# Machine Learning Fundamental Part 1



Prepared by Muhamad Irfan Fadhullahl

Today's Discussion

## Outline of Topics

What is Machine Learning?
Supervised and Unsupervised Algorithm
Linear Regression
Cost Function
Gradient Descent
Logistic Regression
Naive Bayes
Decision Tree
Confusion Matrix

# Machine Learning?

Machine Learning (ML) is basically that field of computer science with the help of which computer systems can provide sense to data in much the same way as human beings do.

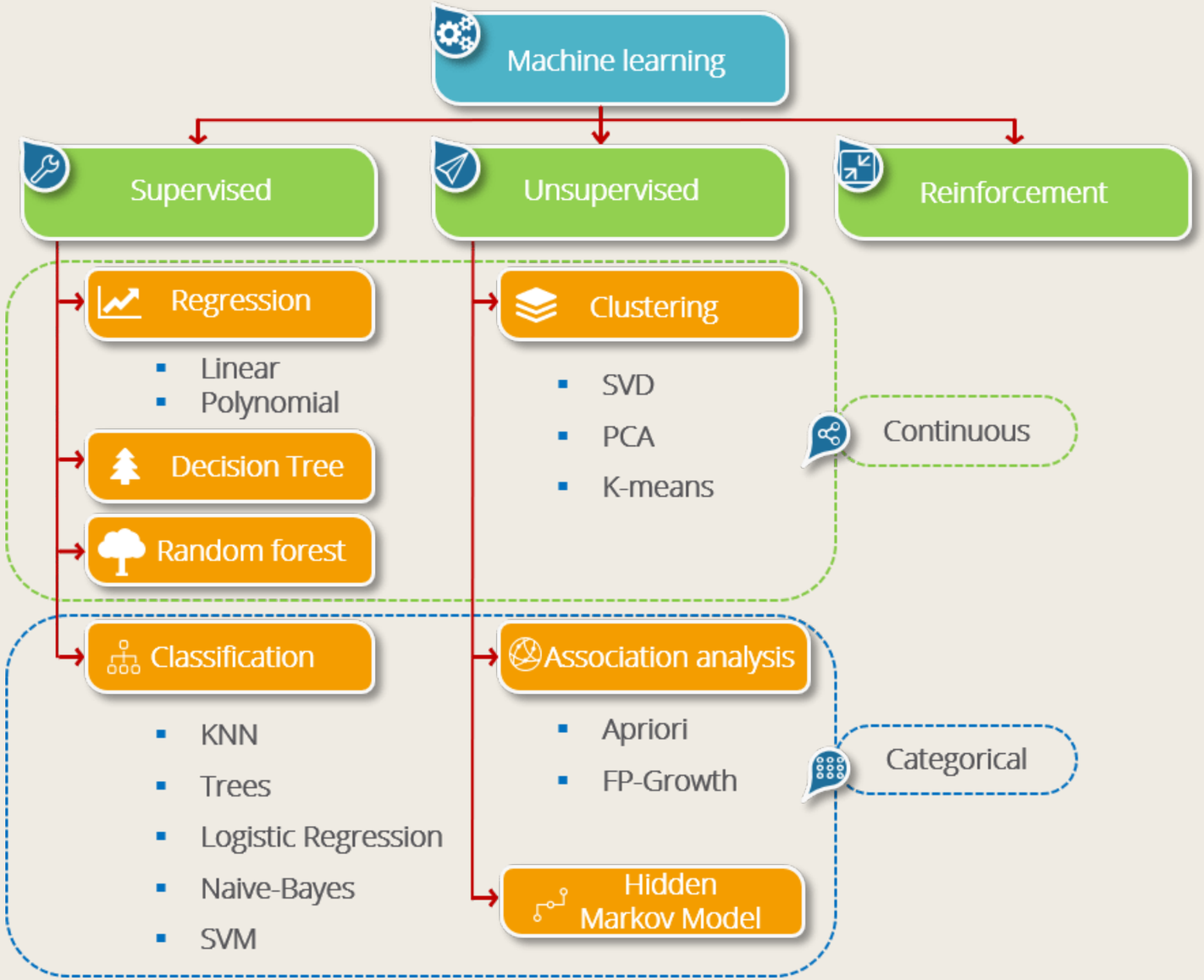
Machine Learning is the science (and art) of programming computers so they canlearn from data.

# Machine Learning vs Traditional Learning

## Traditional Computing


## Machine Learning

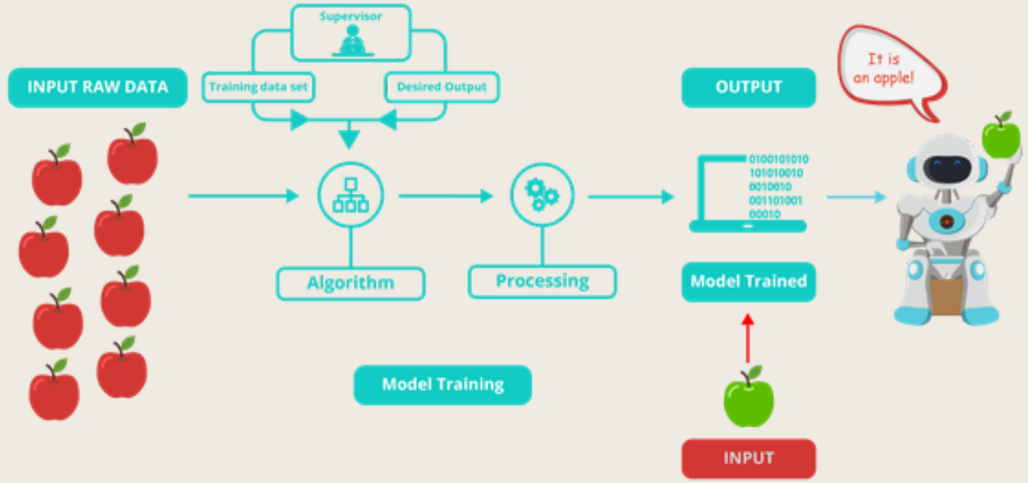
# Machine Learning



# 06 Supervised and Unsupervised

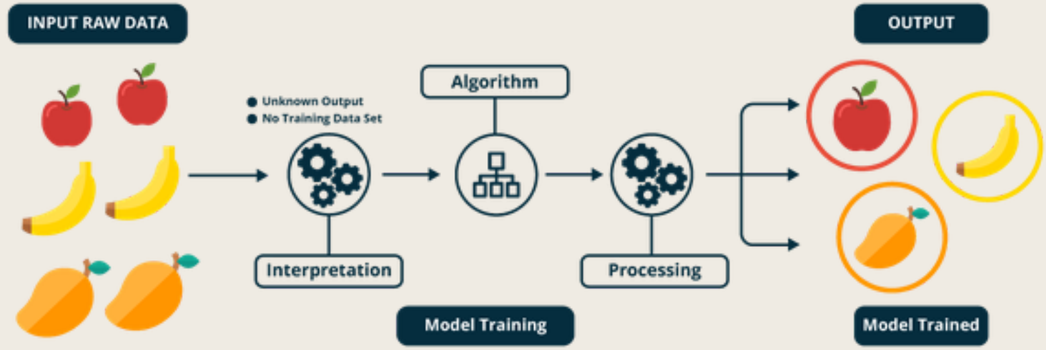
## Supervised Learning

Supervised learning is typically done in the context of classification, when we want to map input to output labels, or regression, when we want to map input to a continuous output.

