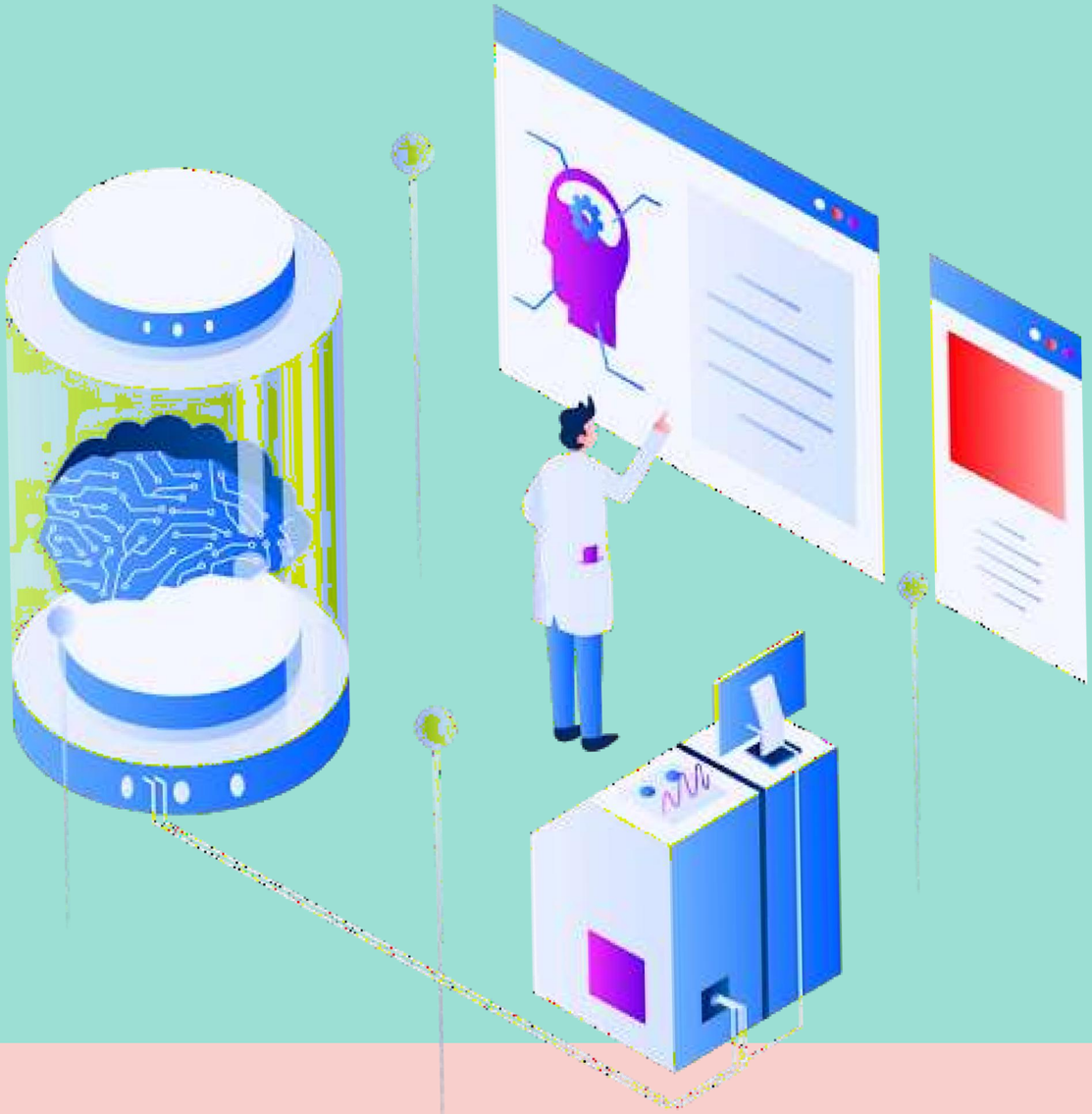
INTRODUCTION TO

Machine Learning & Deep Learning



WHAT IS ARTIFICIAL INTELLIGENCE?

AI

Ability to learn, understand and think” -Oxford Dictionary

Intelligence is also the ability to **learn from the environment** and change our behavior based on inputs we get. The study of how to make computers perform functions which at present humans are good at.

Examples: Voice and Speech Recognition, Face recognition and face identification, Object detection, Intuition, Learning new skills, Decision making,

Introduction

Artificial Intelligence is a branch of Science which deals with helping machines finds solutions to complex problems in a more human-like fashion.

This generally involves borrowing characteristics from human intelligence, and applying them as algorithms in a computer friendly way.

Brief History of A I

- 1941: First electronic computer (technology finally available)
- 1956: Term Artificial Intelligence introduced
- 1960s: Checkers-playing program that was able to play games with opponents
- 1980s: Quality Control Systems
- 2000: First sophisticated walking robot

WHY A I?

Computers are fundamentally well suited to performing mechanical computations, using fixed programmed rules. This allows artificial machines to perform simple monotonous tasks efficiently and reliably, which humans are ill-suited to. `

LIMITATIONS OF HUMAN MIND

- Object recognition. People cannot properly explain how they recognize objects.
- Face recognition. Cannot be passed on to another person by explanation.
- Naming of colours. Based on learning, not on absolute standards.

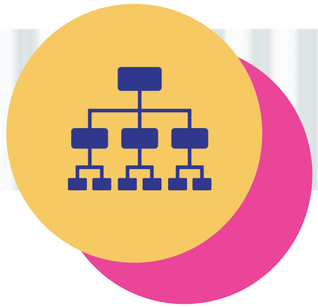
APPLICATIONS OF A I

- Expert systems.
- Natural Language Processing (NLP).
- Speech recognition.
- Computer vision.
- Robotics.

DIFFERENCE :

Data

Any piece of information that can be **stored and processed.**

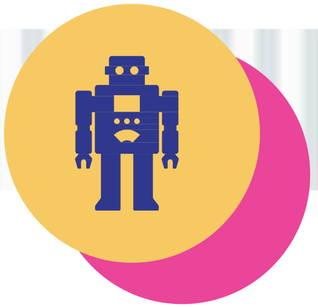


Data Science

A set of methods, processes, heuristics, and **algorithms to extract insights** from data.

Big Data

Extremely **large amounts of data** which traditional data processing fail to handle.



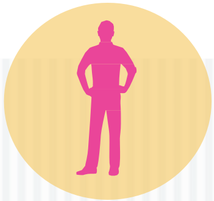
Artificial Intelligence

Simulation of **human intelligence in machines** that are programmed to think like humans and mimic their actions.

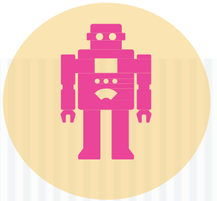
ML/DL

All **computer systems to learn from the data** without explicit programming.

HUMAN vs AI



- We See
- We hear, understand and convert into words
- We speak from a given text
- We smell
- We touch, move and run
- We think, reason, make decisions, solve problems, get info from data
- We move our arms and legs
- We feel



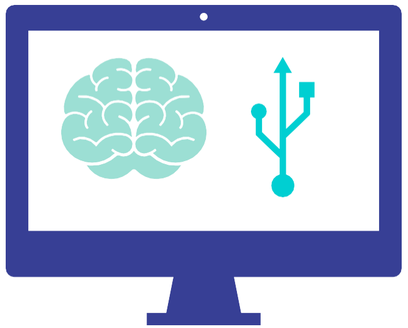
- Computer Vision
- Speech recognition/voice recognition and Natural Language Processing
- Speech Synthesis and Natural Language Processing (NLP)
- Nothing yet , may be possible
- Kinematics and robotics
- General AI - just starting
- Robotics
- Nothing still , is it possible at all ?

