Phishing Email Detection System

Mercidieu Alexis

Objectives

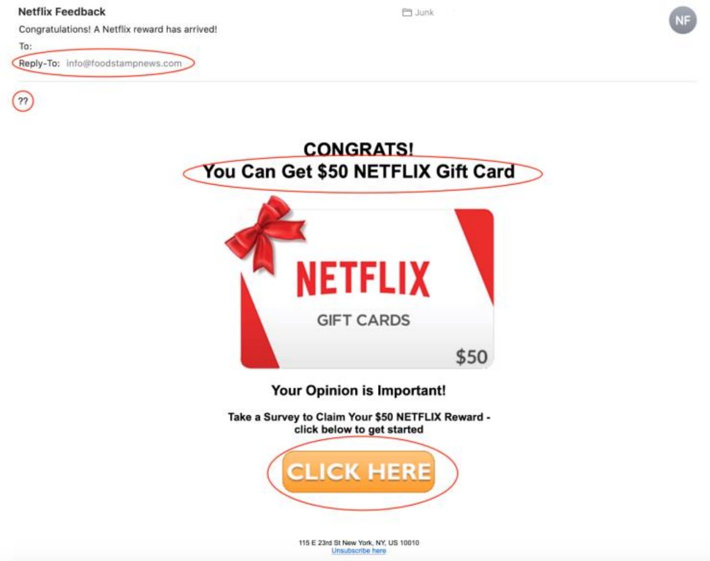
Real Life Scenario

Target

Why I Chose This Approach

Solving the Business Problem

Step-By-Step Walkthrough Of My Code



Netflix Feedback

Congratulations! A Netflix reward has arrived!

To:

Reply-To: info@foodstampnews.com

??

CONGRATS!
You Can Get $50 NETFLIX Gift Card

Your Opinion is Important!

Take a Survey to Claim Your $50 NETFLIX Reward - click below to get started

115 E 23rd St New York, NY, US 10010
Unsubscribe here

$12 billion

In financial losses worldwide

90%

Of cyberattacks

2023

Alone

Target

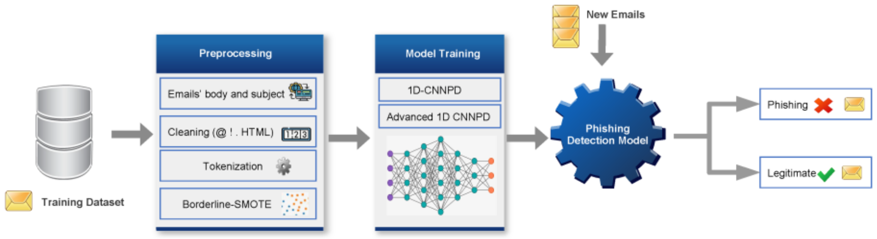
The Business Problem

- Cybercriminals bypass spam filters with evolving tactics
- Users fall for phishing, leading to theft and fraud

Target Market & Stakeholders

- Protect Employees
- Prevent account scams
- Anyone using email is at risk

Why I chose This Approach



Rule-Base Filtering

- Uses predefined rules (flagging emails with certain keywords like “urgent” or “password reset”)
- Attackers easily bypass these tweaking their content

