

```

# Import libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# import data set
data=pd.read_csv(r"C:\Users\sunil\Desktop\DK\vs code\SALARY_PREDICTION\Salary_Data.csv")

#split into dependent and independent variable
x=data.iloc[:,0:1].values
y=data.iloc[:,1].values

# split the data to train test split

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,train_size=0.8,test_size=0.2,random_state=0)

# fit LinearRegression model to x_train, y_train

from sklearn.linear_model import LinearRegression
regressor=LinearRegression()
regressor.fit(x_train,y_train)

#pridict
y_pred=regressor.predict(x_test)
print('The prediction',y_pred)

#comparision
comp=pd.DataFrame({"actual":y_test,"predict":y_pred})
print(comp)

# plot visuals
plt.scatter(x_test,y_test,color="red") # plot data points based on test values
plt.plot(x_train,regressor.predict(x_train),color="blue") # plot regression line with x_train
plt.xlabel("year of experience")
plt.ylabel("Salary")
plt.title("salary VS experience")
plt.show()

#slope
m=regressor.coef_
print("solpe: ",m)

# intercept/slope
c=regressor.intercept_
print("intercept: ",c)

# future prediction y=mx + c of 13 year exp
xp=20
fpred=m*xp+c
print(f"the salary of {xp} year is {fpred}")

# baias and variance score
bais=regressor.score(x_train,y_train)
print("The baise score:-",bais)

variance=regressor.score(x_test,y_test)
print("The variance score :-",variance)

```

```
# Save the trained model to disk
import pickle
filename = 'salary price.pkl'
with open(filename, 'wb') as file:
    pickle.dump(regressor, file)
print("Model has been pickled and saved as salary price.pkl")

import os
print(os.getcwd())
```