Topic - 01
Introduction to Machine Learning:
Concept & Fundamentals

+88 01831-661534

jehadfeni@gmail.com

What is Artificial Intelligence?

Any artificial device or mechanism that is capable of self learning is in reality an Artificial Intelligent being and the technology that empowers that Artificial being is Artificial Intelligence

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

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.

Examples:
Virtual Assistant (Siri, Alexa, Cortana)
Traffic Prediction (Google)
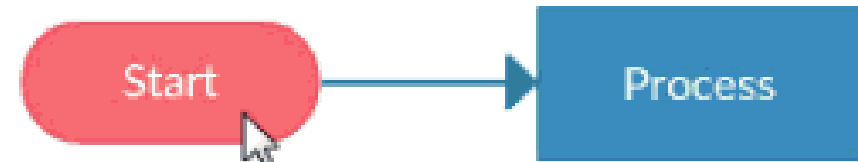Face Recognition (Facebook)
Emails (Spam and not Spam)

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.
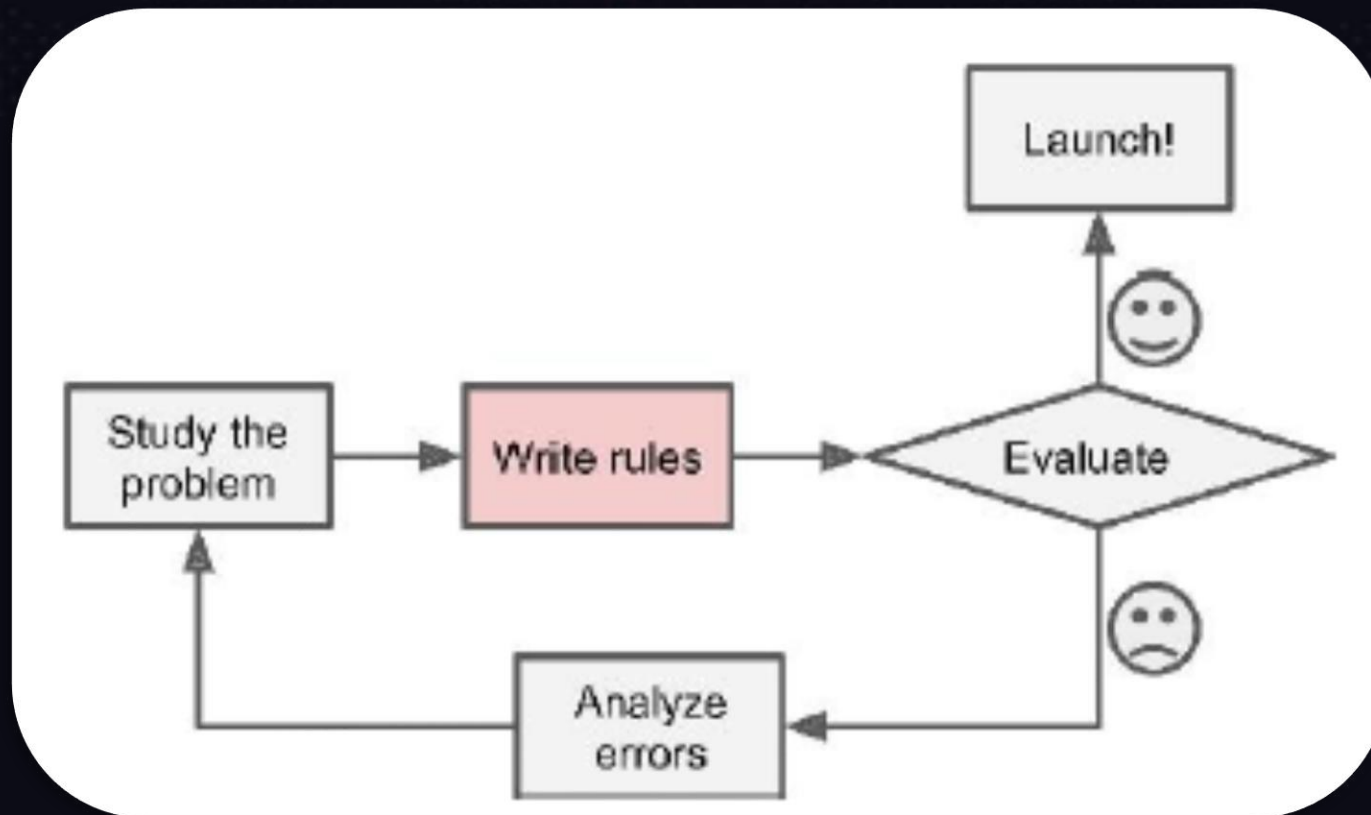
Machine Learning is all about Probability.

