Sant Dynaneshwar Shikshan Sanstha's

ANNASAHEB DANGE COLLEGE OF ENGINEERING AND TECHNOLOGY, ASHTA

Department of Computer Science and Engineering

(Internet of Things and Cyber Security including Block Chain Technology)

Machine learning (1ICPC406)

Lab Manual

Class : Final Year B.Tech

Semester : VII sem

Academic Year : 2025 - 2026

Prepared By –

Prof. Madhuri J Siddharapu

Experiment No. 1

Expt No. 1	Implementation of Simple Linear Regression.
Date :	

Aim:

To implement and understand the working of Simple Linear Regression, a basic predictive modeling technique in Machine Learning.

System Requirements:

- Operating System: Windows 8 or above / Linux / macOS
- Memory (RAM): Minimum 4 GB
- Processor: Minimum 2.33 GHz (Dual Core or higher)

Software/Tools Required:

Jupyter Notebook / Anaconda Navigator / Google Colaboratory / Spyder
Python 3.x with the following libraries:

- NumPy
- Pandas
- Matplotlib
- Scikit-learn

Purpose of the Experiment:

- To understand and implement Simple Linear Regression, one of the most basic yet powerful techniques in supervised learning.
- To explore how a model can learn from data and make predictions based on linear relationships.

- To visualize how the regression line fits the data and observe the accuracy of predictions.
- To build foundational knowledge required for more advanced machine learning models.

Expected Outcomes:

- Understand the working of linear regression algorithm.
- Learn how to fit a regression line using scikit-learn.
- Visualize the relationship between input and output variables.
- Predict output (marks) for new input (hours of study).
- Evaluate the regression model using visual and numerical results.

Theory:

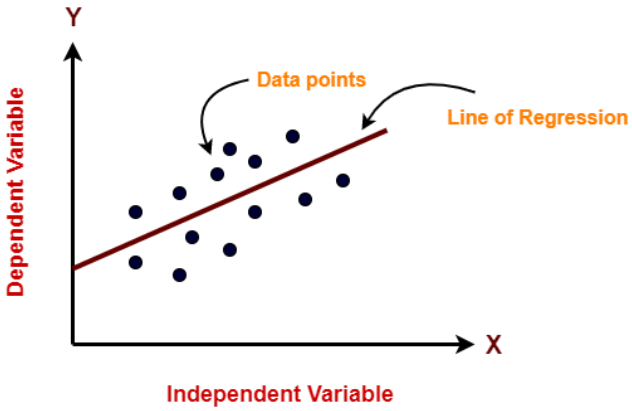
Linear regression is the most commonly used regression model in machine learning. It may be defined as the statistical model that analyzes the linear relationship between a dependent variable with a given set of independent variables. A linear relationship between variables means that when the value of one or more independent variables changes (increase or decrease), the value of the dependent variable will also change accordingly (increase or decrease).

Linear regression is further divided into two subcategories: simple linear regression and multiple linear regression (also known as multivariate linear regression).

In simple linear regression, a single independent variable (or predictor) is used to predict the dependent variable.

When we want to plot the relation between your dependent and independent variable, we need to perform regression analysis. Most importantly our dependent variable must be a numeric value. The relationship between dependent and independent variable in case of linear regression is of linear nature.

The meaning of the term "Linear" is whenever the independent variable(X) increases the dependent variable (Y) also increases.



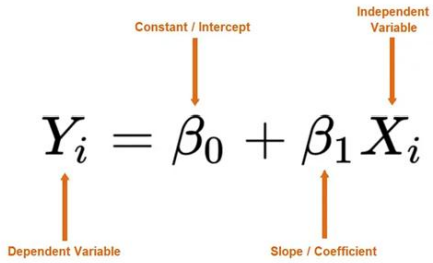
Mathematically, the simple linear regression can be represented as follows –

$$Y = mX + b$$

Where,

Y – represents dependent variable
X – represents independent variable
m – is the slope of line (how much Y changes for a unit change in X)
b – is the intercept (the value of Y when X is 0)

Simple linear regression can be simply represented by the following equation. You can also compare this with equation of a straight line Y = mX + b.



