DEPHIDES: Deep Learning Based Phishing Detection System

Presented By

Nadiya Naaz(4511-23-862-059)

Under the Guidance of

Mrs.Ch.Sudha Rani

Assistance Professor

UNIVERSITY COLLEGE OF ENGINEERING AND TECHNOLOGY

MAHATMA GANDHI UNIVERSITY ,NALGONDA

NALGONDA 500 008, TELANGANA, INDIA.

Contents

Abstract

Phishing attacks trick users into revealing sensitive information like passwords or bank details by imitating trusted websites. To stop such threats, this project introduces DEPHIDES, a phishing detection system that uses deep learning to quickly and accurately identify fake websites based on their URLs. Unlike many other systems, DEPHIDES does not rely on third-party services and can detect even new (zero-day) phishing attacks. It uses five types of deep learning algorithms are ANN, CNN, RNN, Bi-RNN, and Attention-Based Network *. The system is fast, reliable, and language-independent, making it highly effective for cybersecurity.*

Introduction

Digital Revolution: The modern era is defined by the rapid growth of internet-connected devices, transforming how we shop, bank, and communicate online.

Rising Cyber Threats: Cybercriminals exploit this connectivity with sophisticated phishing attacks, targeting sensitive user information like passwords and bank details.

Innovative Defense: Harnessing the power of deep learning, advanced phishing detection system now offer robust solutions to safeguard users from evolving cyber threats.

Existing System

- **Traditional Rule-Based Systems:** Many phishing detection systems rely on predefined rules, such as blacklists and heuristics, which are often unable to adapt to new and evolving phishing techniques.

- **Machine Learning-Based Systems:** These systems improve detection by analyzing features like URL patterns, website content, and metadata but often struggle with scalability and the ability to handle large, complex datasets effectively.

Proposed System

•**Deep Learning-Powered Detection:** The proposed system leverages advanced deep learning algorithms, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Attention Networks, to significantly improve phishing detection accuracy and adapt to evolving cyber threats.

•**Scalable and High-Performance:** Built to handle large datasets (e.g., labeled URLs), this system can efficiently scale to meet the growing demands of cybersecurity, providing high detection accuracy to combat increasingly sophisticated attacks.

Advantages

1.Advanced Cybersecurity Techniques: The rise of deep learning and AI-driven algorithms has led to more accurate and adaptive phishing detection systems that can identify sophisticated cyber threats with high precision.

2.Scalability and Performance: The ability to process large datasets (e.g millions of URLs) allows systems to scale effectively, ensuring they can keep up with the increasing volume and complexity of cyber threats in digital age.

3.Continuous Improvement: The use of machine learning and deep learning models ensures that phishing detection systems can learn and adapt over time, improving their accuracy and responsiveness to new attack strategies.

