import numpy as np *# linear algebra*

import pandas as pd *# data processing, CSV file I/O (e.g. pd.read\_csv)*

*# Input data files are available in the read-only "../input/" directory*

*# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory*

import os

for dirname, \_, filenames **in** os.walk('D:\DNO\minor'):

for filename **in** spam:

print(os.path.join(dirname, filename))

data=pd.read\_csv('D:\DNO\minor\spam.csv')

data

data.columns

data.info()

data['Spam']=data['Category'].apply(lambda x:1 if x=='spam' else 0)

data.head(5)

from sklearn.model\_selection import train\_test\_split

X\_train,X\_test,y\_train,y\_test=train\_test\_split(data.Message,data.Spam,test\_size=0.25)

*#CounterVectorizer Convert the text into matrics*

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.naive\_bayes import MultinomialNB

from sklearn.pipeline import Pipeline

clf=Pipeline([

('vectorizer',CountVectorizer()),

('nb',MultinomialNB())

])

Tarining The Model

clf.fit(X\_train,y\_train)

emails=[

'Sounds great! Are you home now?',

'Will u meet ur dream partner soon? Is ur career off 2 a flyng start? 2 find out free, txt HORO followed by ur star sign, e. g. HORO ARIES'

]

**Predict Email**

clf.predict(emails)

Prediction Of Model

clf.score(X\_test,y\_test)

0.9777458722182341