Training Day-55 Report:

What is Regression?

Regression is a statistical approach used to analyze the relationship between a dependent variable (target variable) and one or more independent variables (predictor variables). The objective is to determine the most suitable function that characterizes the connection between these variables.

It seeks to find the best-fitting model, which can be utilized to make predictions or draw conclusions.

Regression in Machine Learning:-

It is a supervised machine learning technique, used to predict the value of the dependent variable for new, unseen data. It models the relationship between the input features and the target variable, allowing for the estimation or prediction of numerical values.

Regression analysis problem works with if output variable is a real or continuous value, such as "salary" or "weight". Many different models can be used, the simplest is the linear regression. It tries to fit data with the best hyper- plane which goes through the points.

Terminologies Related to the Regression Analysis in Machine Learning:-

Terminologies Related to Regression Analysis:

- **Response Variable**: The primary factor to predict or understand in regression, also known as the dependent variable or target variable.
- **Predictor Variable**: Factors influencing the response variable, used to predict its values; also called independent variables.
- **Outliers**: Observations with significantly low or high values compared to others, potentially impacting results and best avoided.
- **Multicollinearity**: High correlation among independent variables, which can complicate the ranking of influential variables.
- **Underfitting and Overfitting**: Overfitting occurs when an algorithm performs well on training but poorly on testing, while underfitting indicates poor performance on both datasets.

Regression Types

There are two main types of regression:

- **Simple Regression**
  - Used to predict a continuous dependent variable based on a single independent variable.
  - Simple linear regression should be used when there is only a single independent variable. □

- **Multiple Regression**

    o Used to predict a continuous dependent variable based on multiple independent variables.
    o Multiple linear regression should be used when there are multiple independent variables.

- **Non-Linear Regression**
    o Relationship between the dependent variable and independent variable(s) follows a nonlinear pattern.

    o Provides flexibility in modeling a wide range of functional forms.