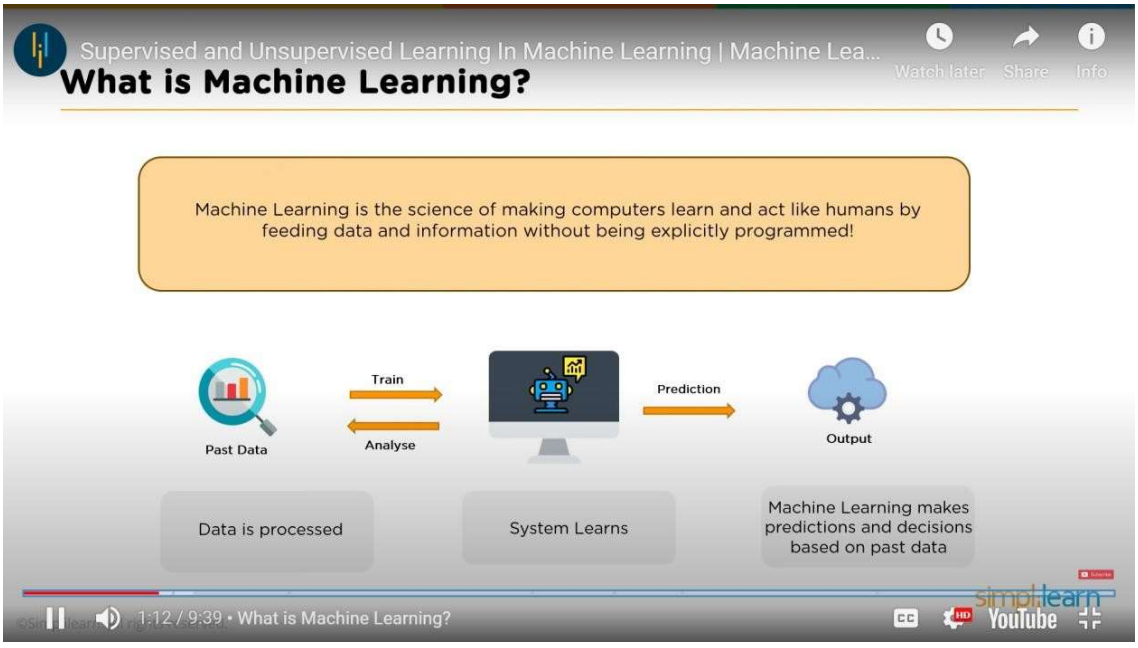
PREPARATION PHASE

PRIOR KNOWLEDGE

DATE	22 AUGUST 2022
TEAM ID	PNT2022TMID15499
PROJECT NAME	EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

PRIOR KNOWLEDGE:

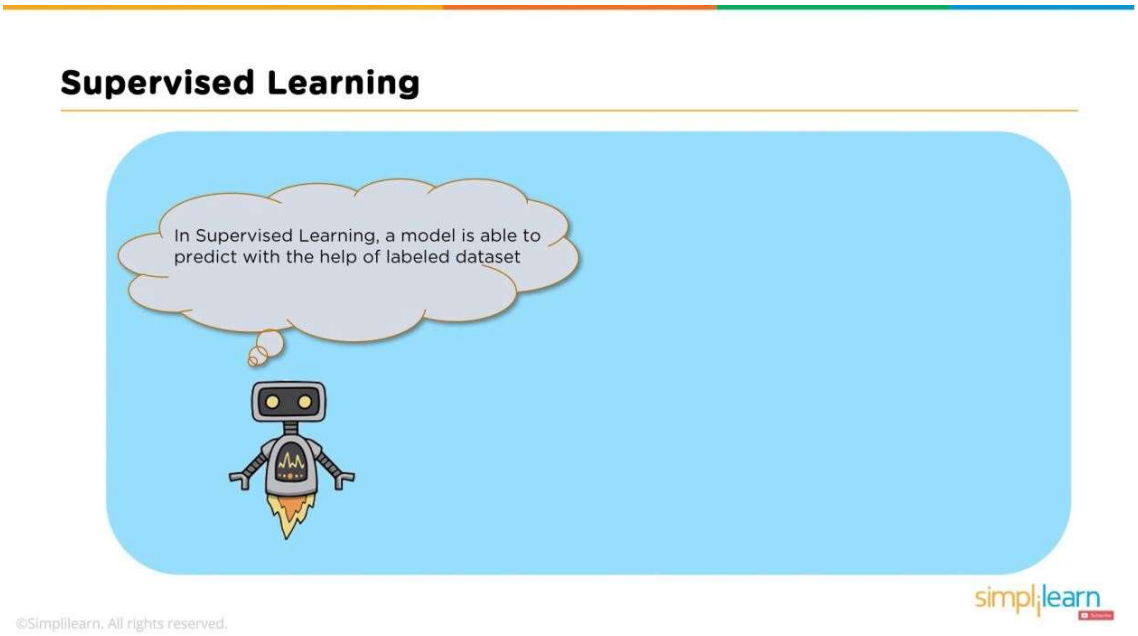
Machine learning:



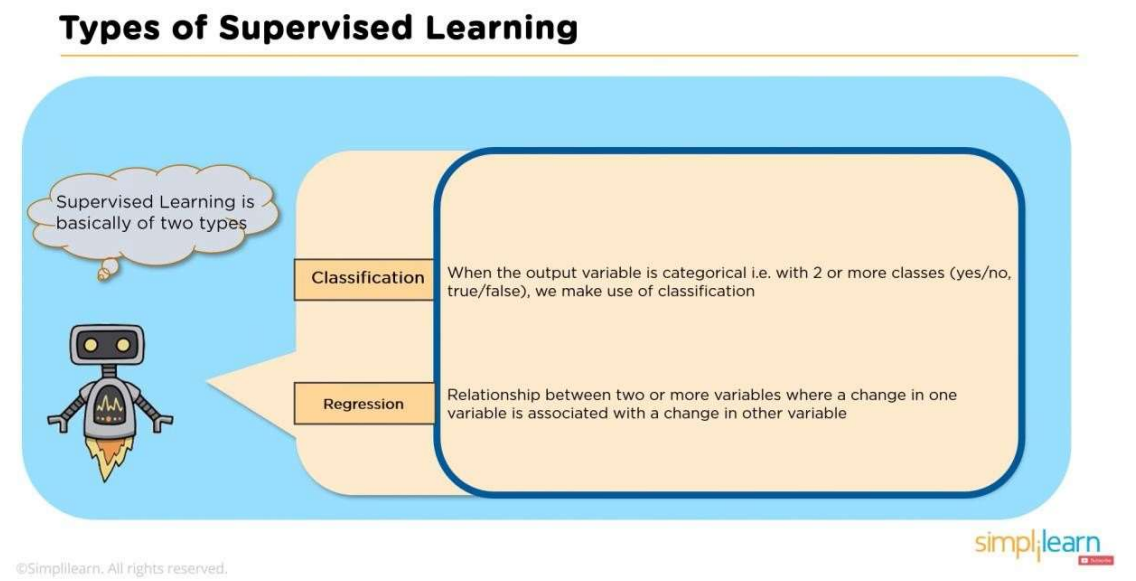
Machine learning is defined as making machines learn and act as humans by feeding them with data.

There are two types of learning in machine learning:

1. Supervised learning:



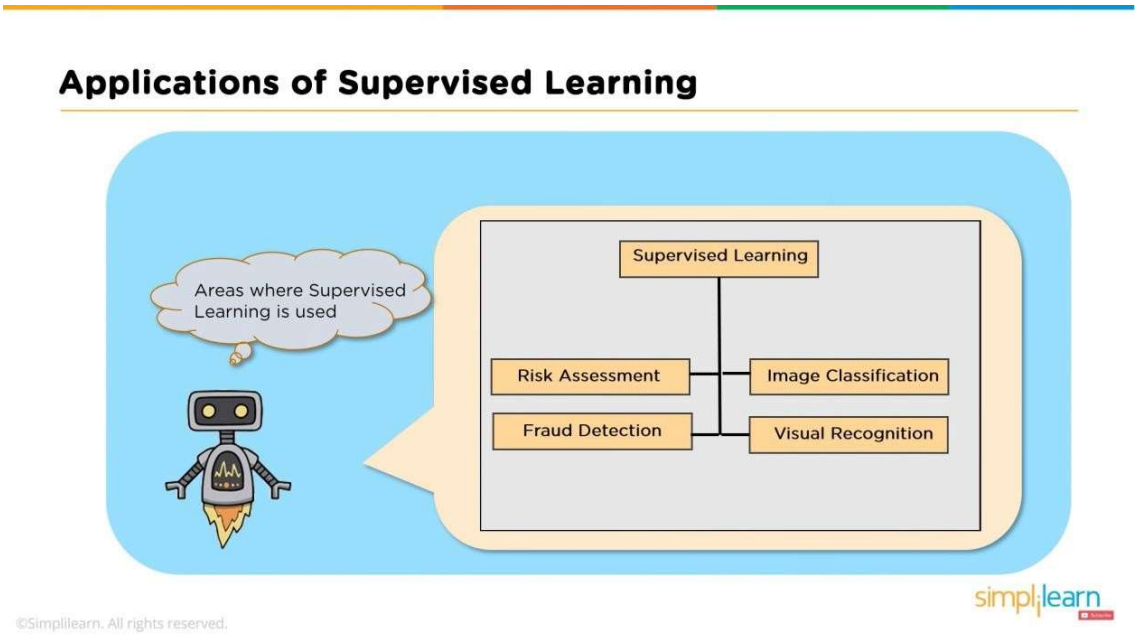
Supervised learning is done with the help of a labelled dataset.



