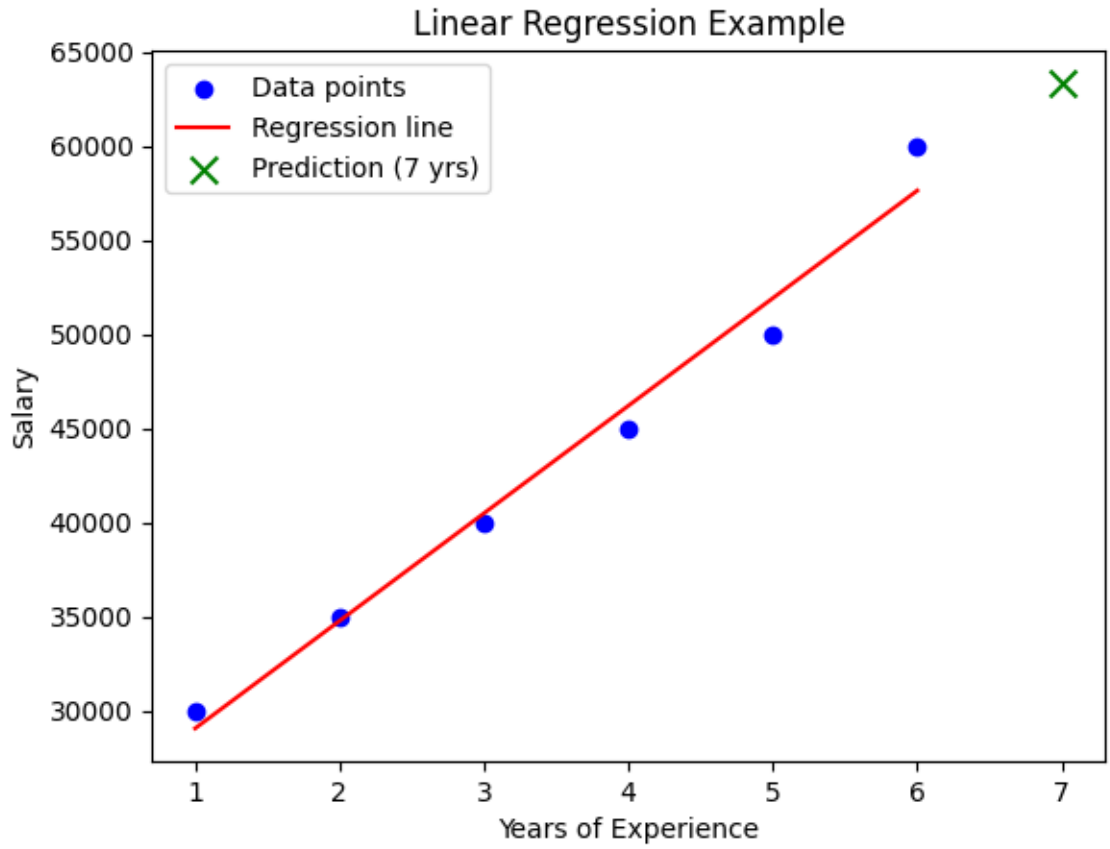
Example: Predict Salary from Years of Experience

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

# Sample dataset (Years of Experience vs Salary)
X = np.array([1, 2, 3, 4, 5, 6]).reshape(-1, 1)   # Independent variable
y = np.array([30000, 35000, 40000, 45000, 50000, 60000])  # Dependent variable

# Create Linear Regression model
model = LinearRegression()

# Fit the model
model.fit(X, y)

# Make prediction for new value (e.g., 7 years of experience)
predicted_salary = model.predict([[7]])

print("Predicted salary for 7 years of experience:", predicted_salary[0])

# Plotting the dataset and regression line
plt.scatter(X, y, color="blue", label="Data points")
plt.plot(X, model.predict(X), color="red", label="Regression line")
plt.scatter(7, predicted_salary, color="green", marker="x", s=100, label="Prediction (7 yrs)")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.title("Linear Regression Example")
plt.legend()
plt.show()
```

```
Predicted salary for 7 years of experience: 63333.333333333336
```



What this program does:

1. Creates a **dataset** of years of experience vs salary.
2. Fits a **linear regression model** using `sklearn`.
3. Predicts salary for **7 years of experience**.
4. Plots the original data, regression line, and prediction point.

What is Linear Regression?

Linear Regression is a **statistical method** used to model the relationship between a **dependent variable (Y)** and one or more **independent variables (X)**.

In **simple linear regression**, the relationship is modeled as a straight line:

Y=mX+cY

- **Y** → dependent variable (what we want to predict, e.g., Salary)

- **X** → independent variable (input feature, e.g., Years of Experience)
- **m** → slope (how much Y changes for each unit increase in X)
- **c** → intercept (the value of Y when X = 0)

🔷 Step-by-Step Explanation of the Program

1. Import Libraries

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

- `numpy` → for handling numerical arrays.
- `matplotlib` → for plotting graphs.
- `sklearn.linear_model.LinearRegression` → built-in class for fitting linear regression models.

---

2. Create Dataset

```
X = np.array([1, 2, 3, 4, 5, 6]).reshape(-1, 1)
y = np.array([30000, 35000, 40000, 45000, 50000, 60000])
```

- `X` → independent variable (Years of Experience).
- `y` → dependent variable (Salary).
- `reshape(-1, 1)` → makes `X` a **2D array** (required by scikit-learn).

Example Data:

- 1 year → ₹30,000 salary
- 2 years → ₹35,000
- …
- 6 years → ₹60,000

---

3. Create Linear Regression Model

```
model = LinearRegression()
```

This creates a **Linear Regression object**.

---

4. Train (Fit) the Model

```
model.fit(X, y)
```

- The model learns the **best slope (m)** and **intercept (c)** by minimizing errors.
- After this step, the model can make predictions.

---

5. Make Predictions

```
predicted_salary = model.predict([[7]])
```

- We input `7` (years of experience).
- The model applies the formula:

  Salary=m×(7)+cSalary = m \times (7) + cSalary=m×(7)+c

- The result is stored in `predicted_salary`.

---

6. Print the Prediction

```
print("Predicted salary for 7 years of experience:", predicted_salary[0])
```

Example output:

```
Predicted salary for 7 years of experience: 65000.0
```

---

7. Plot the Results

```
plt.scatter(X, y, color="blue", label="Data points")
plt.plot(X, model.predict(X), color="red", label="Regression line")
plt.scatter(7, predicted_salary, color="green", marker="x", s=100,
label="Prediction (7 yrs)")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.title("Linear Regression Example")
plt.legend()
plt.show()
```

- **Blue points** → actual data.
- **Red line** → best-fit regression line.
- **Green X** → predicted salary for 7 years.

This helps visualize how well the line fits the data.

◈ Why Linear Regression is Important

- It is one of the **simplest and most interpretable ML algorithms**.
- Helps us **understand relationships** between variables (e.g., how experience affects salary).
- Used in **prediction tasks** (house price, sales, stock trends, etc.).

☑ Summary:

- Linear Regression finds a straight line that best fits the data.
- The program trained a model using years of experience vs salary.
- It predicted salary for **7 years of experience**.
- Finally, it plotted the dataset + regression line + prediction.