TYPES OF AI

Strong AI /

Full AI /

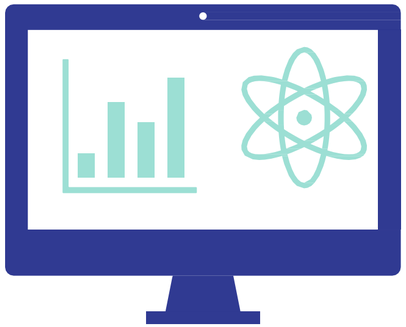
General Intelligence



Narrow AI /

Weak AI /

Applied AI



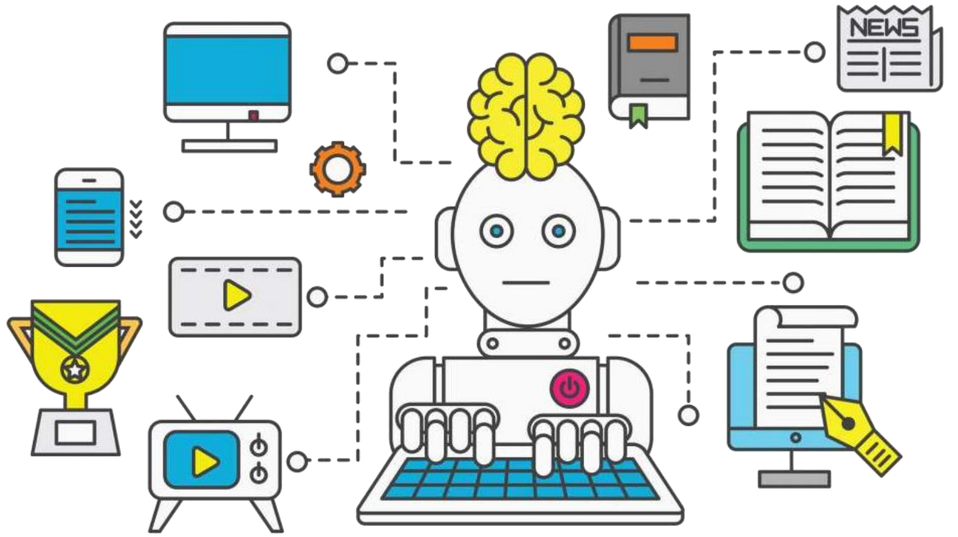
Strong AI

- We make the machines **copy or perform the cognitive functions** that we associate with human minds such as problem solving , learning and decision making. This is what we call as Strong AI or Broad AI.

- In other words the machine **mimics the complete cognitive abilities** and the full range of abilities that homo sapiens have.

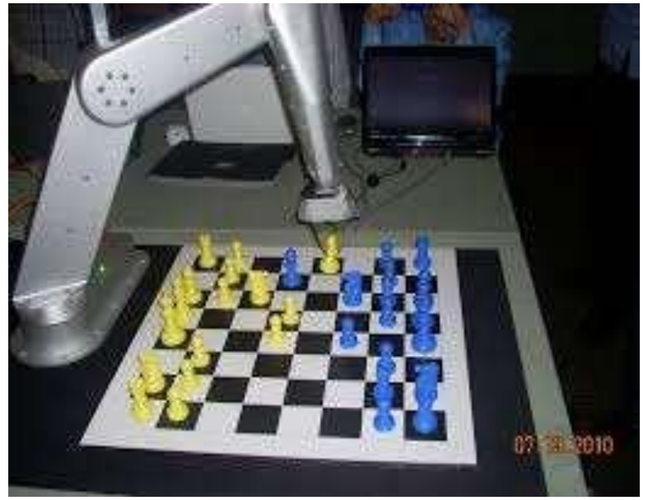
- Till now we do not have any full products in General AI, the likes of which we see in feature films- **think Arnold Schwarzenegger in Terminator series or in superstar in Robot series –** we are many decades away from that as per IBM AI head.

Weak AI



- Amazon’s recommendation Engine.
- Deep blue the first computer-based chess playing champion.
- The Advanced Driver Assistance Systems (ADAS).
- Amazon’s “echo dot” and “echo plus” voice recognition systems .
- Nest home connectivity startup acquired by Google in 2014.
- A music streaming service such as Pandora Internet Radio- It is a music recommendation service – this is another example for Narrow AI.
- Deep mind of google is working in weak AI, but they want to keep their focus also on General AI.

Artificial Intelligence in real world

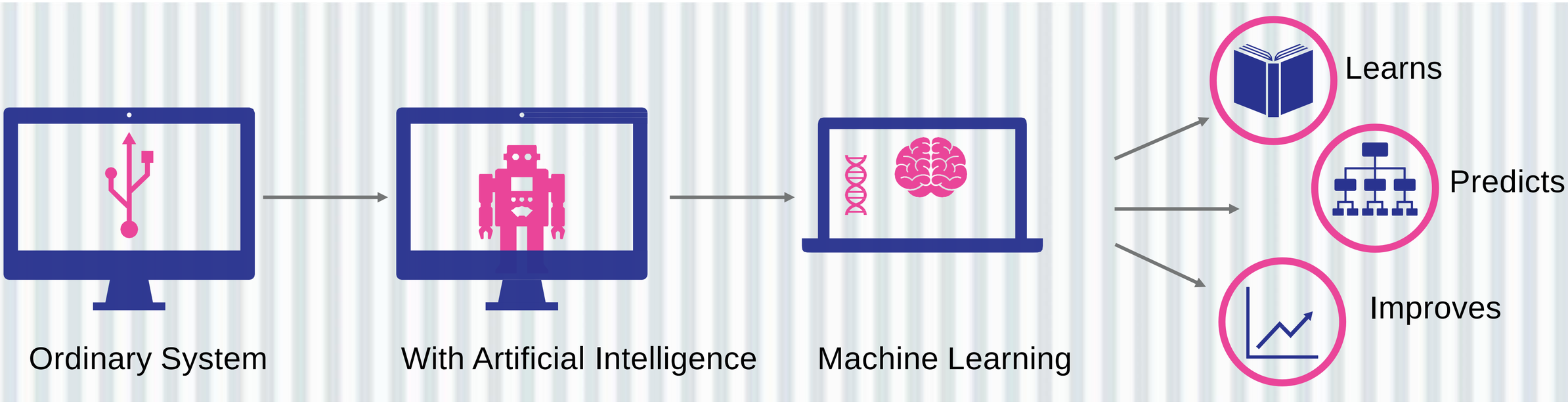


SCI-FI Movies with concept of AI

WHAT IS MACHINE LEARNING ?

MACHINE LEARNING

- Machine Learning is the **application of Artificial Intelligence (AI)** that provides systems the **ability to automatically learn** and improve from experience without being explicitly programmed.



