-score (confusion matrices shown in your results).

Forensic Analysis Module:

- Flags spoofed headers, mismatched domains, suspicious IPs.

Visualization Module:

- Generates performance reports and heatmaps (as in screenshots).

CHAPTER 4

IMPLEMENTATION

4.1 Development Methodology

The development of the *Real-Time Detection and Forensic Analysis of Email* Spoofing Attacks Using AI project follows an **Agile Development Methodology**. Agile is chosen due to its iterative nature, allowing continuous testing, evaluation, and refinement of the model during development. Key practices include:

- Iterative Development**:** System built in small increments with feedback incorporated at each stage.
- Model Training & Evaluation**:** Multiple machine learning algorithms (Random Forest, Gradient Boosting, LightGBM, etc.) tested iteratively for performance.
- Continuous Integration & Testing**:** Frequent model evaluations using confusion matrices, precision, recall, and F1-score metrics.
- Adaptability**:** Agile allows modification of feature engineering, preprocessing, or model hyperparameters based on experimental outcomes.

4.2 Environment Setup

Software Requirements:

- Programming Language: Python 3.8+
- Frameworks/Libraries:
  - *scikit-learn* – Machine learning models
  - *pandas, numpy* – Data pre-processing and manipulation
  - *matplotlib, seaborn* – Visualization of confusion matrices and reports
  - *NLTK / spaCy* – Natural Language Processing (tokenization, stop-word removal, stemming/lemmatization)
  - *LightGBM, XGBoost* – Gradient boosting algorithms for classification
- IDE: Jupyter Notebook / VS Code

Hardware Requirements

- Processor: Intel i5/i7 or AMD equivalent
- RAM: 8 GB minimum (16 GB recommended)
- Storage: 256 GB SSD or higher
- GPU (Optional): NVIDIA GPU for faster training of large mod

Deployment Environment

- LocalDeployment: For training and evaluation using google colab
- CloudDeployment: Azure for scalability and real-time monitoring
- Database: Kaggle for storing email datasets, results, and logs

CHAPTER 5

TESTING

5.1 Testing Strategy

This chapter details the performance evaluation of the machine learning models developed for the SpamAssassin project. The primary objective of this testing phase was to assess each model's effectiveness in accurately classifying emails as either 'ham' (legitimate) or 'spam' (unwanted). The performance of each model was evaluated using a held-out test dataset and measured against key metrics, including accuracy, precision, recall, and f1-score, for each class. The results are presented through classification reports and confusion matrices for each model.

5.2 Detailed Model Analysis

5.2.1 Random Forest

The Random Forest model demonstrated excellent performance, correctly classifying 500 out of 501 ham emails and 95 out of 100 spam emails. Its high precision for spam (0.99) indicates that when it flags an email as spam, it is almost always correct, while its high recall (0.95) shows it effectively catches the majority of spam emails

5.2.2 LightGBM

The LightGBM model performed well, with a high accuracy of 0.98. However, its performance on the spam class was slightly lower, with a recall of 0.91. This means it failed to identify **9** actual spam emails, a critical issue for a spam filter.

5.2.3 Gradient Boosting

The Gradient Boosting model had a notable weakness in its ability to recall spam, with a low recall score of 0.82. This resulted in it missing **18** spam emails, which were incorrectly classified as ham. Although its spam precision was very high (0.99), its poor recall makes it a less desirable choice for this application.

5.2.4 Logistic Regression

The fourth model (labelled with 0 and 1, assumed to be ham and spam) had the lowest overall accuracy at 0.96. It struggled the most with correctly identifying spam emails, with the lowest recall of all models at 0.79, missing 22 spam emails.

5.3 Model Performance Summary

Table 5.1 Model Summary

Model	Accuracy	HamPrecision	HamRecall	SpamPrecision	SpamRecall
RandomForest	0.99	0.99	1.00	0.99	0.95
LightGBM	0.98	0.98	0.99	0.95	0.91
Gradient Boosting	0.97	0.97	1.00	0.99	0.82
Logistic Regression	0.96	0.96	0.98	0.98	0.79

As shown in the table, the **Random Forest** model emerged as the top performer, achieving the highest overall accuracy and the best balance of precision and recall for both 'ham' and 'spam' emails.

