Terna Engineering College

Department of Artificial Intelligence and Data Science

Program : Sem VI

Course: Machine Learning Lab

Experiment No.04

PART A

(PART A: TO BE REFERRED BY STUDENTS)

A.1 Aim: To implement Multiple Linear Regression using Python.

A.2 Theory:

Regression:

Regression in Machine Learning is a supervised learning technique. There are various types of regression like linear, logistic, polynomial, stepwise, ridge, lasso, etc.

Linear Regression:

Linear regression helps us to predict the relationship between two variables by assuming a linear connection between the independent and dependent variables.

It is one of the very simple and easy algorithms that works on regression and shows the relationship between the continuous variables.

It shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis).

There are two types of linear regression- Simple and Multiple.

Multiple Linear Regression:

Multiple linear regression is an approach for predicting a response using a multiple features. It attempts to model the relationship between two or more features and a response by fitting a linear equation to observed data. The steps to perform multiple linear Regression are almost similar to that of simple linear Regression.

The formula for a simple linear regression is:

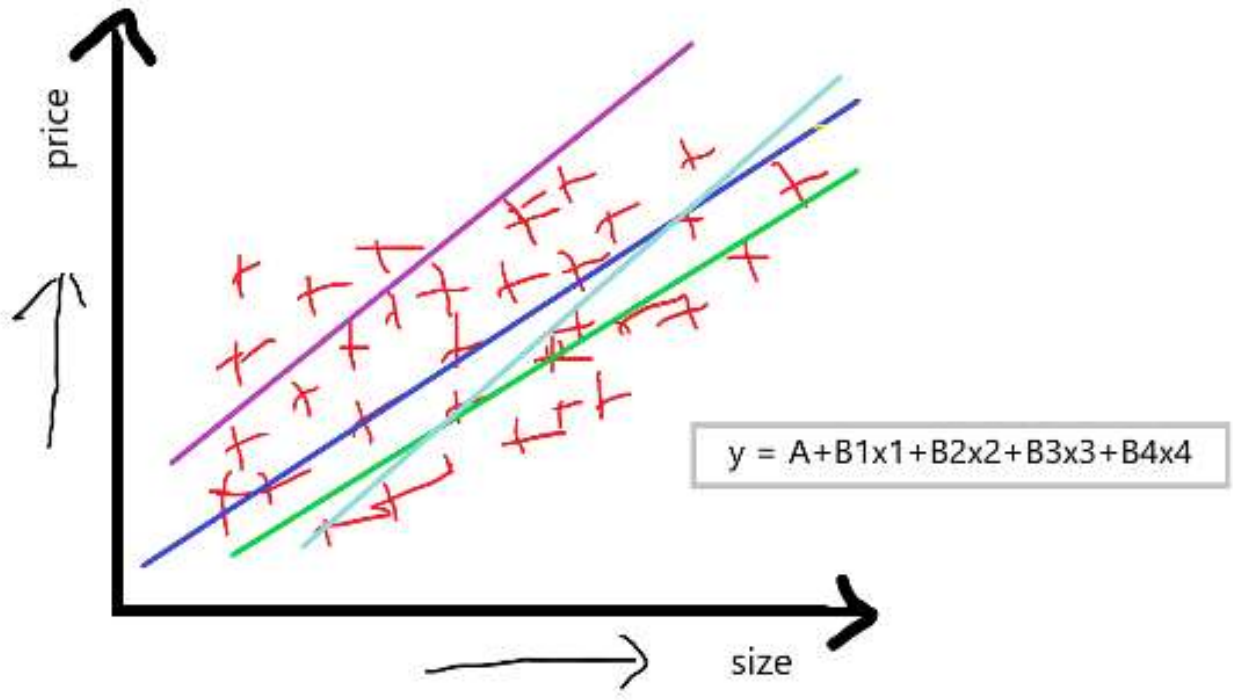
$$Y = b0 + b1 * x1 + b2 * x2 + b3 * x3 + \ldots\ldots bn * xn$$

Where Y is the dependent variable and x1, x2, x3, …… xn are multiple independent variables.

Assumptions for Multiple Linear Regression:

- A linear relationship should exist between the Target and predictor variables.
- The regression residuals must be normally distributed.
- MLR assumes little or no multicollinearity (correlation between the independent variable) in data.

e.g. In the given graph, we can see a sample model for predicting house prices where we have multiple independent variables (x_i) and price is the dependent variable(y).



Our aim in using the multiple linear regression is that we have to compute A, which is an intercept. The key parameters B1, B2, B3, and B4 are the slopes or coefficients concerning this independent feature. This indicates that if we increase the value of x1 by 1 unit, then B1 will tell you how much it will affect the price of the house. The others B2, B3, and B4, also work similarly.

PART B

(PART B: TO BE COMPLETED BY STUDENTS)

(Students must submit the soft copy as per following segments within two hours of the practical. The soft copy must be uploaded on the Blackboard or emailed to the concerned lab in charge faculties at the end of the practical in case the there is no Black board access available)

Roll. No. A56	Name: Shubham Mangaonkar
Class: TE – AI & DS	Batch: A3
Date of Experiment: 08/02/24	Date of Submission: 23/02/24
Grade:	

B.1 Input and Output:

Code :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

dataset = pd.read_csv("D:\College\SEM 6\Machine
Learning\Advertising.csv")

x = dataset[['TV', 'Radio', 'Newspaper']]
y = dataset['Sales']

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size =
0.3, random_state = 100)

mlr = LinearRegression()
mlr.fit(x_train, y_train)
```

```
print("Intercept: ", mlr.intercept_)
print("Coefficients:")
print(list(zip(x, mlr.coef_)))

y_pred_mlr= mlr.predict(x_test)

mlr_diff = pd.DataFrame({'Actual value': y_test, 'Predicted value':
y_pred_mlr})
print('\n')
print(mlr_diff.head())
```

Output:

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  SEARCH ERROR

PS D:\> python -u "d:\College\SEM 6\Machine Learning\MLR.py"
Intercept:  2.652789668879498
Coefficients:
[('TV', 0.04542559602399794), ('Radio', 0.18975772766893614), ('Newspaper', 0.004603078953112072)]


     Actual value  Predicted value
126           6.6        10.621601
104          20.7        20.006253
99           17.2        16.918509
92           19.4        19.170407
111          21.8        20.949741
PS D:\>
```

B.2 Conclusion:

Thus we have successfully implemented multiple linear regression in Python and understood how the model is fitted to the given data.