TYPES OF MACHINE LEARNING

1

SUPERVISED

Task Driven

Predicts Next Value

(Outcome is Labelled)

Eg: **Regression, Classification**

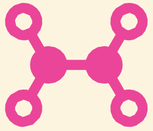
2

UNSUPERVISED

Data Driven

Finds Clusters

(Outcome is not Labelled)

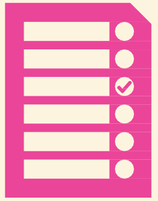


Eg: **Association, Clustering**

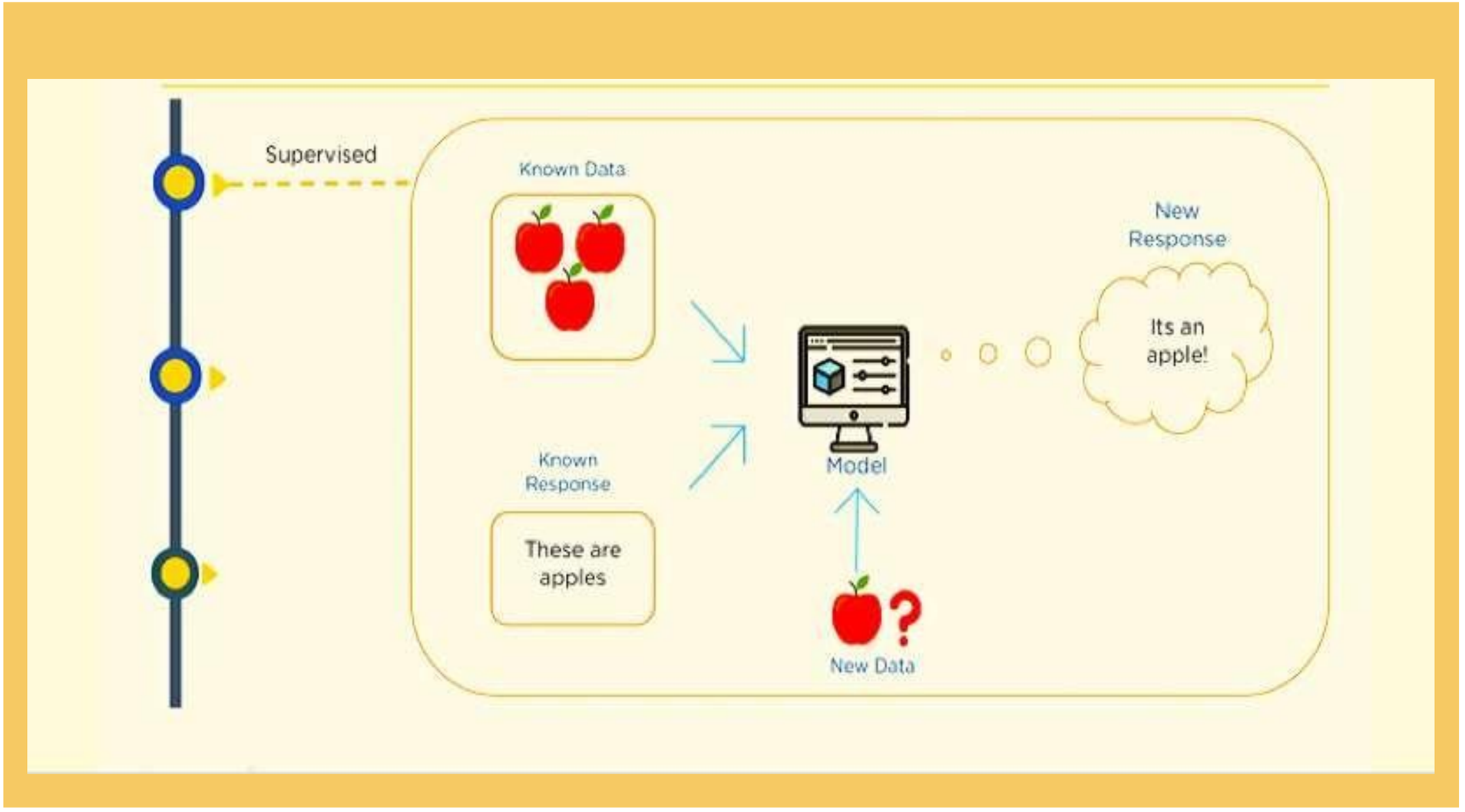
3

REINFORCEMENT

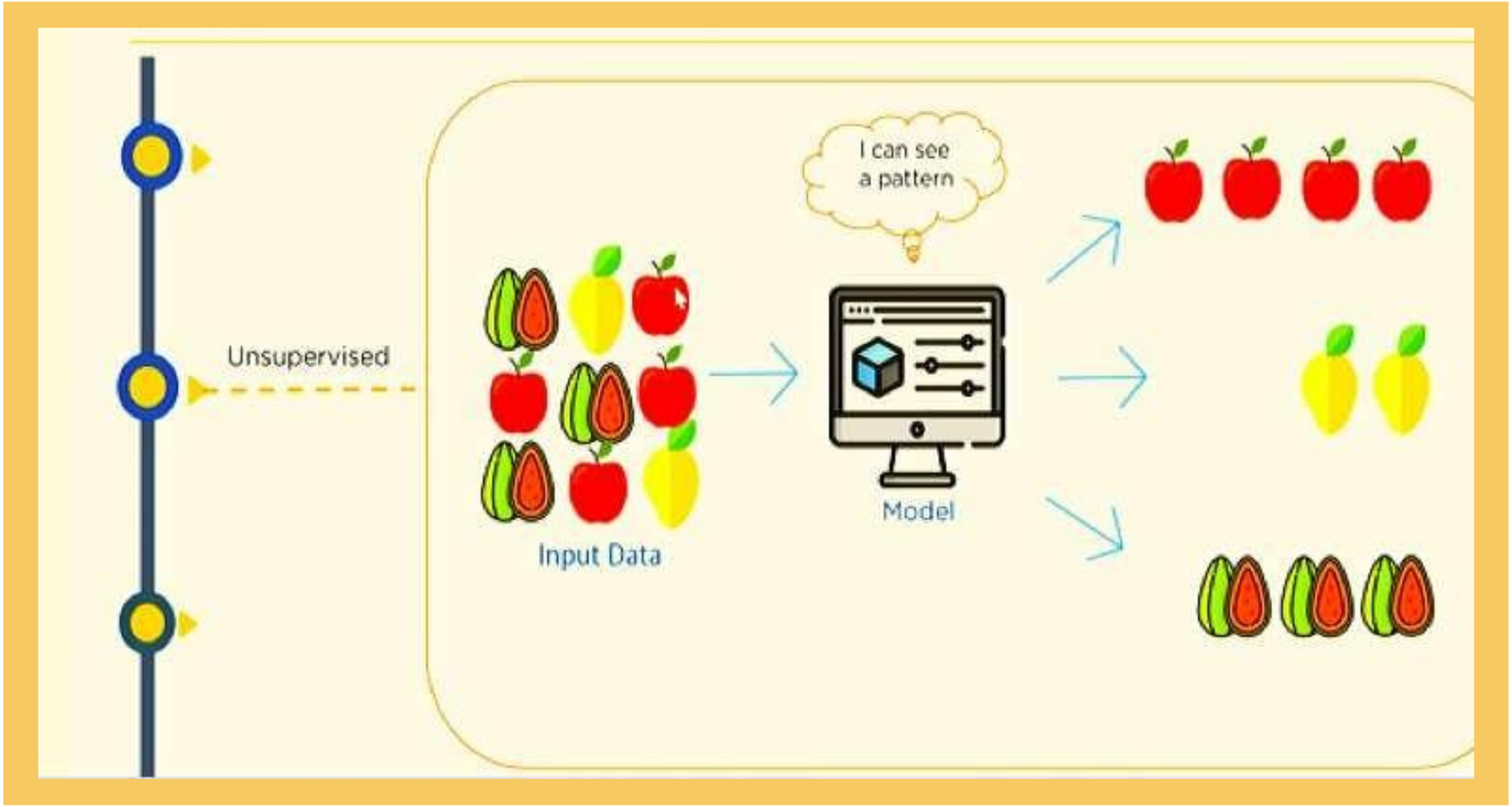
Learn from mistakes



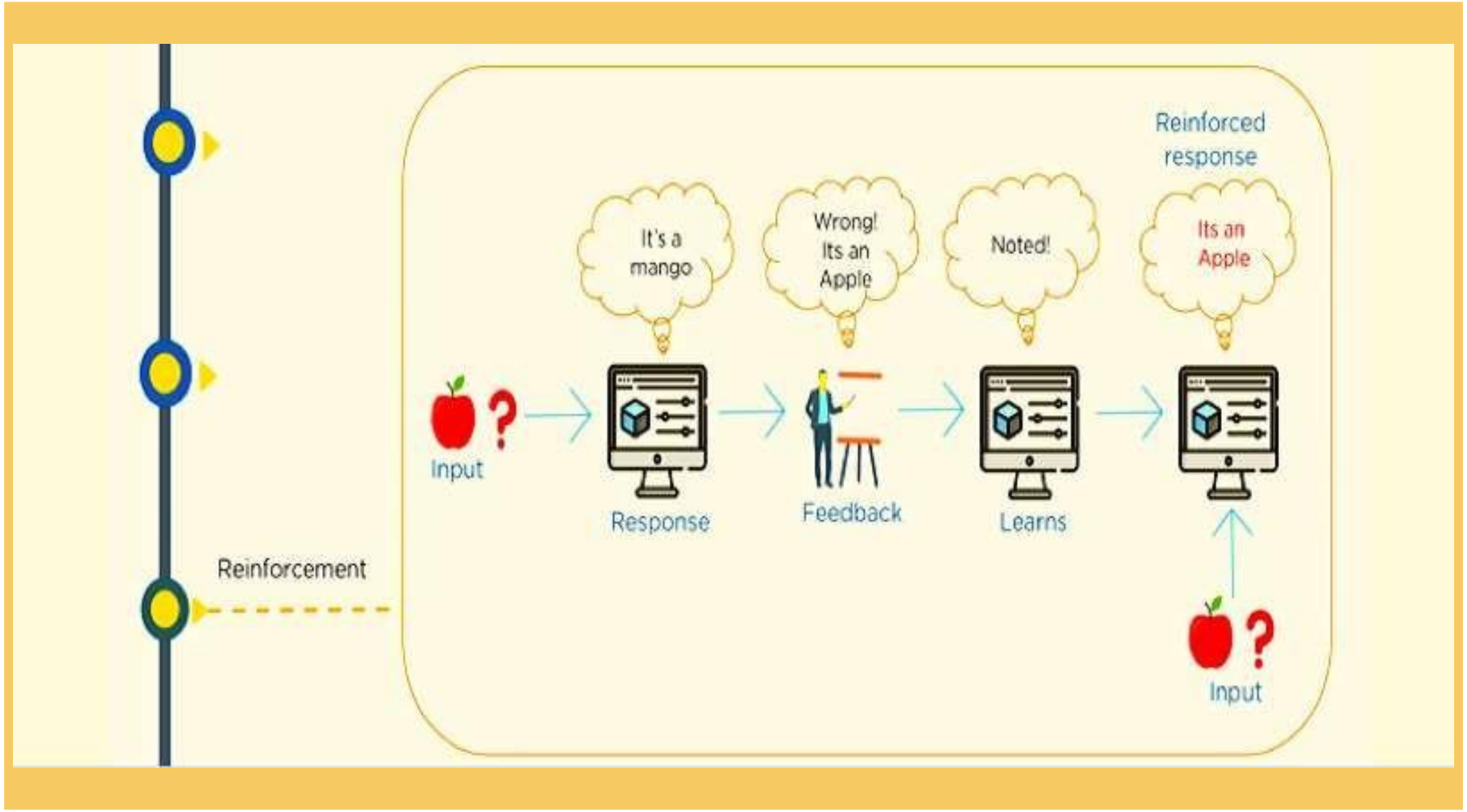
Supervised Learning



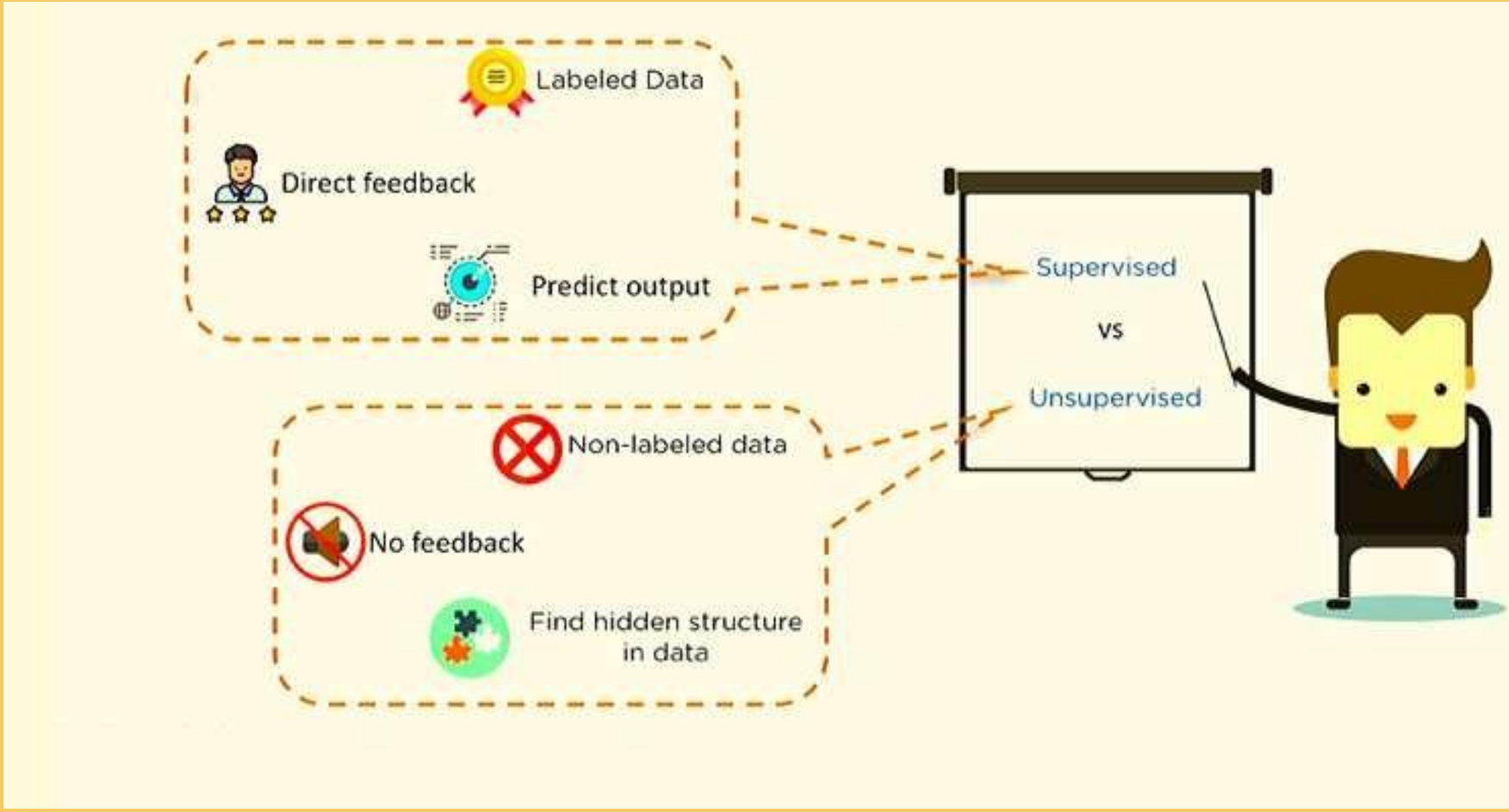
Unsupervised Learning



Reinforcement



Supervised Vs Unsupervised



1. Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.

