WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Deep Learning-Based Email Phishing Detection with NLP-Driven Features

Ms. Tahera Abid[1], Ammar Ahmed[2], Nimra Fatima[3], Mohd Ismail Khan[4]

[1]Assistant Professor, Department of IT, Nawab Shah Alam Khan College of Engineering and Technology, Affiliated to Osmania University, Hyderabad, India.

[2,3,4] Department of IT, Nawab Shah Alam Khan College of Engineering and Technology, Affiliated to Osmania University, Hyderabad, India.

A B S T R A C T

The Increasing volume of phishing emails has become a major cybersecurity concern as attackers use deceptive language and misleading content to steal personal data, passwords and financial information. Traditional email filtering systems and machine learning approaches struggle to detect modern phishing attempts due to evolving tactics and linguistic manipulation. This paper presents a deep learning-based approach that utilizes Natural Language Processing (NLP) techniques to analyze and classify phishing emails with improved accuracy. Models such as CNN, LSTM and BERT are trained on email datasets containing both legitimate and malicious samples. Text preprocessing, tokenization and feature embedding are used to enhance learning. Experimental results show that deep learning models outperform traditional classifiers by achieving higher precision and recall with minimal false positives. This framework offers a scalable and intelligent solution for securing email communication and reducing phishing risks.

Keywords: Phishing Detection, Deep Learning, NLP, Email Security, Cybersecurity, LSTM, BERT, CNN.

1. Introduction:

Email communication is an essential part of everyday digital life. Unfortunately, it has also become one of the main channels for phishing attacks where attackers send fake messages pretending to be from trusted organizations. These phishing emails aim to trick users into revealing private information downloading harmful files or visiting malicious websites **[1]**.

Conventional phishing detection systems depend mainly on rule-based filtering, blacklists and keyword detection. Although these techniques were effective in earlier times, they are now unable to detect modern phishing emails that use realistic language customized content and cleverly disguised links **[2]**. As phishing attacks continue to grow in complexity there is a strong need for intelligent systems that can adapt and understand the real intent behind the message.

Deep learning models combined with Natural Language Processing (NLP) methods have proven to be powerful tools for understanding text and identifying suspicious behavior in emails. They are capable of analyzing both the structure and meaning of text more effectively than traditional models.

This study proposes a deep learning-based phishing detection framework that uses NLP-driven techniques such as tokenization and feature extraction. It evaluates multiple models including CNN LSTM and BERT to achieve high accuracy precision recall and F1-score compared to older methods.

2. Existing System:

Existing phishing email detection techniques primarily depend on static rule-based filters, lexical analysis and traditional machine learning algorithms such as Naïve Bayes, SVM and Random Forest. These systems often rely on manually selected features like suspicious URLs, header data, meta-data or keyword presence.

However, existing systems face multiple limitations:

- **Limited Adaptability:** Static filters cannot detect new phishing patterns or changing email structures.
- **High False Positives:** Many legitimate emails are flagged incorrectly due to keyword mismatches.
- **Lack of Context Understanding:** Traditional methods cannot interpret sentence semantics or intent.
- **URL and Domain Dependence:** Attackers often bypass blacklist systems by using temporary or modified URLs.

- **Manual Feature Engineering:** Human-defined features reduce scalability and effectiveness.

Due to these drawbacks or limitations, there is a grow in need for an intelligent and automated approach that understands email content contextually and adapts to evolving phishing strategies.

3. Problem Statement:

Phishing emails have become one of the most common and dangerous methods used by cybercriminals to deceive users and steal sensitive information such as passwords, credit card details and personal data. Traditional email filtering systems depend mainly on manually created rules, blacklists or keyword- based detection, which fail to identify new or well-crafted phishing attacks. These methods cannot fully understand the meaning or context of email content, allowing many phishing messages to bypass security filters.

There is a growing need for an intelligent, automated system that can accurately analyze email text, understand its linguistic patterns and detect phishing attempts even when attackers use unfamiliar words or techniques. Therefore, this research aims to design a deep learning-based phishing detection framework that uses Natural Language Processing (NLP) to identify and classify phishing emails more effectively than traditional approaches.

4. Proposed System:

The proposed system introduces a deep learning-based phishing detection model enhanced with NLP-driven feature extraction. The framework eliminates the need for manual feature engineering and instead learns from raw email text.

Key Features of the Proposed System:

- **NLP-Based Preprocessing:** Includes tokenization, stop-word removal, stemming, lemmatization, and embedding.
- **Deep Learning Models Used:** CNN, LSTM, and BERT for contextual understanding.
- **Automated Feature Learning:** Extracts semantic and structural patterns directly from email content.
- **Improved Accuracy:** Minimizes false positives and enhances phishing email identification.
- **Scalable and Real-Time:** Can be integrated into email clients or enterprise email security gateways.

