Table Of Contents

✳ Machine Learning

Introduction:-

→ At a high-level, machine learning is simply the study of teaching a computer program/ algorithm how to progressively improve upon a set task that it is given.

So, more practically it is the study of how to build application that exhibit this, iterative improvement.

there are many ways to implement it and largely there are 3 major categories.

1) supervised Learning.

2) unsupervised Learning.

3) Reinforcement learning.

And also there is semi-supervised learning.

machine learning algorithms are the 'engine of machine learning', meaning it is the algorithms that turn a dataset into a model.

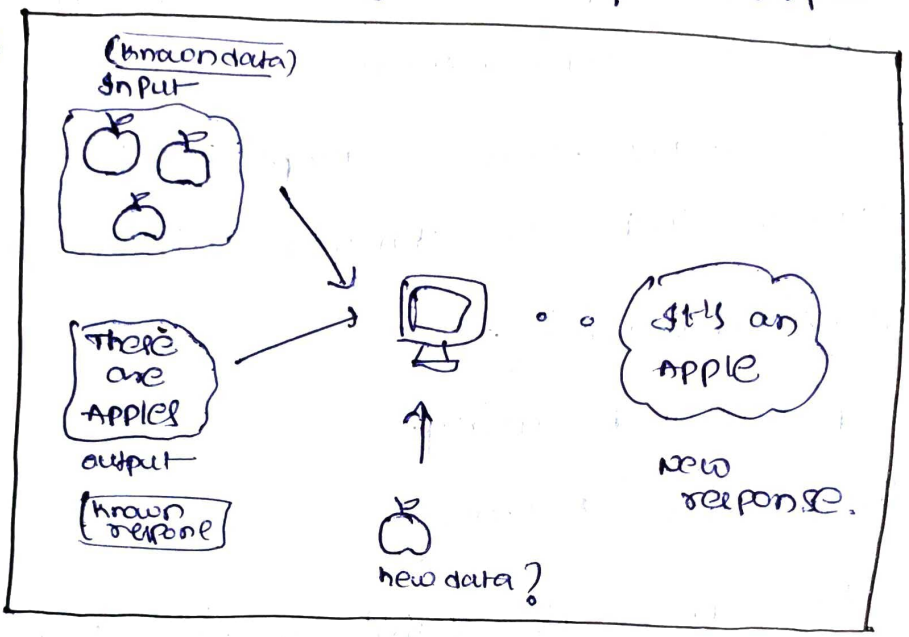
→ which kind of algorithm works best and what to choose (supervised, unsupervised, classification, regression, etc..) depends on the kind of problem you're solving, the computing resources available & the nature of the data.

Briefing the types:-

(1) Supervised learning

Supervised learning is the most popular paradigm for performing machine learning operations. It is widely used for data where there is a precise mapping b/w input-output data.

i.e;



→ Given data in the form of examples with labels, we can feed a learning algorithm these 'example-label' pairs one by one. overtime, the algorithm will learn to approximate the exact nature of the relationship b/w examples & their labels. (It's called training the data)

→ when fully trained, the supervised learning algorithm will be able to observe a new, never-before seen example & predict a good label for it.

→ And so supervised algorithms are called 'task-oriented'. As we provide it with more & more examples, it is able to learn more & more properly so that it can undertake the task & yield us the output more accurately.

It is exhibited in many of following applicatns

Face Recognition:-

Face lock app in our phone has been using a supervised learning algorithm that it is trained to recognize your face. Having a system that takes a photo, finds faces & guesses who that is in the photo (suggesting a tag) is a supervised process.

Spam classifications:-

Spam filter is a supervised learning system. Feed email examples & labels (Spam/not-Spam) these systems learn how to filter out malicious emails so that the user is not harassed.

Advertisement popularity:-

Selecting advertisements that will perform well is often a supervised learning task. Many of ads you see as you browse the are placed there because a learning algorithm said that they would be reasonable popularity (more clicked).

and many examples...

Types of Supervised Learning

1) Regression
2) classification
3) Neural networks
4) Boosting techniques.

→

Regression

→ Regression algorithms predict a 'continuous value' based on the input variables.

Regression algorithms

1) Linear Regression
2) Polynomial Regression
3) Exponential Regression
) Logarithmic Regression

Applications

① Risk Assessment
② Score prediction etc.

