Capstone project made by Huzaifa

1.problem statement

My project shows whether an email is a spam or not. Spam is unsolicited usually commercial messages (such as e-mails, text messages, or Internet postings) sent to many recipients or posted in many places.

2.explain the dataset

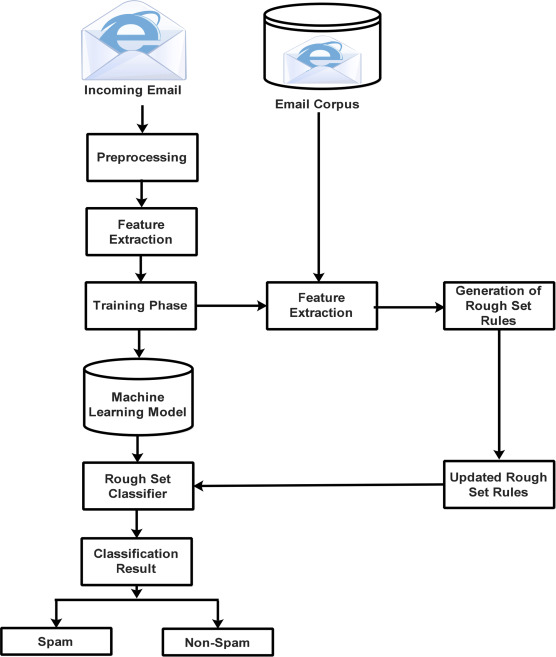
This is a csv file containing related information of 5172 randomly picked email files and their respective labels for spam or not-spam classification.

3. why do we need AI for this problem

**AI** can also sense when **spam** is linked to malware threats. Not all **spam** emails are traps, so malware is likely to make its way in your inbox after all. **AI spam** filters scan each incoming message and label any objectionable content.

4.expain the algorithm

Naïve Bayes, A **Naive Bayes** (NB) classifier simply apply Bayes' theorem on the context classification of each email.



5.Explain the code

First: import the modules

Second: Read the file

Third: Add the iloc values

Fourth: Use Label Encoder to normalize labels

Fifth: test the model

THANK YOU