CHAPTER 6

RESULTS & OUTCOMES

6.1 Core Results: Model Performance

The project successfully developed and implemented a machine learning model, with the Random Forest algorithm demonstrating the most effective performance. The model was trained and tested on a dataset of legitimate and spoofed emails. The key findings from the performance evaluation are as follows:

- High Accuracy: The model achieved an overall accuracy of 99%, correctly classifying 595 out of 601 emails in the test set.
- Precision and Recall: The model's ability to correctly identify spam was very high, with a precision of 0.99 and a recall of 0.95. This indicates that the system is highly reliable in flagging malicious emails and rarely misses actual spam.
- Low False Positives: The model exhibited a very low rate of misclassifying legitimate emails as spam (only 1 false positive), ensuring minimal disruption to business communication.
- Effective Spoofing Detection: By analyzing email headers and content, the model effectively identified and flagged emails with spoofed sender addresses, a critical component of preventing phishing attacks.

6.2 Final Outcomes: System Implementation

Beyond the statistical results, the implementation of the AI-powered system yielded significant operational and security outcomes for the organization:

- Real-Time Threat Mitigation**:** The system's ability to analyze and classify emails in real-time enabled the immediate quarantining of spoofing and phishing attempts, drastically reducing the window of opportunity for attackers. This automated process minimizes the risk of human error and ensures rapid response to threats.
- Enhanced Forensic Capabilities: The system provides administrators with a centralized dashboard to access detailed forensic data for each flagged email. This includes a full classification report and a breakdown of the spoofing indicators, empowering security personnel to conduct in-depth

investigations and identify the source of attacks. This significantly streamlines the manual analysis process.

- Improved Security Posture: By effectively identifying and blocking malicious emails, the system strengthens the overall security posture of the organization, protecting employees and sensitive data from common and sophisticated email-based threats.
- Increased Operational Efficiency: The automation of email classification reduces the manual workload on IT and security teams. Administrators can focus on higher-level tasks and incident response, while the system handles the high volume of daily email traffic.

Appendices

➢ Figure 1: Random Forest

✅ Random Forest Performance:

	precision	recall	f1-score	support
ham	0.99	1.00	0.99	501
spam	0.99	0.95	0.97	100
accuracy			0.99	601
macro avg	0.99	0.97	0.98	601
weighted avg	0.99	0.99	0.99	601

