



Vidyavardhini's College of Engineering and Technology

Department of Artificial Intelligence & Data Science



Vidyavardhini's College of Engineering and Technology

Department of Artificial Intelligence & Data Science

BARI ANKIT (56)

Experiment No.2
Implement R Program for Simple Linear Regression
Date of Performance:
Date of Submission:



EXPERIMENT NO 2

Aim: To write the implementation of linear regression.

Objective:- To understand the use of simple linear regression techniques by implementing user define dataset and importing dataset

Description:

Regression analysis is a very widely used statistical tool to establish a relationship model between two variables. One of these variables is called a predictor variable whose value is gathered through experiments. The other variable is called response variable whose value is derived from the predictor variable.

In Linear Regression these two variables are related through an equation, where the exponent (power) of both these variables is 1. Mathematically a linear relationship represents a straight line when plotted as a graph. A non-linear relationship where the exponent of any variable is not equal to 1 creates a curve.

The general mathematical equation for a linear regression is

$$y = ax + b$$



Following is the description of the parameters used -

y is the
response
variable. x is
the
predictor
variable.

a and b are constants which are called the coefficients.

Procedure:

The steps to create the relationship is

1. Carry out the experiment of gathering a sample of observed values of height and corresponding weights.
2. Create a relationship model using the **lm()** functions in R.
3. Find the coefficients from the model created and create the mathematical equation using these. Get a summary of the relationship model to know the average error in prediction. Also called **residuals**.

To predict the weight of new persons, use the **predict()** function in R.

Program:

```
# Import required libraries

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

# Sample data (for demonstration purposes)

data = {
```



```
'X': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],  
  
'y': [2, 3, 5, 7, 11, 13, 17, 19, 23, 29]  
  
}  
  
df = pd.DataFrame(data)  
  
  
# Features and target variable  
  
X = df[['X']]  
  
y = df['y']  
  
  
# Splitting the dataset into training and testing sets  
  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)  
  
  
# Creating the model  
  
model = LinearRegression()  
  
  
# Fitting the model  
  
model.fit(X_train, y_train)  
  
  
# Making predictions  
  
y_pred = model.predict(X_test)  
  
  
# Plotting the results  
  
plt.scatter(X, y, color='blue', label='Original data')  
  
plt.plot(X_test, y_pred, color='red', linewidth=2, label='Linear Regression line')  
  
plt.title('Linear Regression Example')  
  
plt.xlabel('X')  
  
plt.ylabel('y')
```



```
plt.legend()
```

```
plt.show()
```

```
# Output the model coefficients
```

```
print(f'Coefficient: {model.coef_[0]}')
```

```
print(f'Intercept: {model.intercept_}')
```

OUTPUT:

Coefficient: 2.8181818181818183

Intercept: -1.4545454545454542

Conclusion:

The implementation of linear regression using Python showcases the effectiveness of using libraries such as Scikit-learn for predictive modeling. By fitting a linear model to the data, we are able to identify the relationship between the independent variable XXX and the dependent variable yyy. The coefficient and intercept obtained provide insights into how changes in XXX influence predictions of yyy.



Linear regression serves as a foundational technique in data analytics, applicable in numerous fields such as business, economics, and social sciences. Its simplicity and interpretability make it an excellent starting point for those delving into the world of data science and predictive modeling. Despite its limitations, such as the assumption of linearity and sensitivity to outliers, linear regression remains a vital tool for understanding relationships in data and guiding decision-making.