Email Spam Classifier - Project Report

Project Summary:

This project implements a machine learning-based email spam classifier using Python and scikit-learn. It uses TF-IDF vectorization to extract features from the email text and a Logistic Regression model to predict whether a message is spam or ham (not spam).

Model Performance:

- Accuracy on training data: 97%

- Accuracy on test data: 97%

- Class labels: 0 = spam, 1 = ham

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.96	0.97	130
1	0.97	0.98	0.97	185

Overall, the model demonstrates excellent performance and generalization, making it suitable for basic email filtering tasks.

Libraries Used:

- numpy

- pandas

- scikit-learn

- fpdf

Sample Input:

'Congratulations! You've won a 100% free iPhone - click here now to claim your limited-time offer!'

Prediction Output:

The mail is spam

Developed by:

Mahmud - AI & ML Developer