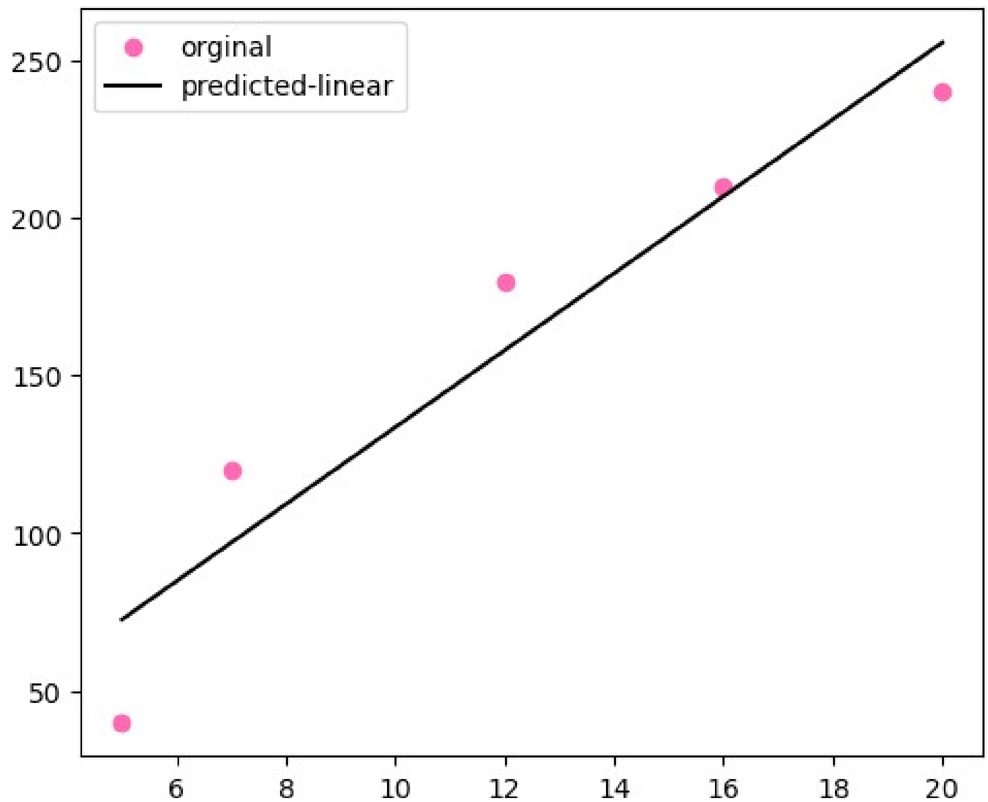
linear regression

- y=a+bx+b1 x1+b2 x2......
- y=>dependent(1) [10]
- x=>independent(0) [20]

```
from sklearn.linear_model import LinearRegression
import numpy as np
from sklearn.metrics import
r2_score,mean_absolute_error,mean_squared_error

time=np.array([5,7,12,16,20]).reshape(-1,1)

mass=np.array([40,120,180,210,240])

mymodel=LinearRegression()
mymodel.fit(time,mass)

LinearRegression()

x=int(input("enter the time in minute:"))
result=mymodel.predict([[x]])
print("if the time is",x,"minututes the mass is",result[0],"grams")

enter the time in minute: 41

if the time is 41 minututes the mass is 512.025974025974 grams

mass_model=mymodel.predict(time)
print(mass_model)

[ 72.54545455  96.96103896 158.         206.83116883 255.66233766]

import matplotlib.pyplot as plt
plt.figure(figsize=(6,5))
plt.scatter(time,mass,label='orginal',color='hotpink')
plt.plot(time,mass_model,label='predicted-linear',color='k')
plt.legend()
plt.show()
```



Evalution

R-square

- lower,the better

```
r2score=r2_score(time,mass_model)
print(r2score)

-816.6925282509699
```

MSE

- LOWER THE BETTER

```
mse=mean_squared_error(time,mass_model)
print(mse)

25184.929870129872
```

MAE

- LOWER THE BETTER

```
mae=mean_absolute_error(time,mass_model)
print(mae)

146.0
```

CASE: Predicating the salary from age,experience,gender,education

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as pt
from sklearn.preprocessing import LabelEncoder
from sklearn.linear_model import LinearRegression
from sklearn.metrics import
r2_score,mean_absolute_error,mean_squared_error
from sklearn .model_selection import train_test_split

x=pd.read_csv(r"C:\Users\DELL\Downloads\Salary_EDA.csv")
x.head()

    Age  Gender Education Level          Job Title  Years of
Experience  \
0  32.0    Male      Bachelor's  Software Engineer
5.0
1  28.0  Female        Master's       Data Analyst
3.0
2  45.0    Male             PhD     Senior Manager
15.0
3  36.0  Female      Bachelor's    Sales Associate
7.0
4  36.0  Female      Bachelor's    Sales Associate
7.0

     Salary
0   90000.0
1   65000.0
2  150000.0
3   60000.0
4   60000.0
```

DATA PREPROCESSING

```
g_e=LabelEncoder()
x['Gender_e']=g_e.fit_transform(x['Gender'])
x['Education_e']=g_e.fit_transform(x['Education Level'])
x.head()

    Age  Gender Education Level          Job Title  Years of
Experience  \
0  32.0    Male      Bachelor's  Software Engineer
5.0
```

```
1  28.0  Female        Master's       Data Analyst
3.0
2  45.0    Male             PhD     Senior Manager
15.0
3  36.0  Female      Bachelor's    Sales Associate
7.0
4  36.0  Female      Bachelor's    Sales Associate
7.0

     Salary  Gender_e  Education_e
0   90000.0         1            0
1   65000.0         0            1
2  150000.0         1            2
3   60000.0         0            0
4   60000.0         0            0
```