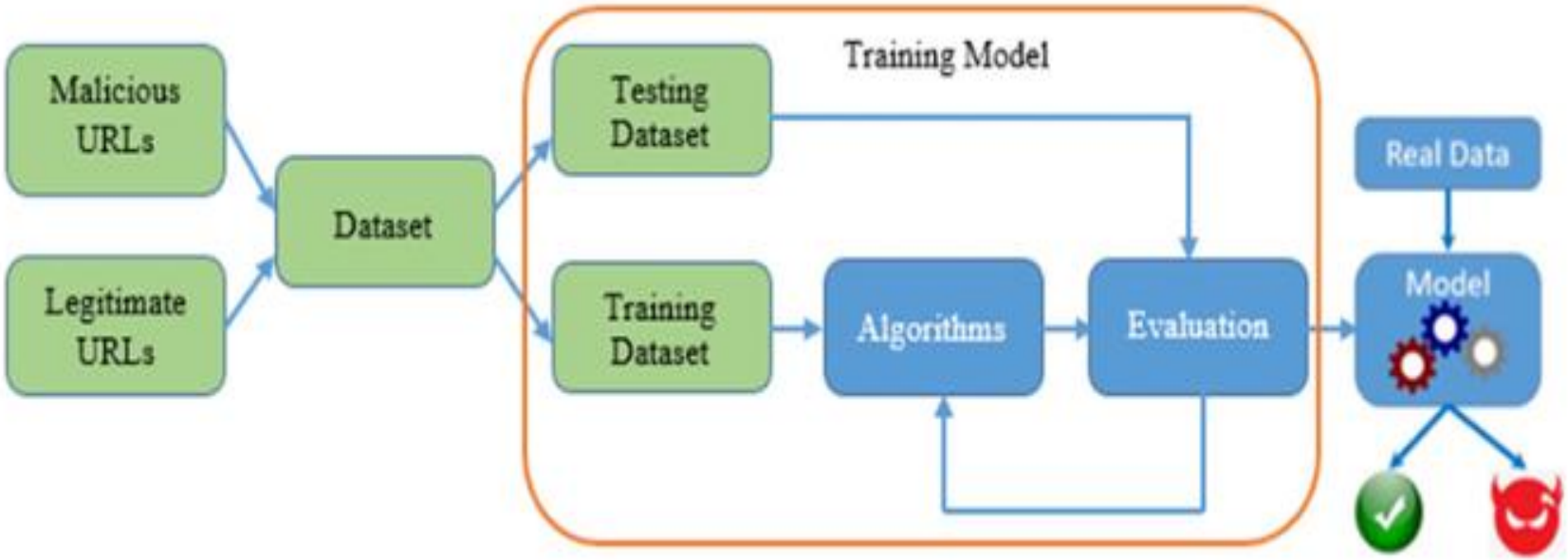
Limitation

1.Increased Cybersecurity Vulnerabilities: As digital devices become more interconnected, the attack surface for cybercriminals expands, making it easier for phishing and other cyberattacks to target users.

2.Ineffective Traditional Defenses: Existing rule-based systems and basic machine learning models often fail to detect new, more sophisticated phishing techniques, leaving users vulnerable to evolving threats.

3.Data Privacy Concerns: The vast amount of personal and financial data being shared online increases the risk of sensitive information being compromised, especially when detection systems fail to keep up with the scale and complexity of attacks.

System Architecture



Module

- **Data Collection**
- **Data Preprocessing & vectorization**
- **Training Models**
- **Evaluation of Model**

Data Collection

- Data collection is the process of gathering information from different sources.
- The data is collected from:
  - **GitHub** : Open-source datasets
  - Urldata set
- The collected data contains attributes are phish_id, url, phish_detail_url, submission_time, verified, verification_time, online target
- Collected data was exported to CSV format for further analysis.

Preprocessing & Vectorization Process

- ✓ Transform raw URLs into a uniform, numerical format that can be used as input for deep learning models.
- ✓ Tokenization of URLs
- ✓ Removing special characters
- ✓ Feature extraction (length, domain age, use of HTTPS, etc.)
- ✓ URL normalization (e.g., removing special characters)
- ✓ Converting text to lowercase Removing whitespace

Training Models

- Artificial Neural Network (ANN)
- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- Bidirectional Recurrent Neural Network (BRNN)
- Attention-Based Network (ATT)

1. Artificial Neural Network (ANN)

- An ANN is the simplest form of a neural network.
- It has an input layer, more hidden layers, and an output layer.
- Each layer consists of neurons that are fully connected to the next layer.

How It Works :

- Used as a baseline model with dense layers.
- Processes **vectorized URLs character embeddings** to learn patterns indicating phishing

2. Convolutional Neural Network (CNN)

- CNNs are designed to detect patterns in data by using convolutional filters that scan the input.
- Originally developed for image recognition, they are now widely used in text and sequence processing.

How It Works :

- Helps in **reducing computation and avoiding overfitting**.
- Applies filters (small sliding windows) across the sequence. These filters detect local patterns like login, secure, @, -, or repeated segments that are common in phishing URLs.

3. Recurrent Neural Network (RNN)

- RNNs are built to process sequential data — where the order matters (like characters in a URL or words in a sentence).
- They maintain a hidden state (memory) of previous inputs.

How It Works:

- Reads the URL character-by-character in order.
- At each character, updates a hidden memory state based on what it has seen so far.
- Remembers previous patterns to understand **context within a URL**.