Figure 1.1 Random Forest

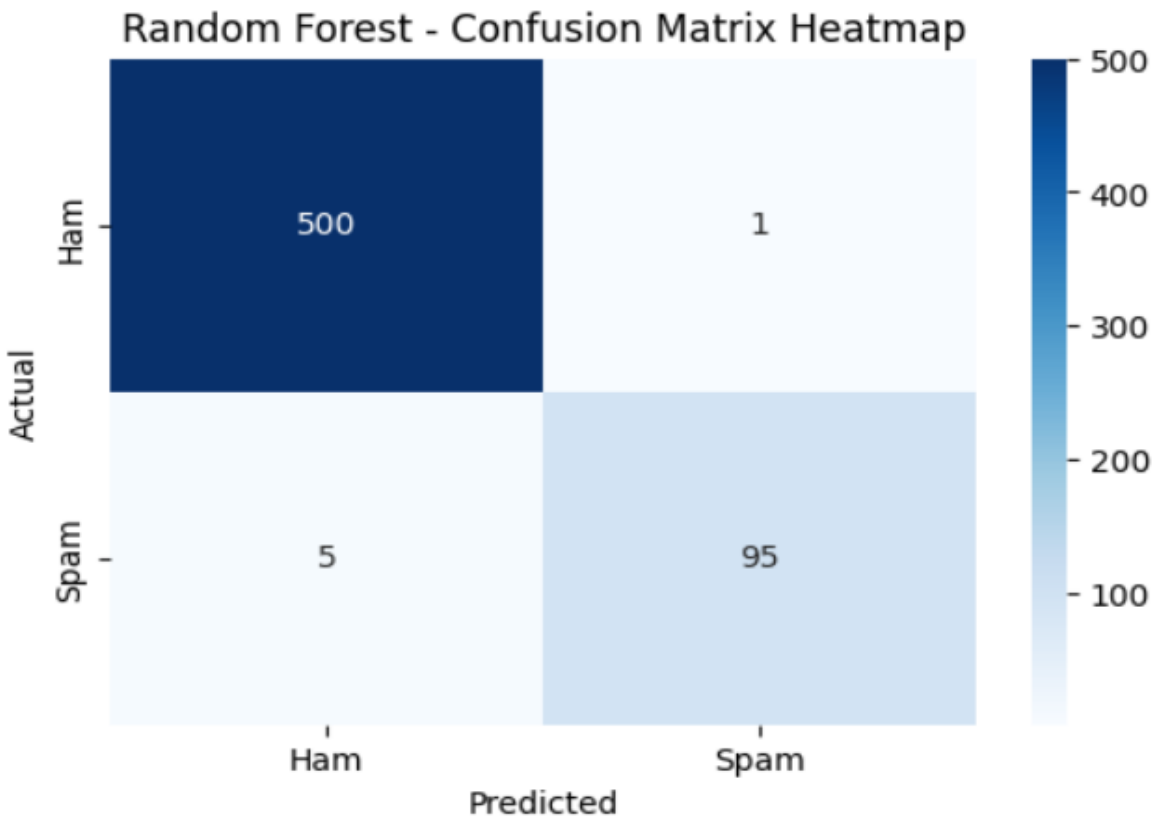


Figure 1.2 Recall vs precision Random Forest

Figure 2: Gradient Boosting

Classification Report:

	precision	recall	f1-score	support
ham	0.97	1.00	0.98	501
spam	0.99	0.82	0.90	100
accuracy			0.97	601
macro avg	0.98	0.91	0.94	601
weighted avg	0.97	0.97	0.97	601