Probability of something happening with lots of conditions!!!

We get lots of Spam emails from time to time from various sources. If we had to write a Spam Filter program with traditional conditional method, the program would look like the following:
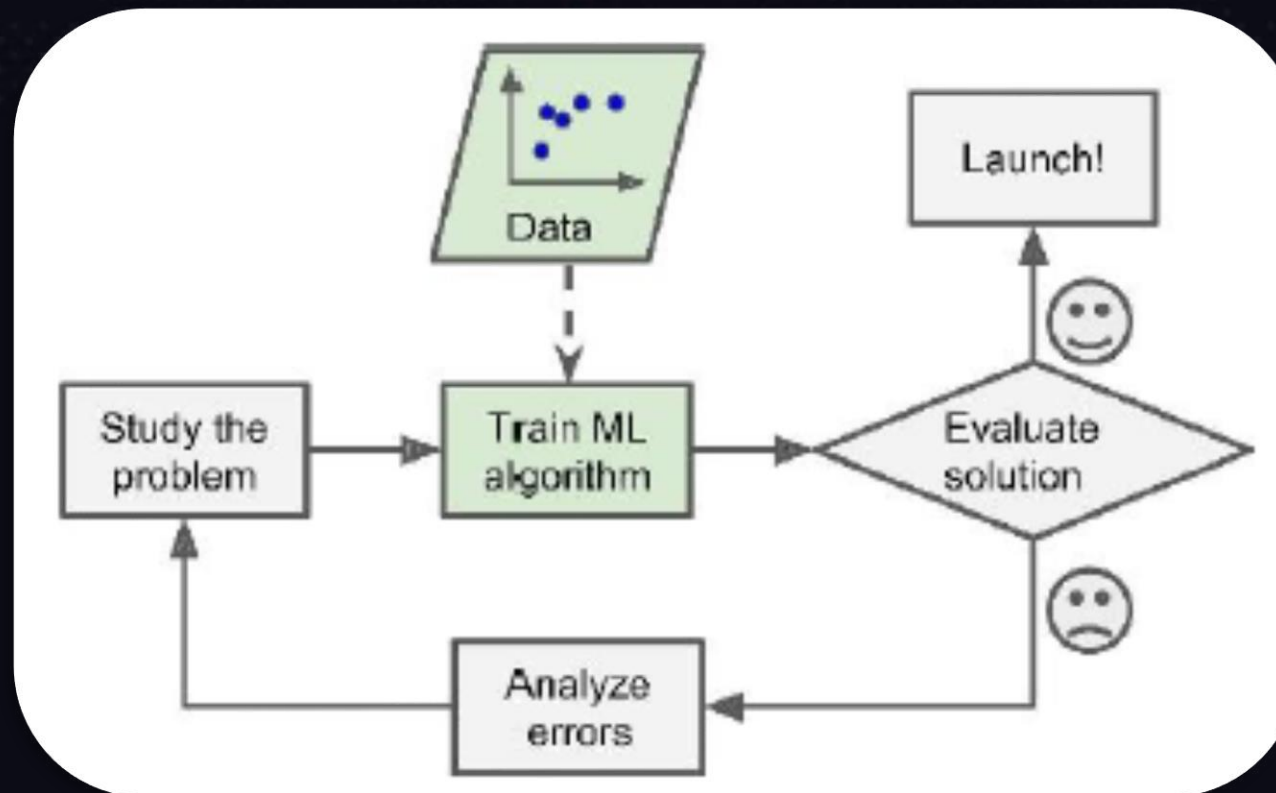

@MJS Global Ltd. by ABK Bhuiyan

- First we will create a list of words or phrases that are commonly available in a Spam Email.
  For example: Offer, Bonus, Free Money, 4U, Lottery, Credit Card, Amazing, Gift, Won etc.
- Then we will create a login when these words are found, we will call an email SPAM, otherwise HAM (Not SPAM)
- Make lots of modifications in conditions and keep repeating the above 2 processes.