4. Bidirectional Recurrent Neural Network (BRNN)

- A BRNN is an extension of RNN that reads the input in both directions — from start to end and end to start.

How It Works:

- One RNN reads the URL from left to right. Another reads it from right to left.
- Useful when phishing indicators appear in **reverse patterns**.

5. Attention-Based Network (ATT)

- An advanced architecture that learns to focus on important parts of the input.
- Inspired by how humans pay attention to key information while ignoring less relevant details.

How It Works:

- The model processes the URL but learns which characters or sections are more important.
- Helps the model focus on **critical parts of the URL** while ignoring irrelevant parts.

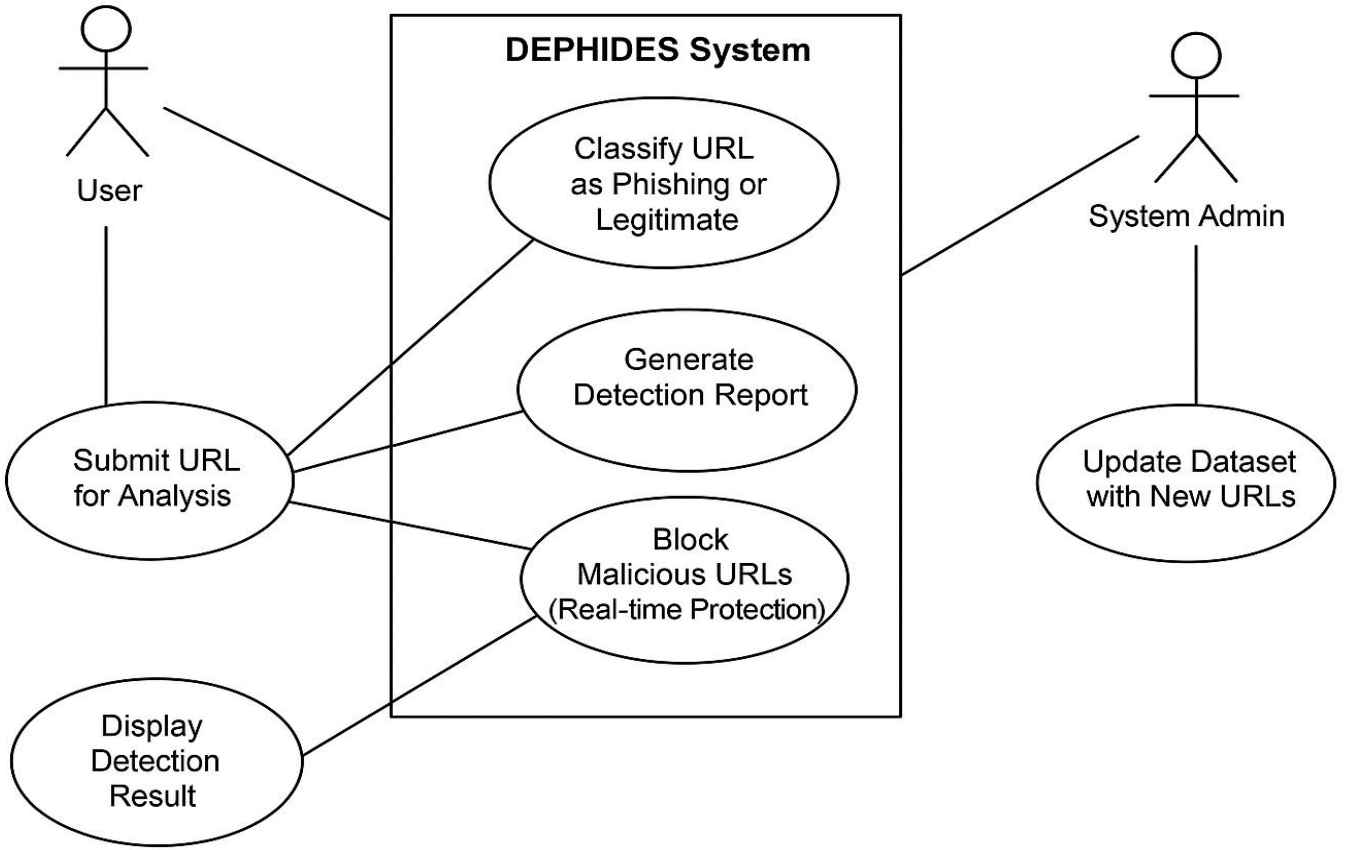
Evaluation Module

• The process of using different evaluation metrics to understand a model's performance.

