EXPERIMENT 3 :

Aim :

Verifying the performance of a over fitting by using choosen database with python code.

Program :

import numpy as np

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

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_squared\_error

import matplotlib.pyplot as plt

np.random.seed(42)

x=2\*np.random.rand(100,1)

y=4+3\*x+np.random.randn(100,1)

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=42)

model=LinearRegression()

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

y\_train\_pred=model.predict(x\_train)

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

mse\_train=mean\_squared\_error(y\_train,y\_train\_pred)

mse\_test=mean\_squared\_error(y\_test,y\_test\_pred)

print(mse\_train)

print(mse\_test)

plt.scatter(x\_test,y\_test, color='yellow',label='Actual')

plt.plot(x\_test,y\_test\_pred,color='pink',label='Prediction')

plt.title('Linear Regression')

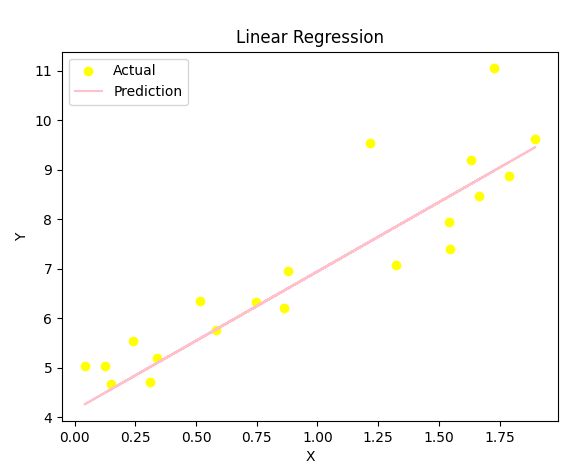
plt.xlabel('X')

plt.ylabel('Y')

plt.legend()

plt.show()

Output :



0.8476788564209705

0.6536995137170021