import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

dataset = pd.read\_csv('Data.csv')

X = dataset.iloc[:, :-1].values

y = dataset.iloc[:, 3].values

#Taking care of missing data

'''from sklearn.preprocessing import Imputer

imputer = Imputer(missing\_values = 'NaN',strategy='mean',axis=0)

imputer=imputer.fit(X[:, 1:3])

X[:, 1:3]=imputer.transform(X[:, 1:3])

#encoding categorical data

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

labelencoder\_x= LabelEncoder()

X[:,0]=labelencoder\_x.fit\_transform(X[:,0])

onehotecoder=OneHotEncoder(categorical\_features=[0])

X=onehotecoder.fit\_transform(X).toarray()

labaelencoder\_y=LabelEncoder()

y=labaelencoder\_y.fit\_transform(y)''''

#Splitting the dataset into the training set and test set

from sklearn.cross\_validation import train\_test\_split

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

#feature scaling

from sklearn.preprocessing import StandardScaler

sc\_x=StandardScaler()

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

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