Project Title: Email Spam Detection Using Naive Bayes

Description:

This project focuses on developing an email spam detection system using the Naive Bayes classification algorithm. The primary goal is to automatically classify incoming emails as either spam or not spam (ham) based on their textual content. Spam emails often contain unsolicited advertisements, phishing links, or harmful attachments, making it essential to filter them out to protect users and improve productivity.

Using a labeled dataset of emails, the system is trained to recognize common patterns and keywords found in spam messages. The Naive Bayes algorithm is particularly effective for this task due to its simplicity, speed, and high accuracy in text classification problems.

Key Features:

- Preprocessing of email text (tokenization, stop word removal, etc.)
- Training using the Naive Bayes classifier
- Evaluation using accuracy, precision, recall, and F1 score
- Real-time prediction of new email messages

Technologies Used:

- Python
- Scikit-learn
- Pandas
- NLTK (for text preprocessing)

This project demonstrates how machine learning can be used to enhance digital communication

security by filtering out unwanted emails automatically.