**Name:** Samruddhi Pandit

**Roll no:** 281014

**Batch:** A1

**Assignment 6**

**Statement:**

**Problem Statement:** Download the temperatures dataset from following link:

<https://www.kaggle.com/datasets/venky73/temperatures-of-india>.

This data consists of temperatures of INDIA averaging the temperatures of all places month-wise.

Temperature values are recorded in CELSIUS.

a) Apply Linear Regression using a suitable library function and predict the Month-wise temperature.

b) Assess the performance of regression models using MSE, MAE and R-Square metrics

c) Visualize a simple regression model.

**Objective:**

This assignment will assist us in understanding the applications of linear regression and how predictions can be made using it.

**Resources Used:**

1. Software: Python 3.x, Google Colab
2. Libraries: NumPy, Matplotlib, Scikit-learn

**Introduction to Linear Regression:**

1. Linear regression is a statistical technique used for forecasting analysis.
2. Predictions are made using linear regression for continuous numerical variables such as sales, earnings, age, and product price.
3. It models a linear relationship between one or more independent variables (X) and a dependent variable (Y).

**Types of Linear Regression:**

1. Simple Linear Regression: Uses one independent variable to predict the dependent variable.
2. Multiple Linear Regression: Uses multiple independent variables to predict the dependent variable.

**Applications of Simple Linear Regression:**

1. Student Grades Prediction: Exam scores depend on the number of study hours (independent variable).
2. Agricultural Yield Estimation: Yield depends on rainfall data.
3. Salary Prediction: Experience determines salary.

**Limitations of Simple Linear Regression:**

1. Assumes Linearity: The model assumes a straight-line relationship, which may not always hold.
2. Sensitive to Outliers: Outliers can significantly impact model performance.
3. Doesn't Establish Causation: Correlation does not imply causation.

**Methodology:**

1. Import Necessary Libraries: Utilize pre-built functionalities to streamline the process.
2. Data Collection: Gather relevant data for the model.
3. Data Preprocessing and EDA: Clean and analyze data before modeling.
4. Splitting the Data: Divide the dataset into training and testing subsets.
5. Model Training: Fit a linear regression model using training data.
6. Making Predictions: Use the trained model to predict dependent variable values.
7. Model Evaluation: Assess performance using metrics like Mean Squared Error (MSE) and R-squared.

**Results:**

|  |  |
| --- | --- |
|  |  |

**Conclusion:**

A simple linear regression model establishes a linear relationship between one independent variable and a dependent variable. It is widely used in predictive analytics to understand trends and make informed decisions in various domains, such as education, agriculture, and finance.