****

**BRACT’s**  
**Vishwakarma Institute of Information Technology  
Department of CSE-AI**

(An Autonomous Institute affiliated to Savitribai Phule Pune University)



# Lab Manual

Data Science and Machine Learning

**Submitted by,**

**Vaibhav Vinod Chavhan**

SY

Semester II Academic Year 2023-24

**ASSIGNMENT NO: 6**

**Problem Statement-**

This data consists of temperatures of INDIA averaging the temperatures of all places month wise. Temperatures values are recorded in CELSIUS.

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

temperature.

b) Assess the performance of regression models

**S/W Packages and Libraries used-**

Software Package: Python

Libraries Used:

pandas: For data manipulation and analysis.

scikit-learn: For implementing machine learning algorithms, including Linear Regression.

matplotlib: For data visualization.

**Theory-**

* Data Preprocessing:
  + Load the temperature data.
  + Prepare the data by separating features (months) and target variable (monthly temperatures).
  + Split the data into training and testing sets.
* Model Training:
  + Initialize a Linear Regression model.
  + Fit the model on the training data.
* Model Evaluation:
  + Predict the month-wise temperatures using the trained model.
  + Evaluate the model's performance using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-Square metrics.

**Application:**

* The application of this regression technique is to predict the month-wise temperatures in India based on historical data.
* It can be used by meteorologists, climatologists, and environmental scientists for various purposes such as weather forecasting, climate trend analysis, and planning for agriculture and infrastructure.

**Limitations:**

* Linear Regression assumes a linear relationship between the independent and dependent variables, which may not always hold true in real-world scenarios.
* It may not capture complex patterns or non-linear relationships present in the data.
* The accuracy of predictions can be affected by outliers and noise in the data.
* The model's performance heavily depends on the quality and representativeness of the training data.

**Working:**

A white surface with a white line

Description automatically generated with medium confidence

A calendar with numbers and a few days

Description automatically generated with medium confidence

A table with numbers and letters

Description automatically generated

A screenshot of a computer

Description automatically generated

A white background with black text

Description automatically generated

A screenshot of a computer

Description automatically generated

A graph with blue dots

Description automatically generated

A screenshot of a computer program

Description automatically generated

A white rectangular object with a black border

Description automatically generated

A graph with green dots and lines

Description automatically generated

**Conclusion:**

* Linear Regression can be effectively used to predict month-wise temperatures in India.
* The model's performance can be assessed using metrics like MSE, MAE, and R-Square.
* Despite its limitations, Linear Regression provides a simple and interpretable approach for temperature prediction, which can be valuable for various applications in climate science and related fields.