2. In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly.

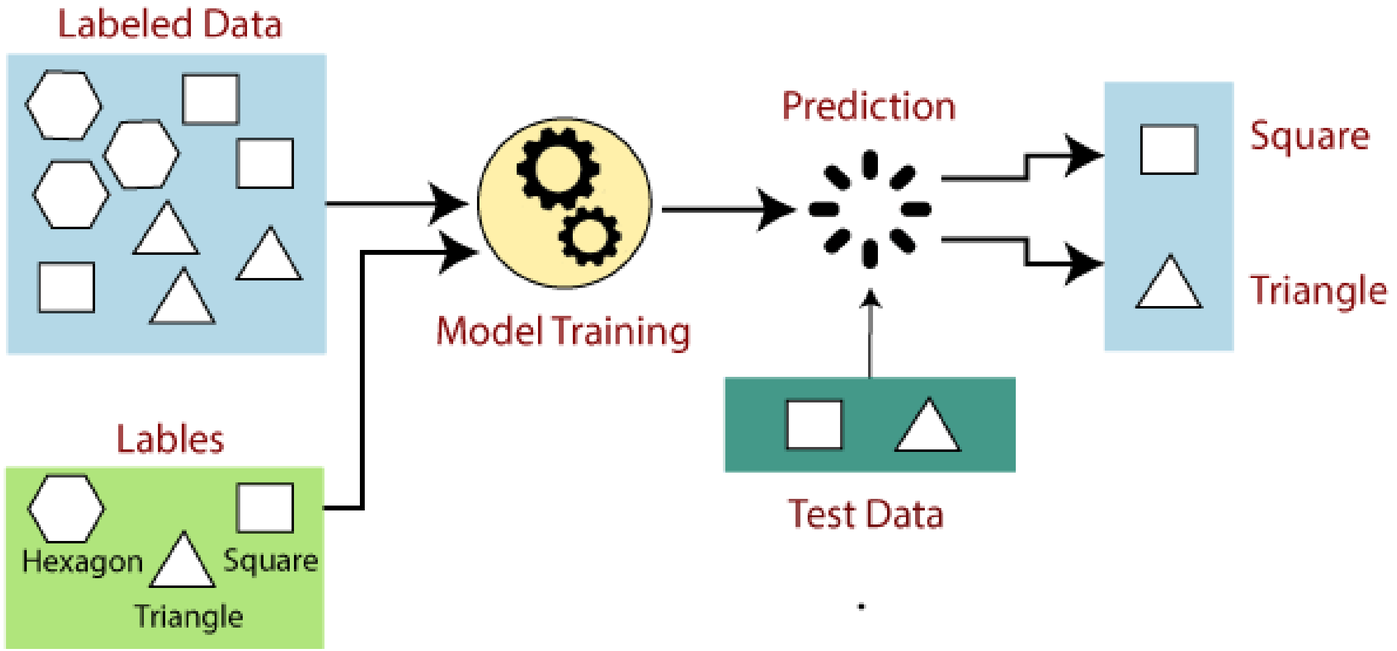
3. Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to **find a mapping function to map the input variable(x) with the output variable(y)**

In the real-world, supervised learning can be used for **Risk Assessment, Image classification, Fraud Detection, spam filtering**, etc.

SUPERVISED LEARNING

- **In supervised learning, models are trained using labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested on the basis of test data (a subset of the training set), and then it predicts the output.**

Example:



How Supervised Learning Works?

Suppose we have a dataset of different types of shapes which includes square, rectangle, triangle, and Polygon. Now the first step is that we need to train the model for each shape.

- If the given shape has four sides, and all the sides are equal, then it will be labelled as a **Square**.
- If the given shape has three sides, then it will be labelled as a **triangle**.
- If the given shape has six equal sides then it will be labelled as **hexagon**.

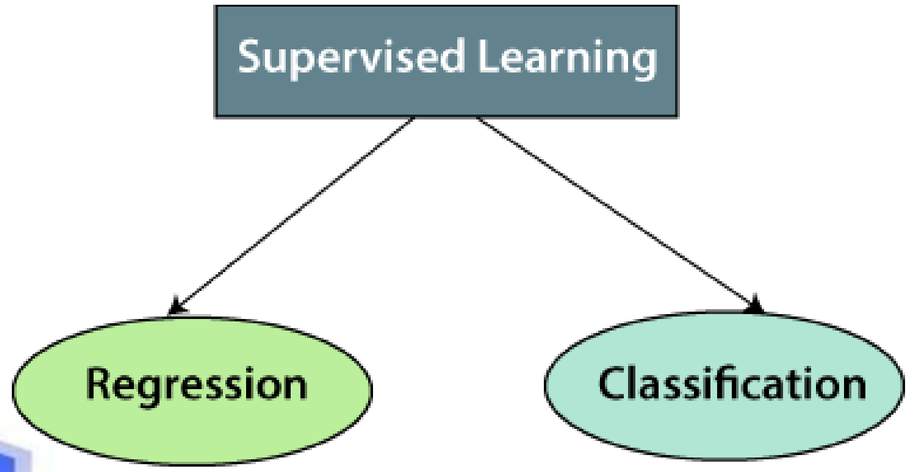
Now, after training, we test our model using the test set, and the task of the model is to identify the shape.

The machine is already trained on all types of shapes, and when it finds a new shape, it classifies the shape on the bases of a number of sides, and predicts the output.

