**Name: Nikhil Sanjay Shinde**

**Roll No: TYITB120**

**Assignment No. 2**

**AIM:** Assignment on Linear Regression

**PREREQUISITE:** Python programming

**THEORY:**

Linear Regression is a fundamental statistical method used in machine learning for predicting an output variable based on one or more input variables. In its simplest form, when only one independent variable is used, it is known as Simple Linear Regression. When multiple independent variables are involved, it becomes Multiple Linear Regression.

**Simple Linear Regression**

Simple Linear Regression aims to establish a linear relationship between an independent variable (input) and a dependent variable (output). The relationship can be represented by a straight-line equation where one variable influences the other.

In this model:

* The dependent variable is what we want to predict.
* The independent variable is what we use to make predictions.
* The intercept is the constant term that determines the starting point of the line.
* The slope defines the degree of impact the independent variable has on the dependent variable.

The objective of linear regression is to find the best-fitting line that minimizes the difference between the actual values and the predicted values.

**Estimating Coefficients**

The coefficients in the linear regression model are determined using statistical calculations based on the dataset. The slope is estimated by assessing how much the dependent variable changes with respect to the independent variable. Once the slope is obtained, the intercept can be computed easily using the average values of the dataset.

**Making Predictions**

After determining the coefficients, we can use the linear regression equation to predict values for new input data. By substituting different values into the equation, we can estimate the expected output values. A graphical representation helps visualize how well the model fits the data points.

**Assessing Model Performance**

The accuracy of a linear regression model is often evaluated using the Root Mean Squared Error (RMSE). RMSE measures the average magnitude of prediction errors and provides insight into how well the model generalizes to new data. A lower RMSE indicates better performance.

**Importance of Linear Regression**

Linear Regression is widely used in various fields such as economics, finance, healthcare, and engineering. It helps in forecasting trends, identifying relationships, and making data-driven decisions. The simplicity of the model makes it a great starting point for understanding machine learning concepts.

**CONCLUSION:**

Linear Regression is a fundamental technique in predictive modeling that establishes a relationship between variables. By understanding and applying this method, we can make informed predictions and analyze patterns in data. This method serves as a building block for more advanced machine learning algorithms and real-world applications.