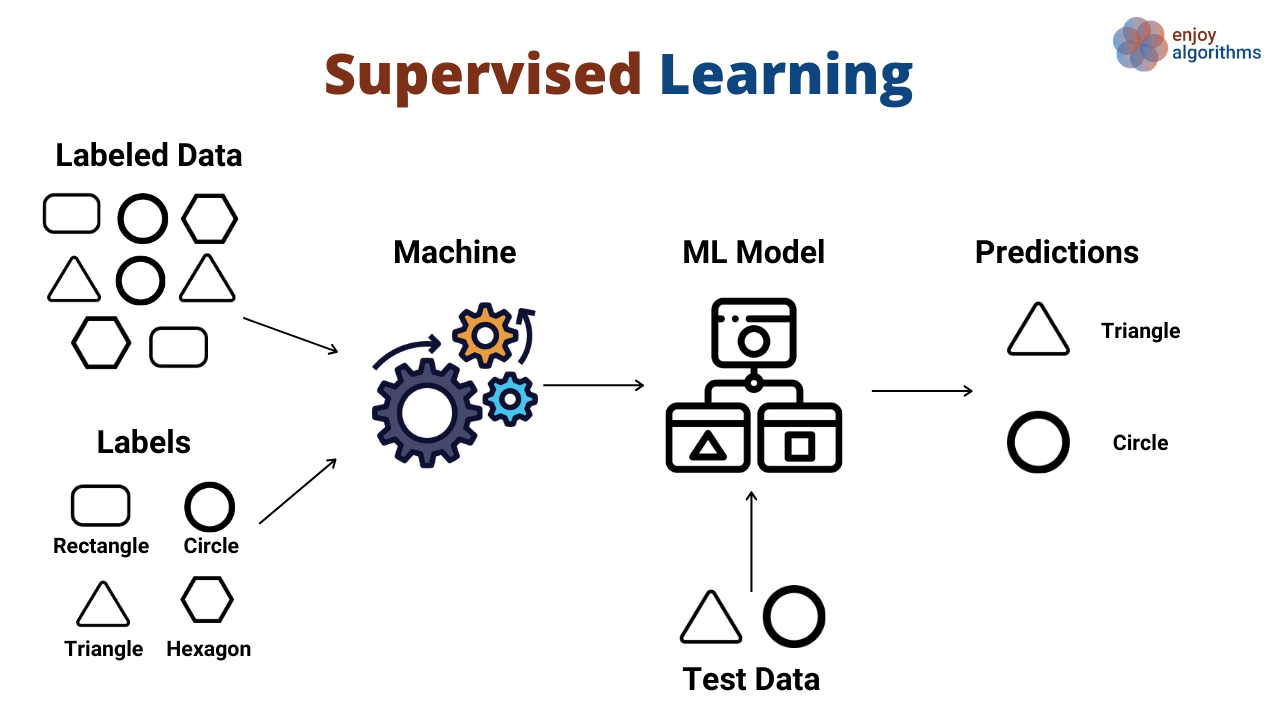
What is Supervised Learning –

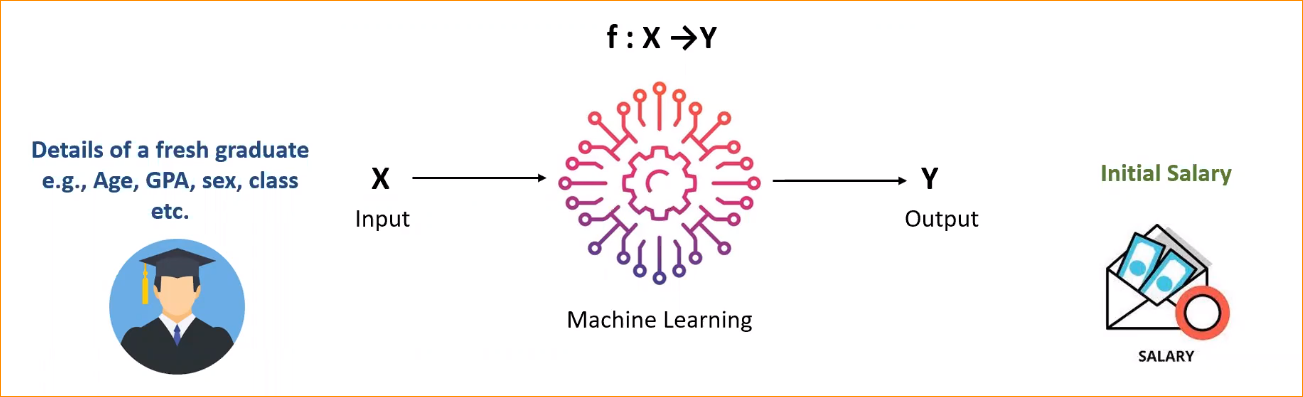
* A Computer algorithm is trained on input data has been labeled fro a particular output.



* ML – Where you have both input data and the output data, so call as **labeled** data
* During the learning process, ML algorithm learn relationship between input data and output lables.
* So that output is already presented, can learn the relationship between input Q1CP24 and output. – call Supervise learning.
* The model is tained until it can detect the underling **patterns** and **relationship** between the input data and the output data.

Example- Predict the initial salary of the fresh s graduate.

Function



Output

Input

Given details - Age,GPA,Sex,Class etc.

In here we have All input data and output lables.

For example 1000 fresh graduated students who got there initial salary, We have their records

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Student** | **Age** | **GPA** | **Sex** | **Class** | **Salary** |