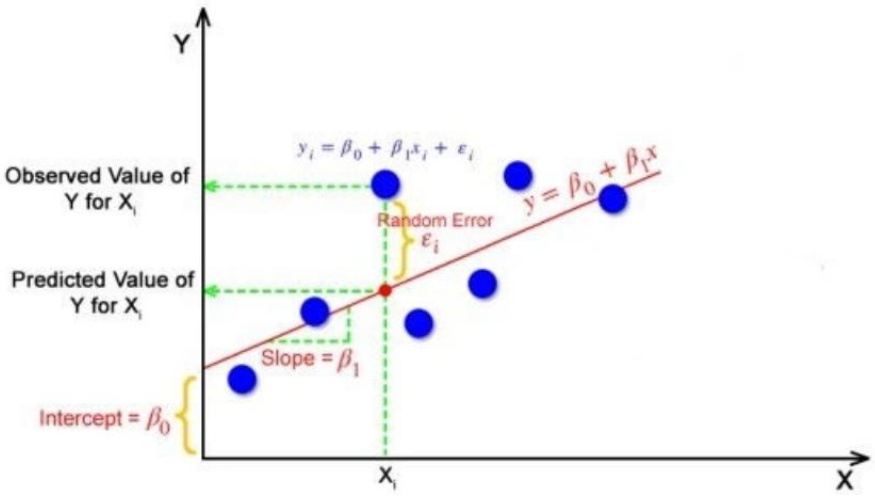
Simple Regression Calculation

To calculate best-fit line linear regression uses a traditional slope-intercept form which is given below,

Y i = β 0 + β 1 X i

- where Y i = Dependent variable, β 0 = constant/Intercept, β 1 = Slope/Intercept, X i = Independent variable.
- This algorithm explains the linear relationship between the dependent(output) variable y and the independent(predictor) variable X using a straight line Y= B 0 + B 1 X.

The goal of the linear regression algorithm is to get the best values for B 0 and B 1 to find the best-fit line. The best-fit line is a line that has the least error which means the error between predicted values and actual values should be minimum.



Real-World Applications of Linear Regression

Linear Regression helps to make data-driven decisions and reliable forecasts, some applications of linear regression are:

1. In real estate, the algorithm can help you predict house prices based on the size of the house, number of bedrooms, location, and access to other facilities.
2. In marketing, we can use linear regression to predict the sales volume based on seasonal shifts, past purchasing trends advertising costs, etc.

Procedure to implement Simple Linear Regression –

Step1: Import required libraries

Step 2: Define and load the dataset

Step 3: Split input (X) and output (Y)

Step 4: Fit the Linear Regression model

Step 5: Obtain slope, intercept, and regression line equation

Step 6: Visualize actual and predicted values

Step 7: Make prediction for a new value (e.g., 3 hours)

Procedure with Sample Code:

Step 1: Import necessary libraries

 # These libraries help in data manipulation, visualization, and model building

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

Step 2: Define and load the dataset

 # Creates a DataFrame with Hours studied and corresponding Marks scored

```
data = {'Hours': [1, 2, 4, 5, 6],
      'Marks': [50, 60, 80, 90, 100]}
df = pd.DataFrame(data)
print(df)
```

Step 3: Separate input and output variables

```
X = df[['Hours']].values                    # Independent variable as 2D arrayY
Y = df[['Marks']].values                    # Dependent variable as 2D array
```

Step 4: Train the Linear Regression model

```
model = LinearRegression()                  # Create model object
model.fit(X, Y)                             # Train the model using the data
```

Step 5: Extract slope, intercept, and form regression equation

 # Model learns coefficients and outputs the line equation

```
slope = model.coef_[0][0]
intercept = model.intercept_[0]
print(f"Equation of line: y = {slope:.2f}X + {intercept:.2f}")
```

