**Name:** Muzammil Rauf Patel

**Class:** SYMCA **Div:** B

**Roll No:** 45 **Batch:** B2

**Course:** ML Lab

------------------------------------------------------------------------------------------------

**Assignment No: 4**

**Problem Statement:**

Apply K-Nearest Neighbor Classifier on Data set. Test for Accuracy and Precision. Classify the email using binary classification method. Email Spam detection has two states: a) Normal State - Not Spam, b) Abnormal State - Spam. Use K-Nearest Neighbors for classification.

**------------------------------------------------------------------------------------------------**

**Code:**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.preprocessing import StandardScaler

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

from sklearn.linear\_model import LinearRegression

from sklearn.neighbors import KNeighborsClassifier

from sklearn.metrics import accuracy\_score

from sklearn.metrics import precision\_score

# Load the dataset from a CSV file

data = pd.read\_csv('email.csv')

print(data.head())

data['Prediction'] = data['Prediction'].replace(0,"spam")

data['Prediction'] = data['Prediction'].replace(1,"ham")

X = data.iloc[:,1:-1]

y = data.iloc[:,-1]

print(X,y)

# Split the data into training and test sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2,random\_state=0)

# Scale the features using StandardScaler

scaler = StandardScaler()

X\_train = scaler.fit\_transform(X\_train)

X\_test = scaler.transform(X\_test)

knn = KNeighborsClassifier(n\_neighbors=5, metric='minkowski', p=2)

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

y\_pred = knn.predict(X\_test)

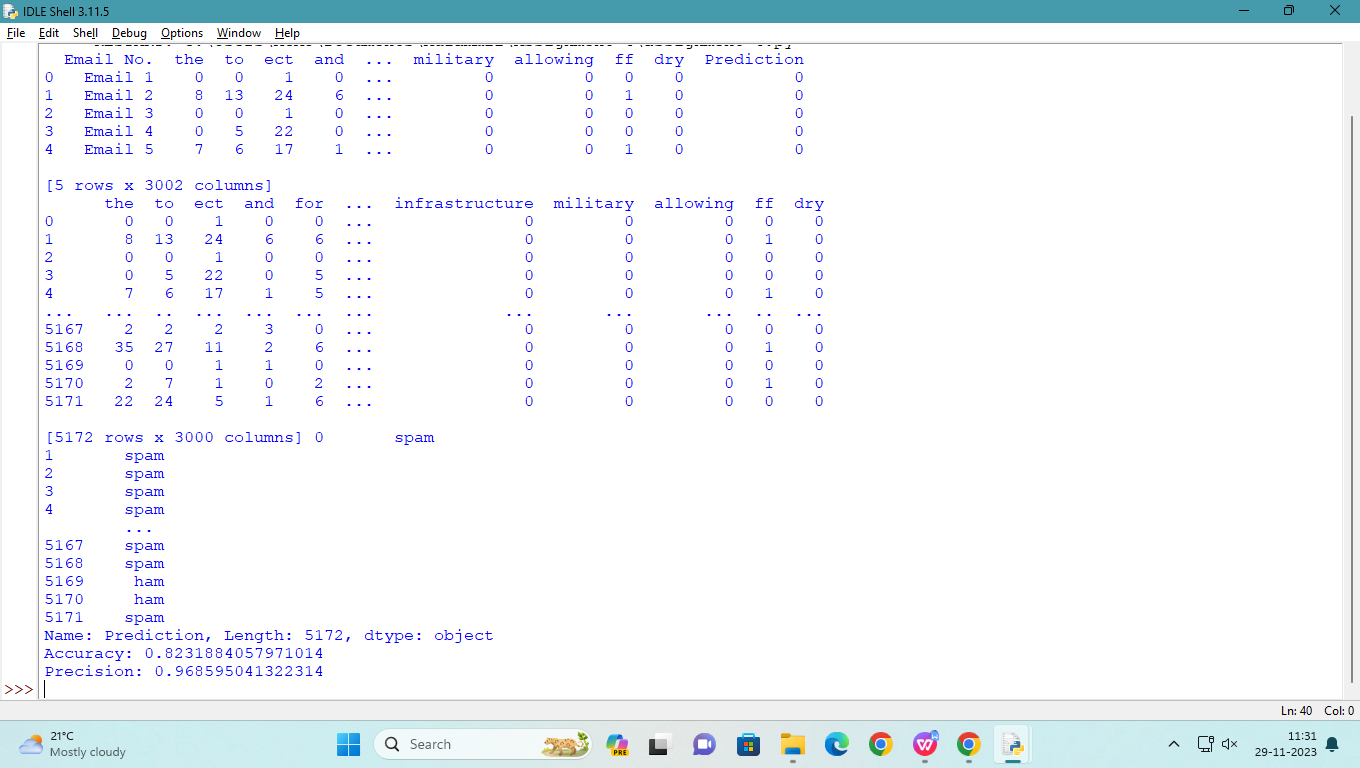
accuracy = accuracy\_score(y\_test, y\_pred)

precision = precision\_score(y\_test, y\_pred, pos\_label='spam')

print("Accuracy:", accuracy)

print("Precision:",precision)

**Output:**

****