Blacklists Heuristic-Based Detection

- Blocks known threats
- Fails against new, unknown phishing tactics

Solving the Business Problem

How the system works

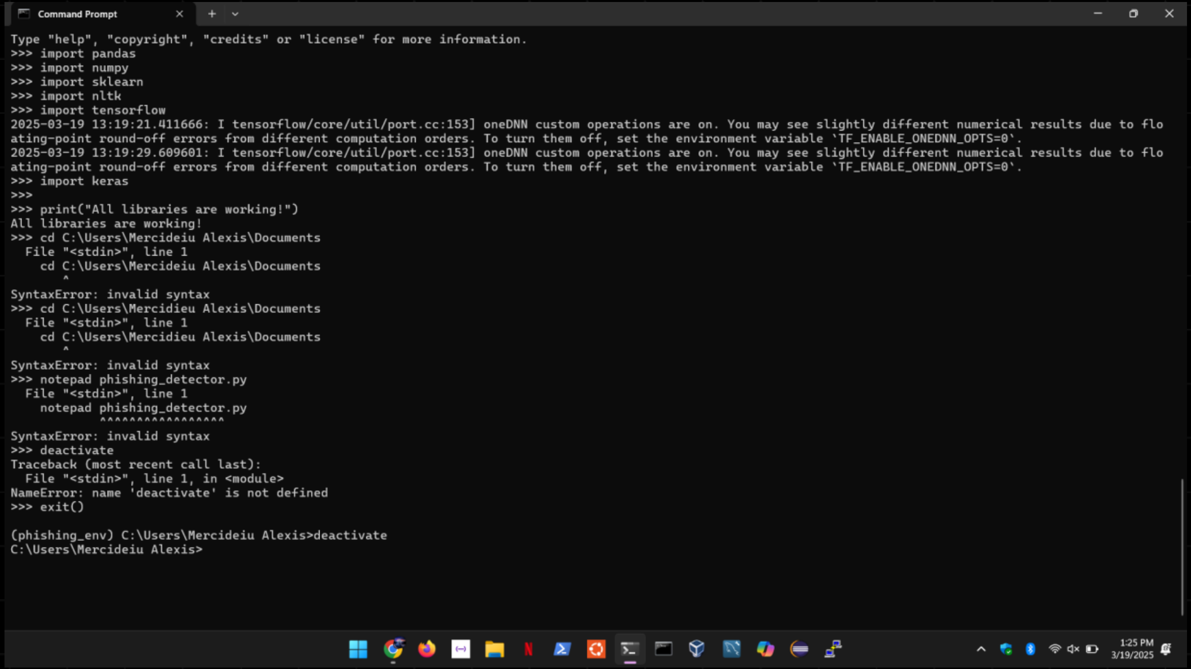
- User uploads or pastes an email
- The System extracts key features (Subjects, sender, email body, links)
- NLP techniques like TF-IDF and word embeddings extracts patterns from the email

- A trained SVM or Random Forest classifier determines if the email is phishing or legitimate
- If phishing is detected, the system provides a warning with an explanation of suspicious elements

