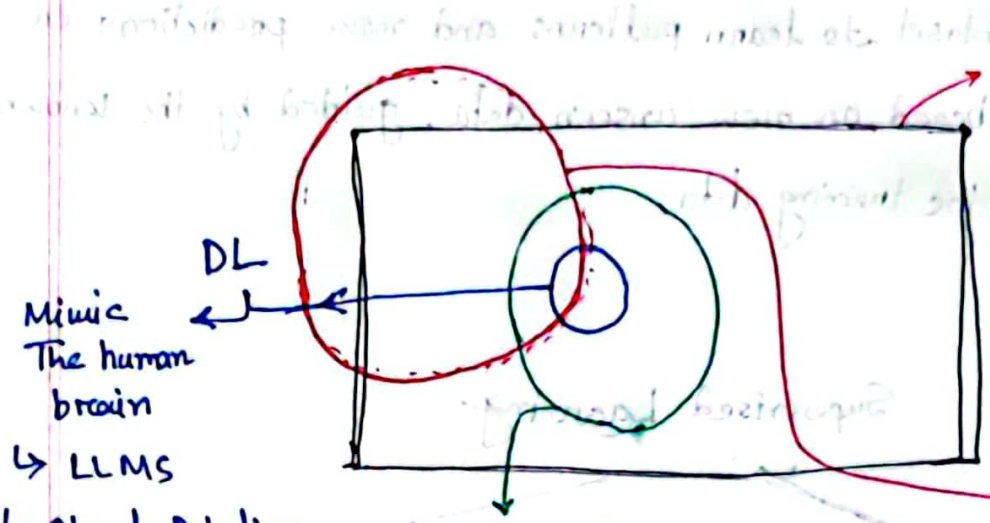
Introduction To Machine Learning

AI vs ML vs DL vs DS



AI
- Self Driving Car
- Robotics
- Alexa

DL
Mimic The human brain
- LLMs
- Object Detection
- Image Recog.
- Chatbots

ML
- Recommendation system
- Sentiment analysis
- Regression

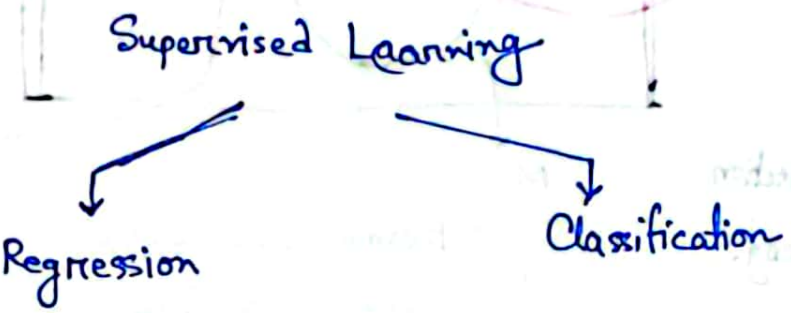
DS
- Works in DA-ML→DL.
- Create AI products

Supervised, Unsupervised and Reinforcement Learning:

There are 4 types of Machine Learning.

1. Supervised ML
2. Unsupervised ML
3. Semi Supervised ML
4. Reinforcement Learning

Supervised ML: It is a type of ml where algorithms are trained on labeled dataset to learn patterns and make predictions or classification based on new unseen data, guided by the known outcomes of the training data.



Example of classification:

Predicting a student will pass or fail depending on his/her given hours in study and in playing

Input Feature

O/p Feature → Dependent Feature

No. of hours Played	No. of study hours	Pass/Fail
8	2	Fail
7	3	Fail
6	4	Fail
5	5	Pass
4	6	Pass

Example of Regression:

Predicting the house price based on the size of the house and no. of rooms.

~~Size~~

Input Features **Output Feature →** (continuous)

Size of House (sq)	No. of Rooms	Price of house
3	2	5000
4	3	7000
2	1	3000
5	3	8000

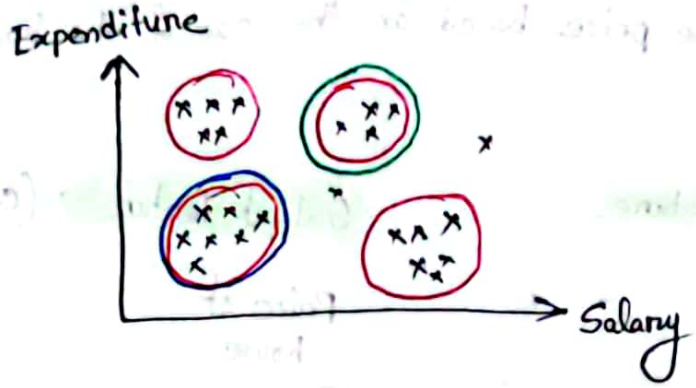
(2) **Unsupervised ML:** It is a type of ml where algorithm explores data patterns and structures without predefined labels. It includes

→ Clustering → Dimensionality Reduction

→ Enabling data driven insights

→ Discovery of hidden relationships in the dataset.