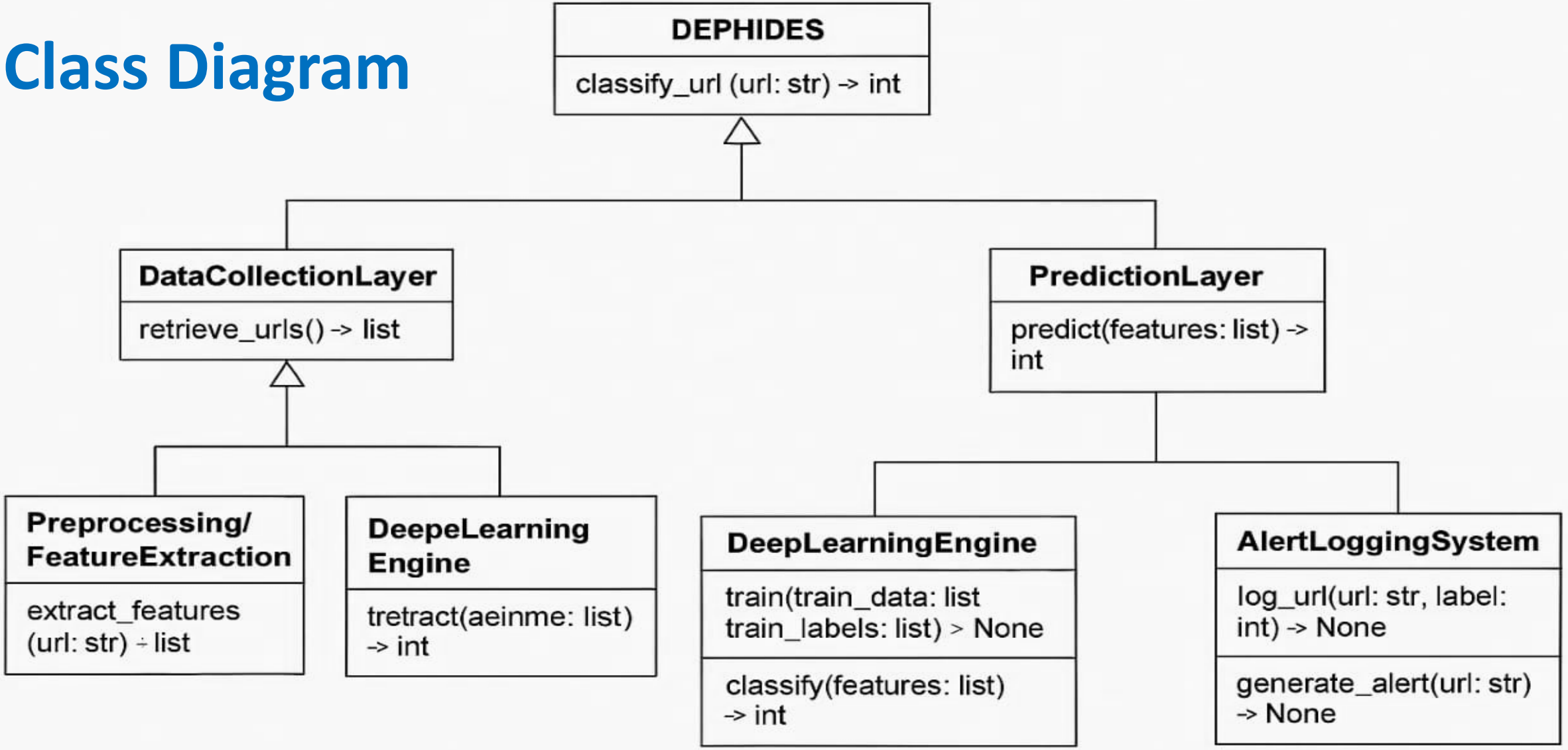
- **Accuracy** – How many total predictions were correct
- **Precision** – Out of predicted phishing URLs, how many were really phishing
- **Recall** – Out of all actual phishing URLs, how many were correctly detected
- **F1-Score** – Balance between precision and recall
- **Speed** – How fast it can process URLs in real time.