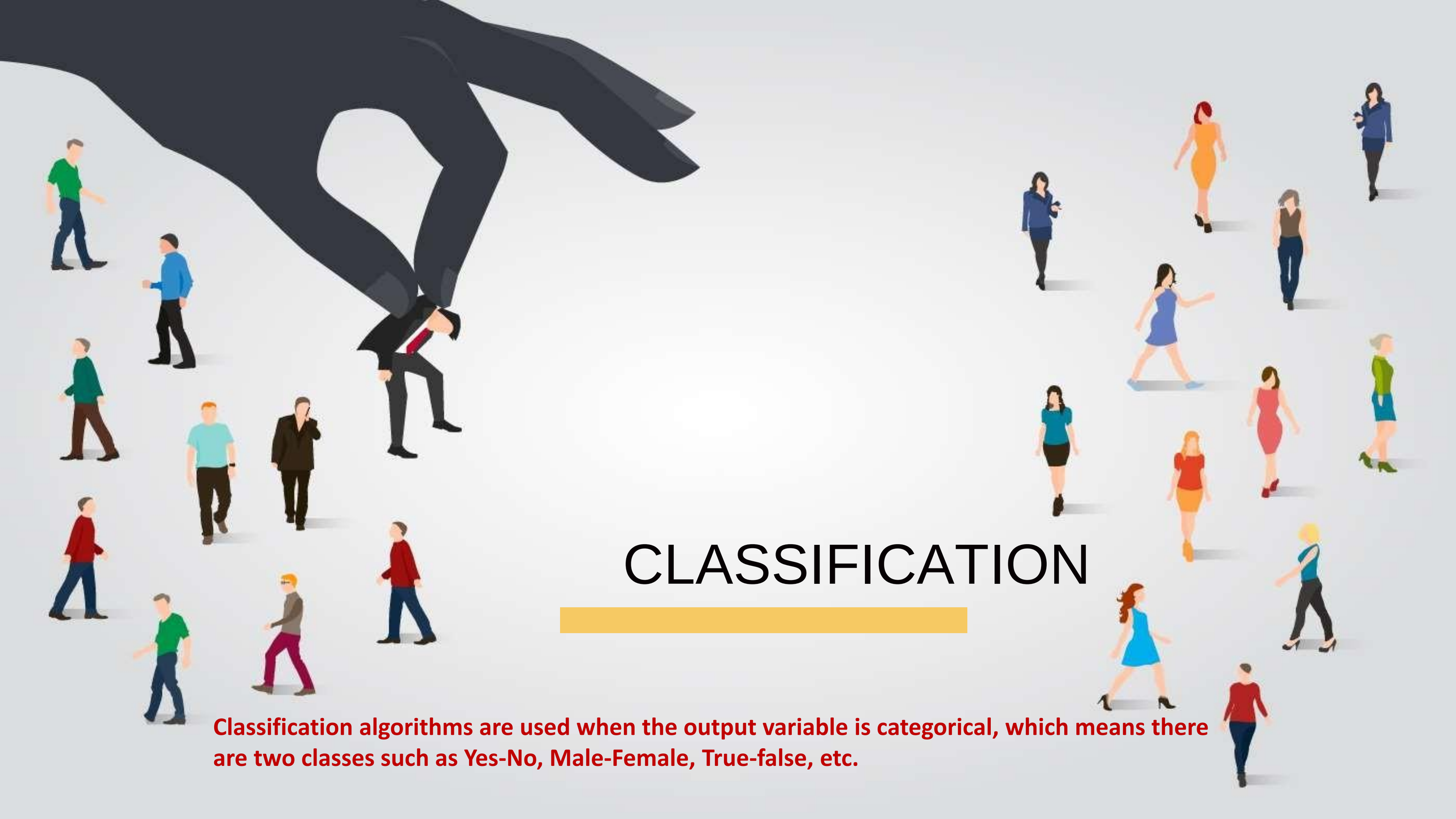
How Supervised Learning Works?

- **First Determine the type of training dataset.**
- **Collect/Gather the labelled training data.**
- **Split the training dataset into training dataset, test dataset, and validation dataset.**
- **Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.**
- **Determine the suitable algorithm for the model, such as support vector machine, decision tree, etc.**
- **Execute the algorithm on the training dataset. Sometimes we need validation sets as the control parameters, which are the subset of training datasets.**
- **Evaluate the accuracy of the model by providing the test set. If the model predicts the correct output, which means our model is accurate.**

Steps Involved in Supervised Learning



TYPES OF SUPERVISED LEARNING



CLASSIFICATION

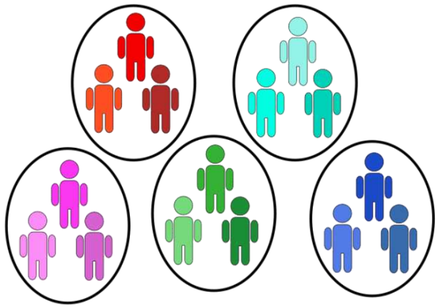
Classification algorithms are used when the output variable is categorical, which means there are two classes such as Yes-No, Male-Female, True-false, etc.

CLASSIFICATION AND ITS TYPES

- **Classification** is the process of predicting the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification predictive modelling is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (y).

Output Variable

Categorical variable



Classification

- Decision Tree
- Logistic regression
- Naïve Bayes
- Random Forest
- Bagging
- Boosting
- K Nearest Neighbour
- Support Vector Machine

Applications:

GMAIL

YOUTUBE

REGRESSION

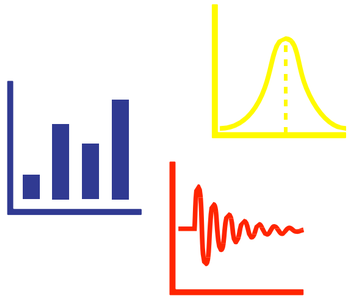
Regression algorithms are used if there is a relationship between the input variable and the output variable. It is used for the prediction of continuous variables, such as Weather forecasting, Market Trends, etc.

REGRESSION

- **Regression** is the output values based on input features from the data fed in the system. This algorithm builds a model on the features of training data and using the model to predict values for new data.

Output Variable

Numerical variable



Regression

Applications:

UBER

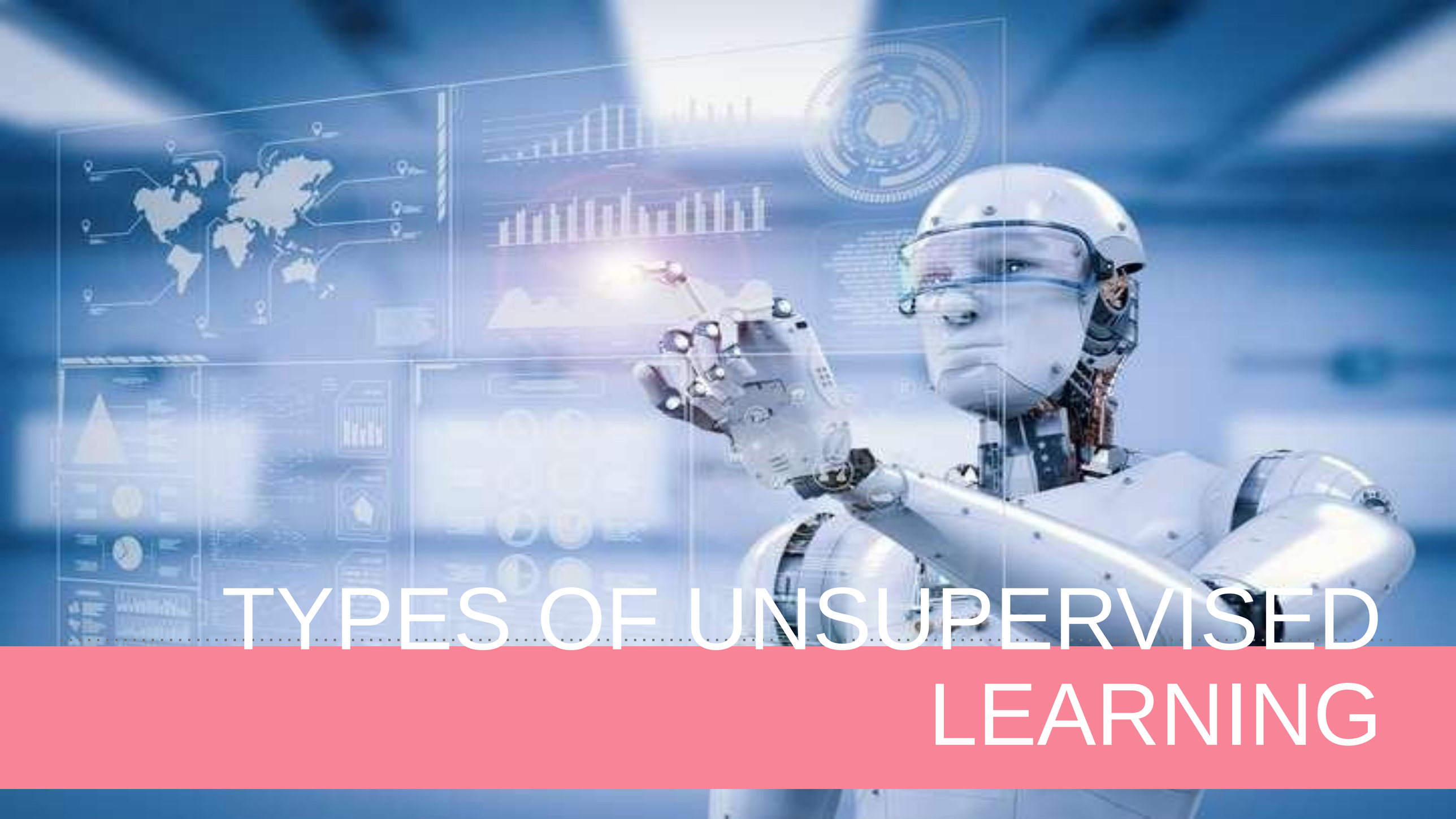
- Simple Linear Regression
- Multiple Linear Regression
- Categorical Regression
- Interaction Regression
- Polynomial Regression
- Ridge Regression
- Lasso Regression
- Elastic Net Regression

