**NAME : MANGESH A. GHADWAJE**

**ROLL NO:24**

**BATCH : B2**

**COURSE: ML PRACTICAL**

**Assginment No. 4**

**Problem Statement :**

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

**Code :**

***# importing all the libraries***

***import numpy as np***

***import matplotlib.pyplot as plt***

***import pandas as pd***

***# Importing the dataset***

***dataset = pd.read\_csv('Social\_Network\_Ads.csv')***

***print(dataset.head())***

***# Filtering out columns to retain age and salary columns***

***X = dataset.iloc[:,[2,3]]. values***

***Y = dataset.iloc[:,4].values***

***# Data split for training and testing (75/25)***

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

***X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X,Y, test\_size = 0.25, random\_state = 0)***

***#Scaling using standard scalar for Normal Distribution***

***from sklearn.preprocessing import StandardScaler***

***sc = StandardScaler()***

***X\_train = sc.fit\_transform(X\_train)***

***X\_test = sc.transform(X\_test)***

***#Building the model using RBF kernel***

***from sklearn.svm import SVC***

***classifier\_rbf = SVC(kernel = 'rbf', random\_state = 0)***

***classifier\_rbf.fit(X\_train, Y\_train)***

***Y\_pred\_rbf = classifier\_rbf.predict(X\_test)***

***#Printing the confusion matrix***

***from sklearn.metrics import confusion\_matrix***

***cm\_rbf = confusion\_matrix(Y\_test, Y\_pred\_rbf)***

***print(cm\_rbf)***

***# Classification Report***

***from sklearn.metrics import classification\_report***

***class\_report\_rbf = classification\_report(Y\_test, Y\_pred\_rbf)***

***print(class\_report\_rbf)***

**Output :**