Step 6: Predict values and visualize

 # This plots actual vs predicted values with a red line showing the model's prediction

```
df['Predicted Marks'] = model.predict(X)  # Add predictions to DataFrame
plt.scatter(df['Hours'], df['Marks'], color='blue', label='Actual Marks')
plt.plot(df['Hours'], df['Predicted Marks'], color='red', label='Regression Line')
plt.xlabel('Hours Studied')
plt.ylabel('Marks Scored')
plt.title('Linear Regression: Hours vs Marks')
plt.legend()
plt.show()
```

Step 7: Make prediction for new data

 # The model predicts score for unseen input (3 hours of study)

```
hours_to_predict = np.array([[3]])
predicted_marks = model.predict(hours_to_predict)[0][0]
print(f"Predicted marks for 3 hours of study: {predicted_marks}")
```

Observation

I. Record the following in observation table:

1. Regression Equation: y = mX + b
2. Value of slope (m): _____
3. Value of intercept (b): _____
4. Predicted Marks for 7.5 hours: _____
5. Show a screenshot or print of the plotted regression line

II. Perform the following tasks and record results by applying Simple Linear Regression. Use the datasets given for each task.

Task 1: Salary Prediction Based on Experience					
Years of Experience	1	2	3	4	5
Salary (₹ in 1000s)	30	35	45	55	60

Answer the Following:

1. Write the equation of the regression line derived from this data.
2. Predict the salary for 6 years of experience.
3. Plot the regression line and interpret the slope.

Task 2: House Price Prediction Based on Area					
House Area (sq.ft)	500	1000	1500	2000	2500
Price (₹ in Lakhs)	30	50	65	85	100

Answer the Following

1. Find the best-fit line and state the intercept and slope.
2. Predict the price of a house with 1800 sq.ft.

Task 3: Calories Burned vs Workout Duration					
Workout Duration (min)	10	20	30	40	50
Calories Burned	60	120	180	240	300

Answer the Following

1. Derive and write the regression equation.
2. Estimate the calories burned for a 35-minute workout.
3. What does the slope of the line represent in this context?

Conclusion :

Hence, the experiment provides students with hands-on experience in implementing and analyzing a Simple Linear Regression model, a foundational technique in Machine Learning. Through real-life datasets and visualization, students gain a practical understanding of how to build predictive models, interpret regression coefficients, and evaluate model performance..

Viva Questions

Sr.No	Question	CO
1.	What is Simple Linear Regression? How is it different from Multiple Linear Regression?	**1ICPC406_1**
2.	Define the terms slope and intercept in a regression equation. What is their significance?	**1ICPC406_1**
3.	What is the mathematical equation of a linear regression line?	**1ICPC406_1**

4.	Which method is used to fit the best line in Linear Regression?	**1ICPC406_1**
5.	What do you understand by the term 'model training' in Machine Learning?	**1ICPC406_1**
6.	What is the role of the fit() function in scikit-learn's Linear Regression?	**1ICPC406_1**
7.	Why do we split data into training and testing sets in supervised learning?	**1ICPC406_1**
8.	How does the model make predictions once it is trained?	**1ICPC406_1**
9.	Which Python libraries are commonly used for implementing regression models?	**1ICPC406_1**
10.	What are the limitations of Simple Linear Regression?	**1ICPC406_1**
11.	How can we measure the performance of a regression model?	**1ICPC406_1**
12.	What kind of relationship between variables does Simple Linear Regression assume?	**1ICPC406_1**
13.	Explain overfitting and underfitting in the context of regression?	**1ICPC406_1**
14.	How would the model behave if the dataset contains outliers?	**1ICPC406_1**
15.	Explain the use of matplotlib.pyplot in visualizing regression results.	**1ICPC406_1**

References –

a. Textbook –

i. Machine Learning with Python – An approach to Applied ML – Abhishek Vijayvargiya, BPB Publications, 1st Edition 2018
ii. Machine Learning, Tom Mitchell, McGraw Hill Education, 1st Edition 1997

b. Online references –

i. https://www.analyticsvidhya.com/blog/2021/10/everything-you-need-to-knowabout-linear-regression/
ii. https://www.javatpoint.com/simple-linear-regression-in-machine-learning
iii. https://www.24tutorials.com/machine-learning/simple-linear-regression/