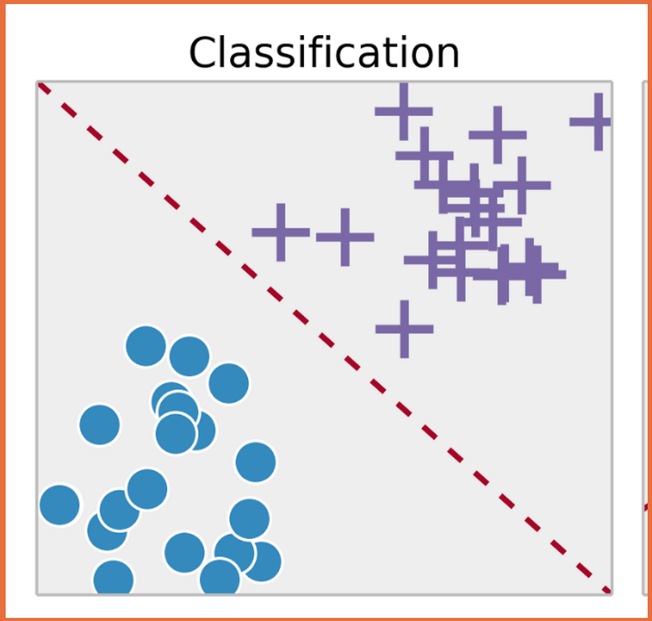



## Unsupervised Learning

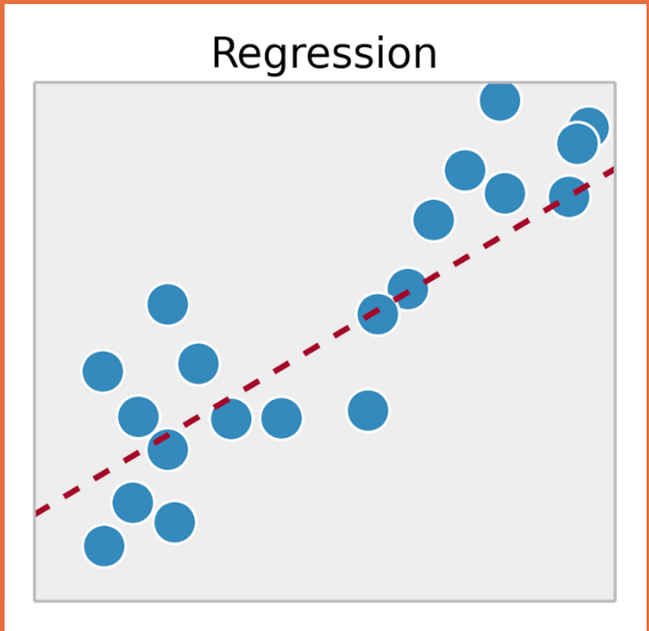
Unsupervised Learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses.



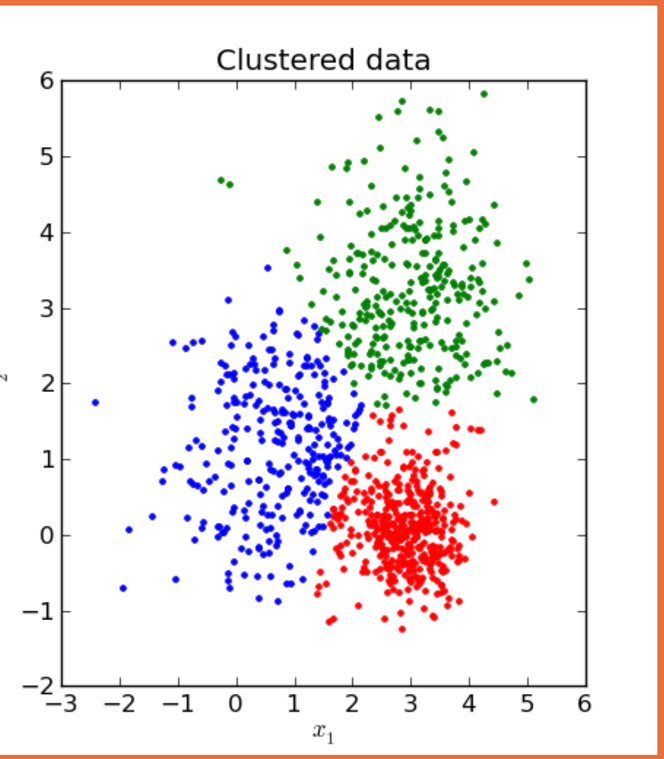
# Supervised Learning

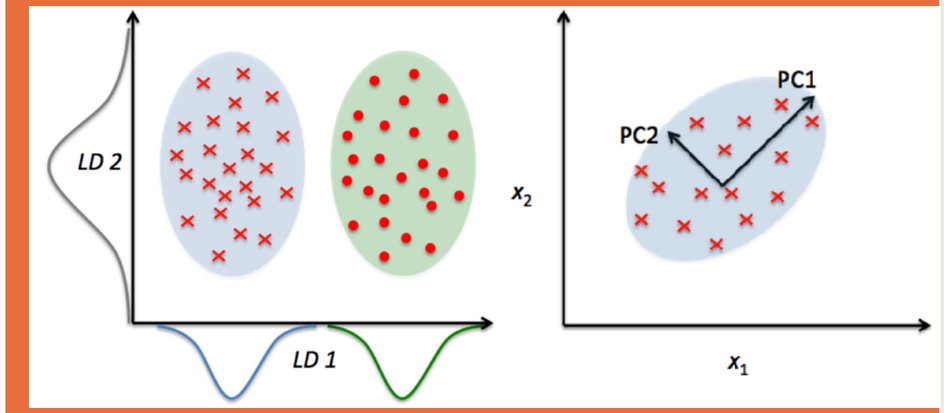



- *Logistics Regression*
- *Naive Bayes*
- *Decision Tree*
- *Random Forest*
- *Support Vector Machine*




