T.DHIRAJ
16BIS0137

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"
```

T.DHIRAJ
16BIS0137

```
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)
```

T.DHIRAJ
16BIS0137

```python
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=0.2,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)

```

T.DHIRAJ
16BIS0137
OUTPUT

```
In [38]: runfile('C:/Users/Dell/Desktop/7-sem/AI/lab/linear_regression/task_2/salary.py',
wdir='C:/Users/Dell/Desktop/7-sem/AI/lab/linear_regression/task_2')

16BIS0137
 T.DHIRAJ

train : 50%
rmse 7196.720730498622
mse 51792789.27278862


train : 70%
rmse 6256.134269625979
mse 39139215.99958858


train : 80%
rmse 7165.055721503387
mse 51338023.49224842

In [39]:
```

train : 50%

rmse 7196.720730498622

mse 51792789.27278862

train : 70%

rmse 6256.134269625979

mse 39139215.99958858

train : 80%

rmse 7165.055721503387

mse 51338023.49224842