But the problem with the Conditional Approach is that, it will contain lots of rules and you will have to change rules each and every time a new kind of SPAM email comes with different word variations.
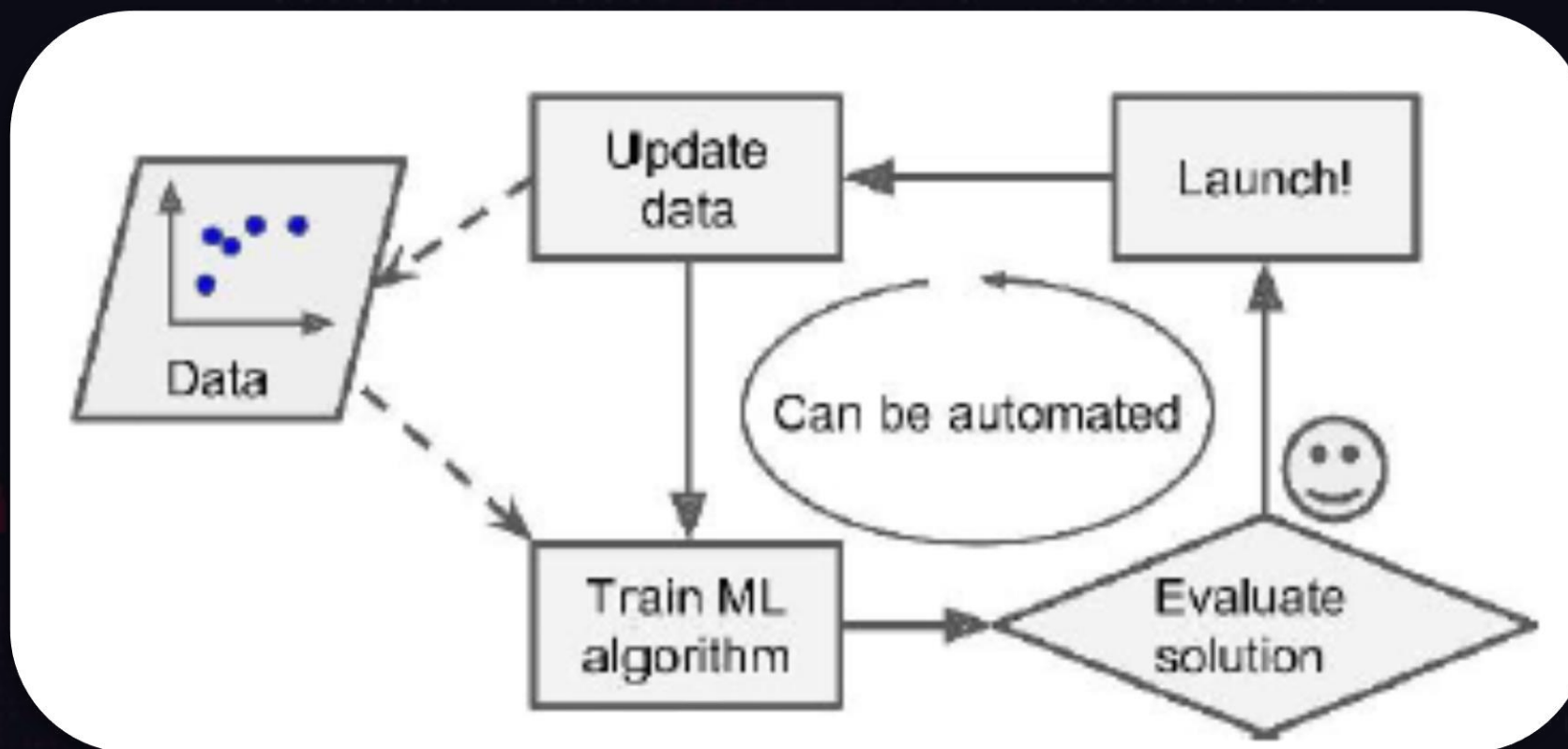**For example**: Instead of **4U**, using **For You**