Demo of the user interface

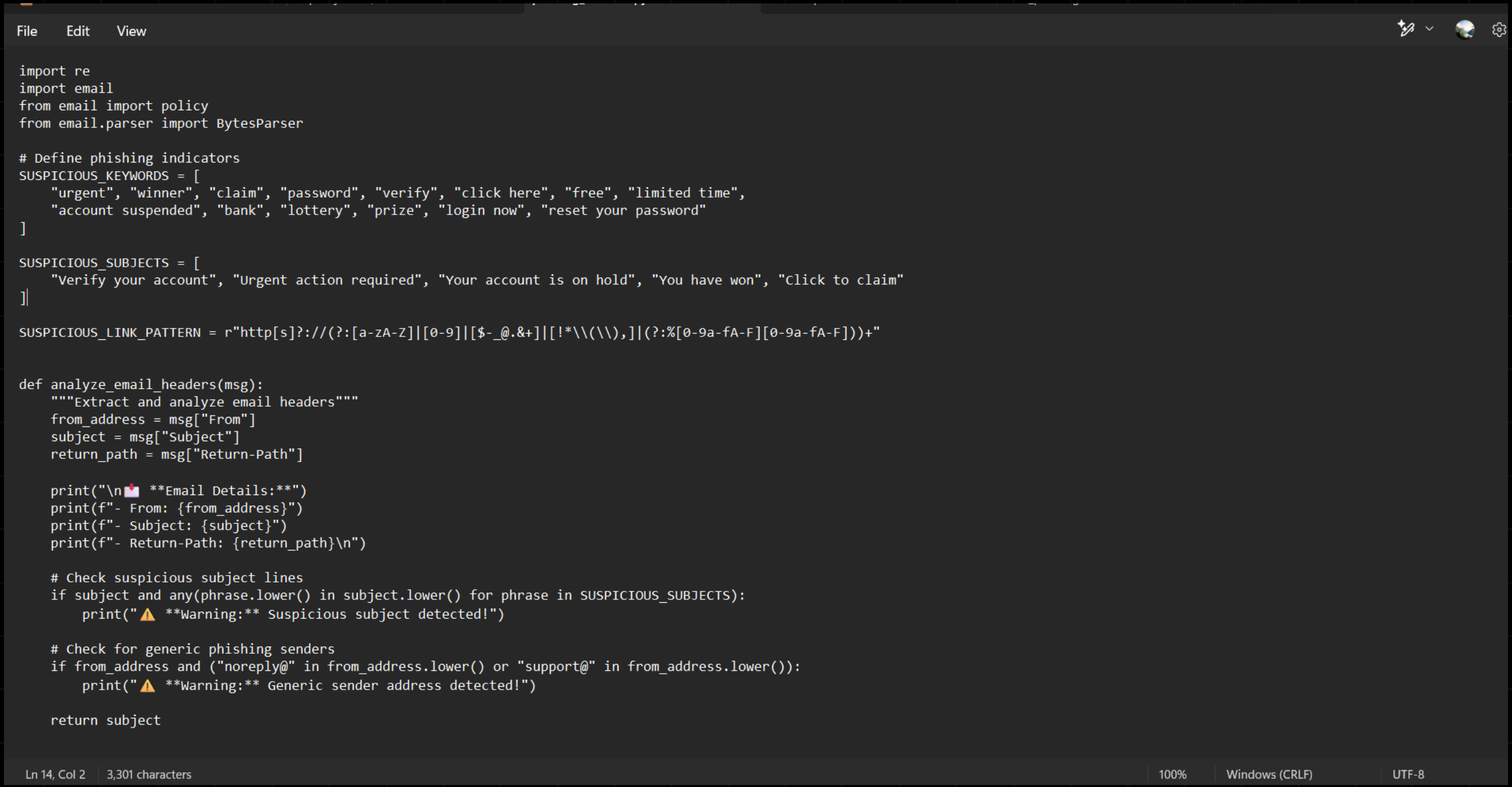
- A simple text box for pasting emails
- Displays a confidence rating
- Highlights red flags like suspicious links or deceptive language to help users understand why an email is dangerous

