

# Lec. IX. Intro to M.L.

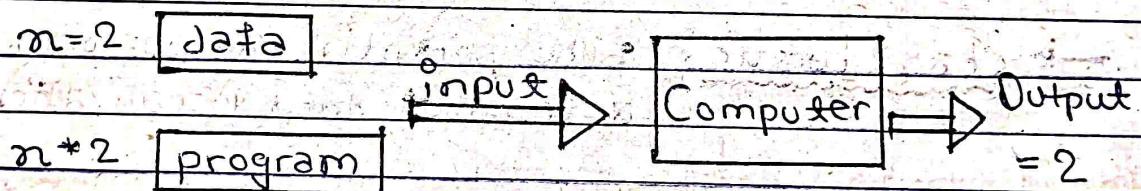
## Introduction to Machine Learning:

→ What is Machine Learning?

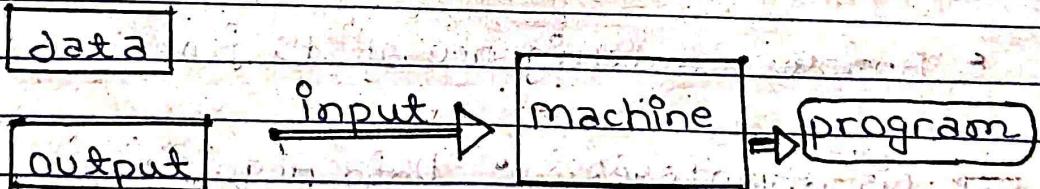
M.L. is the science of getting computers to learn and act like humans do & improve their learning over time in autonomous fashion, by feeding them data & information in the form of observations and real-world interactions.

→ How M.L. is different from traditional programming?

- traditional - programming -  $\log \approx n^{\alpha}$  of a number ( $n$ )