Spam Filter based on Machine Learning can learn from all the marked emails (SPAM & HAM) and when a new email comes, it can automatically detect which email is SPAM and which email is HAM.
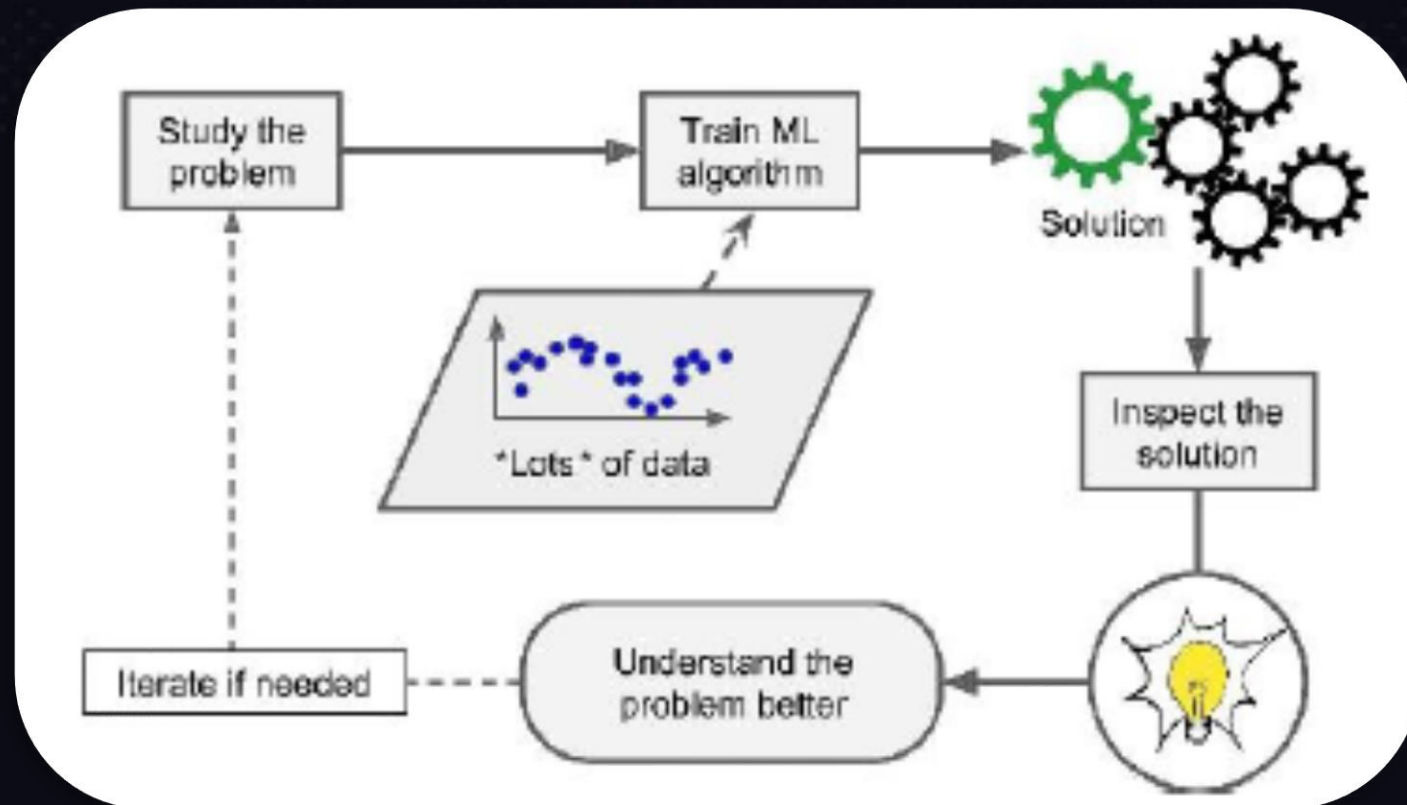
And even though some email type was not marked SPAM earlier, if they are later on Marked SPAM from several users, it can also learn from there! You don't have to write new rules for new SPAMs!!! ☺

As Machine Learning can Flag emails for SPAM, it can help us to also find the words or patterns for which SPAM is happening! Applying Machine learning in data can help finding NEW Patterns!

In case of problems, where there is no available algorithm or are too difficult for traditional approach to overcome, Machine Learning can be applied to solve those problems!

