TOPIC

MACHINE LEARNING (ML) ALGORITHM

* Q. Explain me about the Linear Regression and Multiple Linear Regression with suitable Examples.

Answer

1. Simple Linear Regression :

Linear Regression is a statistical method that models the relationship between two variables by fitting a linear equation to observed data. It is used to predict the value of a dependent variable (target) based on the value of an independent variable (predictor).

* Formula :-

$$y = \beta_0 + \beta_1 x + \epsilon$$

where,

* 'y' is the dependent variable (what we are trying to predict).
* 'x' is the independent variable (the predictor).
* 'β_0' is the y-intercept (the value of Y when $x = 0$).
* 'β_1' is the slope of the line (the change in y for a one-unit change in x).
* 'ϵ' is the error term (the difference between the observed and predicted values of y).

* Example of simple Linear Regression:

* Scenario:
predicting the height of a person based on their age.

* data:

* Age (years): [5, 7, 10, 12, 15]
* Height (cm): [110, 120, 130, 140, 150]

* steps:

(i) plot the data:
plot age on the x-axis and height on the y-axis.

(ii) Fit a line:
draw a line that best fits the data points.

(iii) Determine the Equation:
use statistical methods to determine the value of β_0 and β_1.

* Resulting Equation:

$$\text{Height} = 100 + 3 \times \text{Age}$$

This equation suggests that the height increases by 3 cm for each additional year of age, with a base height of 100 cm at age 0.

* 2. Multiple Linear Regression:

Multiple Linear Regression models the relationship between two or more independent variables and a dependent variable by fitting a linear equation to observed data.

* Formula:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n \beta x + \epsilon$$

where,

* 'y' is the dependent variable.
* $x_1, x_2, \ldots, x_n$ are the independent variables.
* β_0 is the y-intercept
* $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients for the independent variables.
* ϵ is the error term.

* Example of Multiple Linear Regression:

* Scenario:

predicting a house's price based on its size, number of bedrooms & age.

* Data:

* (i) Size (sq ft): [1400, 1600, 1700, 1875, 1100]
* (ii) Number of Bedrooms: [3, 3, 2, 4, 2]
* (iii) Age (years): [20, 15, 18, 10, 25]
* (iv) price (in $1000s): [245, 312, 279, 308, 199]

* steps:

(i) prepare the data:
organize the data into a table.

(ii) Fit a model:
Use statistical software to fit a multiple linear Regression model.

(iii) Determine the Equation:
Calculate the coefficients (β).

* Resulting Equation:

$$price = 50 + 0.1 \times size + 20 \times Bedrooms - 0.5 \times Age$$

* This equation suggests:

(i) The base price is $50,000

(ii) Each additional square foot adds $20,000 to the price.

(iii) Each additional year of age decreases the price by $500.

Summary

* Simple Linear Regression involves one target variable. It fits a straight line to the data.

* Multiple Linear Regression involves multiple predictors and one target variable. It fits a plane (or hyperplane) to the data.

Both methods are used to understand relationships between variables and make predictions.