- machine - Learning - automated process-

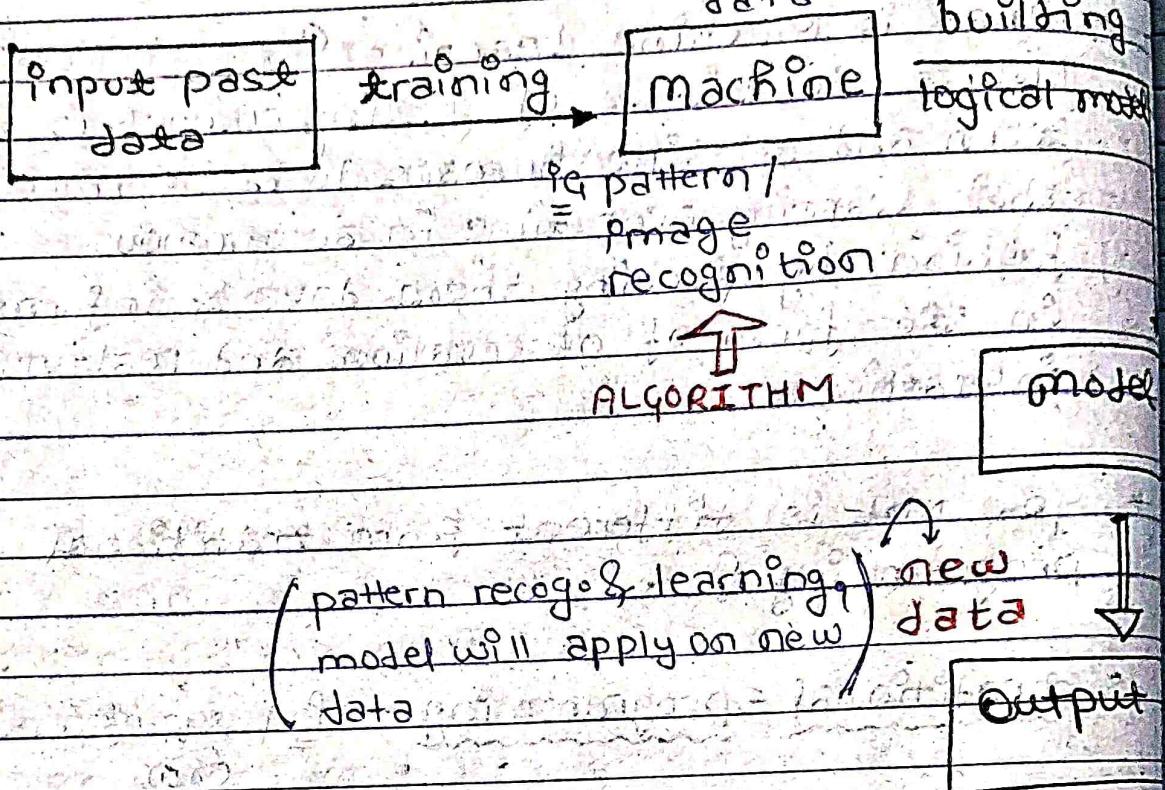


e-commerce → visit some website → next time category visit

(data = fuel of machine)

Show prev  
visit related  
category

## • How it Works?



→ characteristic : • The potential to perform automated data visualization

- precise data analysis (efficiently analysis of heterogenous data)

- Business Intelligence at its finest.

→ Dis-advantages : • Data Acquisition

(collecting hectic data & providing quality data to machine)

- Highly error prone (if encounter, incorrect, invalid, missing-value data)

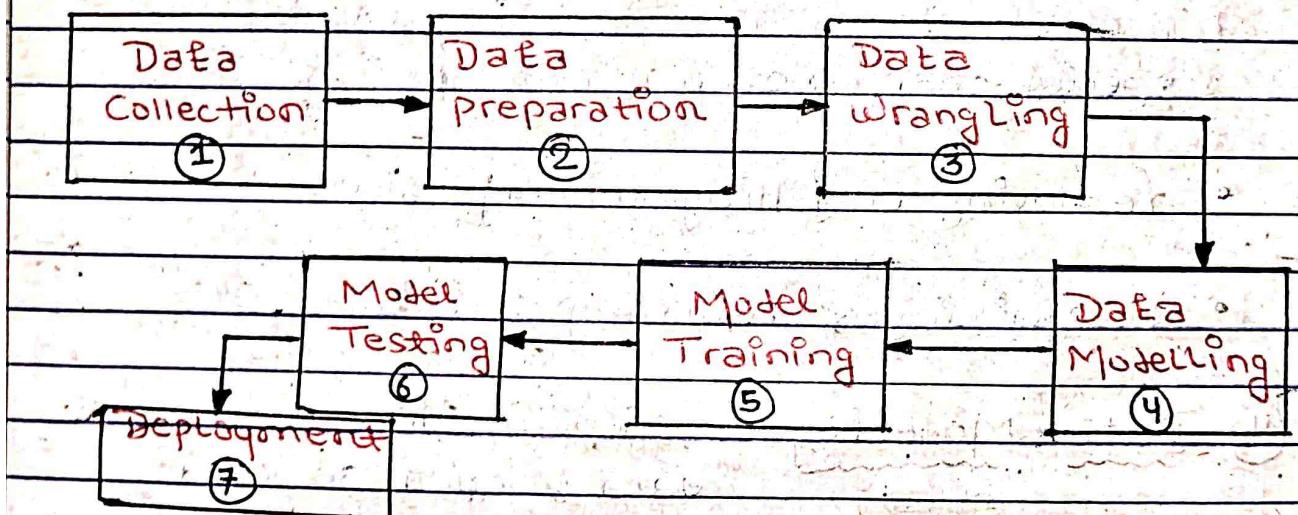
- Algorithm Selection

- Time Consuming (time, money, cost losses if fails)

## → Trending applications of ML:

- Recommendation Engines (ie, Netflix)
- Self-Driving Cars
- Gamified Learning & Education
- E-commerce websites
- Medical Diagnosis
- Getting your right answers (ie, Quora).

## → Machine Learning Life Cycle :



① Data-Collection : • Goal over here is to gather as much as relevant data as possible.