| Pradeep | 25 | 3.11 | Male | 2nd lower | 50000 |
| Nisha | 24 | 3.96 | Female | 1st class | 70000 |
| Kamal | 25 | 2.69 | Male | General | 40000 |

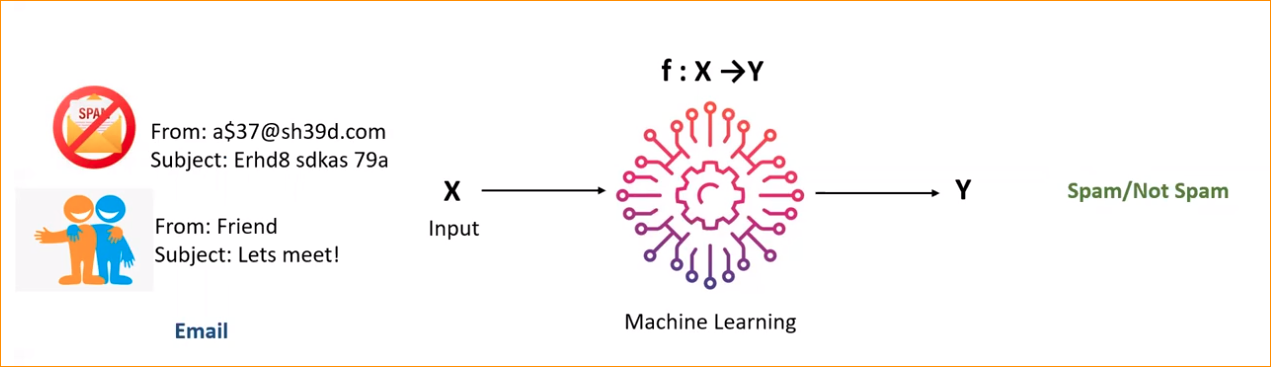
In here we have our input and output.

So now it you build the model that ML model will learn the relationship between input and output.

That relationship we can define as **Function (Relation between X and Y)**

So now given input and predict the output.

Example – Spam or not spam Email classifier.



Input – Email, Email ID, Subject or body

Output – Spam or Ham

If you have data set -

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Email** | **ID** | **Subject** | **Body** | **From** | **Output** |
| dialog@gmail.com | 5589k4dwe1 | Bill reminder | Dear customer… | Dialog | Spam |
| kamal @gmail.com | Dfgh56fe74s | Office Document | Hi, kavinda….. | Kamal | Ham |
| mobitel@gmail.com | Lom75Ul8c | Data package | Dear customer… | Mobile | Spam |

Output is Instances in our data set.

