**Business Intelligence Lab**

**Experiment 07**

**Aim:**

To implement regression model using Rapid Miner and Python

**Theory:**

**Regression:**

Regression analysis is a statistical method that helps us to analyze and understand the relationship between two or more variables of interest. The process that is adapted to perform regression analysis helps to understand which factors are important, which factors can be ignored, and how they are influencing each other.

**Linear Regression**

The simplest of all regression types is Linear Regression where it tries to establish relationships between Independent and Dependent variables. The Dependent variable considered here is always a continuous variable.

Linear Regression is a predictive model used for finding the linear relationship between a dependent variable and one or more independent variables.

Hypothesis function for Linear Regression :



While training the model we are given :

x: input training data (univariate – one input variable(parameter))

y: labels to data (supervised learning)

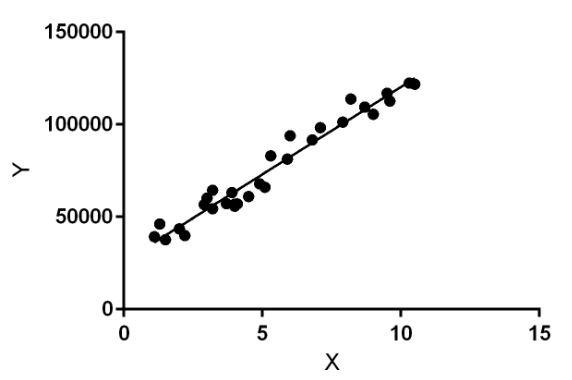
When training the model – it fits the best line to predict the value of y for a given value of x. The model gets the best regression fit line by finding the best θ1 and θ2 values.

θ1: intercept

θ2: coefficient of x

Once we find the best θ1 and θ2 values, we get the best fit line. So when we are finally using our model for prediction, it will predict the value of y for the input value of x.

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting.



**Implementing regression using Python libraries**