```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

```
df = pd.read_csv('/content/Salary_dataset.csv')
```

```
x = df['YearsExperience']
y = df['Salary']
x=np.array(x).reshape(-1,1)
y=np.array(y).reshape(-1,1)
```

```
x_train.shape,x_test.shape,y_train.shape,y_test.shape
```

```
((24, 1), (6, 1), (24, 1), (6, 1))
```

```
from sklearn.model_selection import train_test_split
```

```
x_train, x_test, y_train, y_test = train_test_split(
    x, y,
    test_size=0.2,
    random_state=42
)
```

```
x_train
```

```
array([[10.4],
       [ 8.8],
       [ 4.1],
       [ 1.2],
       [ 2.3],
       [ 5.2],
       [ 3. ],
       [ 4.2],
       [ 4.1],
       [ 8. ],
       [ 1.4],
       [ 1.6],
       [ 9.1],
       [ 2.1],
       [ 7.2],
       [ 9.6],
       [ 6. ],
       [10.6],
       [ 6.9],
       [ 3.3],
       [ 4. ],
       [ 4.6],
       [ 6.1],
       [ 3.1]])
```

```
model = LinearRegression()
```

```
model.fit(x_train,y_train)
```

▾ LinearRegression

```
LinearRegression()
```

```
y_pred = model.predict(x_test)
```

```
print("Actual Salary:",y_test[:5])
print("Predicted Salary:",y_pred[:5])
```

```
Actual Salary: [[112636.]
 [ 67939.]
 [113813.]
 [ 83089.]
 [ 64446.]]
Predicted Salary: [[115791.21011287]
 [ 71499.27809463]
 [102597.86866063]
 [ 75268.80422384]
 [ 55478.79204548]]
```

```
from sklearn.metrics import mean_squared_error
```

```
mse = mean_squared_error(y_test,y_pred)
print("Mean Squared Error:",mse)
```

```
Mean Squared Error: 49830096.855908394
```

```
rmse = np.sqrt(mse)
print("Root Mean Squared Error:",rmse)
```

```
Root Mean Squared Error: 7059.04362190151
```

```
from sklearn.metrics import r2_score
```

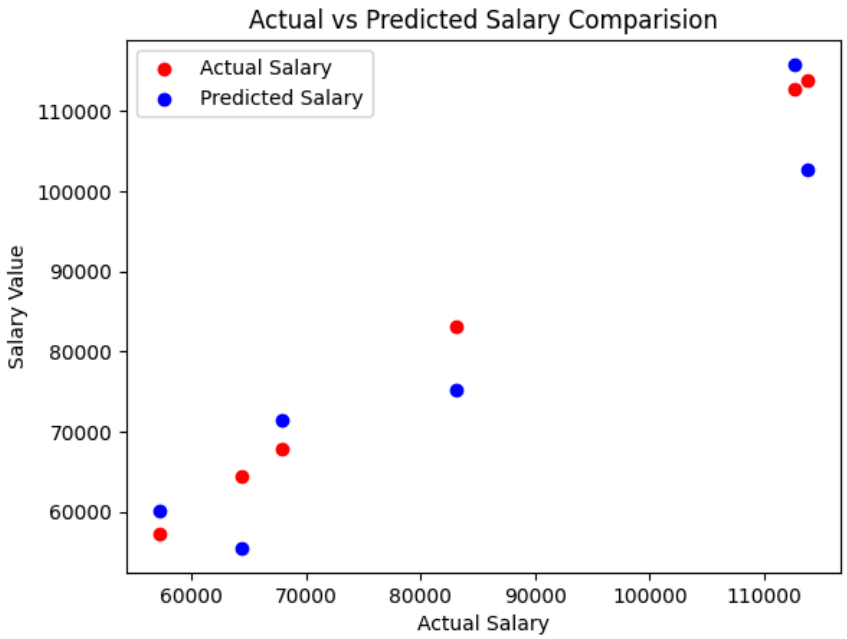
```
R2 = r2_score(y_test,y_pred)
print("R2 Score:",R2)
```

```
R2 Score: 0.9024461774180497
```

```
plt.scatter(x_test,y_test,color='red')
plt.plot(x_test,y_pred,color='blue')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.title('Salary vs Experience')
plt.show()
```


```
plt.scatter(y_test,y_test,color='red',label="Actual Salary")
plt.scatter(y_test,y_pred,color='blue',label="Predicted Salary")
plt.xlabel('Actual Salary')
plt.ylabel('Salary Value')
plt.title('Actual vs Predicted Salary Comparision')
```

```
plt.legend()
plt.show()
```




```
slope = model.coef_[0]
intercept = model.intercept_
print("Slope:(m)",slope)
print("Intercept(c):",intercept)
```

```
Slope:(m) [9423.81532303]
Intercept(c): [24380.20147947]
```

```
experience=[[5]]
```

```
predicted_salary = model.predict(experience)
print("Predicted Salary:",predicted_salary[0])
```

```
Predicted Salary: [71499.27809463]
```

```
Start coding or generate with AI.
```