④ ——— Quiz ——— Temp Predict

classification

→ we use classification algorithms for predict set of items class or category.

classification algorithms

Logistic Regression
1) K-nearest neighbours
2) Decision trees
3) Random forest
4) Support vector machine (SVM)
5) Naive Bayes
6) Ada boost
7) XG boost

Applications

① Fraud detection
② Email spam detection
③ Image classification
④ Diagnosis etc...

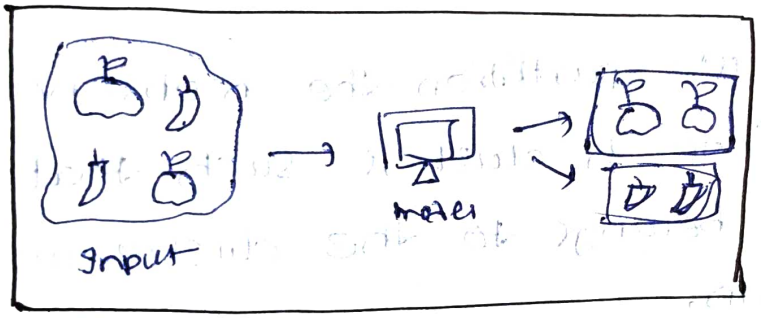
Hot | cod

② unsupervised Learning

(Based on Hebbian learning)

→ To easy understand, in case of unsupervised learning algorithm, the data is not explicitly labeled into diff. classes, that is no labels. The model is able to learn from the data by finding implicit patterns.

→ unsupervised learning algorithms identify the data based on their density / structures, similar segments & other similar features.



Types of unsupervised learning :-

1) clustering
2) Dimensionality reduction & visualization
3) Anamomaly detection
4) Association
5) Auto encoders

1) Clustering :-

Clustering, also known as cluster analysis, is a technique of grouping similar sets of objects in the same group that is diff from the objects in other group.

→ Some of the essential clustering techniques are as follows:-

a) K-means

In this we partition the n observations in the data into K cluster such that each observation belongs to the cluster with nearest mean.

b) DBSCAN

It groups the data based on density. It groups together the points that are given in space & mark the outliers in low density regions.

c) Hierarchical clustering

A hierarchy of cluster is built.

d) Mean Shift

2) Dimensionality Reduction & visualization:

a) Principal Component Analysis (PCA):

Reduce data from more dimensions to lower dimensions while attempting to preserve the variance.

→ we Reduce the size of data to extract useful information.

b) t-Distributed Stochastic Neighbor Embedding (t-SNE):

It is used to visualize high-dimensionality data in a 2D / 3D space.

---

3) Anomaly detection:

Anomaly detection techniques detect outliers in the unlabeled data under an assumption that most of the data examples are normal. by observing the instances that fit the remainder dataset.

one-class classification:

Train a model on only one-class, if anything lay outside of this class, it may be an anomaly.

Algorithms for doing so include..

(i) one-class k-means
(ii) one-class SVM
(iii) Isolation forest

4) Association

We use association algorithm for associating co-occurring items / events.

Associate algorithms

- Apriori

5) Autoencoders (Dimensionality Reduction)

Autoencoders take input data, compress data into a code, then tried to recreate the input data from summarized code.

→ Autoencoders can remove noise from visual data like images, video / medical scans to improve quality.

Note:

→ Autoencoders can also be used to find outliers (Anomaly).

③ Reinforcement Learning

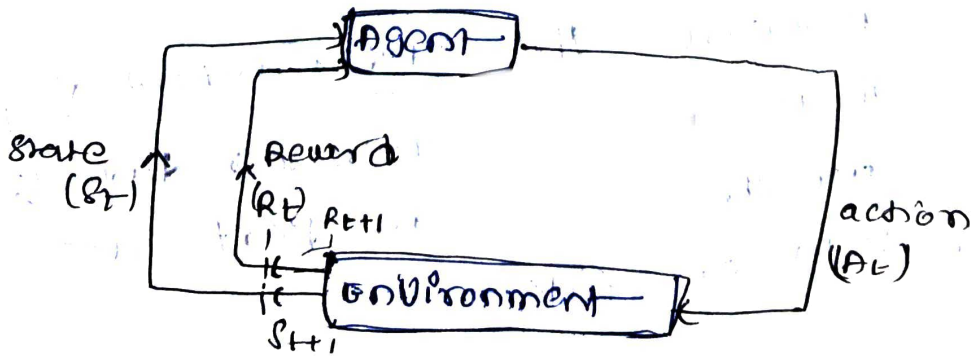
It doesn't require any data as labeled input or labeled output. All it has is the main data.