- Identify various sources of information
- Gather data (Data Engg. makes data useful for Data Scientists)
- Combine the data acquired from various data sources

② Data-Preparation: exploration of your data to generate better result

③ Data-Wrangling: cleaning & converting data into usable format (json, csv, audio etc)

- filtering | cleaning up of raw data

- filtering noise.

(incorrect data-values, missing values, special char, not required repetitions)

- removing or filling missing values

- recognizing & removing Outliers

④ Data-Modeling: step in which we take data & build select a ML algo. to build model.

- selecting ML algo.

- building the model

- validating the results

⑤ Model-Training: A ML training model is a process in which a ML algo. is fed with sufficient training data to learn (pattern engineering)

⑥ Model-Verification: Checking for the accuracy of model by providing with inputs that are unseen,

⑦ Model-deployment is ready to go for production and modelling

AI vs ML vs DL

A.I. is the science & engineering of making computers capable of performing tasks that typically requires Human Intelligence.

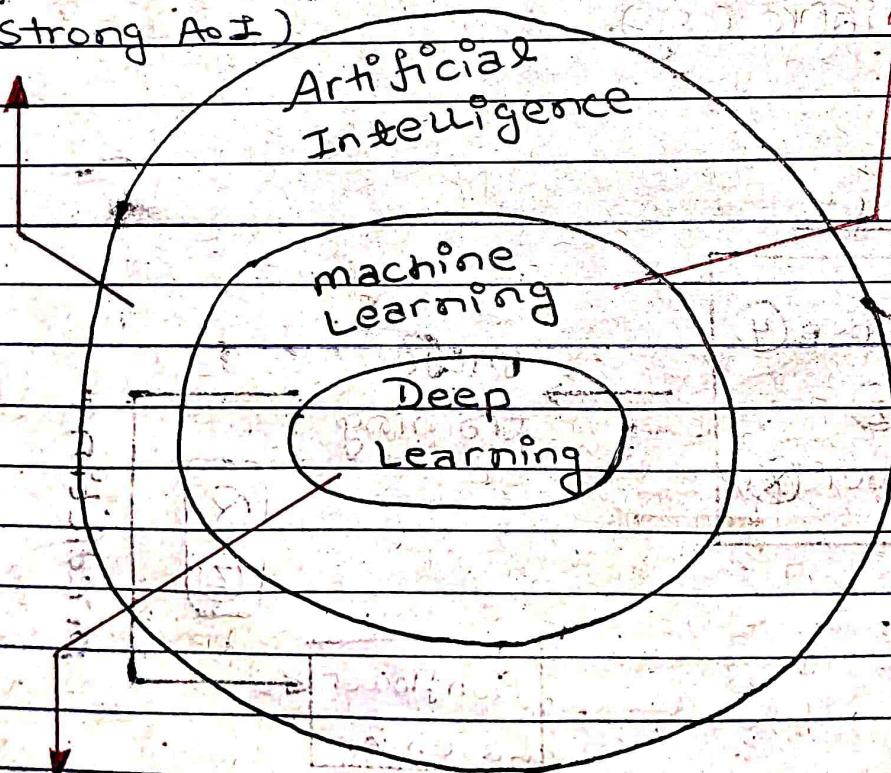
• generalized A.I. (Weak A.I.)  
• Strong A.I.)

Learning is a subset of A.I. which enable machine to

learn from

past data or experience without being explicitly programmed

- supervised
- unsupervised
- reinforced



Lib →  
scikit learn  
seaborn  
scipy etc

Deep learning is a subset of ML concerned with the algos. Inspired by the function & structure of Human Brain.

(Libraries - PyTorch,  
Tensorflow)

## → What is Supervised Learning?

In supervised learning you train the machine using data which is well labelled i.e., some input data is already tagged with correct answer and this algorithm learns from labelled training data that helps you to predict the further outcomes.

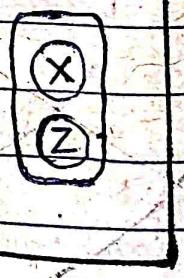
## ○ How supervised learning works?