Example 1: Speech Recognition, where the whole thing is measured by Voice Spectrogram (ছয় and নয় – have some similarities but the Pitch gets lot intensified in case of ছয়) and different people speak in different way in different situations

Example 2: Generating Automated Answers from Questions. Even though human understands which question is what, for Machine to understand that is very difficult – আমার ব্যালেন্স কত, আমার বালান্স, বালান্স কত, balance koto, blance koto, balance kto, belance koto, valance khoto, vlance koto – lot many variations can be obtained for the same problem

- Solving problems where you need to write lots of rules – less code, less time and better performance

- Solving complex problems where there is no good solution available

- Getting adapted to new kind of data – it can be known data, it can be unknown data

- Getting a good idea on problems and measure the data that are being used

- Getting insights about problems

- Solving problems which were impossible before

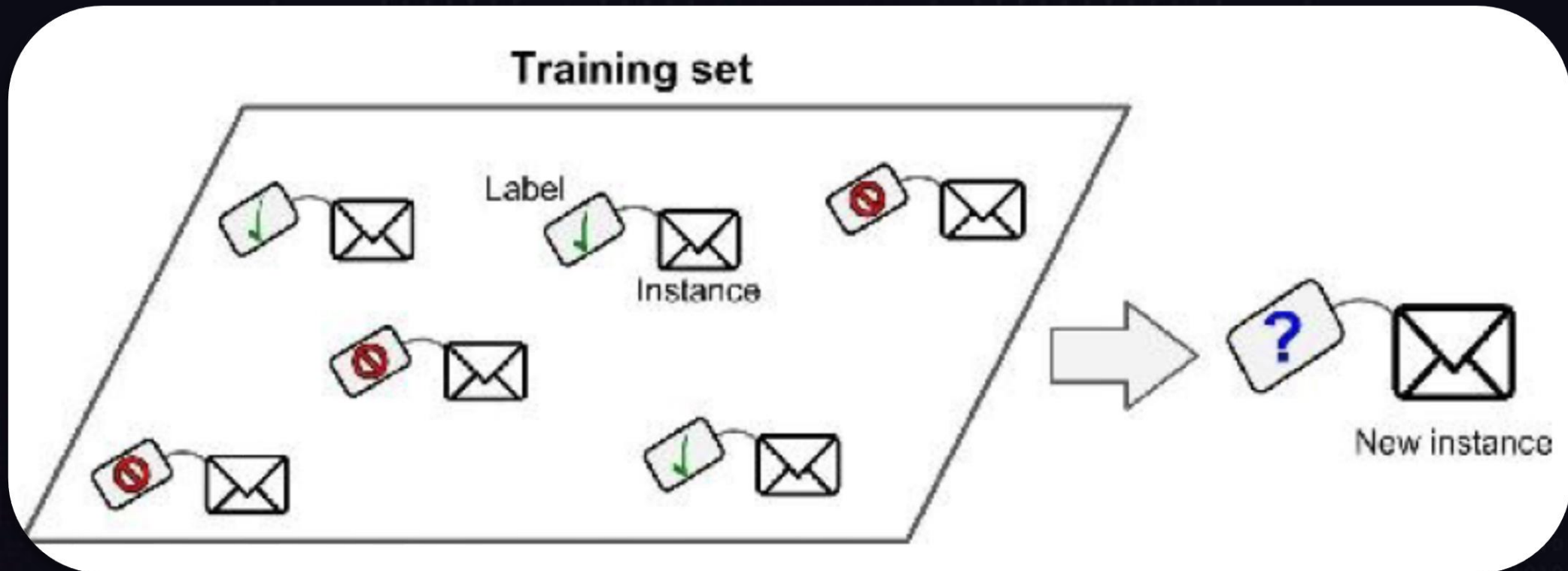Based on Data Supervision, Machine Learning can be divided into:

Supervised
Learning

Unsupervised
Learning

Semi-supervised
Learning

Reinforcement
Learning

In case of supervised learning, the solution of the problem is also given to the algorithm in the time of Training. Organizing data with their desired value is called "Labelling" and the data as "Labelled Data"

Let's say we want to measure the Price of houses in an area based on:
Number of rooms, House rent, Age, Neighborhood and few other matrices.
In order to apply the Supervised Learning process, we will give lots of House
samples (data) with all the matrices as well as the Price as input and then
Train the System (Training Data).

