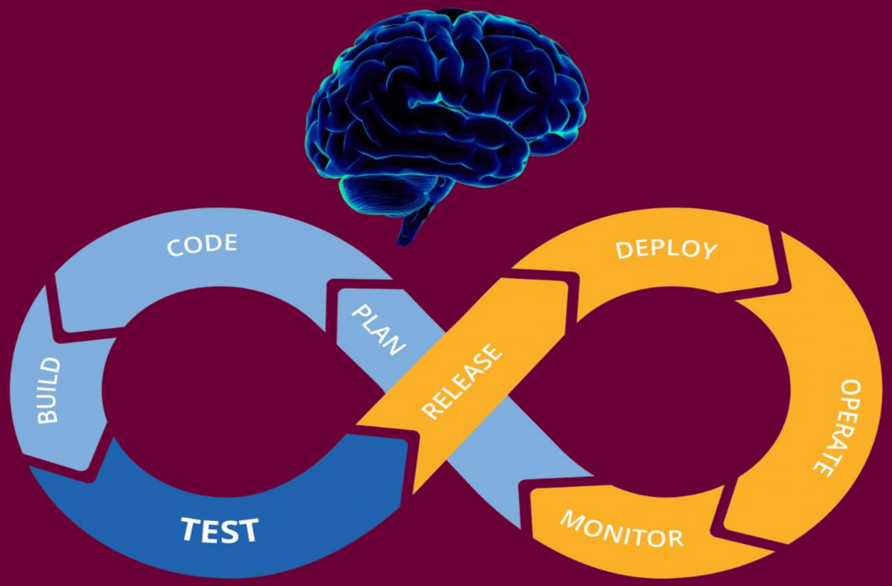
CSED!

"A Place to Invent and Learn"



Machine Learning Course

Forecasting and Learning Theory

Topics

- Forecasting
- Linear Regression
- Regression model using Gradient Descendent.
- Project: Pridicting salary of an employee
- Project: Aircraft Fuel Consumption Prediction

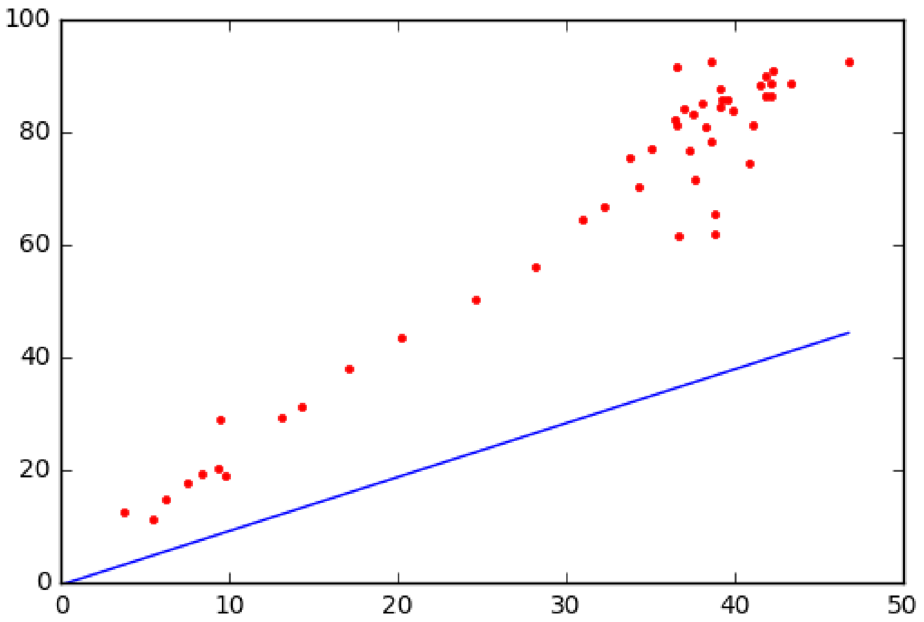
Forecasting and Learning Theory

Forecasting is the process of making predictions about future events based on historical data and analysis. It is widely used in various fields such as finance, economics, supply chain management, and weather prediction.

