AI-Powered Spam Email Classifier

1. Project Overview

The goal of this project is to build an intelligent email spam detection system using Machine Learning (ML) and Natural Language Processing (NLP) techniques. Emails are automatically categorized as Spam or Ham. This helps protect users from phishing, fraud, and irrelevant content.

2. Objectives

- Analyze and clean raw email data.
- Convert email text into numerical features using NLP.
- Train ML models to classify emails.
- Evaluate using accuracy, precision, recall.
- Develop a real-time filtering system.

3. Dataset Description

The dataset contains labeled emails with 'text' (email content) and 'label' (spam or ham).

Initial analysis includes:
- Total number of emails
- Spam vs Ham count
- Missing values
- Duplicates

4. Data Preprocessing

- Text cleaning (remove punctuation, lowercase)
- Tokenization
- Stopword removal
- Lemmatization
- Text vectorization (TF-IDF/BoW)

5. Exploratory Data Analysis (EDA)

- Visualization of Spam vs Ham distribution
- Word clouds and bar plots of frequent spam words

6. Feature Engineering

- Label encoding (spam=1, ham=0)
- Data splitting (train/test)
- Feature extraction (TF-IDF)

7. Model Building

Model: Multinomial Naive Bayes

Steps:

1. Vectorize text
2. Train on labeled data
3. Predict on test set

8. Model Evaluation

- Accuracy: % of correct predictions
- Precision, Recall, F1-score
- Confusion Matrix for performance visualization

9. Spam Detection System

The classifier can:

- Take raw email input
- Predict spam/ham
- Be deployed in applications, chatbots, or web services

10. Conclusion

A high-performance spam filter was created using classic NLP and ML techniques. The system is accurate

and scalable.

11. Future Work

- Use Word2Vec/BERT
- Build web app with Flask
- Train with more data
- Real-time incremental learning