Advantages of Supervised learning:

- **With the help of supervised learning, the model can predict the output on the basis of prior experiences.**
- **In supervised learning, we can have an exact idea about the classes of objects.**
- **Supervised learning model helps us to solve various real-world problems such as fraud detection, spam filtering, etc.**

Disadvantages of supervised learning:

- **Supervised learning models are not suitable for handling the complex tasks.**
- **Supervised learning cannot predict the correct output if the test data is different from the training dataset.**
- **Training required lots of computation times.**
- **In supervised learning, we need enough knowledge about the classes of object.**



TYPES OF UNSUPERVISED LEARNING

CLUSTERING

In basic terms, the objective of **clustering** is **to** find different groups within the elements in the **data**.

To do so, **clustering algorithms** find the structure in the **data** so that elements of the same **cluster** (or group) are more similar **to** each other than **to** those from different **clusters**.

ASSOCIATION

Association Rule Mining is one of the ways to find patterns in data. It finds:

Features (dimensions) which **occur together**

Features (dimensions) which are **“correlated”**

CLUSTERING

- K Means
- Hierarchical

Application:

Clustering similar products

Myntra amazon Flipkart

ASSOCIATION

- Association Mining
- A-priori Algorithm
- Market Basket Analysis

Application:

Recommending similar products

Myntra amazon Flipkart

AI

Rule Based AI

- Rules are known
- Rules are given to machine [if-then-else]
- Flawless

ML Based AI

- Rules are Unknown
- Lots of data is given to the machine and machine will figure out the rules.
- Error Prone.

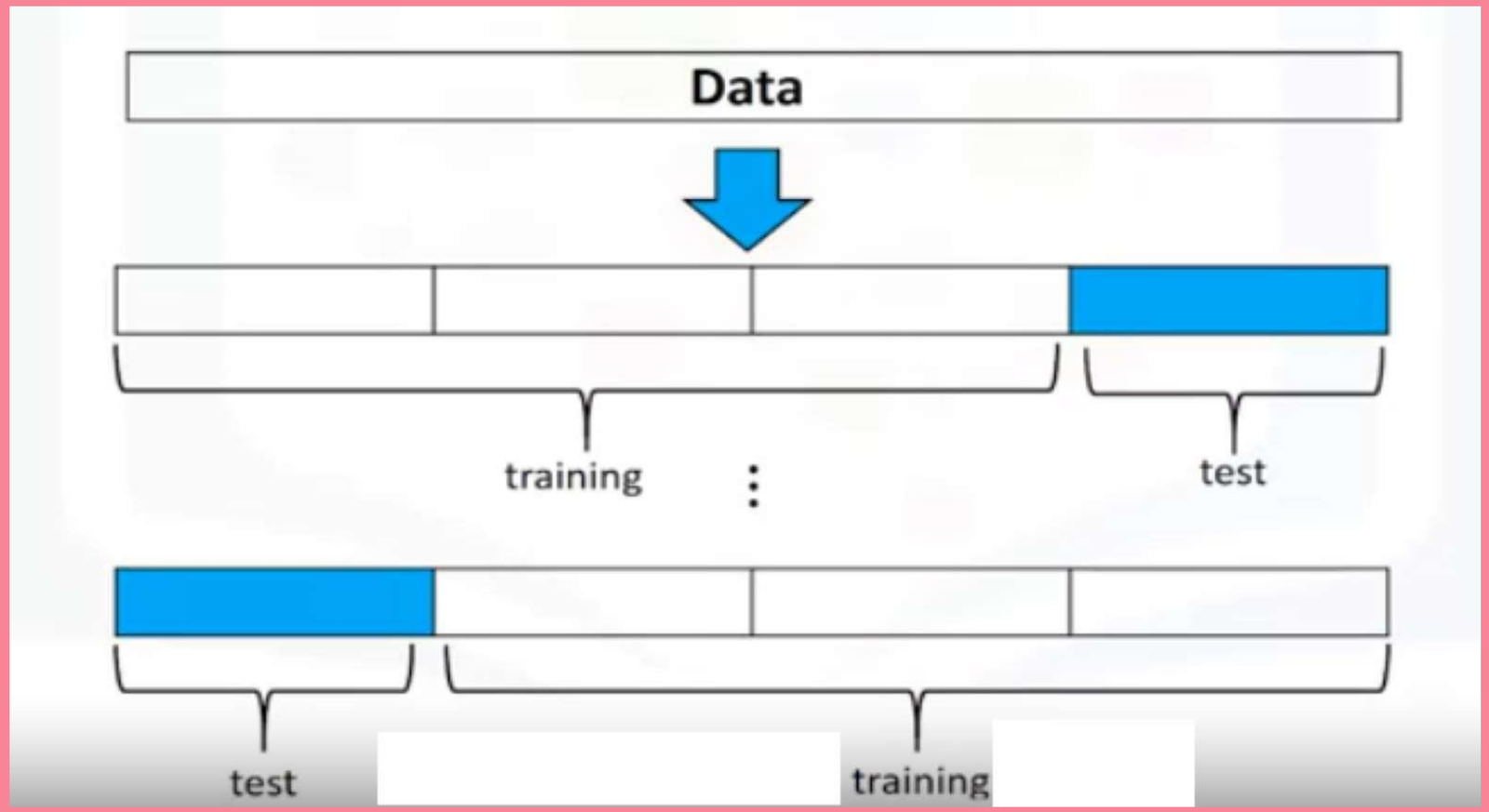
Algorithm

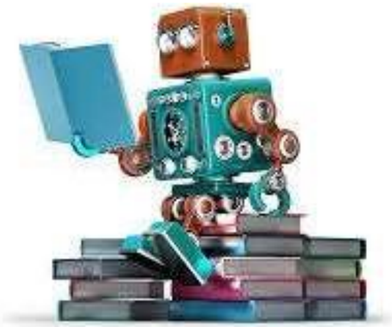
Data————————> Model /Rules/Functions

CROSS VALIDATION

A technique to **evaluate predictive models** by partitioning the original sample into a training set to train the model, and a testing set to evaluate it.