And Out-put (Spam or not) – Call label data / Class label

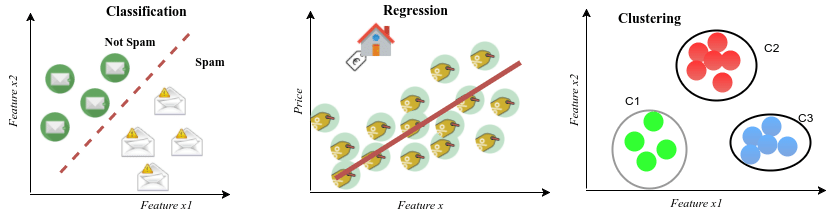
Ground Truth –

"Ground truth" refers to the absolute, true, or correct data or labels for a given problem or dataset. It serves as the reference against which predictive models are evaluated. The ground truth is essential for training supervised learning models and for evaluating the accuracy, precision, recall, and other performance metrics of those models.

Supervise learning has two major types-

1. **Classification – Learning a discrete function (or prediction discrete value)**

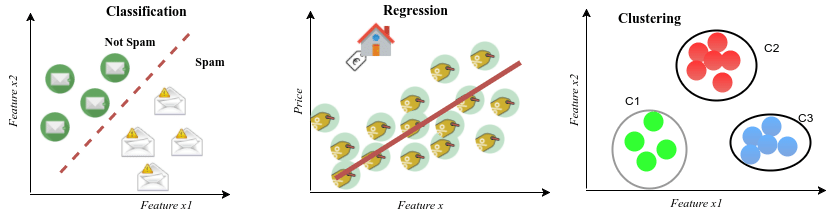
Example – Spam mail or not



Example – Grete of student A B C S F

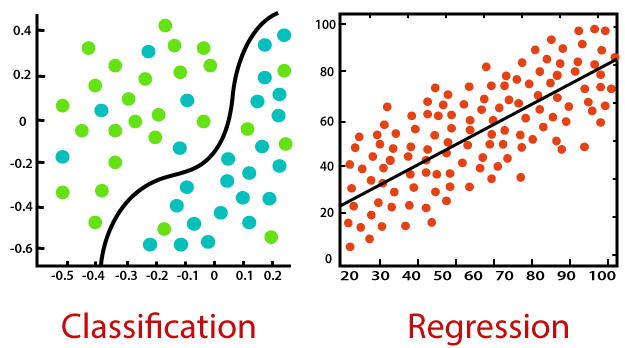
1. **Regression – Learning a continuous function (or predict continuous values)**

Example – Age or revenue of the organization or price of the goods or salary of the persion