→ Like, trial & error method. If it fails in a trial then it learns & tries & it learns etc.

→ Games are a real example of RL.
Take a chess game model. All it has is the environment. First it takes a step & sees the outcomes & learns from it and so on.

→ In this kind of RL, AI agents are attempting to find the optimal way to accomplish a particular goal, or improve performance on specific task. As the agent takes action that goes toward the goal, it receives a reward.

overall aim: Predict the best next step to earn the big reward.



when there is no train dataset, it learns from experience.

④ Semi-supervised data:

→ Semi-supervised learning is, for the most part, just what it sounds like: a training dataset with both labeled & unlabeled data.

→ This method is particularly useful when extracting relevant features from the data is difficult & labelling examples is a time-intensive task for experts.

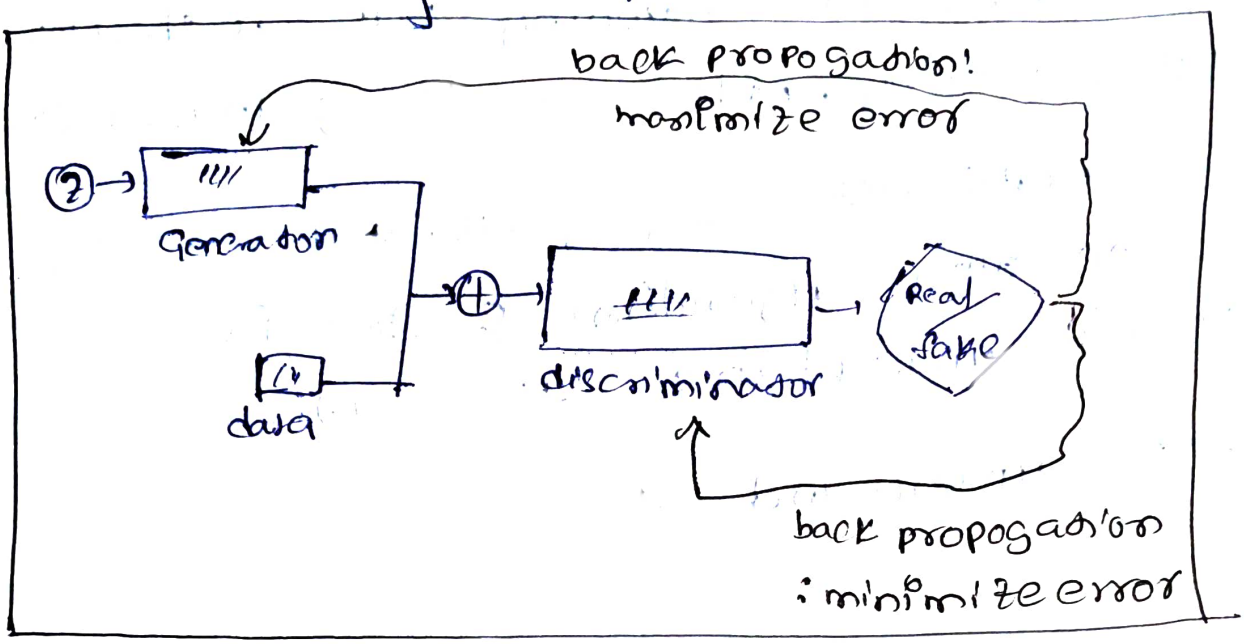
→ A popular training method that starts with a very small set of labelled data is using 'general adversarial networks' / GANs

Like two deep learning networks in competition, each trying to outsmart other. That's a GAN.

→ one of the networks called generator, tries to create new data points that mimic the training data.

→ the other network the discriminator, pulls in these newly generated data & evaluates whether they are part of training data or fake.

→ the networks improve in a positive feedback loop - as the discriminator gets better at separating the fakes from originals, the generator improves its ability to create convincing fakes.



eg:-

→ common situations for this kind of learning are medical images like CT Scan (or) MRI's. A trained radiologist can go through & label a small subset of scans for tumors or diseases.

→ It would be too time-intensive & costly to manually label all the scans — but the deep learning network can still benefit from the small proportion of labeled data & improve its accuracy compared to fully unsupervised models.

→ It's the basic knowledge about algorithms.
→ Before going deep into algorithms. better to know the flow chart of machine learning. I feel ML is like a christopher Nolan movie, first time you don't understand anything but u should keep watching, at last everything makes sense. Same works with ML projects, take a small proj on a algorithm & see it. you will understand the path.

So, let's understand the path.