Application of Forecasting and Learning Theory

- Estimating economic indicators such as GDP growth.
- Predicting disease progression or patient outcomes based on patient data.
- Analyzing the impact of advertising expenditure on sales.
- Forecasting demand for inventory management.
- Modeling pollution levels based on various factors.
- Predict students' academic performance based on their previous scores, attendance, and other factors.
- Forecast the success rate of recruitment drives based on historical data Etc.

What is Linear Regression

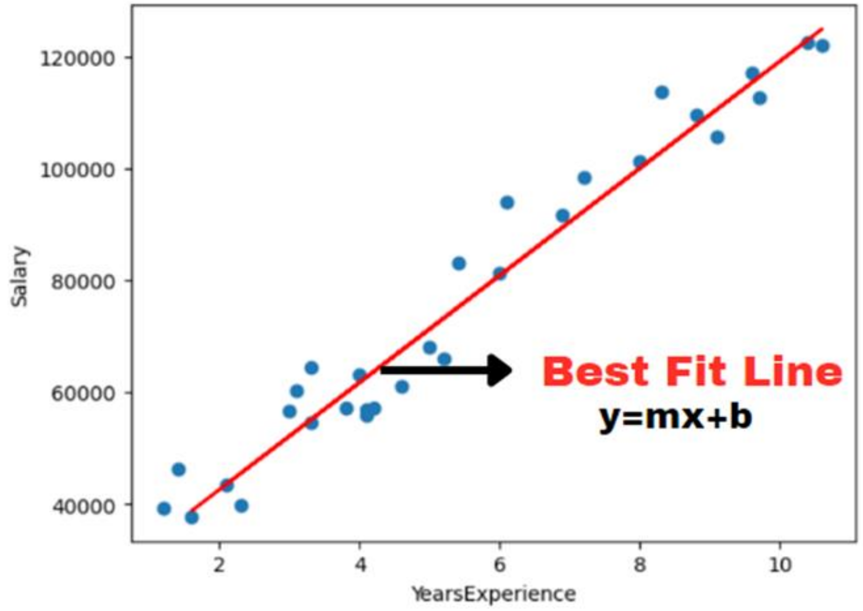


Linear Regression is a supervised learning algorithm in machine learning, which is widely used for solving regression problems. Regression is a type of machine learning problem where the goal is to predict a continuous output variable based on one or more input variables.

In Linear Regression, the goal is to find the best-fitting linear equation to describe the relationship between the input variables (also known as predictors or features) and the output variable (also known as the response variable).

Simple Linear Regression

Simple linear regression models is the relationship between two variables by fitting a linear equation to observed data. The linear equation can be written as:



y = mx + b

Where:

•y is the dependent variable.

•x is the independent variable.

