Email Phishing Detection Using Machine Learning

Presented by:- Abhijit Bhekare (155) ,
Yash Jangali (172) ,
Rushikesh Mahadik (179)

Academic Year:2025-26

- INTRODUCTION

- Email phishing is a major cybersecurity threat.
- Attackers trick users into revealing sensitive information.
- Manual detection is difficult and time-consuming.
- Machine Learning helps automate phishing detection.

- PROBLEM STATEMENT

- Increase in phishing attacks via emails.
- Users find it hard to identify fake emails.
- Traditional security methods are not always effective.
- Need for an automated phishing detection system.

- OBJECTIVES

- To detect phishing emails automatically.
- To classify emails as Phishing or Legitimate.
- To use machine learning for accurate prediction.
- To provide a simple command-line interface

- DATASET DESCRIPTION

- Dataset name: emails.csv
- Contains phishing and legitimate emails.
- Stored in the dataset folder.
- Used for training and testing the model.

- TECHNOLOGIES USED

- Programming Language: Python
- Libraries:
- Pandas
- NumPy
- Scikit-learn
- Concepts:
- Machine Learning
- Natural Language Processing (NLP)

- SYSTEM ARCHITECTURE

- Input Email Text
- Text Preprocessing
- Feature Extraction
- Machine Learning Model
- Prediction Output

- WORKING OF THE SYSTEM

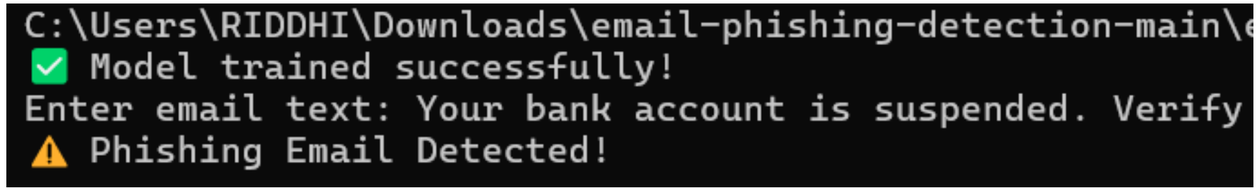
- Dataset is loaded and cleaned.
- Text is converted into numerical features.
- Machine learning model is trained.
- User enters an email.
- System predicts phishing or safe email.

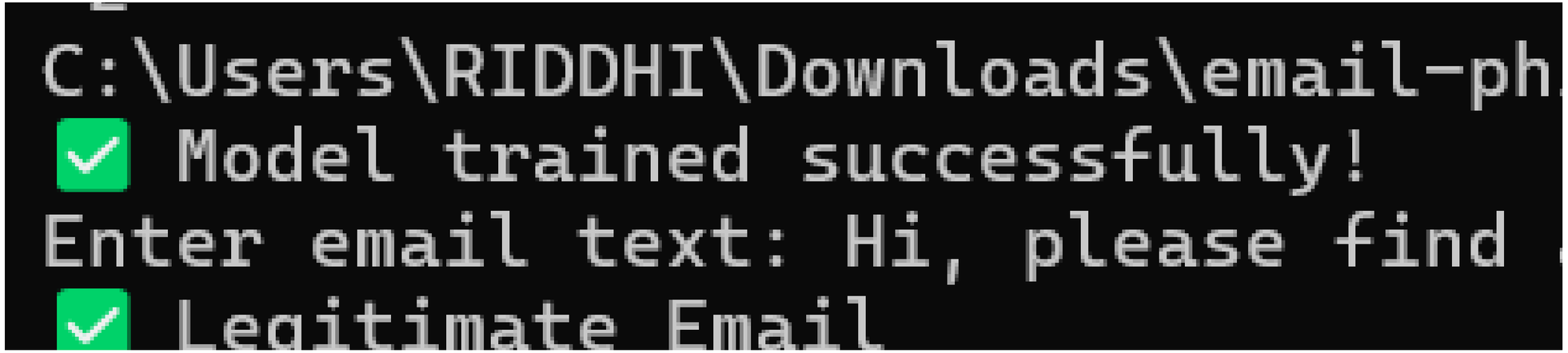
- MODULES USED

- preprocess.py – cleans email text
- model.py – handles model training
- phishing_detection.py – main execution file

- OUTPUT SCREENSHOTS

- Model trained successfully
- User enters email
- Prediction result shown





- ADVANTAGES

- Fast and automated detection
- Reduces risk of cyber fraud
- Easy to use
- Improves email security

- LIMITATIONS

- Accuracy depends on dataset quality
- Cannot detect completely new phishing patterns
- Command-line based interface

- FUTURE SCOPE

- Web-based application
- Real-time email scanning
- Deep learning integration
- Browser or email client extension

- CONCLUSION

- Email phishing is a serious security threat.
- Machine learning helps detect phishing effectively.
- The system successfully classified emails.
- Project achieves its objectives.