- *Linear Regression*



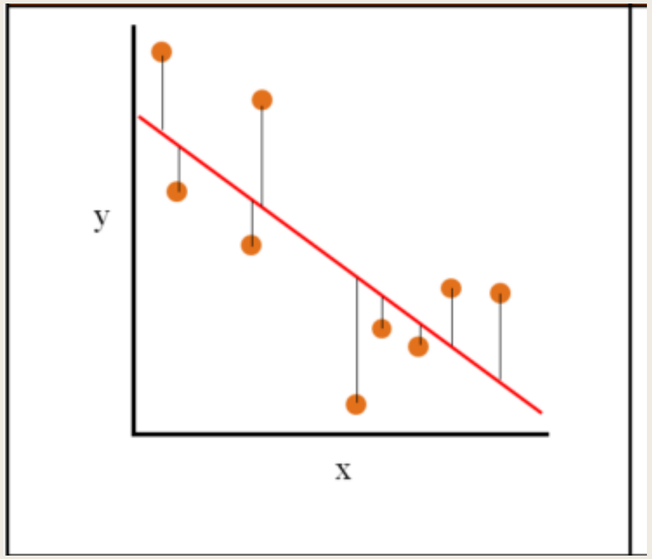

# Unsupervised Learning

- ***SVD***
- ***PCA***
- ***K-Means***

# 09 Linear Regression

## Explanation

Linear regression may be defined as the statistical model that analyzes the linear relationship between a dependent variable with given set of independent variables.




## Equation

**Simple Linear Regression**

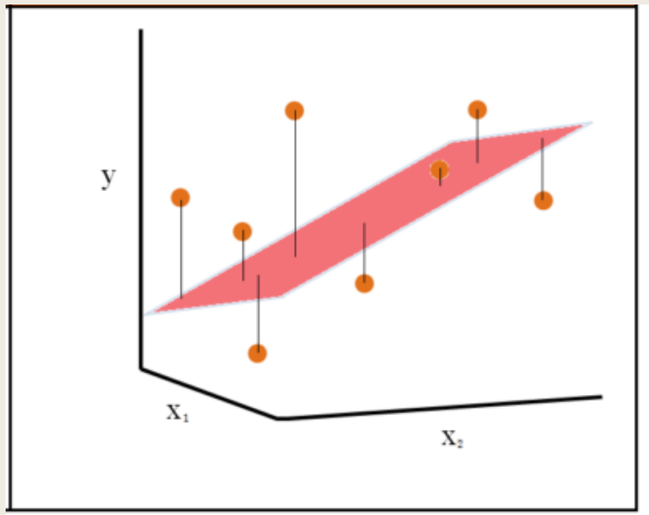
$$y = b_0 + b_1 * x_1$$

**Multiple Linear Regression**

Dependent variable (DV) — Independent variables (IVs)

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \ldots + b_n * x_n$$

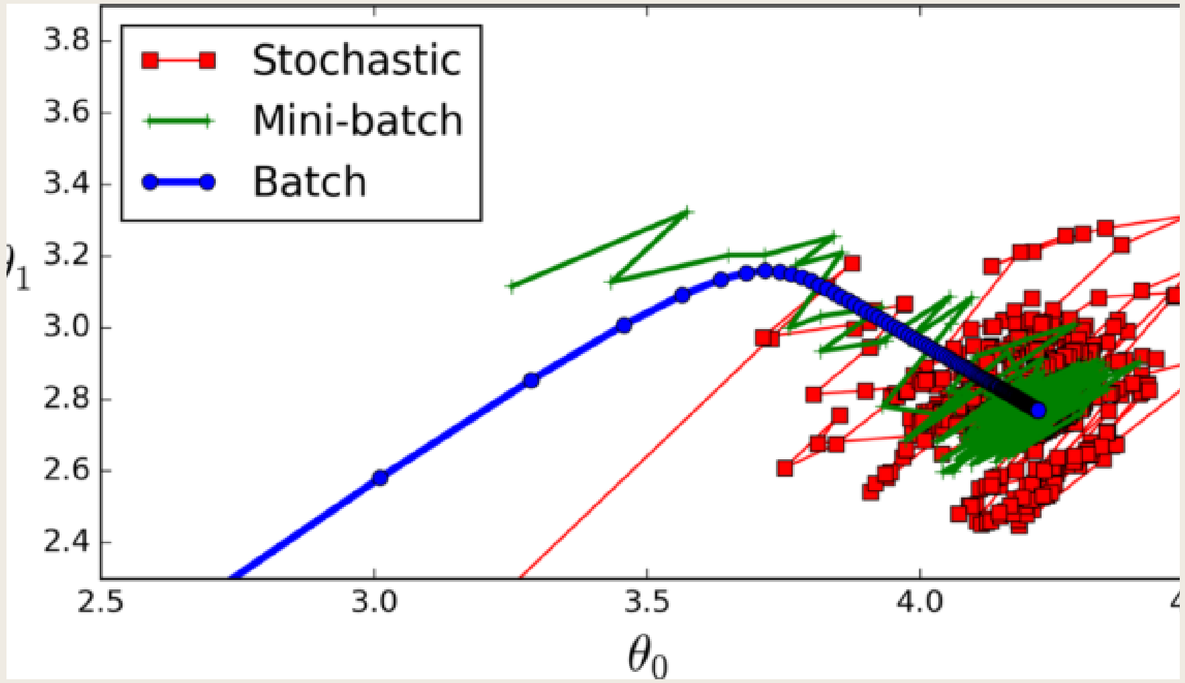
Constant — Coefficients



# Gradient Descent

## Explanation

Gradient descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient.




## Types

- Batch Gradient Descent
- Stochastic Gradient Descent
- Mini-batch Gradient Descent

## Batch Gradient Descent

calculates the error for each observation in the dataset but performs an update only after all observations have been evaluated. Batch gradient descent is not often used, because it represents a huge consumption of computational resources, as the entire dataset needs to remain in memory.

## Stochastic Gradient Descent

Performs a parameter update for each observation. So instead of looping over each observation, it just needs one to perform the parameter update. SGD is usually faster than batch gradient descent, but its frequent updates cause a higher variance in the error rate, that can sometimes jump around instead of decreasing.