Step-By-Step Walkthrough Of My Code



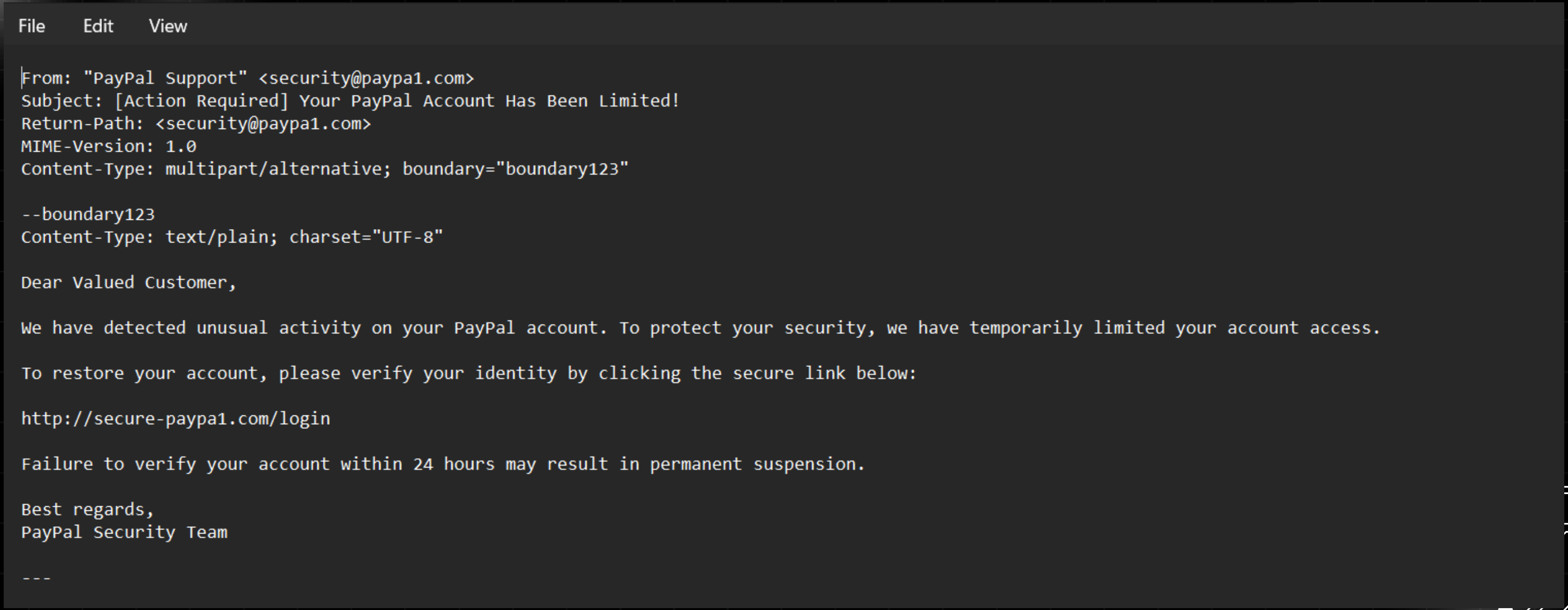
```
Type "help", "copyright", "credits" or "license" for more information.
>>> import pandas
>>> import numpy
>>> import sklearn
>>> import nltk
>>> import tensorflow
2025-03-19 13:19:21.411666: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to flo
ating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2025-03-19 13:19:29.609601: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to flo
ating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
>>> import keras
>>>
>>> print("All libraries are working!")
All libraries are working!
>>> cd C:\Users\Mercideiu Alexis\Documents
  File "<stdin>", line 1
    cd C:\Users\Mercideiu Alexis\Documents
       ^
SyntaxError: invalid syntax
>>> cd C:\Users\Mercidieu Alexis\Documents
  File "<stdin>", line 1
    cd C:\Users\Mercidieu Alexis\Documents
       ^
SyntaxError: invalid syntax
>>> notepad phishing_detector.py
  File "<stdin>", line 1
    notepad phishing_detector.py
            ^^^^^^^^^^^^^^^^^
SyntaxError: invalid syntax
>>> deactivate
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
NameError: name 'deactivate' is not defined
>>> exit()

(phishing_env) C:\Users\Mercideiu Alexis>deactivate
C:\Users\Mercideiu Alexis>
```

Python Code



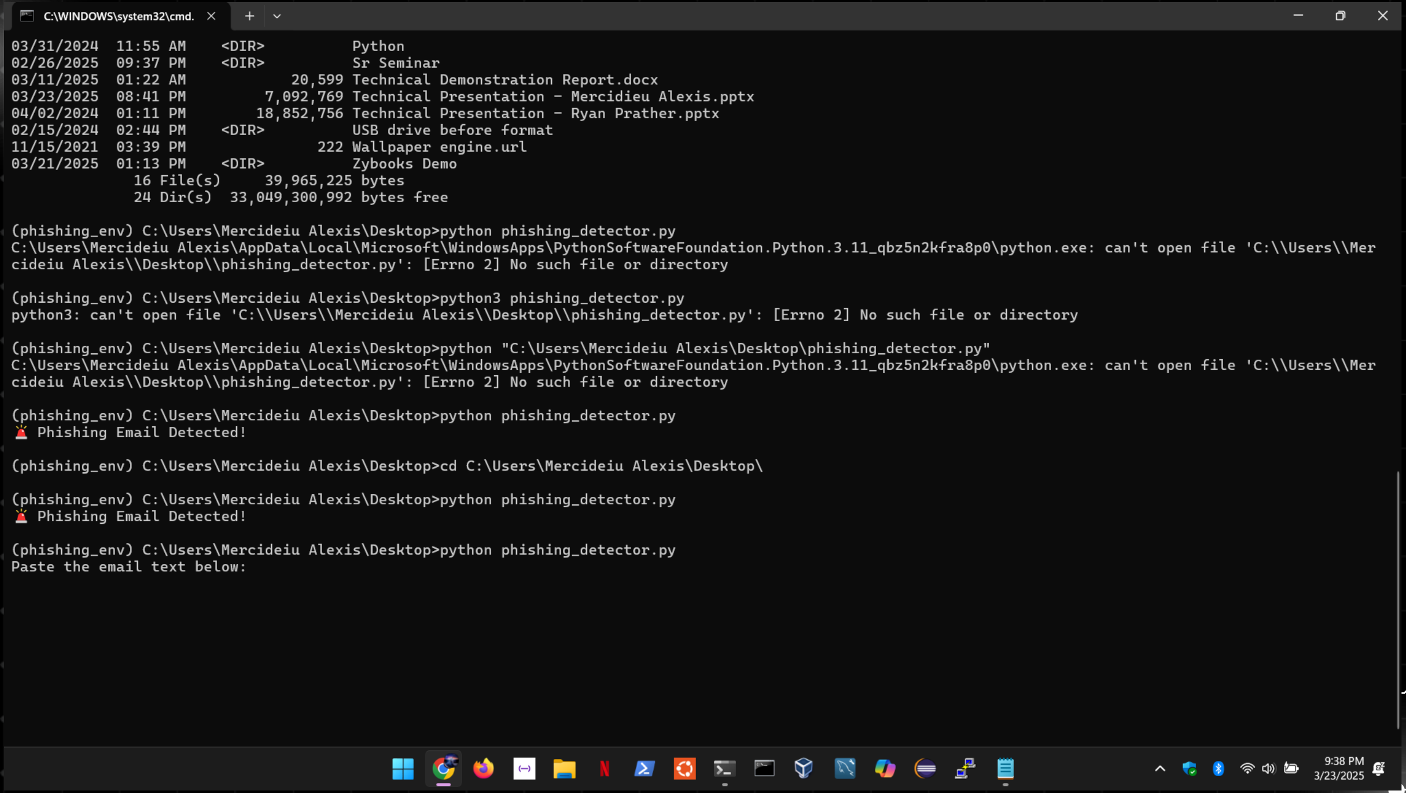
```
import re
import email
from email import policy
from email.parser import BytesParser

# Define phishing indicators
SUSPICIOUS_KEYWORDS = [
    "urgent", "winner", "claim", "password", "verify", "click here", "free", "limited time",
    "account suspended", "bank", "lottery", "prize", "login now", "reset your password"
]

SUSPICIOUS_SUBJECTS = [
    "Verify your account", "Urgent action required", "Your account is on hold", "You have won", "Click to claim"
]

SUSPICIOUS_LINK_PATTERN = r"http[s]?://(?:[a-zA-Z]|[0-9]|[$-_@.&+]|[!*\\(\\),]|(?:%[0-9a-fA-F][0-9a-fA-F]))+"


def analyze_email_headers(msg):
    """Extract and analyze email headers"""
    from_address = msg["From"]
    subject = msg["Subject"]
    return_path = msg["Return-Path"]

    print("\n📩 **Email Details:**")
    print(f"- From: {from_address}")
    print(f"- Subject: {subject}")
    print(f"- Return-Path: {return_path}\n")

    # Check suspicious subject lines
    if subject and any(phrase.lower() in subject.lower() for phrase in SUSPICIOUS_SUBJECTS):
        print("⚠️ **Warning:** Suspicious subject detected!")

    # Check for generic phishing senders
    if from_address and ("noreply@" in from_address.lower() or "support@" in from_address.lower()):
        print("⚠️ **Warning:** Generic sender address detected!")

    return subject
```

Fake phishing email



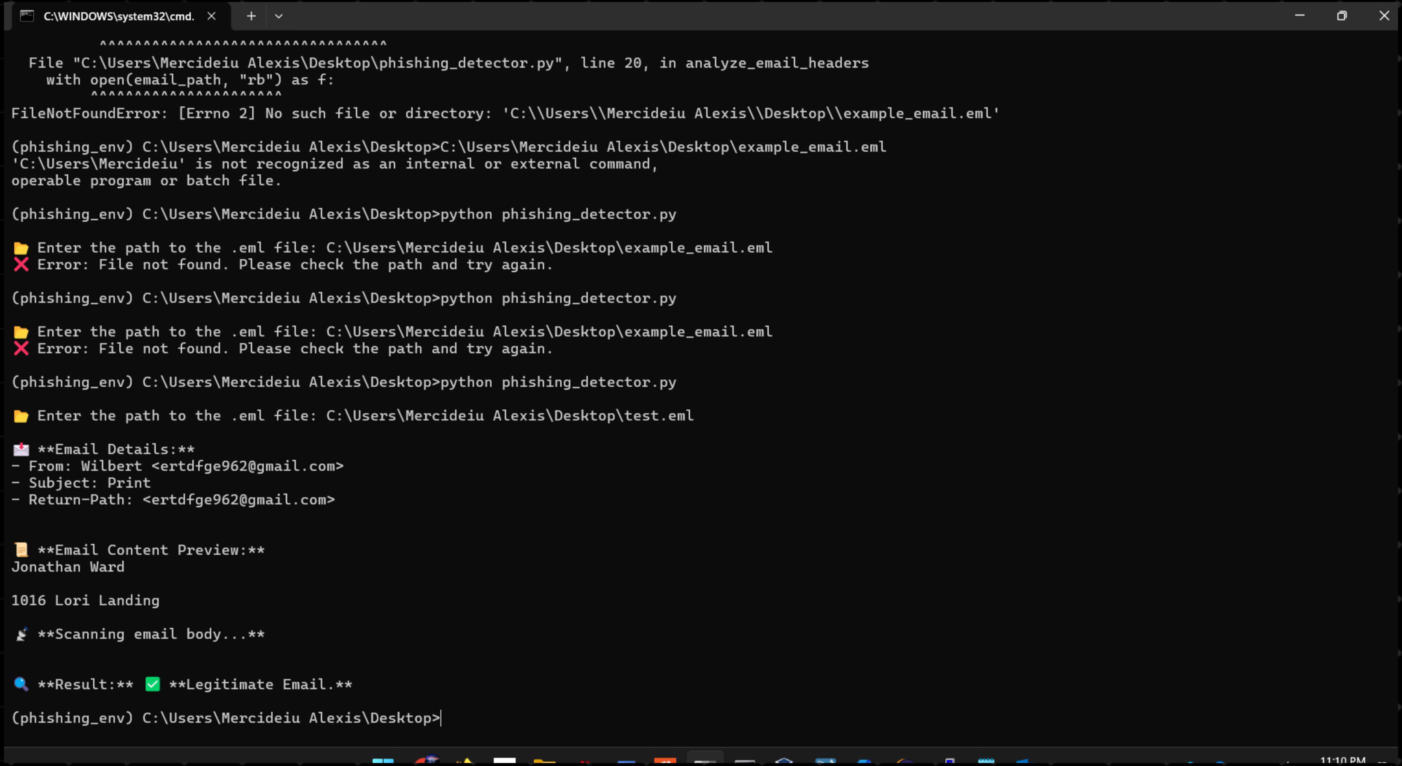
```
From: "PayPal Support" <security@paypa1.com>
Subject: [Action Required] Your PayPal Account Has Been Limited!
Return-Path: <security@paypa1.com>
MIME-Version: 1.0
Content-Type: multipart/alternative; boundary="boundary123"

--boundary123
Content-Type: text/plain; charset="UTF-8"

Dear Valued Customer,

We have detected unusual activity on your PayPal account. To protect your security, we have temporarily limited your account access.

To restore your account, please verify your identity by clicking the secure link below:

http://secure-paypa1.com/login

Failure to verify your account within 24 hours may result in permanent suspension.

Best regards,
PayPal Security Team

---
```

