```
from google.colab import drive
drive.mount('/content/drive')
```

    Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

```
!ls /content/drive/MyDrive/SpamFilter/Results/emails | wc -l
```

    5728

```
!pip install -r /content/drive/MyDrive/SpamFilter/requirements.txt
```

    Requirement already satisfied: pandas==2.1.0 in /usr/local/lib/python3.10/dist-packages (from -r /content/drive/MyDrive/SpamFilter/requirements.
    Requirement already satisfied: cvxopt==1.3.1 in /usr/local/lib/python3.10/dist-packages (from -r /content/drive/MyDrive/SpamFilter/requirements.
    Requirement already satisfied: matplotlib==3.7.0 in /usr/local/lib/python3.10/dist-packages (from -r /content/drive/MyDrive/SpamFilter/requiremen
    Requirement already satisfied: prettytable==3.8.0 in /usr/local/lib/python3.10/dist-packages (from -r /content/drive/MyDrive/SpamFilter/requireme
    Requirement already satisfied: nltk==3.8.1 in /usr/local/lib/python3.10/dist-packages (from -r /content/drive/MyDrive/SpamFilter/requirements.txt
    Requirement already satisfied: numpy>=1.22.4 in /usr/local/lib/python3.10/dist-packages (from pandas==2.1.0->-r /content/drive/MyDrive/SpamFilter
    Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas==2.1.0->-r /content/drive/MyDrive/S
    Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.1.0->-r /content/drive/MyDrive/SpamFilter/
    Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.1.0->-r /content/drive/MyDrive/SpamFilte
    Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/Spa
    Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/SpamFi
    Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/Sp
    Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/Sp
    Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/Spar
    Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/SpamF:
    Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.0->-r /content/drive/MyDrive/Spa
    Requirement already satisfied: wcwidth in /usr/local/lib/python3.10/dist-packages (from prettytable==3.8.0->-r /content/drive/MyDrive/SpamFilter/
    Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk==3.8.1->-r /content/drive/MyDrive/SpamFilter/requireme
    Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk==3.8.1->-r /content/drive/MyDrive/SpamFilter/requirer
    Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk==3.8.1->-r /content/drive/MyDrive/SpamFilter
    Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk==3.8.1->-r /content/drive/MyDrive/SpamFilter/requiremen
    Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas==2.1.0->-r /content/drive

```
import nltk
nltk.download('stopwords')
nltk.download('wordnet')
```

    [nltk_data] Downloading package stopwords to /root/nltk_data...
    [nltk_data]   Package stopwords is already up-to-date!
    [nltk_data] Downloading package wordnet to /root/nltk_data...
    [nltk_data]   Package wordnet is already up-to-date!
    True

```
!cat /content/drive/MyDrive/SpamFilter/main.py
```

    from generate_wordlist import *
    from generate_frequency import *
    from svm_kernel import *
    from svm_trainer import *
    from svm_predictor import *
    from svm_utils import *

    if __name__ == '__main__':

        generate_wordlist()

        generate_frequency()

        train_and_test()

```
!python3 /content/drive/MyDrive/SpamFilter/main.py
```

```
Finished email files: 4100
Finished email files: 4200
Finished email files: 4300
Finished email files: 4400
Finished email files: 4500
Finished email files: 4600
Finished email files: 4700
Finished email files: 4800
Finished email files: 4900
Finished email files: 5000
Finished email files: 5100
Finished email files: 5200
Finished email files: 5300
Finished email files: 5400
Finished email files: 5500
Finished email files: 5600
Finished email files: 5700
Time for segregating data and forming input vector(word frequency): 1400.86 seconds
Finished files: 1
Finished files: 2
Finished files: 3
Finished files: 4
Finished files: 5
Finished files: 6
Finished files: 7
Finished files: 8
Finished files: 9
Finished files: 10
Finished files: 11
Finished files: 12
Finished files: 13
Finished files: 14
Finished files: 15
Finished files: 16
Finished files: 17
Finished files: 18
Finished files: 19
Finished files: 20
Finished files: 21
```