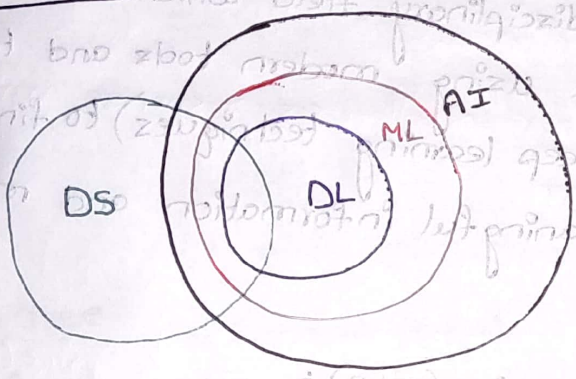
Machine Learning

Def:- Machine learning is defined as "the field of study that gives computers ability to learn without being explicitly programmed.

AI v/s ML v/s DL v/s DS



AI - Artificial Intelligence

ML - Machine Learning

DL - Deep Learning

DS - Data science

Artificial Intelligence:-

It is nothing but creating an application where it performs all its tasks without any human intervention.

Ex:- ① self Driving car
② chatGpt
③ Alexa
④ SIRI
⑤ Google Home.

Machine learning:-

Machine learning provides stats tools to explore, analyze, visualize and perform prediction with the help of data.

Ex:- ① Recommendation system.
② Weather prediction.
③ Spam detection.
④ Disease prediction.

Deep Learning :-

→ Deep Learning is subset of Machine Learning where artificial neural networks - algorithms modeled to work like human brain, learn from large amounts of data.

Data Science :-

Data science is inter disciplinary field which deals with vast volumes of data using modern tools and techniques (ml algorithms and deep learning techniques) to find unseen patterns, derive meaningful information and make business decisions.

Natural language processing (NLP) :-

Natural language processing (NLP) is a form of artificial intelligence that allows computers to understand human language, whether it be written, spoken or even scribbled. (NLP is a technique which works with text data, it falls in both the categories ML & DL).

Ex :- Email filtering, chatbots, Language translation, sentiment analysis.

Computer vision :-

Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs and take actions or make recommendations based on that information.

Ex :- ① parking occupancy detection.
② Traffic flow analysis.
③ X-Ray Analysis.
④ Cancer detection.

Machine Learning

- Supervised
  - Regression
    1) Linear
    2) polynomial
    3) SVR
    4) Decision Tree
    5) Random forest
    6) XG Boost
    7) KNN
  - classification
    1) Logistic Regression
    2) SVM
    3) Decisiontree
    4) Random forest
    5) Naive Bayes
    6) KNN
- unsupervised
  - clustering Algorithms
    1. DB scan
    2. K-Means
    3. Hierarchial clustering
    4. silhouette score.

Supervised Machine Learning:

→ Works on labelled data.

→ Each input has corresponding labelled output.

→ The goal of supervised machine learning is to learn a mapping from the input to the output.

→ The input data is called attributes, features or predictors or Independent features.

→ The output variable is also called as response variables or dependent feature.

Ex:

Brand	car model	Miles driven	price($)
Honda	Brio	70000	19000
Toyota	corolla	45000	20000
Honda	city	25000	30000
Nissan	sunny	90000	10000

Brand
car model
Miles driven
} Independent features

price → Dependent feature.

→ From the given dataset, the machine learning algorithm learns from mapping from input variable to output variable. This learning is represented in the form of a model. when new data is given to the model as shown below, it can predict its output value.

Brand	car model	Miles driven	price ($)
Honda	city	11000	?

→ supervised Machine Learning can be further classified into:

① Regression.

② classification.

① Regression:- When the output variable takes continuous numerical values.

Ex:- price of a car,

House price prediction.

Flight fare prediction.

② classification:- When the output variable takes categorical or discrete (non-continuous) values.

Ex:- ① whether email is spam or not.

② whether the transaction is fraudlent.

③ whether student is pass/fail in exam.

No. of playing hours	No. of study hours	Exam Result
9	1	0
7	2	0
3	3	1
4	5	1

No. of playing hours, No. of study hours } Independent features

Exam Result ↓ Dependent feature.