UML Diagrams

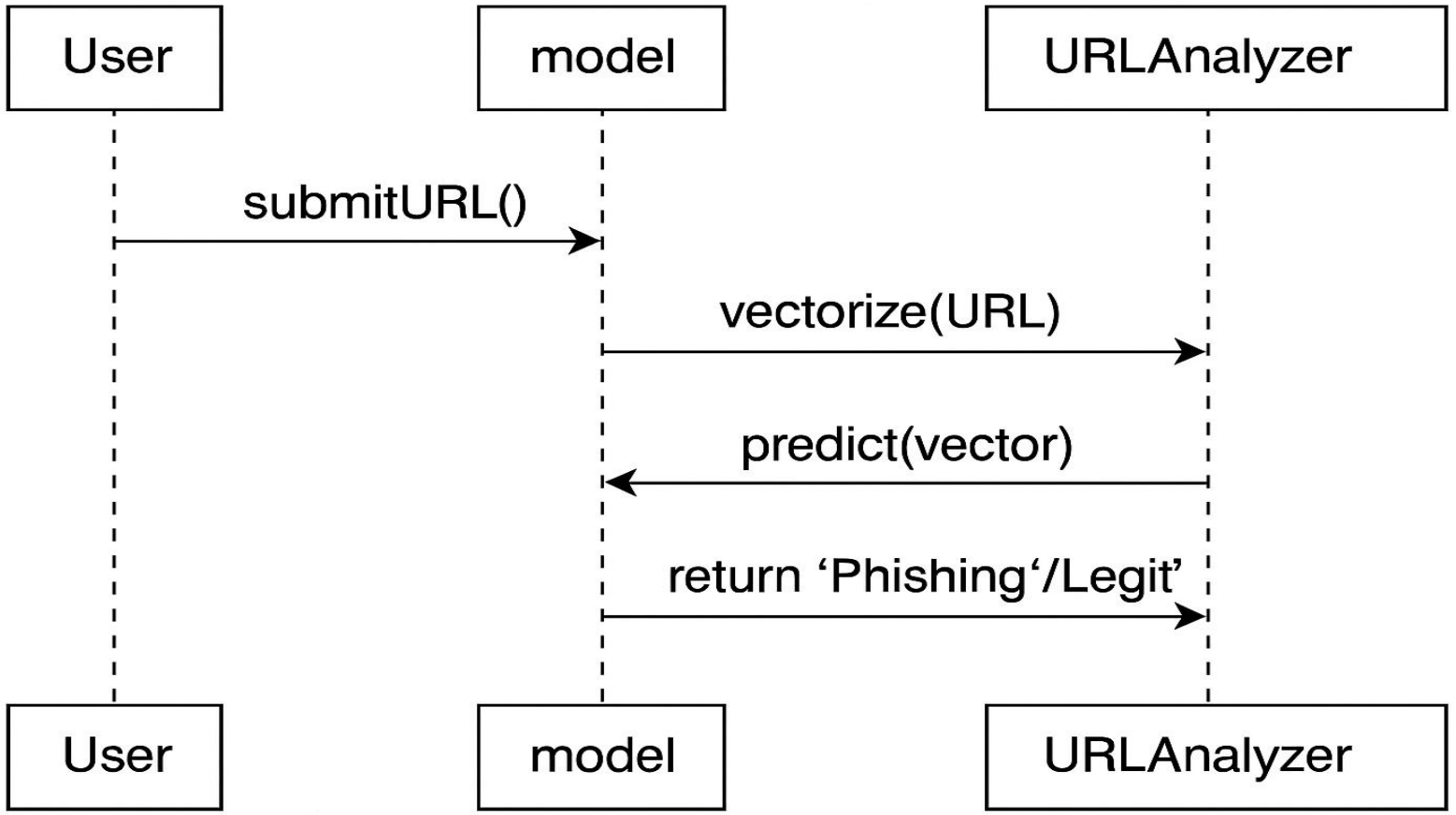
Use case Diagram



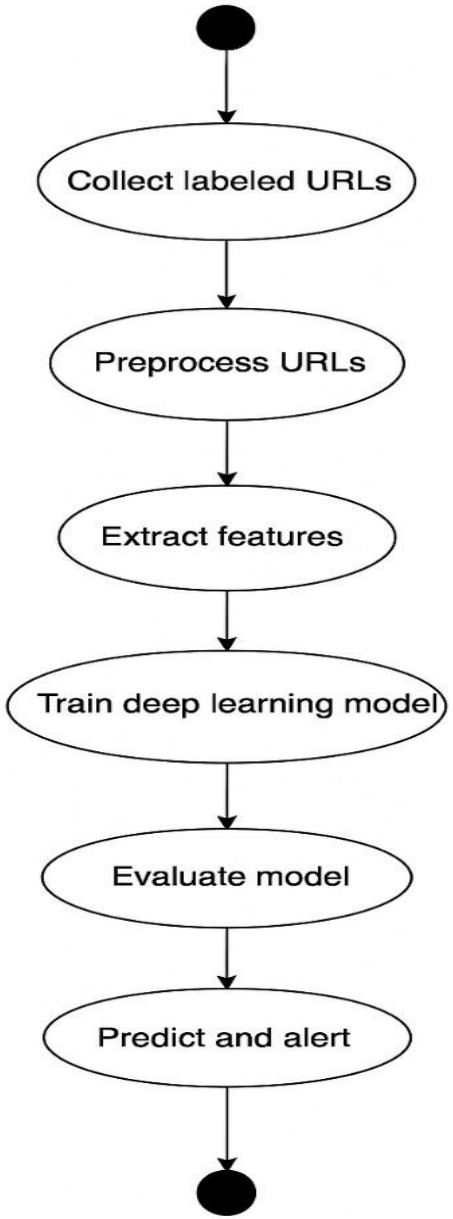
Class Diagram



Sequence Diagram



Activity Diagram



System Requirements

HARDWARE REQUIREMENTS :

- Processor – Intel i5 or higher
- RAM - 4Gb or higher
- Hard disk- 256Gb or higher

SOFTWARE REQUIREMENTS :

- Coding Platform- Jupyter Notebook
- Operating System-Windows 7 or higher
- Programming Language- Python

Literature Survey

Reference-1

Cloudflare, Inc. (2023)

Title: Cloudflare's Phishing Threats Report

•Top Phishing Techniques:
• Deceptive Links (35.6%)
• Identity Deception (↑ to 14.2%)

•Brand Impersonation:
• 1,000+ brands mimicked
• 51.7% targeted top 20 brands
• Most impersonated: Microsoft, Google, Salesforce, Notion

Recommendations:

• Implement AI-based detection beyond traditional filters
• Monitor new domains proactively

Reference-2

M. Alshehri, A. Abugabah, A. Algarni, and S. Almotairi,

Title:" Character-level word encoding deep learning model for combating cyber threats in phishing URL detection**"**

Introduction:

- Cyber threats cause data theft, system disruption, and digital compromise.
- **Phishing**: Mimics genuine websites/URLs to deceive users.

Methodology:

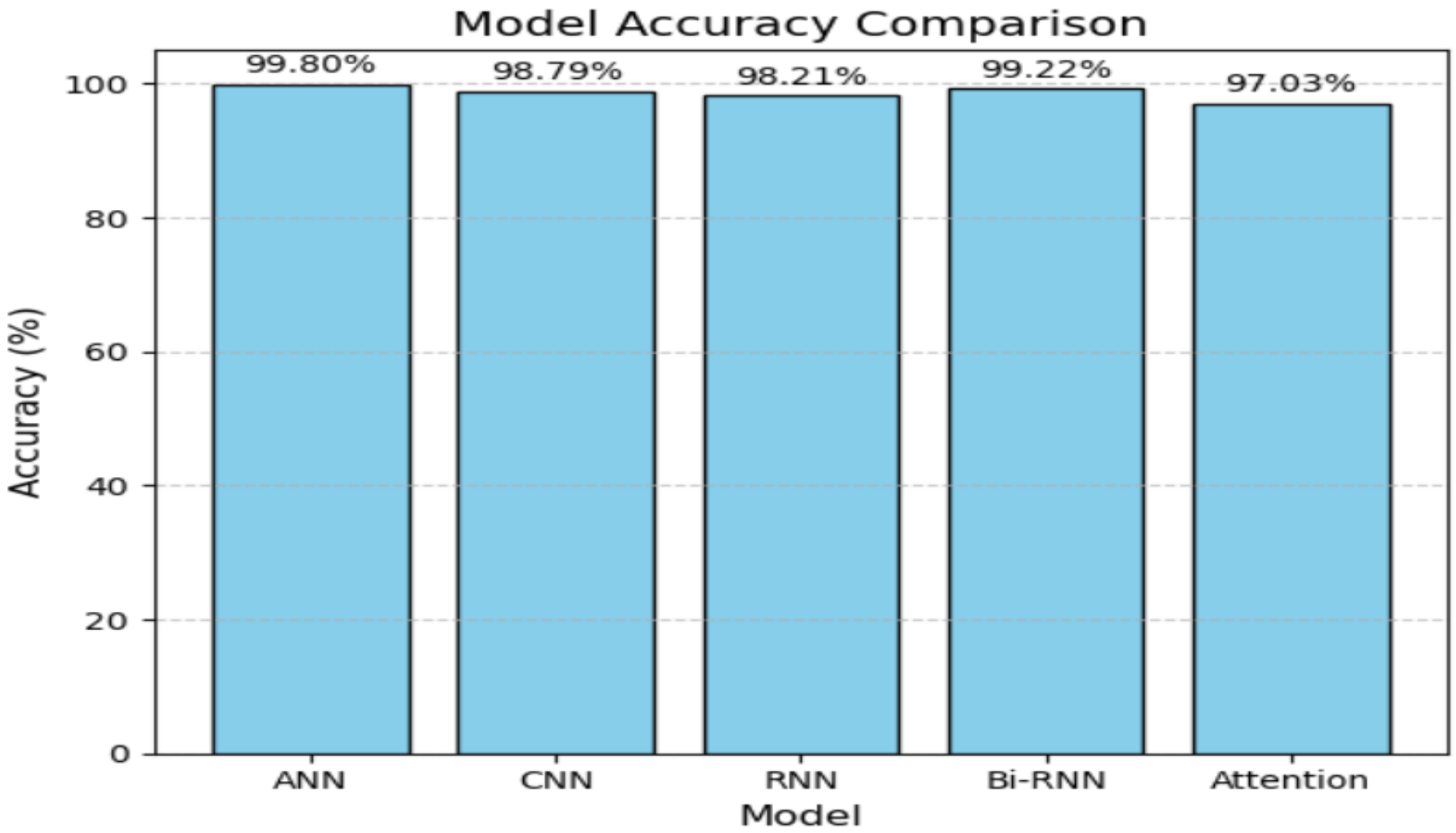
- Developed a **deep learning model** using character-level features.
- Trained on phishing and legitimate URLs.

Future Work:

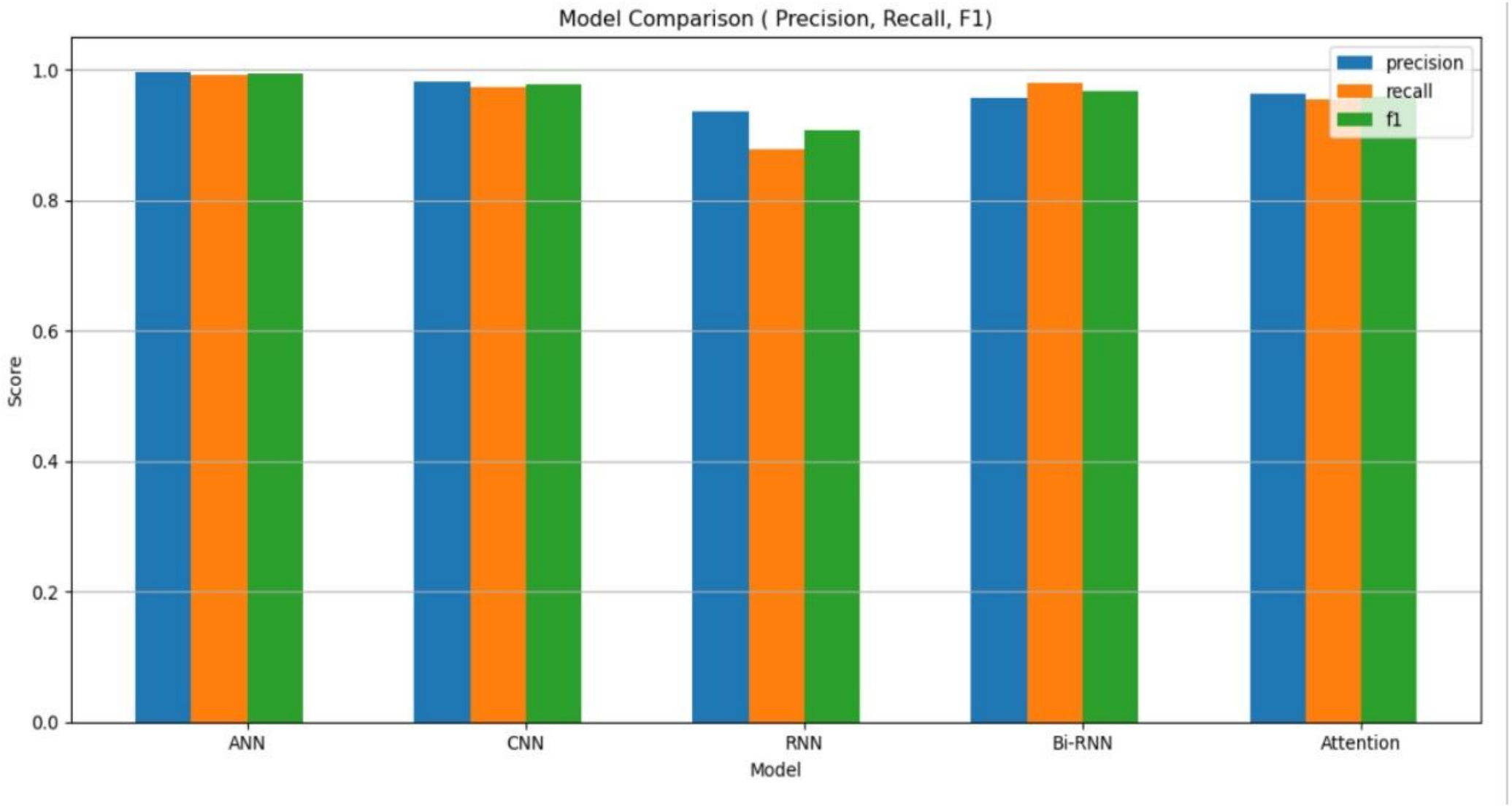
- Enhance the model with multilingual support, real-time deployment, hybrid features, adversarial attack resistance, continuous learning, and lightweight mobile/IoT integration for broader and more robust phishing detection.

Visualisation

Model Accuracy Comparison



Model performance Comparison



Result

Model performance

	accuracy	precision	recall	f1	time
ANN	0.9974	0.9953	0.9960	0.9956	21.1818
CNN	0.9833	0.9828	0.9603	0.9714	67.4729
RNN	0.7042	0.0000	0.0000	0.0000	440.8610
Bi-RNN	0.9893	0.9858	0.9778	0.9818	302.9172
Attention	0.9779	0.9784	0.9462	0.9620	1064.7988

Conclusion

- **High Accuracy and Performance:**
  DEPHIDES achieved **99.80% phishing detection accuracy** using ANN, outperforming traditional ML methods and other deep learning models, while ensuring **low false positives** for reliable real-world application.
- **Large and Balanced Dataset Utilization:**
  The project constructed and utilized **one of the largest phishing detection datasets (URLs)**, ensuring the system learns diverse phishing patterns for **robust and scalable detection**.


Thank
you