Anti-Phishing Email Analyzer

Objective

The Anti-Phishing Email Analyzer is designed to enhance email security by detecting phishing attempts. It identifies potential threats using keyword analysis and URL validation techniques, ensuring safe email communication for users.

Key Features

1. **Keyword Analysis**
    - Scans email content for common phishing keywords like *"verify your account," "click here,"* and *"account locked."*
    - Alerts users if any potentially malicious keywords are detected.
2. **URL Extraction and Validation**
    - Extracts all URLs from the email content.
    - Verifies extracted URLs against a trusted domains list to identify suspicious links.
3. **File-Based Email Input**
    - Accepts email content as a .txt file for analysis, making it easy to integrate with various workflows.
4. **Detailed Results**
    - Provides clear feedback, highlighting phishing keywords and suspicious URLs in the email.

How It Works

1. **Input**
    - User provides an email file (e.g., email_content.txt) containing the email content to be analyzed.
    - Trusted domains and phishing keywords are stored in separate files for easy customization.
2. **Analysis Process**
    - **Step 1: Keyword Detection**
      The analyzer scans the email for phishing keywords that often indicate fraudulent intent.

- **Step 2: URL Validation**
  Extracted URLs are compared against a pre-approved list of trusted domains. Any mismatched domain is flagged as suspicious.

3. **Output**
   - Results include:
     - A list of detected phishing keywords (if any).
     - Suspicious URLs flagged for further investigation.

Technologies Used

- **Programming Language**: Python
- **Key Libraries**:
  - Regular expressions for keyword detection.
  - tldextract for domain extraction and validation.

Benefits

- **Enhanced Security**: Prevents phishing attacks by analyzing email content and links.
- **Customizable**: Easily update trusted domains and phishing keywords.
- **User-Friendly**: Works with simple text files, making it accessible to all users.
- **Scalable**: Can be expanded to include metadata analysis, machine learning models, and GUI interfaces.

Applications

- **Personal Email Protection**: Helps users detect suspicious emails in their inbox.
- **Corporate Email Security**: Assists organizations in monitoring phishing attempts in business communications.
- **Educational Use**: Serves as a foundational project for cybersecurity students and professionals.

Future Scope

1. **Machine Learning Integration**
   - Leverage ML models to predict phishing attempts with higher accuracy.
2. **Real-Time Scanning**
   - Integrate with email services for real-time analysis of incoming emails.

3. **Metadata Analysis**
    - Examine sender details and email headers to identify spoofing or impersonation attempts.
4. **Report Generation**
    - Provide comprehensive, user-friendly reports for detailed threat analysis.
5. **Web-Based Interface**
    - Develop a user-friendly interface for easier email uploads and real-time monitoring.

Conclusion

The Anti-Phishing Email Analyzer offers a straightforward yet powerful solution to detect phishing threats in emails. With its modular design and potential for future enhancements, this project is a valuable tool for improving email security while serving as a strong foundation for further research in cybersecurity.