0 - Fail

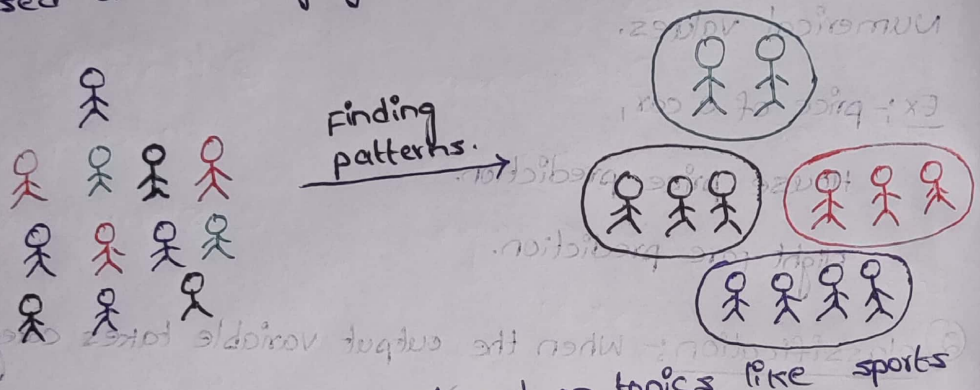
1 - pass

(2) Unsupervised Machine Learning:

→ The unsupervised machine learning has no explicitly defined output.

→ The idea is to discover knowledge or structure in data.

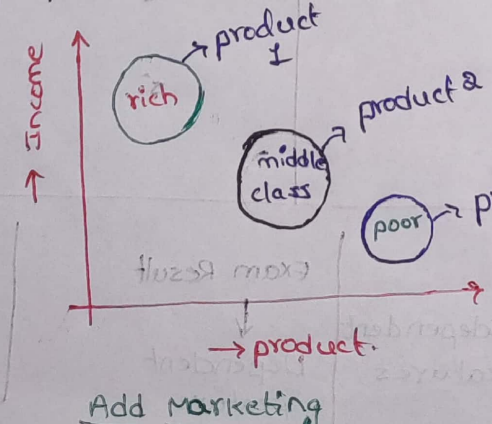
Ex:- For example an online retailer will have data about all items that the customers purchased. Unsupervised learning algorithms can be applied on this data to group customers based on their buying patters.



Ex:- Grouping news articles based on topics like sports politics, business etc., is another example of unsupervised learning.

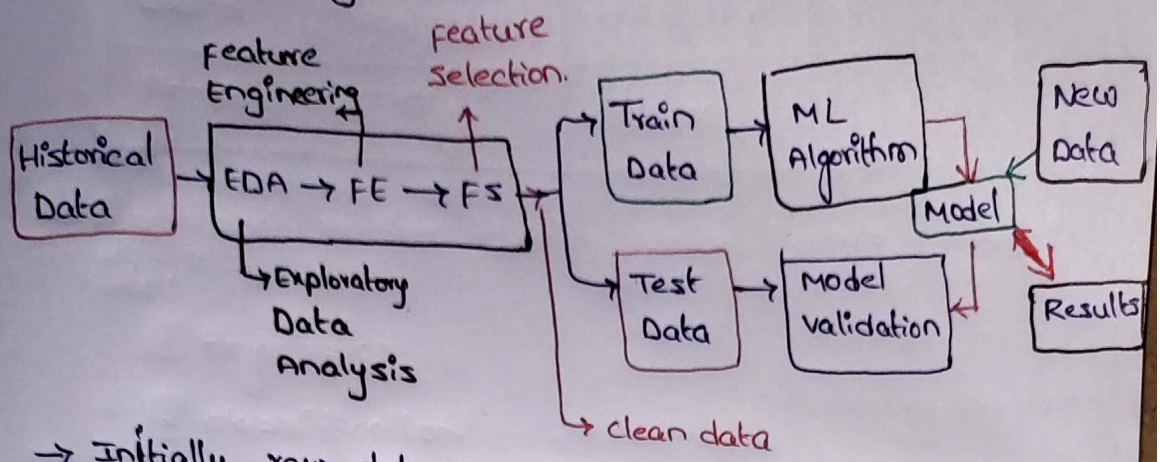
→ This task of finding inherent clusters or groups in the data is known as clustering.

Ex:- Let's say a company launched three different products targeting the rich, Middle class, poor people in the country. Now they try to send those product adds to selected group of people based on their income.



Add Marketing

Machine Learning process :-



→ Initially raw data is gathered and perform "EDA" on the data, after that data is cleaned through a process called feature engineering.

→ Through feature engineering we can handle missing values, handling outliers, creating new features out of existing ones are some of the common tasks through "FE"

→ After performing "FE" we go for feature selection. while building a machine learning model for real-life dataset we come across lot features in the dataset and not all these features are important every time. Adding unnecessary features while training the model and decrease the generalisation capability of the model, increase the complexity and makes the model biased (Accuracy is less in this case) Hence feature selection is one of mandatory steps.

→ After "FS" the data is split into Train and Test data

→ The "train data" is used for training the machine learning model.

→ once the model is built it is validated against the "test data" for accuracy.

→ This accuracy helps us in estimating the performance on previously unseen data. If the model performance on both Train and Test data is satisfactory, the model may be deployed.

→ once deployed, the model makes predictions on new data, these predictions/Insights are used to take business decisions.