## Mini Batch Gradient Descent

It is a combination of both bath gradient descent and stochastic gradient descent. Mini-batch gradient descent performs an update for a batch of observations. It is the algorithm of choice for neural networks, and the batch sizes are usually from 50 to 256.

## Cost Function

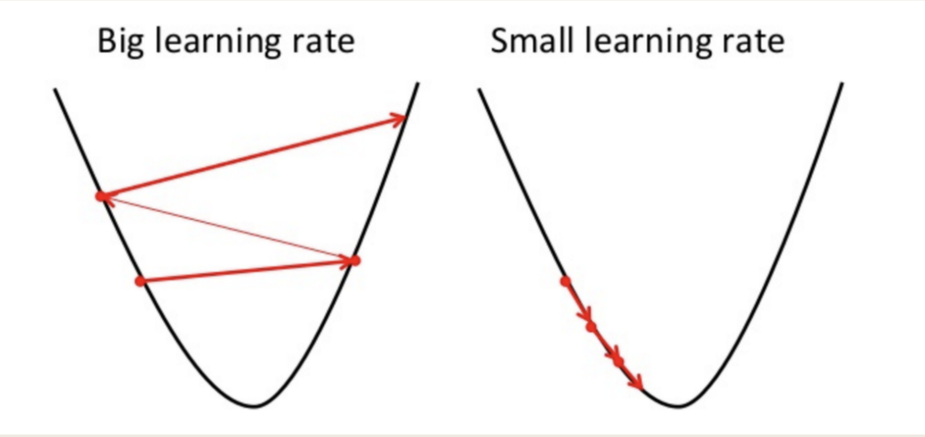
Loss functions are used to determine the error between the output of our algorithms and the given target value. The loss function expresses how far off the mark our computed output is.

$$I(y, \hat{y}) = \begin{cases} 1 & y \neq \hat{y} \\ 0 & y = \hat{y} \end{cases}$$

$$(y,\hat{y}) = C(y- \hat{y})$$

## Learning Rate

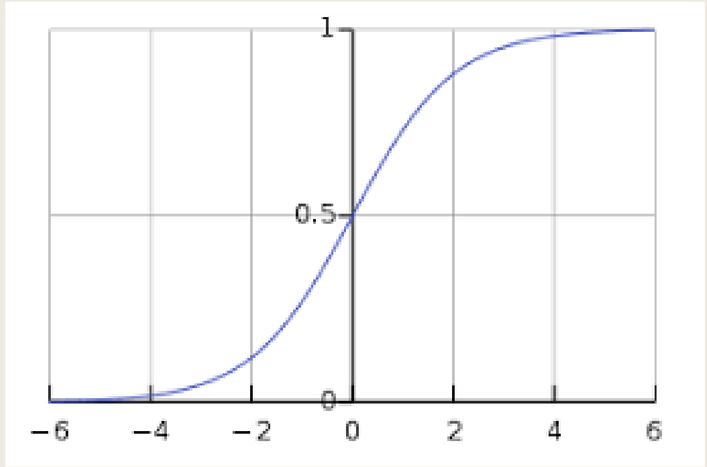
The size of these steps is called the learning rate. High Learning Rate causing overshooting, low learning rate is time consuming.

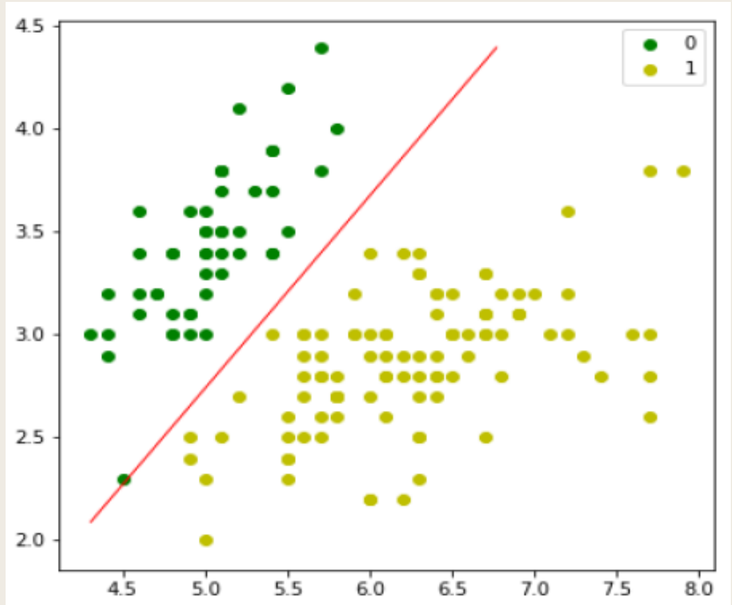


# Logistics Regression

## Explanation

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.