After successful training, we can test our Accuracy with new Data samples
(Testing Data/Validation Set).

Based on the data and everything, we can measure how good our
Supervised Model is performing.

**Among all the Machine Learning Approaches based on Data,
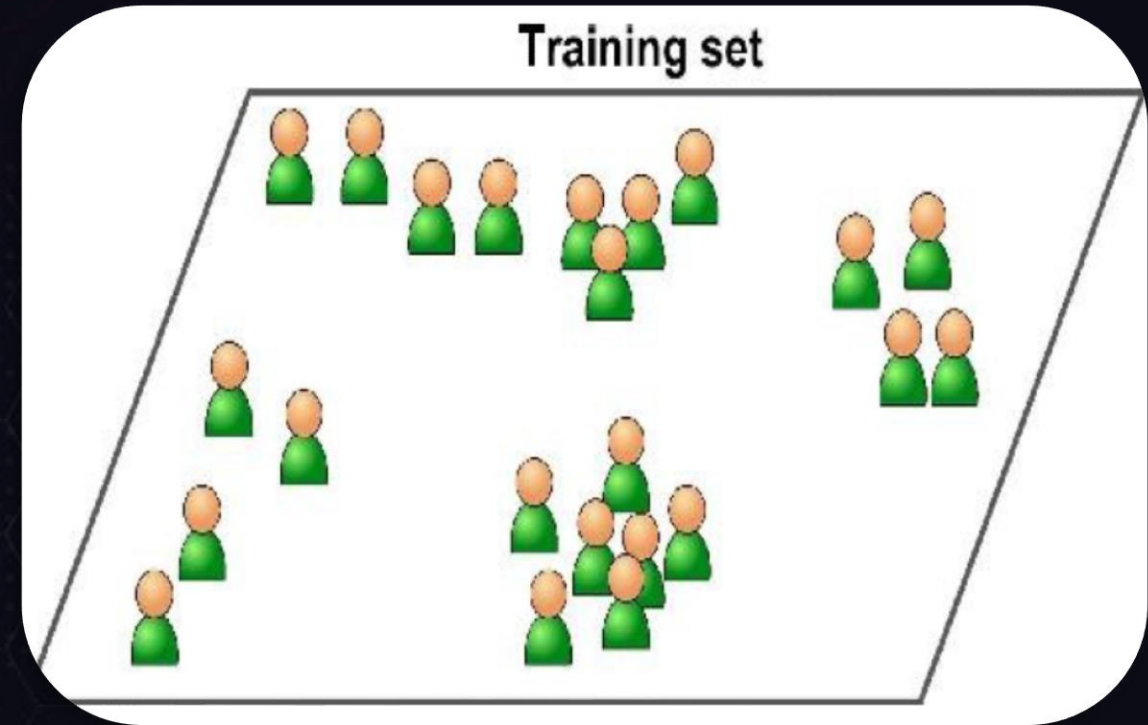Supervised Learning is widely used.**

- k-Nearest Neighbors

- Linear Regression

- Logistic Regression

- Support Vector Machines (SVMs)

- Decision Trees

- Random Forests

- Neural networks

In case of unsupervised learning, the solution of the problem is **NOT** given to the algorithm in the time of Training. As a result, no pre-defined labelled data is given into the System. The system finds pattern from BIG Data available and tries to solve the problem.

Human Example: You have joined into a library but you don't know anything about Library and book structures. Now, your supervisor has given you the task of sorting the books and putting them properly in the shelf. Now, you need to learn the right way of doing that.

Training set

- Clustering based:
  - ✓ k-Means
  - ✓ Hierarchical Cluster Analysis (HCA)
  - ✓ Expectation Maximization
- Visualization and dimensionality reduction
  - ✓ Principal Component Analysis (PCA)
  - ✓ Kernel PCA
  - ✓ Locally-Linear Embedding (LLE)
  - ✓ t-distributed Stochastic Neighbor Embedding (t-SNE)
- Association rule learning
  - ✓ Apriori
  - ✓ Eclat

Self Task

- Watch Movie:

  ✓ HER (by 19.01.2019)

- Watch TV Series

  ✓ Silicon Valley (No time frame)