SPAMEMAILDETECTION

Introduction

In today's digital age, email communication is integral to personal and professional Interactions. However, the rise of spam emails poses significant threats such as phishing attacks, malware distribution, and information theft. This project focuses on leveraging machine learning techniques to develop a robust spam email detection system, enhancing email security for users and organizations.

Objective

a. Develop an efficient and accurate machine learning model for classifying emails as spam or ham.
b. Enhance email security by automatically filtering out spam emails, reducing the risk of phishing attacks.
c. Implement the trained model into a real-time spam detection system for practical deployment.

Methodology

a. Data Collection: Gather a diverse datasets of labeled emails, including spam and ham examples.
b. Preprocessing: Clean and preprocess the data, including text normalization and feature extraction.
c. Feature Engineering: Extract relevant features such as email content, sender information, and metadata.
d. Model Selection: Experiment with various machine learning algorithms and deep learning techniques for classification.
e. Evaluation: Assess the models' performance using metrics like accuracy, precision, recall and F1-score.
f. Deployment: Implement the best-performing model into real-time spam detection system integrated with email clients or servers.

Facilities Required

a. Access to a diverse and labeled datasets of spam and ham emails.
b. Computational resources for training machine learning models and running evaluations.
c. Development environment with programming tools (e.g.,Python).

Bibliography

https://colab.research.google.com/drive/192O3OpJu7jnstN852u7P6b77QbkETVr#scrollTo=wgaf1CxcS9ox
https://www.kaggle.com/datasets/ashfakyeafi/spam-email-classification