

Key Features:

1. Dataset Processing:

- Loads a dataset of labeled email messages (spam and ham).
- Renames columns and converts labels into binary format (spam = 1, ham = 0).
- Cleans the text by removing punctuation, converting to lowercase, and removing numbers.

2. Feature Extraction:

- Uses **TF-IDF Vectorization** to convert text messages into numerical features, ignoring common stop words.

3. Model Training & Evaluation:

- Splits the data into training and testing sets.
- Trains a **Multinomial Naïve Bayes classifier**, a widely used algorithm for text classification.
- Predicts labels for the test set and evaluates performance using **accuracy score and classification report**.

4. Data Visualization:

- Plots a bar chart to show the distribution of spam and ham emails in the dataset.

5. Spam Prediction Function:

- Allows users to input an email message to check whether it is **Spam** or **Not Spam** based on the trained model.

This program provides an effective solution for detecting spam messages, making it useful for email filtering systems and text classification tasks.