The goal of supervised learning is to map the input variable ( $x$ ) with the output variable ( $y$ ).

### Labelled dataset

sunflower (X)	rose (Y)
lavender (Z)	
rose (Y)	sunflower (X)

model  
training



Sunflower  
Lavender

- Model got trained, what kind is of data (training-phase)

- how, we provide new data-set and analyze model through result from labelled dataset,

# Supervised ML Algos

## Regression

- linear regression
- polynomial regression
- regression trees

## Classification

- random forest
- decision trees
- logistic reg.
- support vector machines

• Advantages — • you have full control over what the machine is learning.

- you can easily test and debug your model
- you can determine the nos of classes

• Disadvantages — • have limited scope

- Collecting labelled dataset is expensive and time-consuming
- wrong prediction

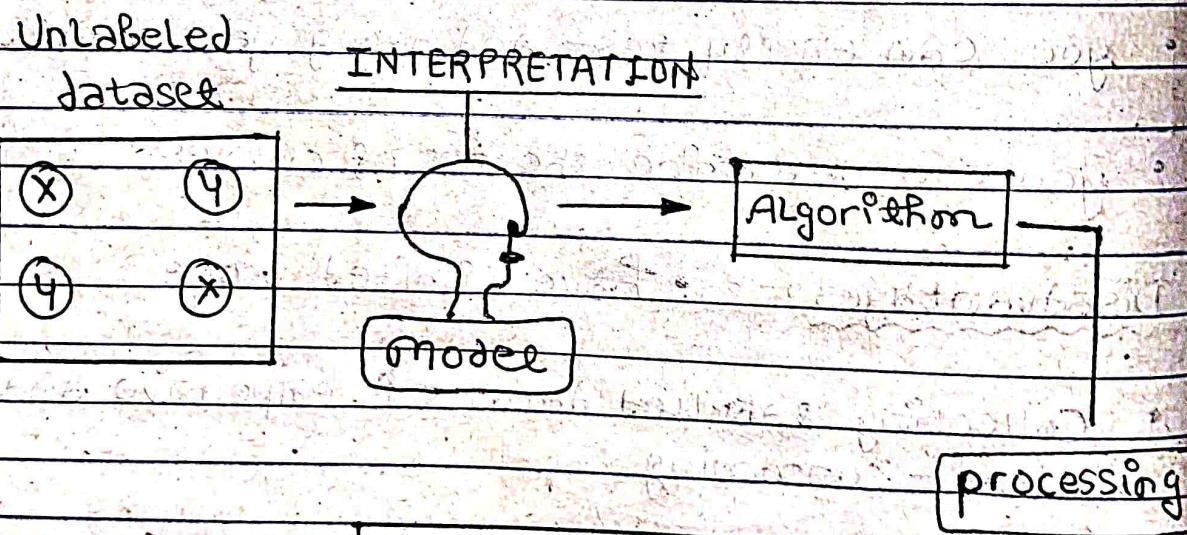
In supervised learning you train the Machine using data which is well labelled

## → What is Unsupervised learning?

In unsupervised Learning, you train the machine using data itself which is "unlabelled" and model itself finds the hidden patterns and insights from the given data.

### ○ How does Unsupervised learning works?

The goal of Unsupervised learning is to group unlabelled data according to their similarities, patterns and differences without any prior training of data.



Learn from objects, similarities & differences & group them

(e.g roses → red, sunflower → yellow etc).

Sunflowers	X	X
roses	Y	Y

## Unsupervised M.L Algos.

### Clustering

raw data

(X)	(Y)	
(Z)	(X)	
(Y)	(Z)	(Y)
(X)	(Z)	

### Algorithm

Output

(Y, Y, Y) (Z, Z) (X, X)

B + Bu + F  
E + B + M  
C + Co + Ch  
B + Bu + M

B + Bu

most frequent items

- Advantages -
  - It's used for more complex task.
  - helpful for finding patterns in data.
  - save human errors as no training by them.
  - save a lot of manual work & expense.
- Disadvantages -
  - Less Accuracy
  - Time Consuming
  - more the features, more the complexity