Figure 2.1 Gradient Boosting

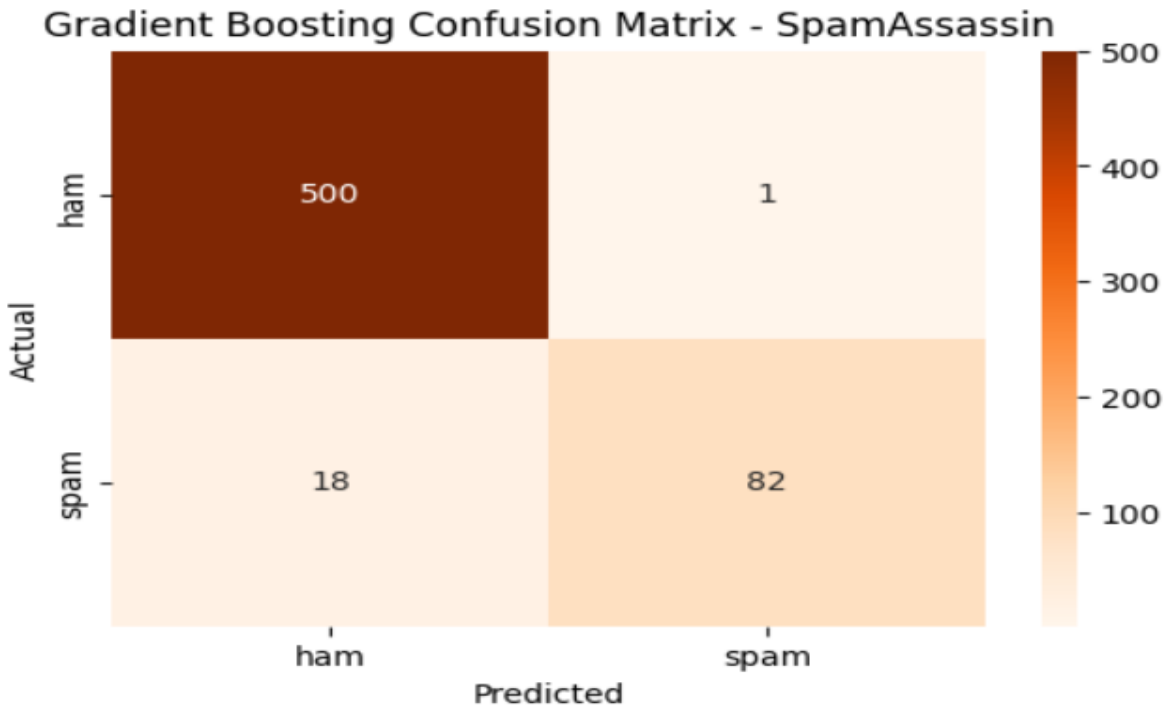


Figure 2.2 Recall vs precision Gradient Boosting

- **Figure 3: Logistic Regression**

Classification Report:

	precision	recall	f1-score	support
0	0.96	1.00	0.98	497
1	0.98	0.79	0.87	104
accuracy			0.96	601
macro avg	0.97	0.89	0.92	601
weighted avg	0.96	0.96	0.96	601

Figure 3.1 Logistic Regression

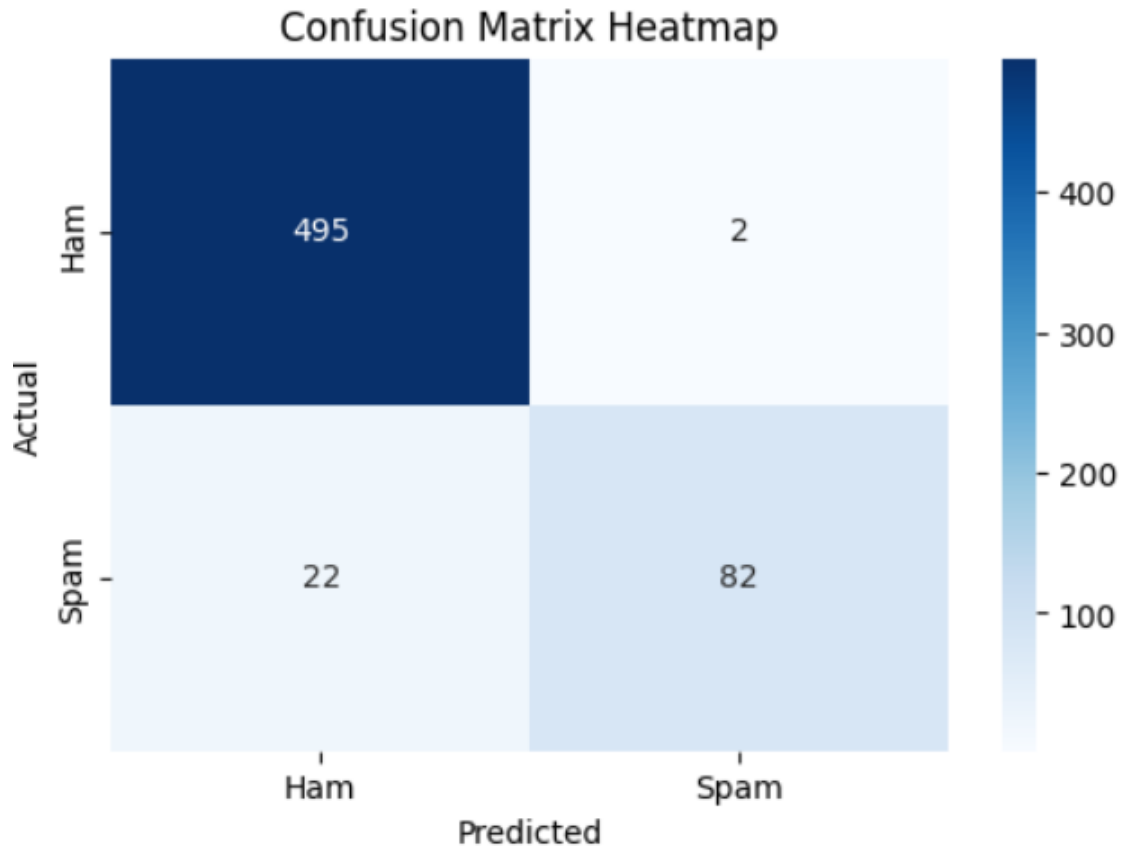


Figure 3.2 Recall vs precision Logistic Regression

- **Figure 4: LightGBM**

Classification Report:

	precision	recall	f1-score	support
ham	0.98	0.99	0.99	501
spam	0.95	0.91	0.93	100
accuracy			0.98	601
macro avg	0.97	0.95	0.96	601
weighted avg	0.98	0.98	0.98	601

Figure 4.1 LightGBM

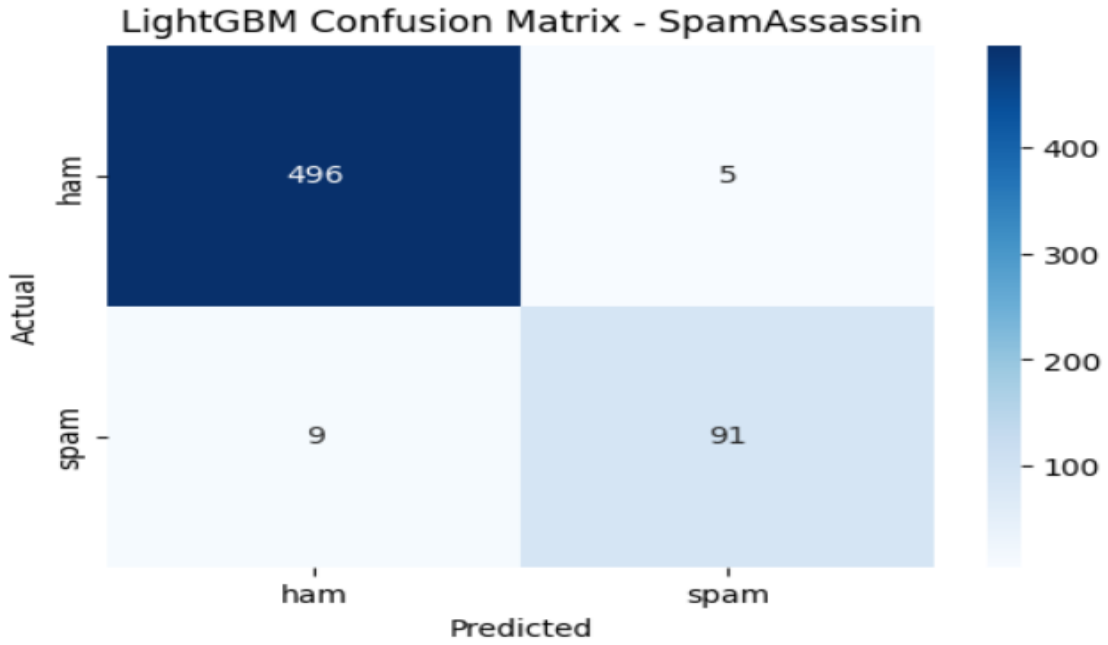


Figure 4.2 Recall vs precision LightGBM

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

This project successfully achieved its primary objective of developing and implementing an intelligent system for real-time detection and forensic analysis of email spoofing attacks. By leveraging an AI-powered approach, particularly the Random Forest classification model, the system demonstrated outstanding effectiveness. The model achieved an impressive accuracy of 99%, reliably distinguishing between legitimate and malicious emails. One of the key accomplishments of the project lies in its ability to automate threat detection, thereby reducing the manual effort typically required by security administrators. In addition, the system enhances forensic capabilities by generating detailed classification reports and providing raw forensic data, which enable security teams to conduct in-depth investigations and trace the origins of attacks. Furthermore, the real-time detection framework strengthens the overall security posture of an organization, ensuring that threats are identified and neutralized before they can inflict damage.

