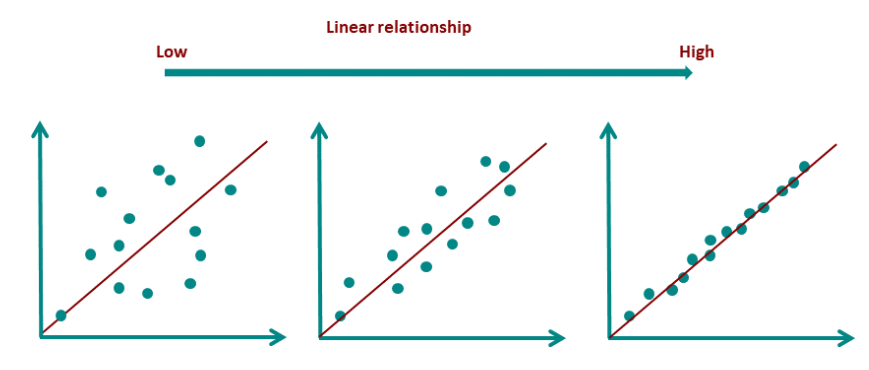
# Linear Regression and Logistic Regression

## Introduction

Regression models are fundamental supervised learning techniques in machine learning. They estimate the relationship between input features (X) and an output target (Y). Two widely used types are Linear Regression and Logistic Regression.

## Linear Regression

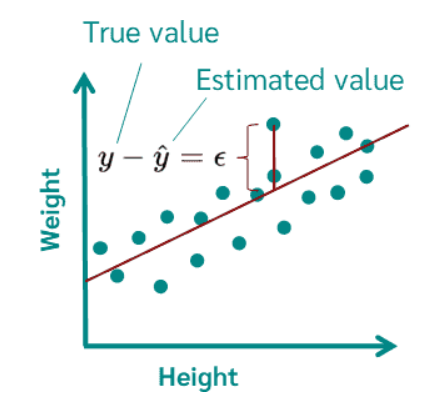
Linear regression is a type of supervised machine-learning algorithm which maps data points with the best linear functions after learning from labeled datasets. This method can be applied to new datasets to make predictions. It makes the assumption that the input and output have a linear relationship, which means that when the input changes, the output changes at the same rate. A straight line is used to depict this relationship.



**Figure 1: Linear Relationship**

The straight line that most accurately depicts the relationship between the independent variable (input) and the dependent variable (output) is known as the best-fit line in linear regression. It is the line that reduces the discrepancy between the actual data points and the model's predicted values.

The goal of linear regression is to find a straight line that minimizes the error between the observed data points and the predicted values. This line helps us predict the dependent variable for new, unseen data.



**Figure 2: Linear Regression Graph**

### Equation of a straight line



Where:

* y is the predicted value (dependent variable)
* x is the input (independent variable)
* m is the slope of the line or weight (feature of the dataset that dominates the result)
* b is the bias (randomness of data/ outside data) it’s also known as intercept

The best-fit line will be the one that optimizes the values of m (slope) and b (intercept) so that the predicted y values are as close as possible to the actual data points.

## Loss function

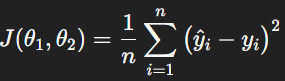
The difference between the predicted value ŷ and the true value y and it is called loss function or the cost function.

In Linear Regression, the Mean Squared Error (MSE) cost function is employed, which calculates the average of the squared errors between the predicted values ŷ*i​* and the actual values y*i*​. The purpose is to determine the optimal values for the intercept *θ1*​ and the coefficient of the input feature *θ2*​ providing the best-fit line for the given data points.

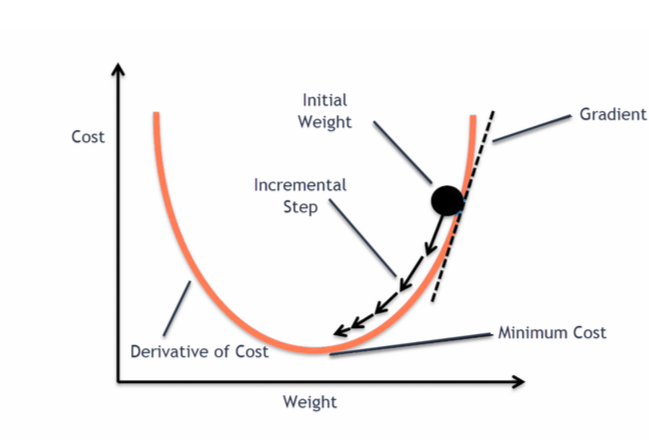
The linear equation expressing this relationship is:



MSE function can be defined as:



## Gradient Descent



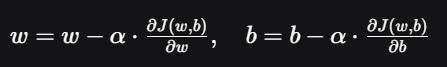
**Figure 3: Gradient Descent Graph**

Gradient descent minimizes the Mean Squared Error (MSE) which is used as the loss function to determine the best-fit line. Since MSE is a convex function gradient descent guarantees convergence to the global minimum if the learning rate is appropriately chosen.

The algorithm calculates the MSE's gradient in relation to the biases and weights.

It uses the following formula to update the weights (w) and bias (b):

* determining the log-loss gradient in relation to the weights.
* Iteratively updating weights and biases to increase the probability of the right classification:



## Mean Square Error (MSE)

Mean Squared Error (MSE) is an evaluation metric that determines the average of the squared discrepancies between the actual and expected values for each and every data point. To prevent positive and negative differences from canceling each other out, the difference is squared.



* *n* is the number of observations
* yi represents the actual value for the ith data point.
* ŷ*i​* represents the predicted value for the ith data point.

Better model performance is indicated by a lower MAE value. Since we are looking at absolute differences, it is not affected by outliers.

## Mean Absolute Error (MAE)

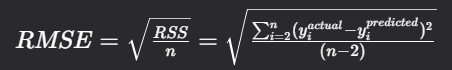
Mean Absolute Error is an evaluation metric used to calculate the accuracy of a regression model. MAE measures the average absolute difference between the predicted values and actual values.



* *n* is the number of observations
* yi represents the actual values
* ŷ*i​* represents the predicted values

## Root Mean Squared Error (RMSE)

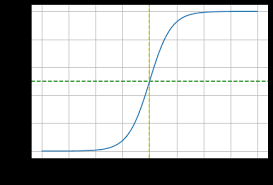
The Root Mean Squared Error is the square root of the variance of the residuals. It explains the absolute fit of the model to the data or how well the observed data points match the expected values.



# Logistic Regression

Logistic regression is one of the supervised machine learning technique for classification issues. It predicts the likelihood that an input belongs to a particular class, as opposed to linear regression, which predicts continuous values. In binary classification, the output can fall into one of two categories, such as True/False, Yes/No, or 0/1. It converts inputs into a probability value between 0 and 1 using the sigmoid function. The fundamentals of logistic regression and its key ideas will be covered in this article.

Logistic regression model transforms the linear regression function continuous value output into categorical value output using a sigmoid function which maps any real-valued set of independent variables input into a value between 0 and 1. This function is known as the logistic function.



**Figure 4: Sigmoid Funcation Graph**

The "S" shape can be broken down into three main sections:

* Left-hand flat region: For highly negative input values, the output is very close to 0. The curve is nearly flat, meaning that large changes in the input produce only negligible changes in the output.
* Steep central region: In the middle of the graph, small changes to the input produce the largest changes in the output. This is where the function is most sensitive.
* Right-hand flat region: For highly positive inputs, the output is very close to 1. Like the left side, the curve is nearly flat, and it becomes "saturated" as it approaches its maximum value.