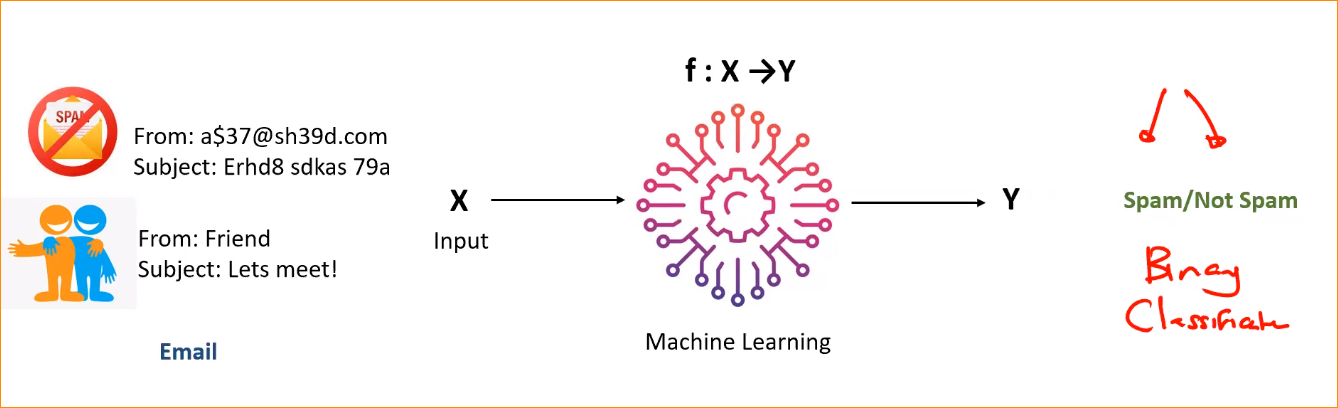


**F(X) 🡪 Y**

F(x) – have function, And We give X So it predicts the output Y



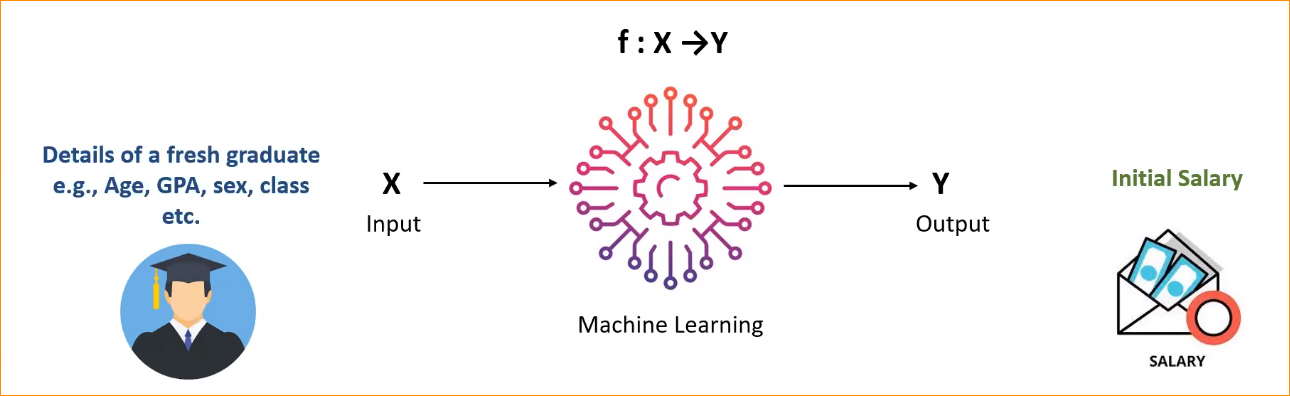
1. **Classification -**



In here we have only two discrete values so this is classification.

In here we have 2 discrete – called Binary classification

1. Regression



In here our function going to continuously.

Supervise Learning cont.

* In here We should have enough labeled data.
* That mean your input and output already available.
* If you train the model using input and labeled output, now you have F (Function- Trained model) so If we use that model in real time, model predicts the output that it learns in past.

**How is the process-**

* In here we have Labeled data
* Divide that data in to two section.
* Training and Testing
* In Training data 🡪 send to the Feature Engineering 🡪 Features (Actual input to the model, How data input to your model. And the model can understand only the numbers)
* Once the features are ready
* Features are actual inputs
* Then Apply Machine leaning algorithm and we will do the training process.
* Now we will receive the ML model
* Now model know what is the relationship between X and Y.

How much evaluate or have accuracy of the model ?

* Final setp the any ML is (not specific to the supervise learning ) that is evaluate the model.
* In here we provide the small percentage of data.
* Now the F () – Function is ready
* Now in testing data we also have X and Y, lets assume 100 records are there.
* Now give the X, on our input data to the learnt model.
* That model or Function give you the prediction
* Now we have actual Y 🡪 Call Y^ that the predicted model.
* Can measure how close they are.
* Now we check Y^ and actual Y of Test data. If 80 output are correct out of 100
* That means 80% accuracy you have.
* Once the traing model is completed , We have to evaluate the model. We have to