"""

@author: Eswar

"""

"Importing required libraries"

import pandas as pd

import numpy as np

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

from sklearn.preprocessing import StandardScaler

from sklearn.neighbors import KNeighborsClassifier

from sklearn.metrics import confusion\_matrix

from sklearn.metrics import f1\_score

from sklearn.metrics import accuracy\_score

"Read dataset file in CSV format"

dataset=pd.read\_csv("data\_banknote.csv")

print(dataset.head())

print(dataset.describe())

print(dataset.groupby('class').size())

"Store independent variables in X and dependent variable in Y"

X=dataset.drop(['class'],axis=1)

Y=dataset['class']

"Splitting the data into train and test data"

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X,Y,test\_size=0.20,random\_state=0)

"Feature Scaling"

sc\_X = StandardScaler()

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

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

"Calculating value of k for KNN"

import math

k=math.sqrt(len(Y\_test))

"Define the model: Init K-NN"

classifier= KNeighborsClassifier(n\_neighbors=15, p=2, metric='euclidean')

"Fit the model"

Y\_pred = classifier.fit(X\_train,Y\_train)

"Predict the test-set result"

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

print(Y\_pred)

"Evaluation of model using confusion matrix"

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

print(cm)

print(f1\_score(Y\_test,Y\_pred))

"Number of '1's in predicted set"

print((Y\_pred==1).sum())