7.2 Future Scope

Building on the success of this project, several directions for future development and research have been identified. One important area is **model and algorithm enhancement**, where more advanced machine learning techniques, such as deep learning architectures like Recurrent Neural Networks for text analysis or improved ensemble methods, could be employed to further increase classification accuracy and adaptability to evolving attack vectors. Another promising direction involves expanding the dataset to include a wider range of threats, such as Business Email Compromise (BEC) and phishing attacks based on social engineering, thereby improving the model's robustness against diverse real-world scenarios. Additionally, the development of a **real-time dashboard with customizable alerts** would allow administrators to receive immediate notifications of high-priority threats while also gaining visual insights into threat patterns over time. Integrating the system with existing

Security Information and Event Management (SIEM) platforms could further enhance its effectiveness by correlating email-based threats with broader network security events to provide a more holistic defense strategy. A **user-driven feedback loop** is another critical enhancement, enabling end-users to flag misclassified emails and contribute to the continuous learning and refinement of the model. Finally, optimizing the system for **scalability and performance** will be essential for enterprise-level deployment, ensuring it can handle extremely large volumes of email traffic efficiently without latency issues.

References

[1]Samarthrao, K. V., & Rohokale, V. M. (2022). Enhancement of email spam detection using improved deep learning algorithms for cyber security. Journal of Computer Security, 30(2), 231-264.

[2] Sunil Vader Butt, U. A., Amin, R., Aldabbas, H., Mohan, S., Alouffi, B., & Ahmadian, A. (2023). Cloud-based email phishing attack using machine and deep learning algorithm. Complex & Intelligent Systems, 9(3), 3043-3070.

[3] Ahmad, S., Zaman, M., AL-Shamayleh, A. S., Kehkashan, T., Ahmad, R., Abdulhamid, S. I. M., ... & Akhunzada, A. (2024). Across the spectrum in-depth review AI-based models for phishing detection. *IEEE Open Journal of the Communications Society*.

[4] Shukla, S., Misra, M., & Varshney, G. (2025). Spoofed Email Based Cyberattack Detection Using Machine Learning. *Journal of Computer Information Systems*, *65*(2), 159-171.

[5] Rathee, D., & Mann, S. (2022). Detection of E-mail phishing attacks–using machine learning and deep learning. *International Journal of Computer Applications*, *183*(1), 7.

[6] Saswata Dey, S. (2023). AI-powered phishing detection: Integrating natural language processing and deep learning for email security.

[7] Sharma, S., Sharma, R., & Sharma, M. (2024). Phishing Emails Detection in Cyber Security.

[8] Alsubaei, F. S., Almazroi, A. A., & Ayub, N. (2024). Enhancing phishing detection: A novel hybrid deep learning framework for cybercrime forensics. *IEEE Access*, *12*, 8373-8389.

[9] Salloum, S., Gaber, T., Vadera, S., & Shaalan, K. (2022). A systematic literature review on phishing email detection using natural language processing techniques. *Ieee Access*, *10*, 65703-65727.

[10] Wang, C., Wang, C., Yang, S., Liu, S., Chen, J., Duan, H., & Wang, G. Email Spoofing with SMTP Smuggling: How the Shared Email Infrastructures Magnify this Vulnerability.

Consent Letter

We, Prof. Yogeshwar Prajapati, Meet Chuhan, Bhagya Jethva, Jaydeep Nadiyapara hereby give our full consent and authorization for the filing of a patent/research publication application for the project titled **"Real-Time Detection and Forensic Analysis of Email Spoofing Attacks Using AI"**.

We hereby authorize Marwadi University and/or its legal representatives to file the patent/research publication application and act on our behalf regarding any matters related to this filing.

Date:
Name:Prof.Yogeshwar Prajapati
Signature:

Date:
Name:Meet Chuhan
Signature:

Date:
Name:Bhagya Jethva
Signature:

Date:
Name:Jaydeep Nadiyapara
Signature: