**LAB TASK-2**

**CODE-**

import numpy as np

import pandas as pd

from sklearn.linear\_model import LinearRegression

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

def mse(y\_t,y\_pre):

su=0

for i in range(len(y\_t)):

su=su+(y\_t[i]-y\_pre[i])\*\*2

return(su/len(y\_t))

def test\_train(x,y,tes):

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=tes,random\_state=1)

reglin=LinearRegression()

reglin.fit(x\_train,y\_train)

y\_pred=reglin.predict(x\_test)

print('rmse',(mse(list(y\_test),list(y\_pred)))\*\*(0.5))

print('mse',(mse(list(y\_test),list(y\_pred))))

#data\_link="https://www.kaggle.com/karthickveerakumar/salary-data-simple-linear-regression"

data=pd.read\_csv(data\_link)

features=['YearsExperience']

x=data[features]

y=data.Salary

print('\n16BIS0137 \n T.DHIRAJ\n')

print('train : 50%')

test\_train(x,y,0.5)

print('\n')

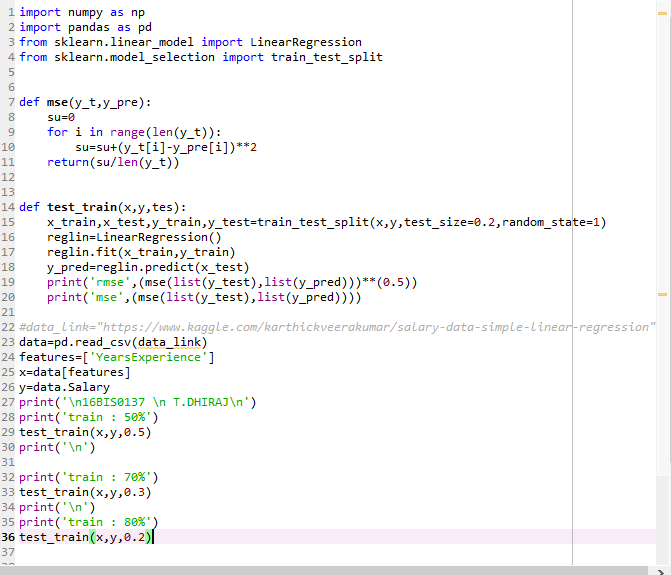
print('train : 70%')

test\_train(x,y,0.3)

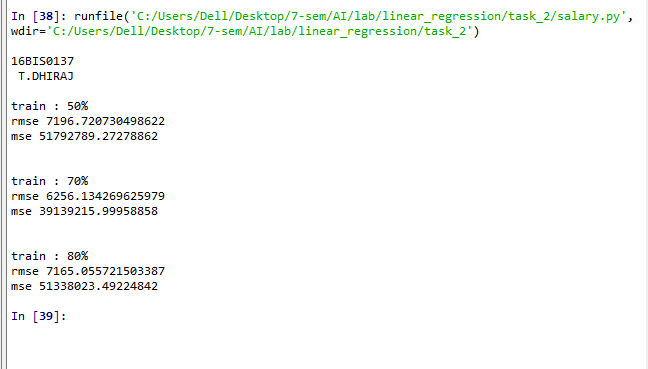
print('\n')

print('train : 80%')

test\_train(x,y,0.2)



**OUTPUT**



train : 50%

rmse 7196.720730498622

mse 51792789.27278862

train : 70%

rmse 6256.134269625979

mse 39139215.99958858

train : 80%

rmse 7165.055721503387

mse 51338023.49224842