## Types

- Binary or Binomial
- Multinomial
- Ordinal

# Logistics Regression Assumption

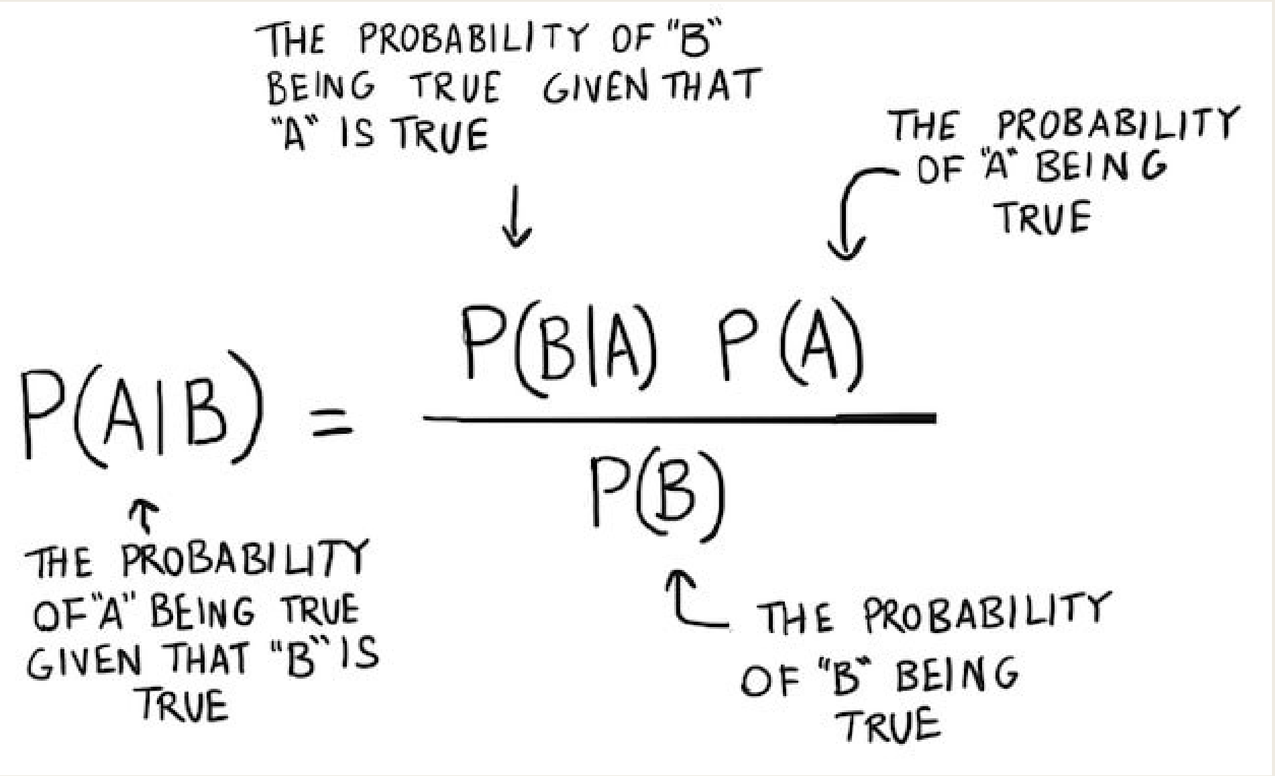


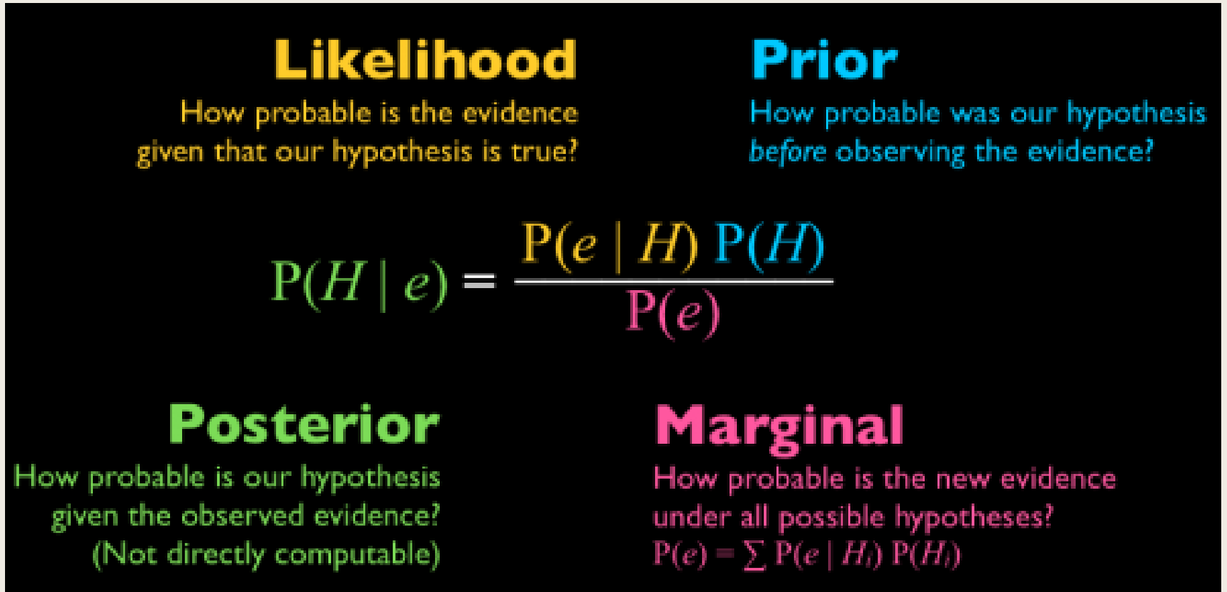
- In case of binary logistic regression, the target variables must be binary always and the desired outcome is represented by the factor level 1.
- There should not be any multi-collinearity in the model, which means the independent variables must be independent of each other.
- We must include meaningful variables in our model.
- We should choose a large sample size for logistic regression.

# Naive Bayes

## Explanation

Naïve Bayes algorithms is a classification technique based on applying Bayes' theorem with a strong assumption that all the predictors are independent to each other.

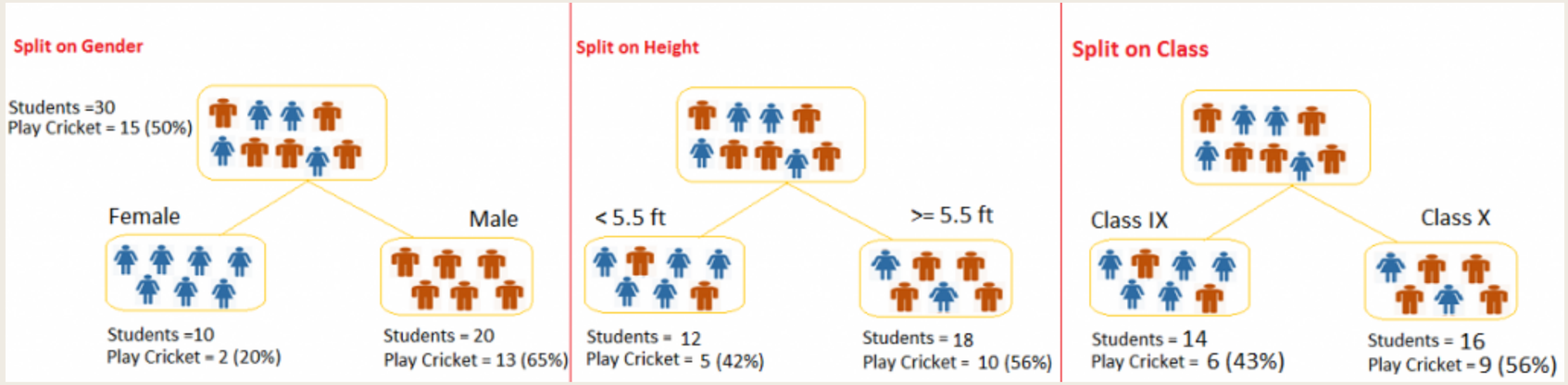


# Decision Tree



A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements.