The proposed system focuses on understanding intent, tone and structure rather than relying solely on static indicators like URLs or keywords. This leads to higher accuracy and adaptability compared to conventional approaches.

5. System Architecture:

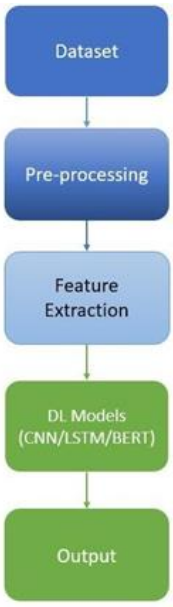

Fig. 1 – System Architecture

6. Methodology:

The methodology consists of the following steps:

Step 1: Dataset Collection

Emails with phishing and legitimate labels are gathered from trusted repositories such as Enron, PhishTank cand Nazario datasets.

Step 2: Data Preprocessing

- Lowercasing text
- Noise removal (symbols, HTML tags, special characters)
- Tokenization
- Stop-word removal
- Lemmatization/Stemming

Step 3: Feature Engineering

- TF-IDF vectorizer for weighted term representation
- Word2Vec embedding for contextual mapping
- BERT tokenizer and embedding for semantic depth

Step 4: Model Training

Three models were implemented and evaluated:

- **CNN:** Extracts local patterns and n-grams
- **LSTM:** Learns long-term dependencies
- **BERT:** Captures context-aware semantics across sentences

Step 5: Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-Score

6.1 Results:

Table 1: Performance Comparison of Deep Learning Models

Model	Accuracy	Precision	Recall	F1-Score
CNN	97.6 %	97.2 %	97.0 %	97.1 %
LSTM	98.3 %	98.0 %	98.1 %	98.05 %
BERT	99.1 %	98.8 %	99.0 %	98.9 %

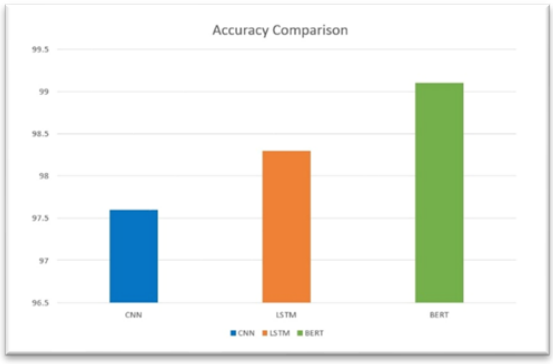

Fig. 2 – Accuracy Comparison Graph

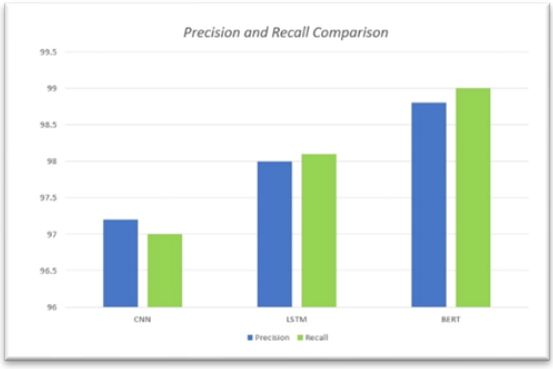

Fig. 3 – Precision and Recall Graph

7. Conclusion:

This paper presents a deep learning-based email phishing detection system using NLP-driven feature extraction. Traditional detection methods struggle to interpret linguistic patterns and evolving attack strategies. The proposed system utilizes CNN, LSTM, and BERT models to analyze email content contextually.

Experimental results demonstrate that deep learning models especially BERT achieve higher accuracy and reduced false positives compared to traditional machine learning techniques. The framework is scalable, adaptable, and suitable for integration into email filtering systems and enterprise security platforms.

8. Future Enhancements:

Future improvements may include:

- Integration of URL and attachment analysis.
- Real-time deployment in email service providers.
- Transformer-based hybrid architectures.
- Multilingual phishing detection.
- Browser extensions or cloud-based security APIs.

These enhancements will further improve detection robustness and adaptability to evolving phishing threats.

References

1. A. Sahu et al., “Deep Learning Approaches for Phishing Email Detection,” Journal of Cybersecurity, 2022.
2. N. Gupta and R. Sharma, “NLP-Based Email Threat Analysis,” IEEE Access, 2021.
3. S. Jain et al., “BERT for Cyber Threat Detection,” ACM Digital Library, 2023.
4. PhishTank Dataset – https://phishtank.org
5. Enron Email Dataset – Carnegie Mellon University