Where to paste your email



```
03/31/2024  11:55 AM    <DIR>          Python
02/26/2025  09:37 PM    <DIR>          Sr Seminar
03/11/2025  01:22 AM            20,599 Technical Demonstration Report.docx
03/23/2025  08:41 PM         7,092,769 Technical Presentation - Mercidieu Alexis.pptx
04/02/2024  01:11 PM        18,852,756 Technical Presentation - Ryan Prather.pptx
02/15/2024  02:44 PM    <DIR>          USB drive before format
11/15/2021  03:39 PM               222 Wallpaper engine.url
03/21/2025  01:13 PM    <DIR>          Zybooks Demo
              16 File(s)     39,965,225 bytes
              24 Dir(s)  33,049,300,992 bytes free

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py
C:\Users\Mercideiu Alexis\AppData\Local\Microsoft\WindowsApps\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\python.exe: can't open file 'C:\\Users\\Mer
cideiu Alexis\\Desktop\\phishing_detector.py': [Errno 2] No such file or directory

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python3 phishing_detector.py
python3: can't open file 'C:\\Users\\Mercideiu Alexis\\Desktop\\phishing_detector.py': [Errno 2] No such file or directory

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python "C:\Users\Mercideiu Alexis\Desktop\phishing_detector.py"
C:\Users\Mercideiu Alexis\AppData\Local\Microsoft\WindowsApps\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\python.exe: can't open file 'C:\\Users\\Mer
cideiu Alexis\\Desktop\\phishing_detector.py': [Errno 2] No such file or directory

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py
🚨 Phishing Email Detected!

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>cd C:\Users\Mercideiu Alexis\Desktop\

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py
🚨 Phishing Email Detected!

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py
Paste the email text below:
```

Final Result of my Code



```
          ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "C:\Users\Mercideiu Alexis\Desktop\phishing_detector.py", line 20, in analyze_email_headers
    with open(email_path, "rb") as f:
         ^^^^^^^^^^^^^^^^^^^^^^
FileNotFoundError: [Errno 2] No such file or directory: 'C:\\Users\\Mercideiu Alexis\\Desktop\\example_email.eml'

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>C:\Users\Mercideiu Alexis\Desktop\example_email.eml
'C:\Users\Mercideiu' is not recognized as an internal or external command,
operable program or batch file.

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py

📂 Enter the path to the .eml file: C:\Users\Mercideiu Alexis\Desktop\example_email.eml
❌ Error: File not found. Please check the path and try again.

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py

📂 Enter the path to the .eml file: C:\Users\Mercideiu Alexis\Desktop\example_email.eml
❌ Error: File not found. Please check the path and try again.

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>python phishing_detector.py

📂 Enter the path to the .eml file: C:\Users\Mercideiu Alexis\Desktop\test.eml

📩 **Email Details:**
- From: Wilbert <ertdfge962@gmail.com>
- Subject: Print
- Return-Path: <ertdfge962@gmail.com>


📜 **Email Content Preview:**
Jonathan Ward

1016 Lori Landing

🔬 **Scanning email body...**


🔍 **Result:** ✅ **Legitimate Email.**

(phishing_env) C:\Users\Mercideiu Alexis\Desktop>
```

Thanks!