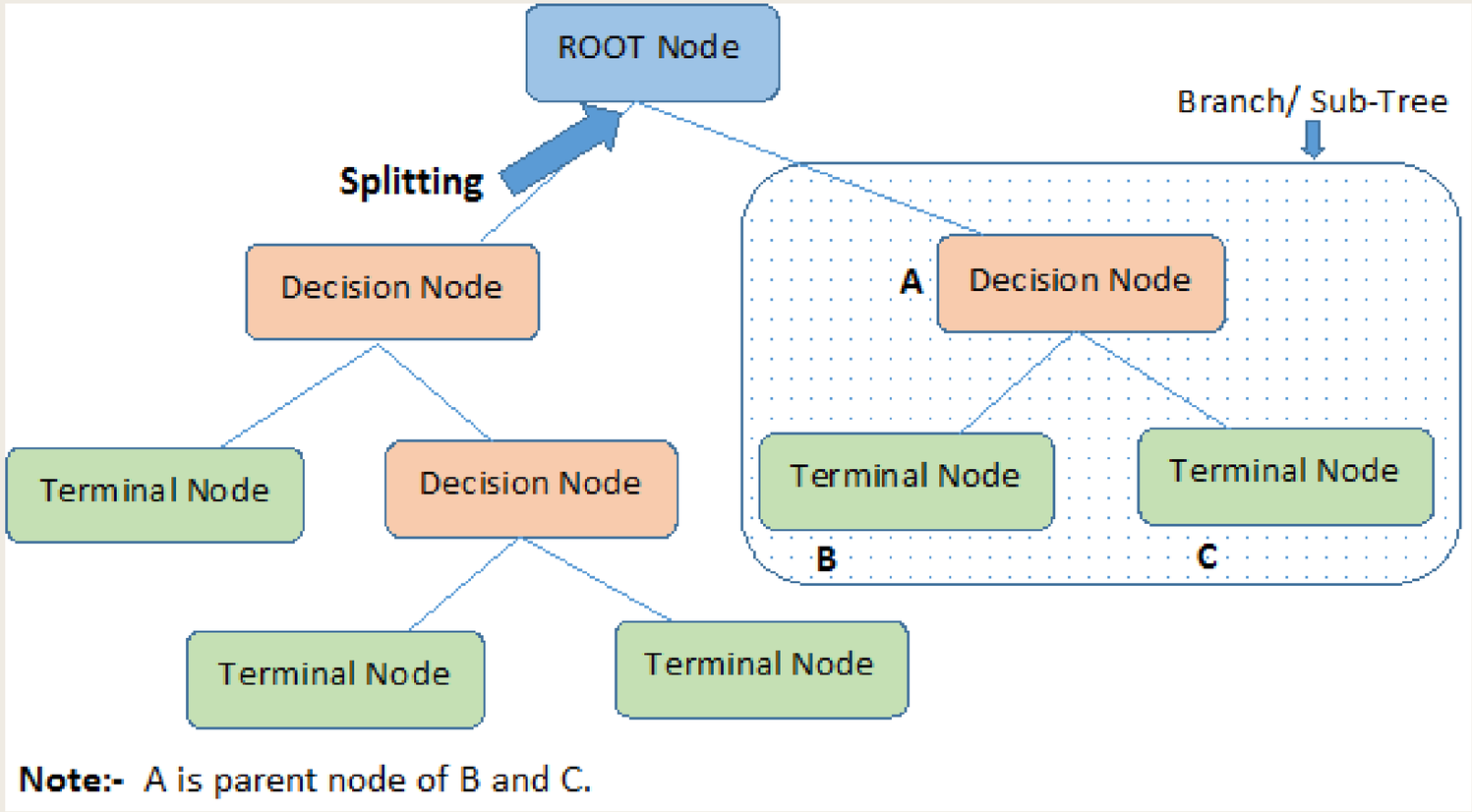



## Benefits

- It does not require any domain knowledge.
- It is easy to comprehend.
- The learning and classification steps of a decision tree are simple and fast.

# Terminology in Decision Trees






- **ROOT Node**: It represents entire population or sample and this further gets divided into two or more homogeneous sets.
- **SPLITTING:** It is a process of dividing a node into two or more sub-nodes.
- **Decision Node**: When a sub-node splits into further sub-nodes, then it is called decision node.
- **Leaf/ Terminal Node**: Nodes do not split is called Leaf or Terminal node.

# Terminology in Decision Trees



- **Pruning**: When we remove sub-nodes of a decision node, this process is called pruning. You can say opposite process of splitting.
- **Branch / Sub-Tree**: A sub section of entire tree is called branch or sub-tree
- **Parent and Child Node**: A node, which is divided into sub-nodes is called parent node of sub-nodes where as sub-nodes are the child of parent node.

**Tree Pruning Approaches**

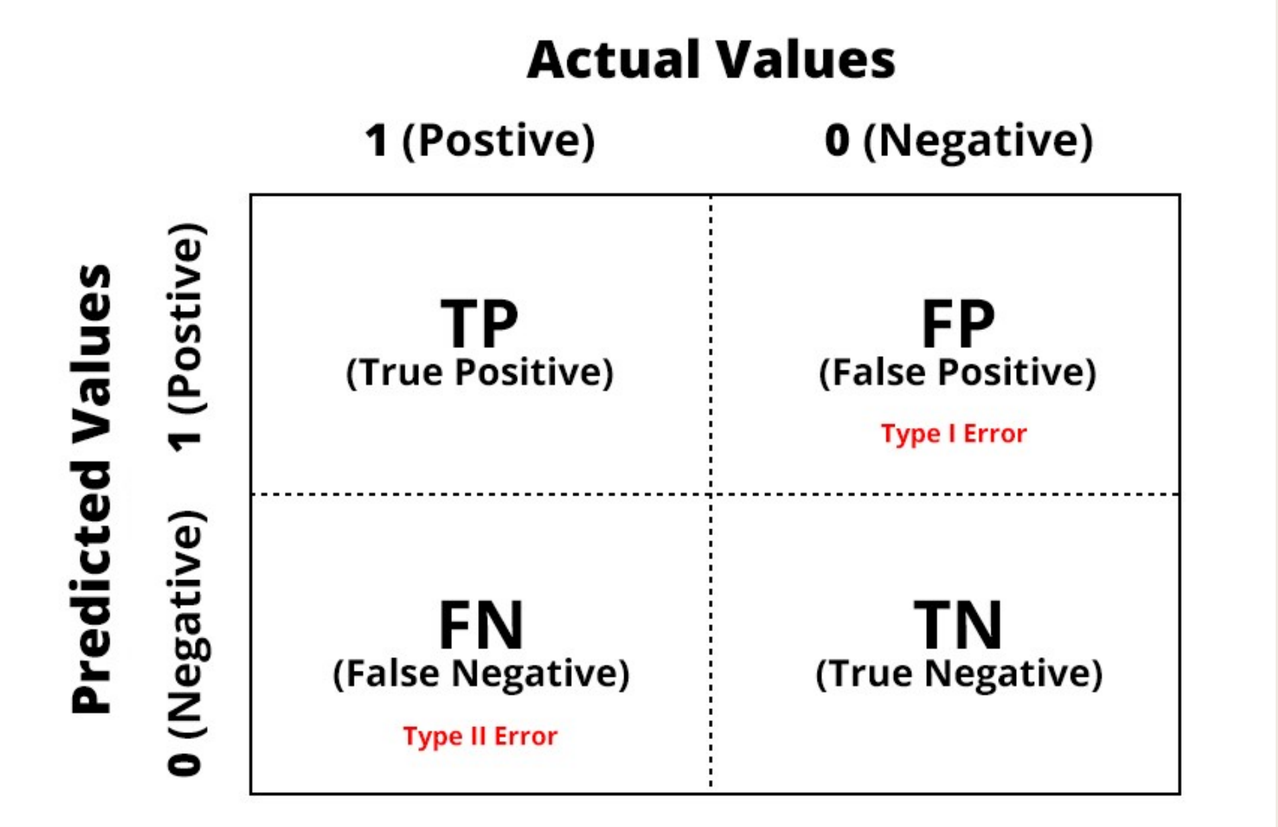
There are two approaches to prune a tree :