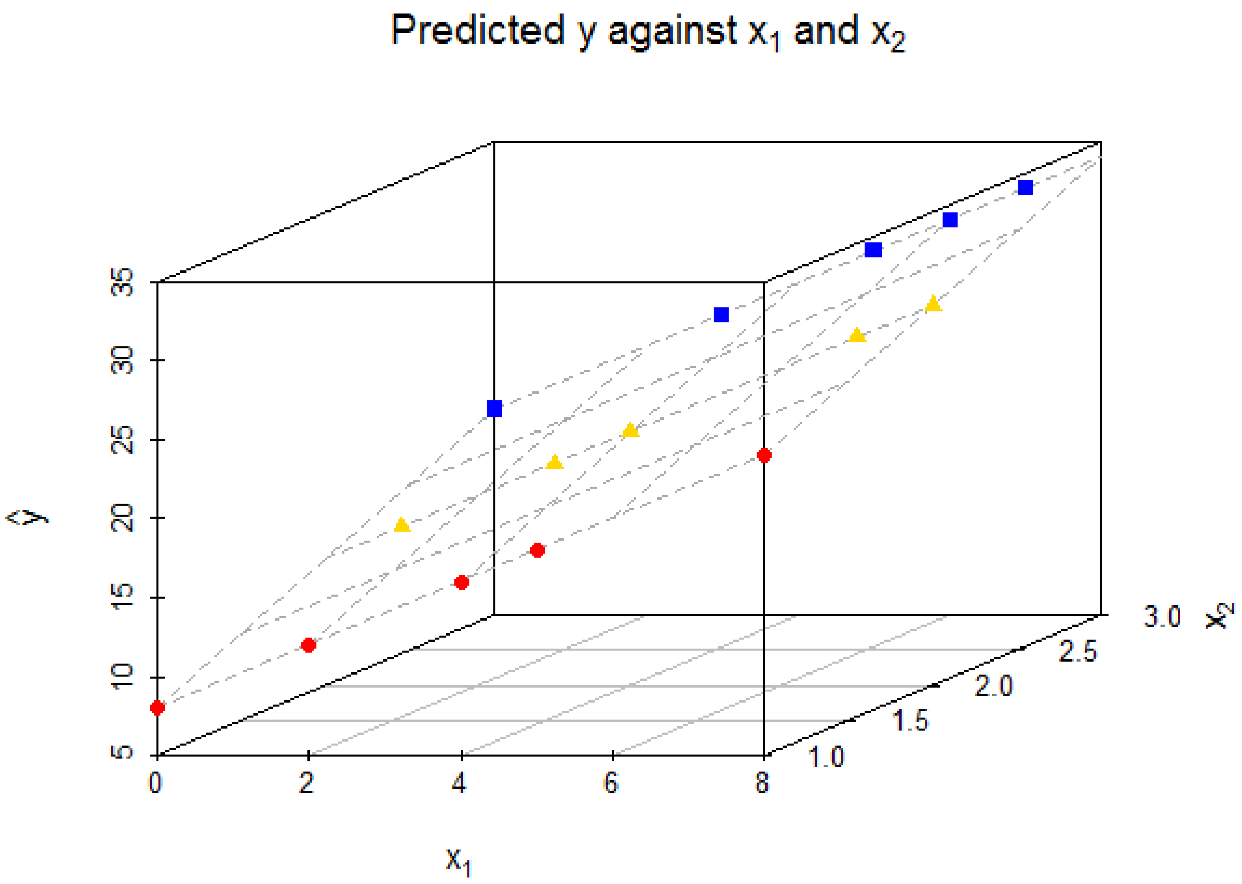
•b is the y-intercept.

•m is the slope of the regression line.

The goal is to find the best-fitting line through the data points that minimizes the sum of squared residuals (errors). This method is called Ordinary Least Squares (OLS)

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 the observed data.



