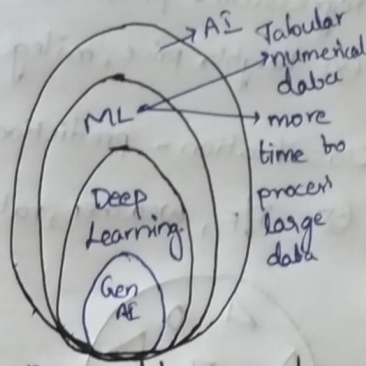# Machine Learning

Introduction:

Machine Learning is a branch of Artificial Intelligence (AI) that enables computers to learn from past data & make predictions or decisions without being explicitly programmed.

* Instead of writing code to solve a specific problem we feed the machine by data & examples & it learns patterns from that.

→ AI Tabular →numerical data

→ more time to process large data

Future Advancements:
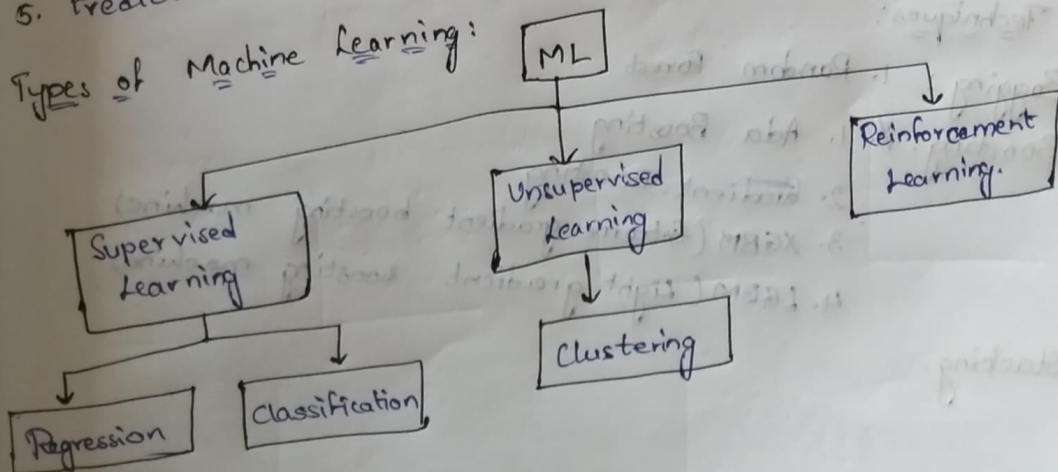Quant AI
Explainable AI.

Traditional Programming Vs Machine Learning:

- In traditional programming, programmer write rules → computer applies them → output is generated.

Ex: Writing an if-else program to identify whether an email is spam or not.

- In Machine Learning, we give the computer thousands of emails labelled as as "spam" or "not spam". The computer learns the rules automatically it predicts if a new email is spam or not.
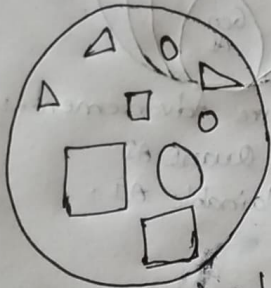
key Components of Machine Learning:

1. Data - The info that machine learns.
2. Features - Input variables used to make prediction
3. Model - The algorithm that learns from data.
4. Training - The process of teaching the model using data.
5. Prediction - Using the trained model on new unseen data.
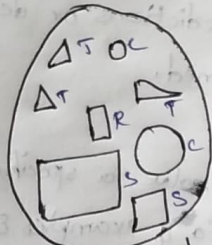
Types of Machine Learning:

```
                    ┌──────┐
                    │  ML  │
                    └──────┘
      ┌────────────────┼────────────────────┐
      ↓                ↓                      ↓
┌──────────┐    ┌─────────────┐      ┌──────────────┐
│Supervised│    │Unsupervised │      │Reinforcement │
│ Learning │    │  Learning   │      │  Learning.   │
└──────────┘    └─────────────┘      └──────────────┘
   ┌───┴────┐          ↓
   ↓        ↓    ┌────────────┐
┌──────────┐ ┌──────────────┐ │ Clustering │
│Regression│ │Classification│ └────────────┘
└──────────┘ └──────────────┘
```

Regression - prediction value is continuous

Ex: Height, price, milege

Classification - prediction value is categorical / discrete.

Ex: T/F, weather, rating

Unsupervised.   Supervised.

Clustering - creates groups.

Topics To be Covered

Supervised Learning:

Pure Regression:

1. Simple Linear Regression
2. Multi Linear Regression
3. Polynomial Regression
4. Lasso & Ridge Regression

Pure Classification:

1. Logistic Regression.
2. KNN (K-nearest neighbour).
3. Naive bias bayes.

Both Regression & Classification:

1. Decision Tree
2. SVM (Support Vector Machine)

Ensemble Techniques:

Bagging: 1. Random Forest

Boosting: 1. Ada Boosting
2. Gradient Boosting
3. XGBM (Extreme gradient boosting machine)
4. LGBM (Light gradient boosting machine)

Stacking.

Unsupervised Learning

Clustering :

1. k-means clustering
2. Hierarchical clustering
3. ~~DBMS~~ DBSCAN (Density based spatial clustering application with Noise)

Learn by urself

Self Learning - Semi supervised Algorithms.

public
private

1. Recommendation System
2. Time - Series Analysis.
3. Anamoly Detection.