- Pre-pruning − The tree is pruned by halting its construction early.
- Post-pruning - This approach removes a sub-tree from a fully grown tree.

# Confusion Matrix



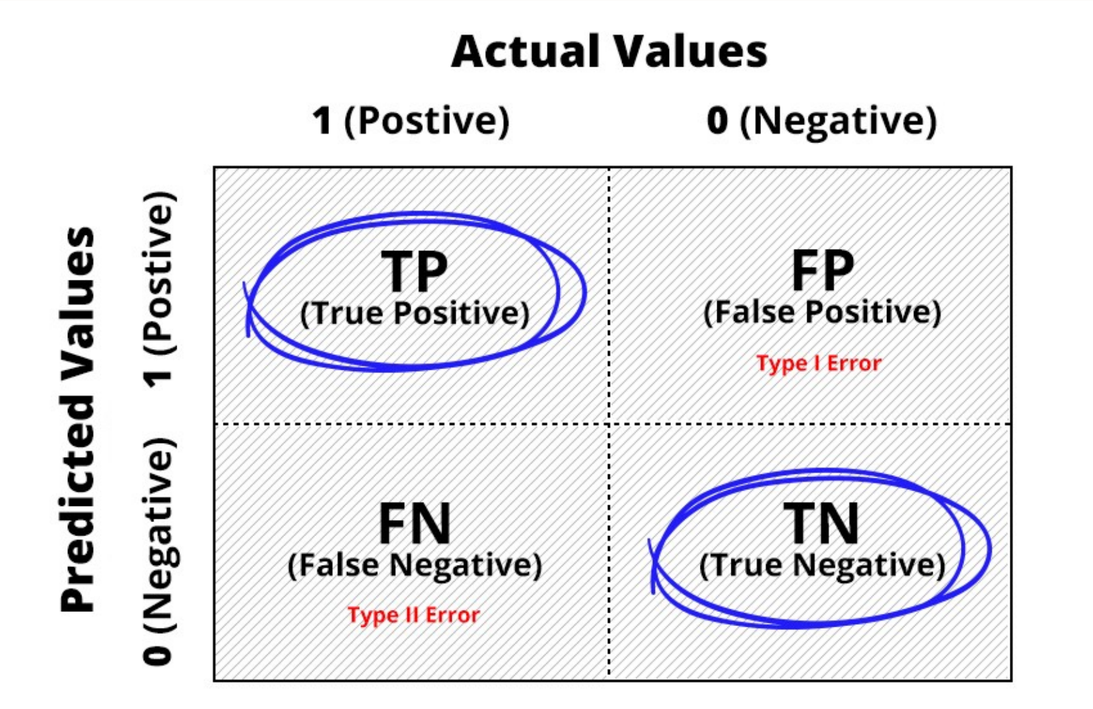
it is a performance measurement for machine learning classification problem where output can be two or more classes. It is a table with four different combinations of predicted and actual values.



# Accuracy



Accuracy describes how accurately a model can classify correctly. accuracy is the ratio of correct predictions (positive and negative) to the whole data. In other words, accuracy is the level of proximity of the predicted value to the actual (actual) value.

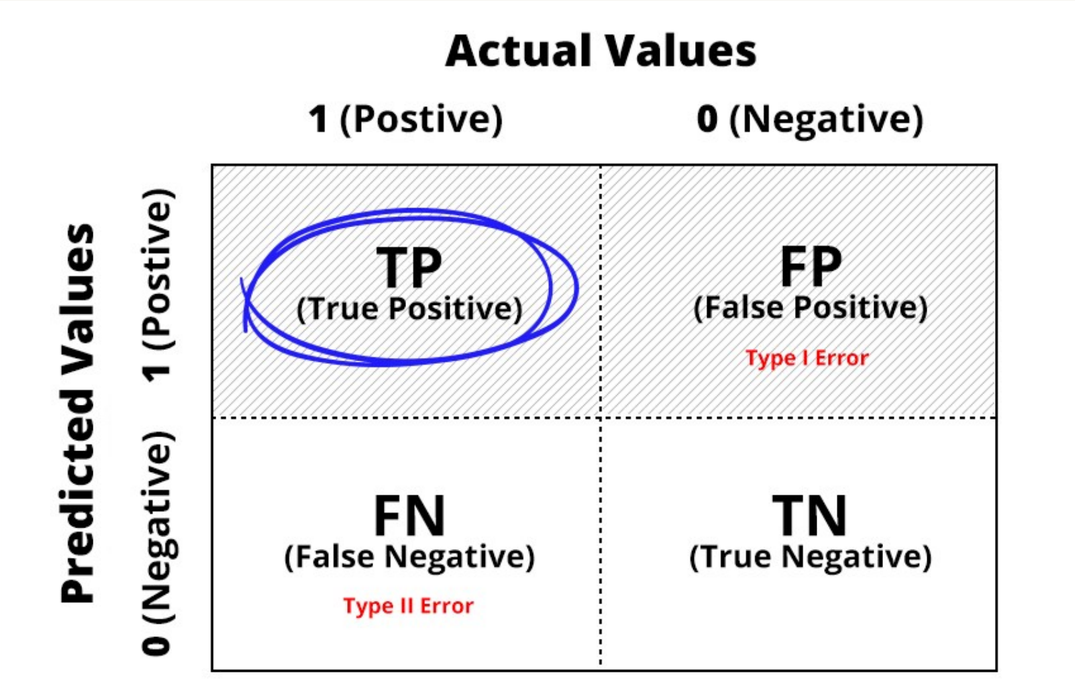



$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

# Precision



Precision describes the level of accuracy between the requested data and the predicted results provided by the model. Then, precision is the ratio of true positive predictions compared to overall positive predicted results. Of all the positive classes that have been correctly predicted, how much data is really positive.