Multiple Linear Regression

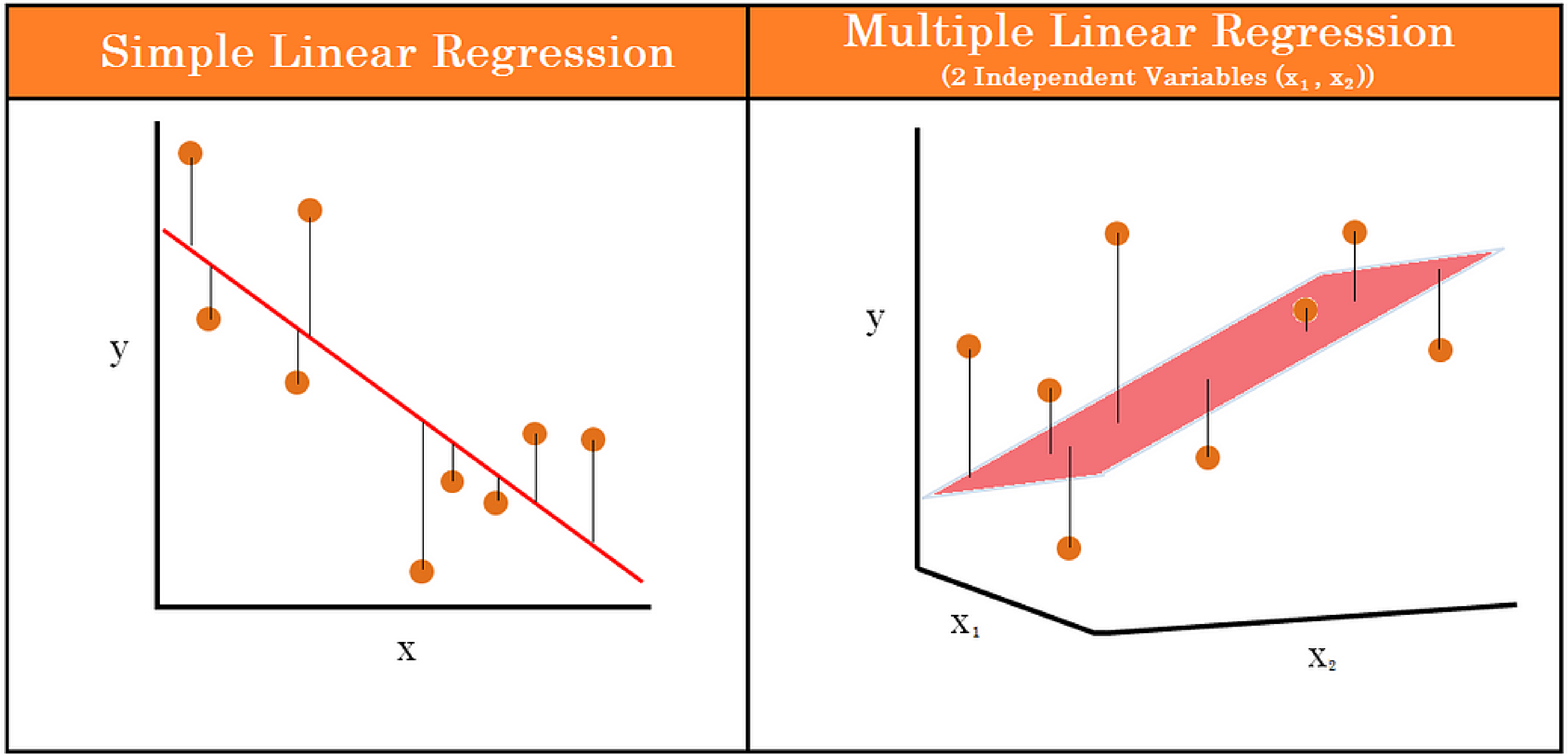
y=b+m1x1+m2x2+……..+mn*xn

Where:

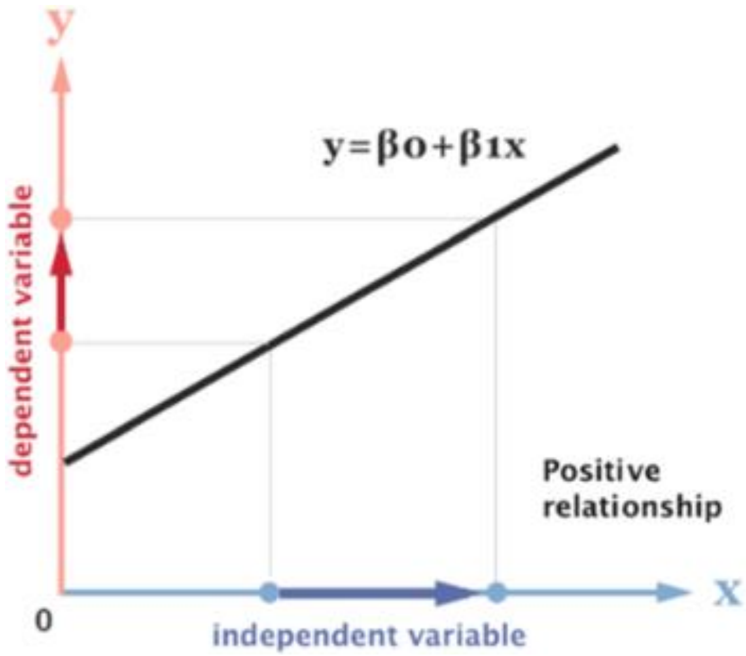
- y is the dependent variable.
- x1,x2,…,…,xn are the independent variables.
- b is the y-intercept.
- m1,m2,………mn are the coefficients.

The goal is to find the best-fitting plane (or hyperplane) that minimizes the sum of squared residuals.

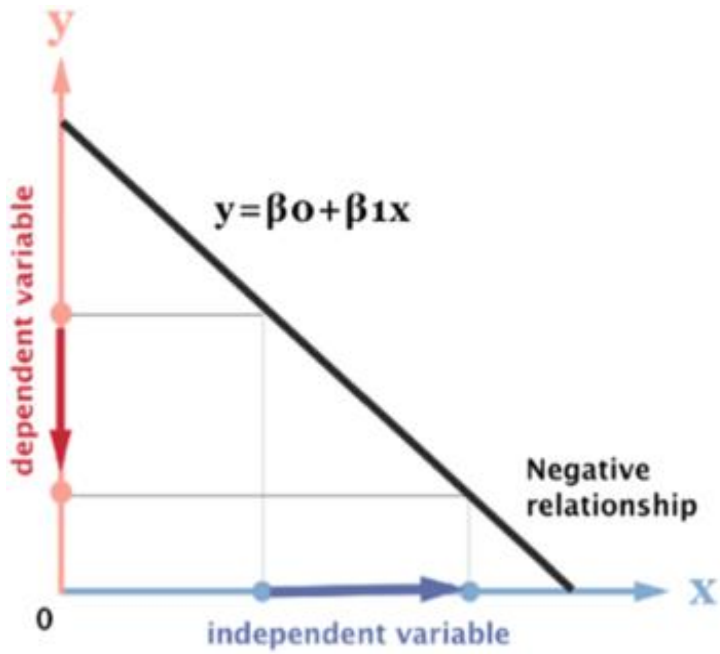
Single vs Multiple Linear Regression



Relationship between Input and Output Variable



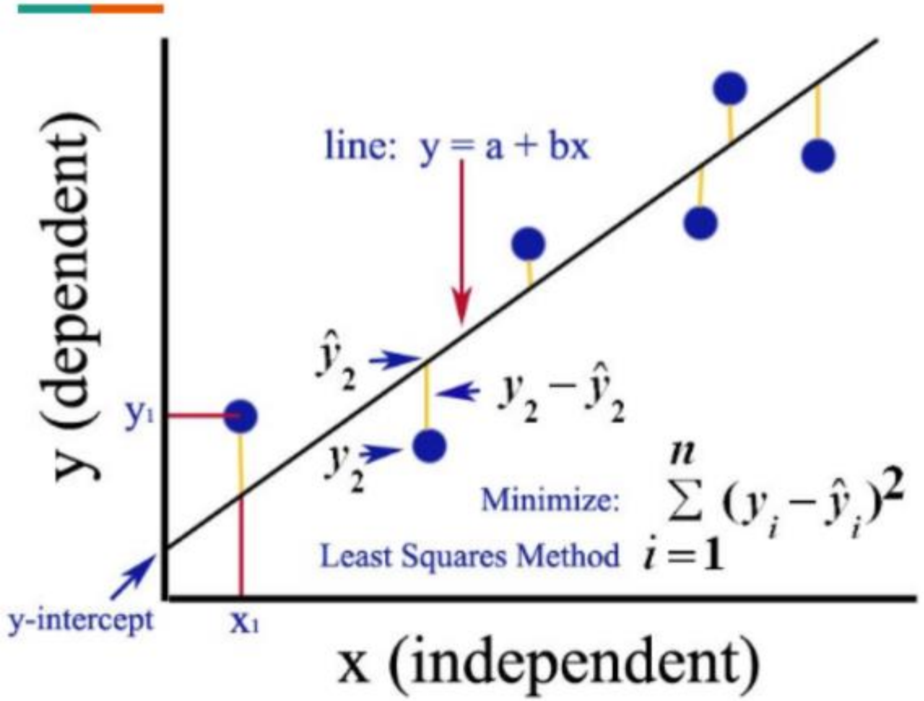
Positive



Negative

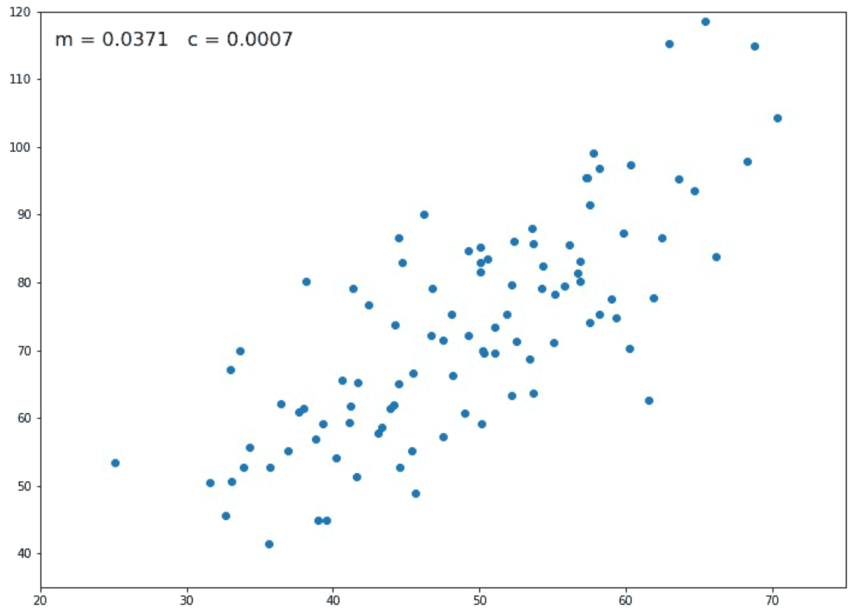
Ordinary Least Square (OLS) Regression

To find the best-fitting line in a linear regression model, we use a process called “ordinary least squares (OLS) regression”. This process involves calculating the sum of the squared differences between the predicted values and the actual values for each data point, and then finding the line that minimizes this sum of squared errors.



