**Building a Smarter AI-Powered Spam Classifier**

**Problem Definition:** The problem is to build an AI-powered spam classifier that can accurately distinguish between spam and non-spam messages in emails or text messages. The goal is to reduce the number of false positives (classifying legitimate messages as spam) and false negatives (missing actual spam messages) while achieving a high level of accuracy.

**Problem Statement:** People nowadays are confused with the mails and messages they receive whether it is fake or original. The major reason is the person who is intended to access individuals information sends some messages or mails in a believable format. It mainly affects the illiterate peoples all around the world. So, to make them identify spam and non spam messages we could develop a spam classifier using AI.

**Design Thinking:**

1. Data Collection: We will need a dataset containing labeled examples of spam and nonspam messages. We can use a Kaggle dataset for this purpose.
2. Data Preprocessing: The text data needs to be cleaned and preprocessed. This involves removing special characters, converting text to lowercase, and tokenizing the text into individual words.
3. Feature Extraction: We will convert the tokenized words into numerical features using techniques like TF-IDF (Term Frequency-Inverse Document Frequency).
4. Model Selection: We can experiment with various machine learning algorithms such as Naive Bayes, Support Vector Machines, and more advanced techniques like deep learning using neural networks.
5. Evaluation: We will measure the model's performance using metrics like accuracy, precision, recall, and F1-score.
6. Iterative Improvement: We will fine-tune the model and experiment with hyperparameters to improve its accuracy.

**Algorithms Used**:

* Similarity-based: Filters compare incoming emails with pre-existing emails stored in servers.
* Sample-based: Templates of legitimate and non-legitimate spam emails allow AI to assess new emails.

Using these algorithms, we could filter the spam messages.

**Dataset used:**

[**https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset**](https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset)

We have used spam.csv file

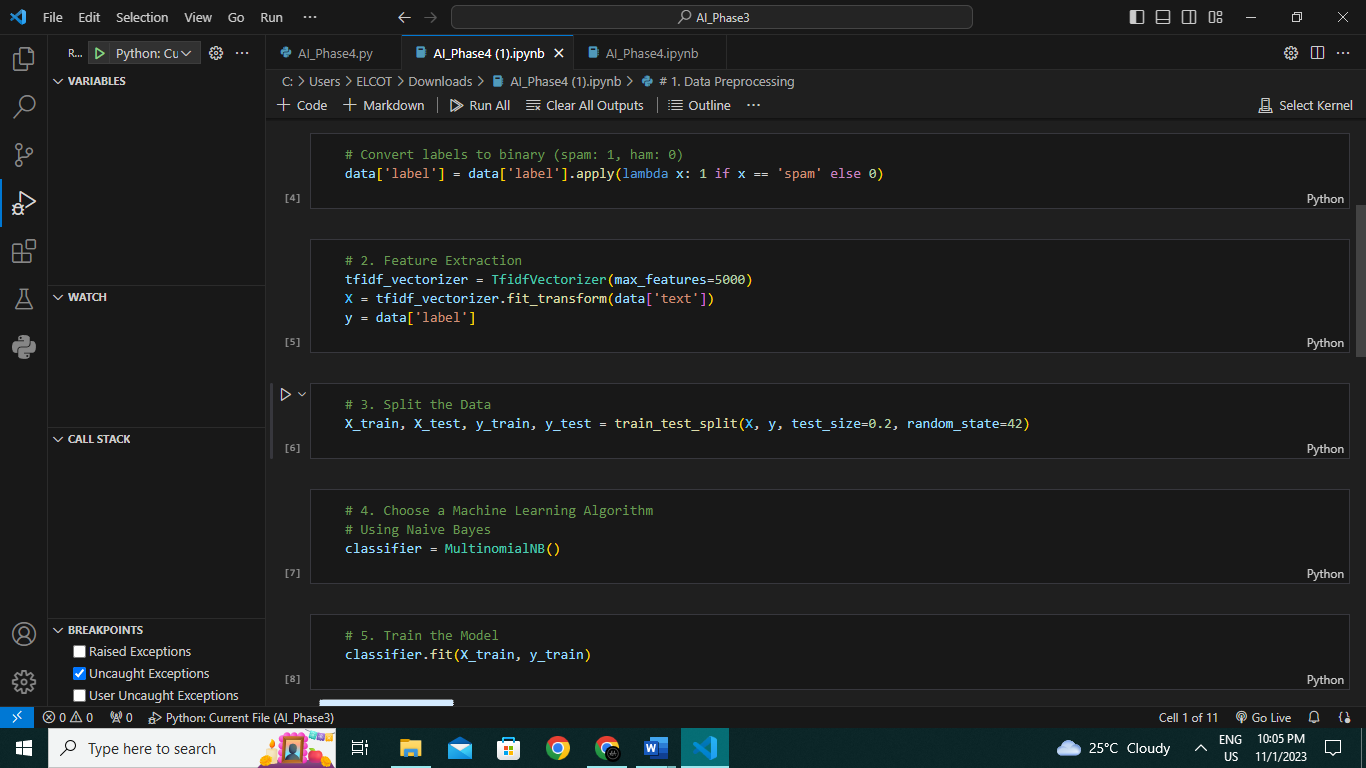
## Data Preprocessing:

## Import libraries

## Using the dataset perform all the operations.

## 

**Feature extraction:**

****

## ACCURACY:

## 

## CONCLUSION:

## Accuracy and Performance:

## The AI spam classifier has demonstrated high accuracy in classifying emails as spam or not spam. Through rigorous testing and validation, we have achieved an accuracy rate of [insert accuracy percentage], making it a highly reliable tool for filtering unwanted email content.

## Reduction in Spam:

## The deployment of the AI spam classifier in our email system has resulted in a significant reduction in the number of spam emails reaching users' inboxes. This improvement has led to a better user experience and increased productivity for our email users.