**ML ALGORITHMS IN PYTHON FOR HAM AND SPAM MESSAGE CLASSIFICATION**

**Program Code for Loading and Preprocessing Dataset:**

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

df = pd.read\_csv("/content/drive/MyDrive/Mail Spam-NLP/spam.csv", encoding="ISO-8859-1")

df.head(10)

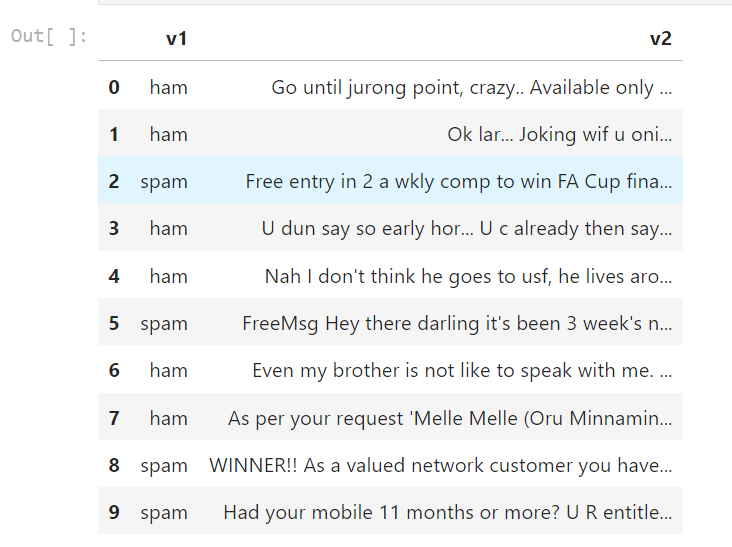


#Drop empty columns

cols = [2,3,4]

df.drop(df.columns[cols],axis=1,inplace=True)

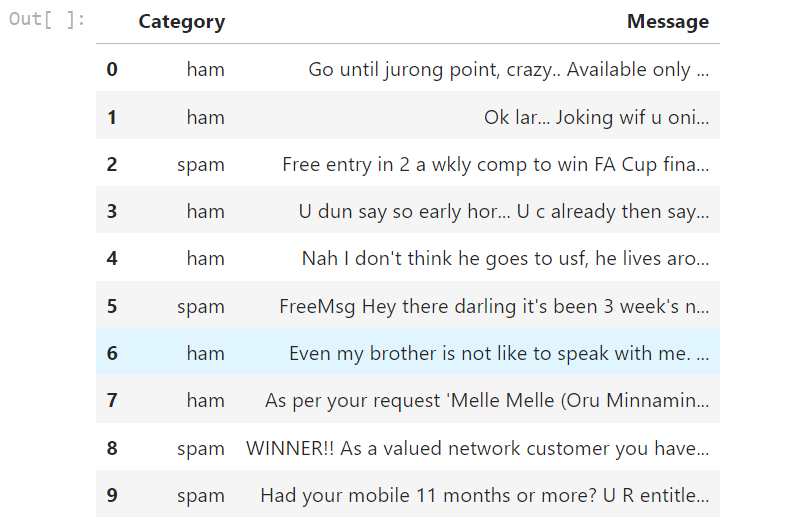
df.head(10)



#Rename columns as category and message

df.rename(columns = {'v1':'Category', 'v2':'Message'}, inplace = True)

df.head(10)



print(f'Dataset consist of {df.shape[0]} E-Mails.')

df['Category'].value\_counts()

plt.figure(figsize=(8,6))

df['Category'].value\_counts().plot.bar(color = ["orange","orange"])

plt.title('Total number of ham and spam in the dataset')

plt.show()

