# AI-Based Email Phishing Det

Created by Vikendra Chaudhary, Sujal Bansode & satyam yadav

## Introduction

Phishing emails attempt to steal sensitive information such as passwords and cred They are often disguised as trustworthy messages.

#### Goal:

- Detect phishing emails using machine learning.
- Provide both a GUI tool for easy access.

# Why Is Detection Necessary?

- Phishing emails look legitimate.
- Attack methods evolve rapidly.
- Manual detection is slow and error-prone.

#### Need:

- Real-time, intelligent, automated phishing detection.

## **Data Preparation**

#### Dataset:

- Label 1: Phishing (e.g., 'Win a free iPhone now!')
- Label 0: Legitimate (e.g., 'Reminder: meeting at 3 PM')
- Stored in pandas DataFrame with 'message' and 'label' columns.

#### **Feature Extraction**

- Used TF-IDF Vectorizer from scikit-learn.
- Measures word relevance across the corpus.
- Reduces impact of common words.
- Outputs sparse matrix suitable for classification.

# **Model Training**

- Classifier: Multinomial Naive Bayes (MultinomialNB).
- Combined with TF-IDF using scikit-learn Pipeline.
- Trained on the labeled dataset.
- Model saved using joblib as 'spam\_detector\_model.pkl'.

# **GUI Using Tkinter**

- Simple desktop interface with Tkinter.
- Paste email → Click 'Check' → Shows result.
- Labels prediction as 'SPAM Email X' or 'Not Spam ✓'.
- Clean layout with dark theme.

### **Real-World Use Cases**

- Enterprise email filtering.
- Outlook/Gmail plugin.
- Training tool for employees.
- Emergency email screening (incident response).

## **Future Scope**

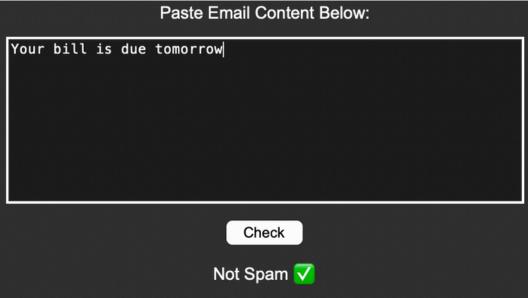
- Add more real-world emails to improve accuracy.
- Use deep learning models (LSTM, BERT).
- Add feedback mechanism for retraining.
- Web-based or browser extension interface.

#### Conclusion

- Built a lightweight AI-based phishing email detector.
- TF-IDF + Naive Bayes for fast, reliable detection.
- GUI for end-user convenience.
- Modular design for future upgrades.

#### **Live GUI Prediction**





# il\_detecto