PREPARATION PHASE
PRIOR KNOWLEDGE

DATE	22 AUGUST 2022
TEAM ID	PNT2022TMID15566
PROJECT NAME	EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

PRIOR KNOWLEDGE:
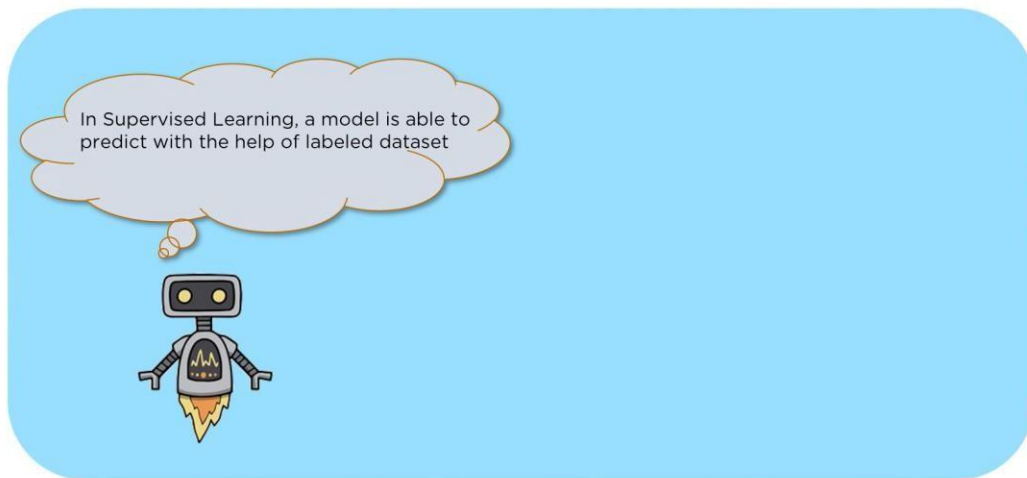
Machinelearning:

Machinelearningisdefinedasmakingmachineslearnandact ashumansbyfeedingthem withdata.

Therearetwotypesoflearninginmachinelearning:
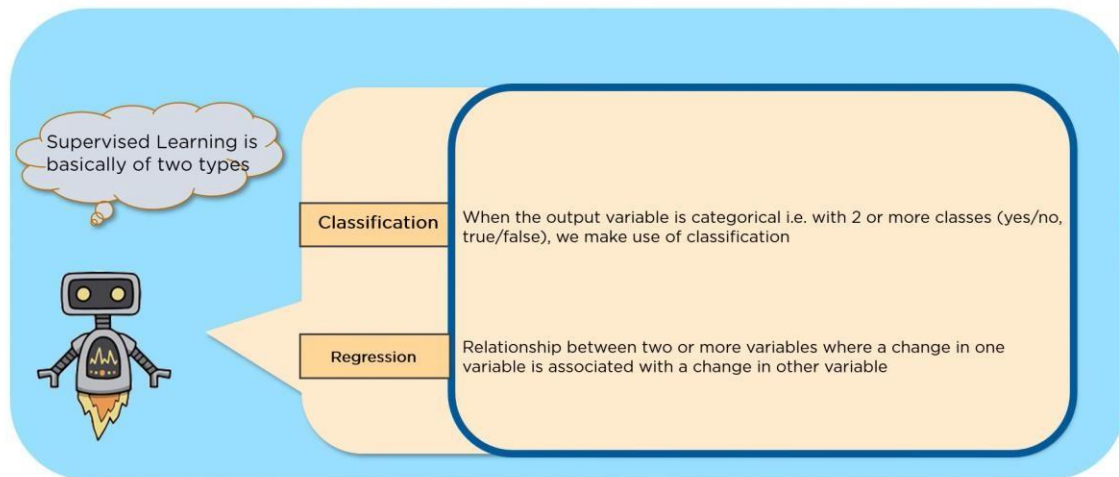1. Supervisedlearning:

Supervised Learning

In Supervised Learning, a model is able to predict with the help of labeled dataset

Supervisedlearningisdonewiththe helpofalabelleddataset.

Types of Supervised Learning

Supervised Learning is basically of two types

Classification — When the output variable is categorical i.e. with 2 or more classes (yes/no, true/false), we make use of classification

Regression — Relationship between two or more variables where a change in one variable is associated with a change in other variable
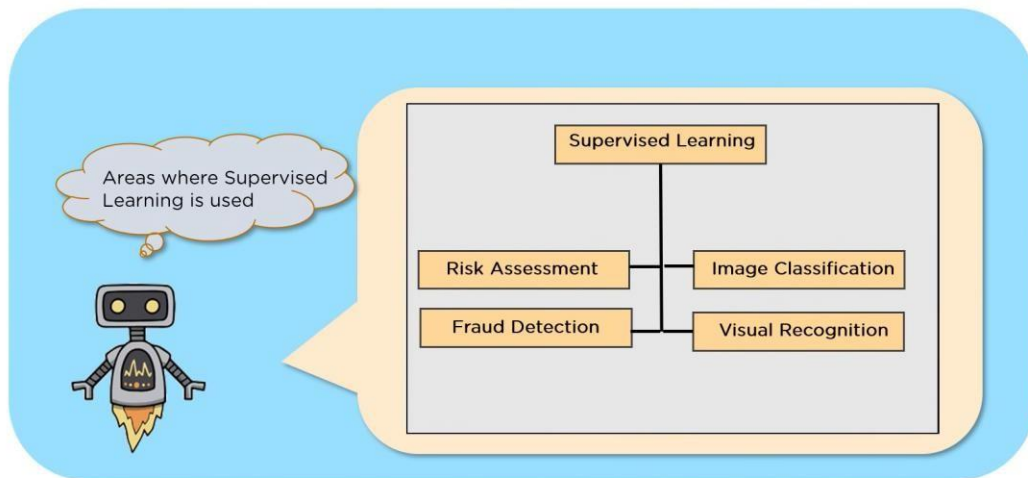
Therearetwotypes ofsupervisedlearning:

a) Classification

b) Regression

Applicationsofsupervisedlearning:

Applications of Supervised Learning

2. Unsupervisedlearning:

Unsupervised Learning
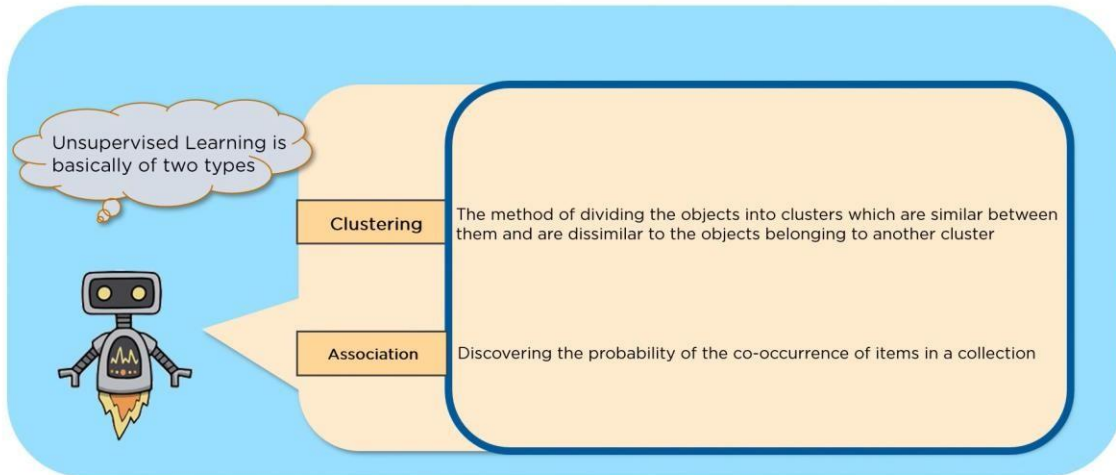
Unsupervised learning is defined as using unlabeled data to train themodel.

Typesof unsupervisedlearning:

      a) Clustering
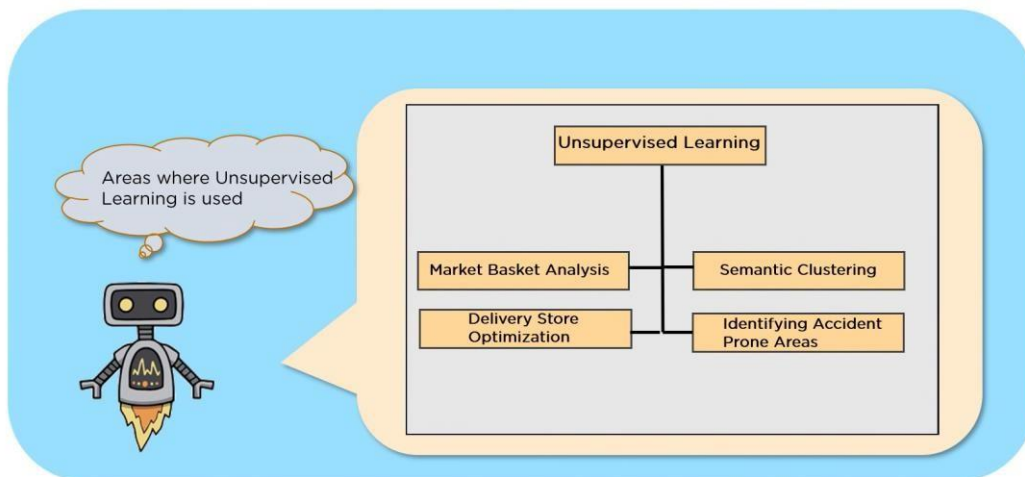   b) Association

Applicationsofunsupervisedlearning:

Classification:

Classification Problem

Goal: predict category of new observation

Earlier Observations   **Estimate** →   CLASSIFIER

Unseen Data   **CLASSIFIER** →   Class

Classification Applications

- **Medical Diagnosis**
- **Animal Recognition**

Important:

- Qualitative Output
- Predefined Classes

Regression:

Regression

PREDICTORS → REGRESSION FUNCTION → RESPONSE
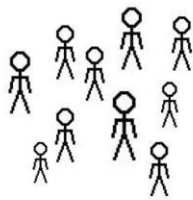
- Relationship: **Height** - **Weight**?
- Linear?
- Predict: **Weight** ⟶ **Height**

Regression Model

Fitting a **linear** function

$$\text{Height} \approx \beta_0 + \beta_1 \times \text{Weight}$$

- **Predictor:**   Weight
- **Response:**   Height
- **Coefficients:** β_0, β_1

Estimate on previous input-output

```
> lm(response ~ predictor)
```

Regression Applications

- Payments ⟶ Credit Scores
- Time ⟶ Subscriptions
- Grades ⟶ Landing a Job

- Quantitative Output
- Previous **input-output** observations

Clustering:

Clustering

- **Clustering:** grouping objects in clusters
  - *Similar* within cluster
  - *Dissimilar* between clusters
- **Example:** Grouping similar animal photos
  - No labels
  - No **right** or **wrong**
  - Plenty possible clusterings