SPLITTING OF DATA SET





DATA SET

- **Training Set:** Used for making the algorithm learn

- **Testing Set:** Used to give input to the algorithm and check against the actual output.

Evaluation Metrics

Some of the **evaluation metrics** for classification are:

1. **Confusion Matrix**
2. **Precision (Specificity)**
3. **Recall (Sensitivity)**
4. **Accuracy**
5. **F1-Score**
6. **TPR (True Positive Rate)**
7. **FPR (False Positive Rate)**
8. **ROC-AUC Curve**

Confusion Matrix

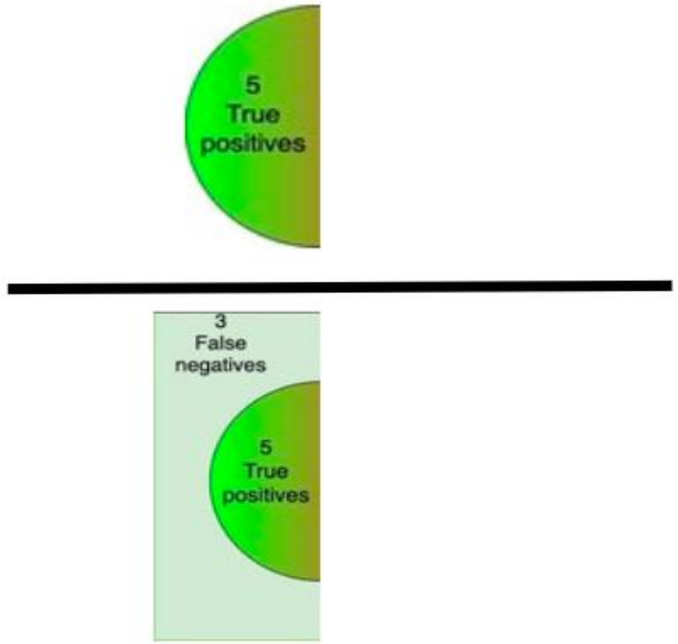
Predicted Values	Actual Values: Positive (1)	Actual Values: Negative (0)
Positive (1)	TP	FP
Negative (0)	FN	TN

Predicted values	Actual values: Apple	Actual values: Orange
Apple	5 True positives	2 False positives
Orange	3 False negatives	3 True negatives

Recall (Sensitivity)

Recall is another important metric, which is defined as the fraction of samples from a class which are correctly predicted by the model. More formally:

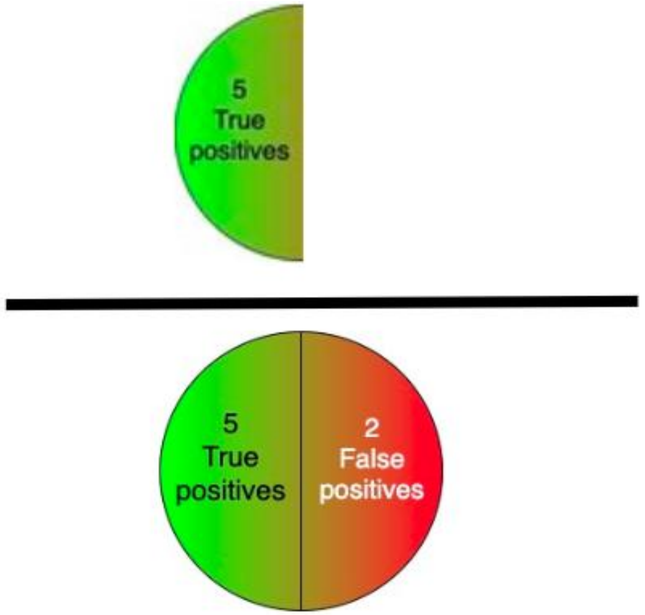
Recall = TruePositive / (TruePositive + FalseNegative)



Precision (Specificity)

There are many cases in which classification accuracy is not a good indicator of your model performance. One of these scenarios is when your class distribution is imbalanced (one class is more frequent than others).

Therefore we need to look at class specific performance metrics too. Precision is one of such metrics, which is defined as:



Precision = TruePositive / (TruePositive + FalsePositive)

Other popular distance measures

- **Hamming Distance:**

Calculate the distance between binary vectors.

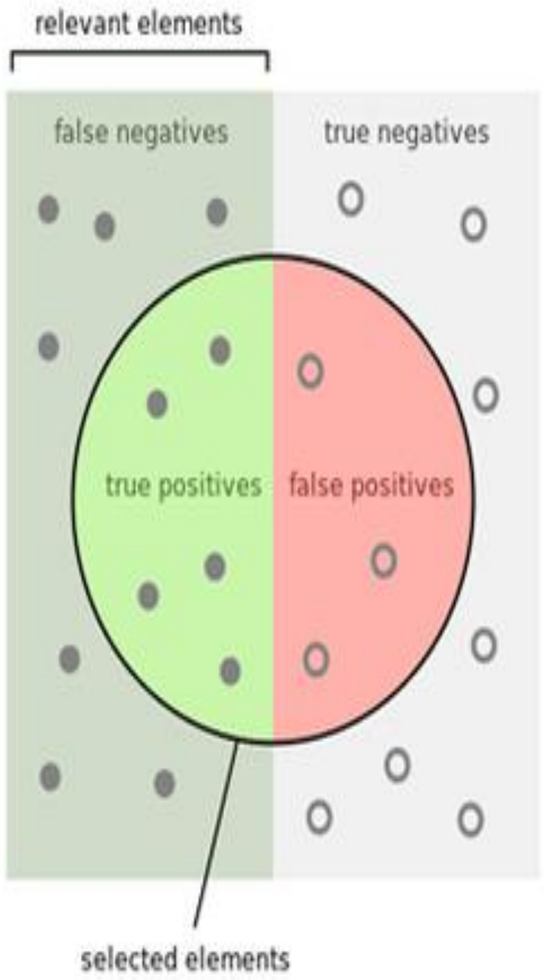
- **Manhattan Distance:**

Calculate the distance between real vectors using the sum of their absolute difference. Also called City Block Distance.

- **Minkowski Distance:**

Generalization of Euclidean and Manhattan distance

Precision and Recall



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{selected elements}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{relevant elements}}$$

Accuracy and F1-Score

Classification accuracy is perhaps the simplest metrics one can imagine, and is defined as the **number of correct predictions divided by the total number of predictions,** multiplied by 100**.**

Accuracy = (TruePositive + TrueNegative) / Total

Depending on application, you may want to give higher priority to recall or precision. But there are many applications in which both recall and precision are important. Therefore, it is natural to think of a way to combine these two into a single metric. **One popular metric which combines precision and recall is called F1-score**, which is the harmonic mean of precision and recall defined as:

F1-score = 2*Precision*Recall / (Precision+Recall)

True Positive Rate and False Positive Rate

TPR = TruePositives / Number of Positives

= TruePositives / (TruePositive + FalseNegative)

FPR = FalsePositives / Number of Negatives

= FalsePositives / (FalsePositive + TrueNegative)