How to achieve best fit line in Linear Regression

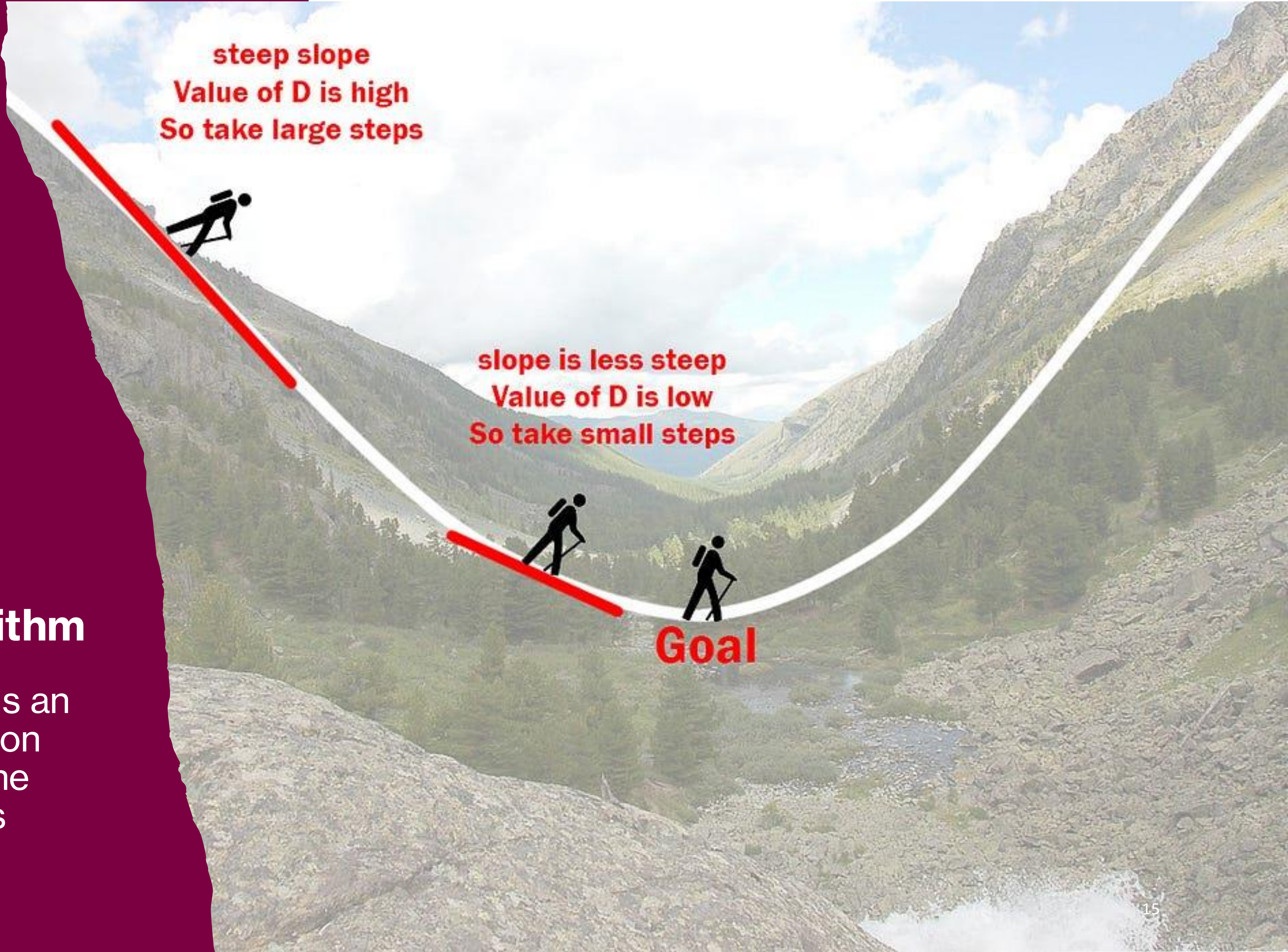
The best-fitting line is found by minimizing the residual sum of squares (RSS), which is the sum of the squared differences between the predicted values and the actual values. This is achieved by adjusting the values of the intercept and slope coefficients, also known as c and m, respectively.



Loss Function

$$E = \frac{1}{n}\sum_{i=0}^{n}(y_i - \bar{y}_i)^2$$

$$E = \frac{1}{n}\sum_{i=0}^{n}(y_i - (mx_i + c))^2$$



The Gradient Descent Algorithm

Gradient descent is an iterative optimization algorithm to find the minimum of a Loss function.

Gradient Descent Steps

1.Initially let m = 0 and c = 0. Let L be our learning rate. This controls how much the value of **m** changes with each step. L could be a small value like 0.0001 for good accuracy.

2.Calculate the partial derivative of the loss function with respect to m, and plug in the current values of x, y, m and c in it to obtain the derivative value **D**.

$$D_m = \frac{1}{n}\sum_{i=0}^{n} 2(y_i - (mx_i + c))(-x_i)$$

$$D_m = \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - \bar{y}_i)$$

$$D_c = \frac{-2}{n}\sum_{i=0}^{n} (y_i - \bar{y}_i)$$

Gradient Descent Steps

3. Now we update the current value of **m** and **c** using the following equation:

$$m = m - L \times D_m$$

$$c = c - L \times D_c$$

4. We repeat this process until our loss function is a very small value or ideally 0 (which means 0 error or 100% accuracy). The value of **m** and **c** that we are left with now will be the optimum values.