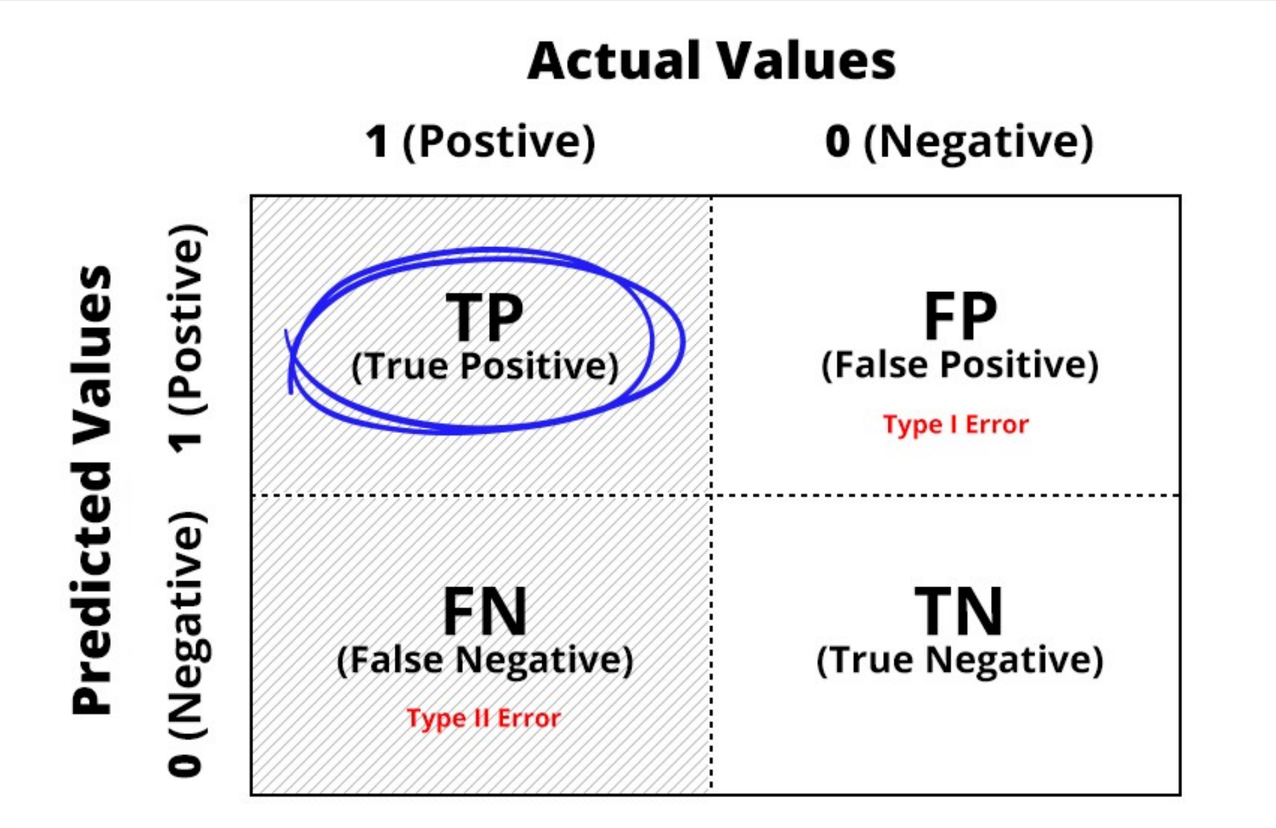



$$precision = \frac{TP}{TP + FP}$$

# Recall



Recall describes the success of the model in rediscovering information. Then, recall is a true positive prediction ratio compared to overall true positive data.

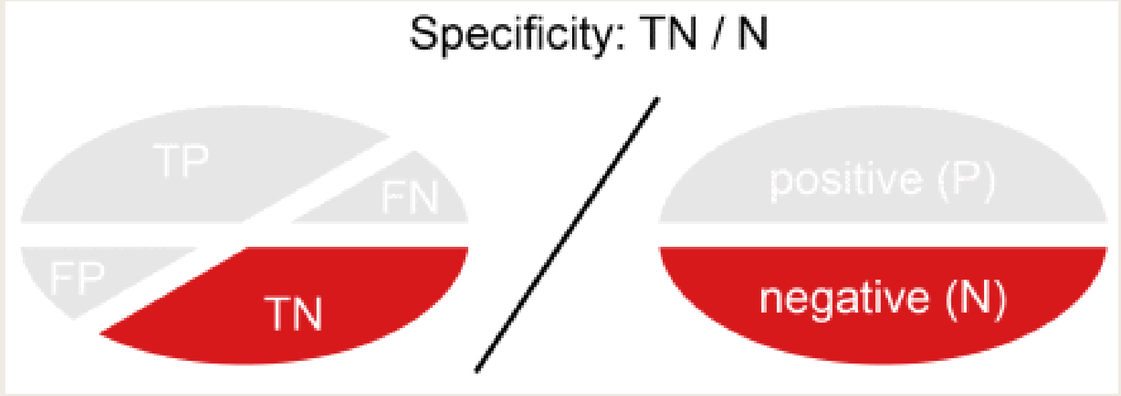



$$Recall = \frac{TP}{TP + FN}$$

# Specificity



Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. It is also called true negative rate (TNR). The best specificity is 1.0, whereas the worst is 0.0.




$$specificity = \frac{TN}{TN + FP}$$

# F1 Score



F1-score is a metric which takes into account both precision and recall and is defined as follows:

$$F1 = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

# To Be Continue..

***Next Week!***

- OTHER CLASSIFICATION ALGORITHM
- UNSUPERVISED LEARNING ALGORITHM
- MODEL PERFORMANCE ,ETC

# References


- www.tutorialspoint.com
- www.iykra.com
- Hands-OnMachine Learningwith Scikit-Learn& TensorFlow
- https://www.datavedas.com/linear-regression/
- www.towardsdatascience.com
- www.medium.com
- https://ml-cheatsheet.readthedocs.io
- https://www.geeksforgeeks.org
- https://www.analyticsvidhya.com

# Thank You!

KEEP LEARNING, STAY HUNGRY!