An example: Custome Segmentation



Here, you can cluster the data and take valuable desicisions.

Suppose you are a watch shop owner. You have done the clustering in your customers. Where you have found 4 categories →

1. Customer with high salary and high expenditure
2. Customers with low salary and high expenditure
3. Customers with low salary and ~~high~~ low expenditure
4. Customers with high salary but low expenditure.

So, by clustering them. you can now decide that what model of watch you can sell to which customer. Who needs discounts or who only need the quality watch. That would be really beneficial to your business.

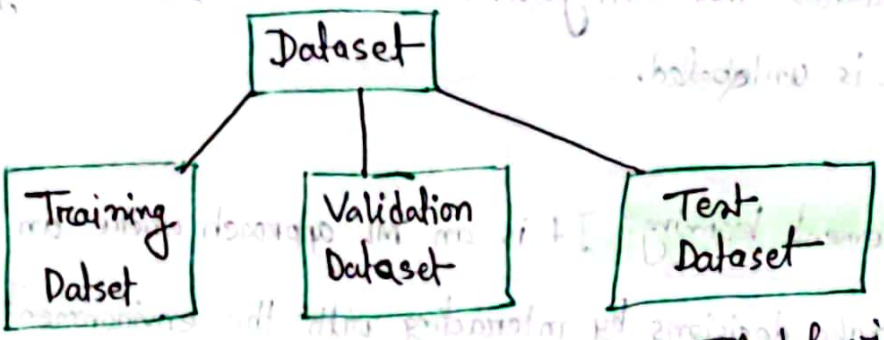
③ **Semi Supervised ML:** Combination of (Supervised + unsupervised) where the model is traine on both labeled and unlabeled data. It leverages the limited labeled data along with the larger pool of unlabeled data to improve performance.

Example: If someone have small set of labeled imegen of cats and dogs (1000) and a much larger set of unlebeled images (10000) for image classification, Semi supervised Learning can be used to train a model to distinguish between cats and dogs more accurately by using the additional unlabeled data to refine its understanding of the features that distinguish the two classes, even though most of the data is unlabeled.

④ **Reinforcement learning:** It is an ML approach where an agent learns to make decisions by interacting with the environment. It relies on trial and error, where the agent recieves rewards or penalties based on actions, allowing it to learn optimal approaches over time.

Example: In training of a self driving car, reinforcement learning can be used. The car explores various driving actions, recieve rewards (positive for safe and efficient driving and negative for accidents) and learn to navigate road autonomously by maximizing cumulative rewards. The process is iterative and make the car to take better decisions and improve its driving performance without any explicit supervision.

Train, Test & Validation:



→ The dataset By which the model get trained

→ Hyper tuning of the model (We check here, by changing the parameter values if the accuracy can be changed or not)

→ Model will be test with this dataset.

Examples: Suppose X is a student and He has exam soon.

So, if we take him as a model →

His training dataset would be his books

The hyper parameter tuning will be when he will read the same topic books but written with some other author. Doing that he can get more detailed concepts and knowledge which will make him more capable to do good in the exam.

And last the test dataset will be the questions that will come in the exam hall. He can show his accuracy by providing right answers of the questions.

Variance, Bias, Overfitting and Underfitting:

if

Trained Dataset → Model is trained → Accuracy ↑↑ (95)%

Test Dataset → Model is tested → Accuracy ↓↓ (60%)

→ This would be overfitting
→ Low Bias, High variance

Means the model is properly trained but can't show proper expected outcome.

Training Accuracy ↑ (High) → Low Bias

Test ~~Training~~ Accuracy ↕ (low) → High Variance

if

Trained Dataset → Model is trained → Accuracy (55%)

Test Dataset → Model is tested → Accuracy (50%)

→ This will be underfitting
→ High Bias, High Variance

Means you have trained your model with low accuracy and so your model is showing outcome with low accuracy.

if

Trained dataset → Model is trained → Accuracy (95%)

Test dataset → Model is tested → Accuracy (91%)

→ This will be Generalised Model
→ Low Bias, Low Variance

Means you have trained your model with highly accurate data and your model is also performing with high accuracy.

Training Accuracy ↓ (low) → High Bias

Test Accuracy ↓ ↑ (High) → Low Variance