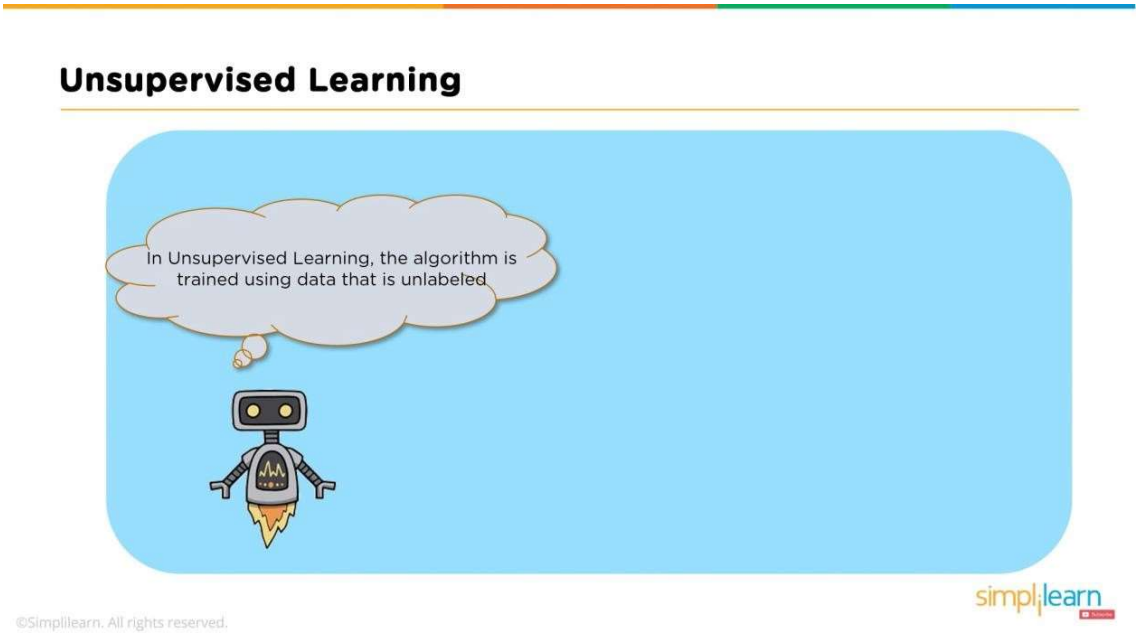
There are two types of supervised learning:
a) Classification

b) Regression

Applications of supervised learning:



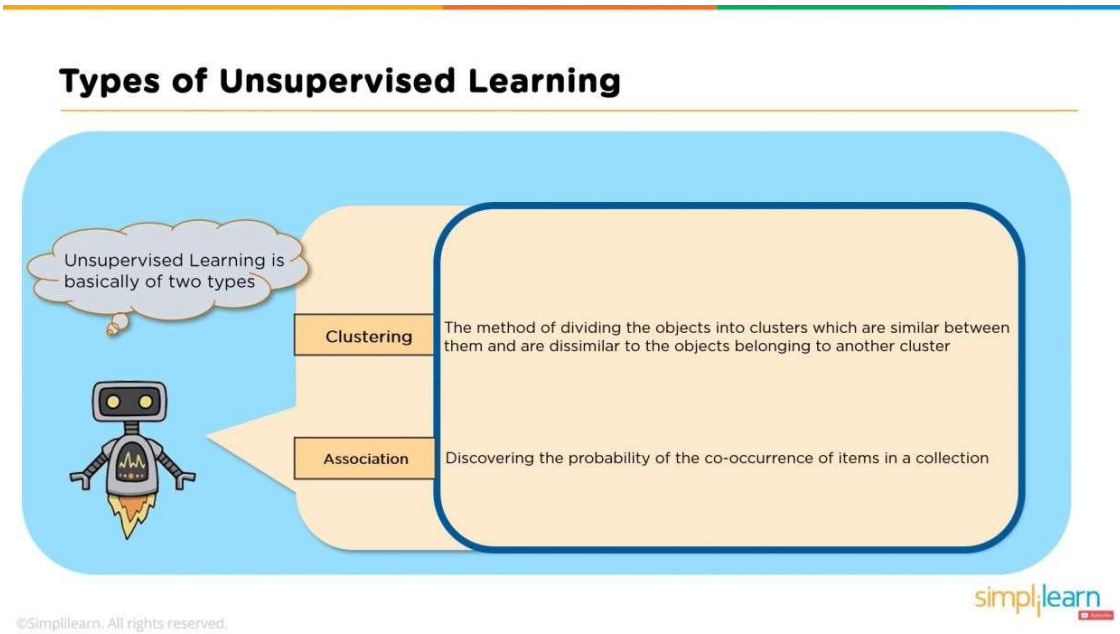
2. Unsupervised learning:



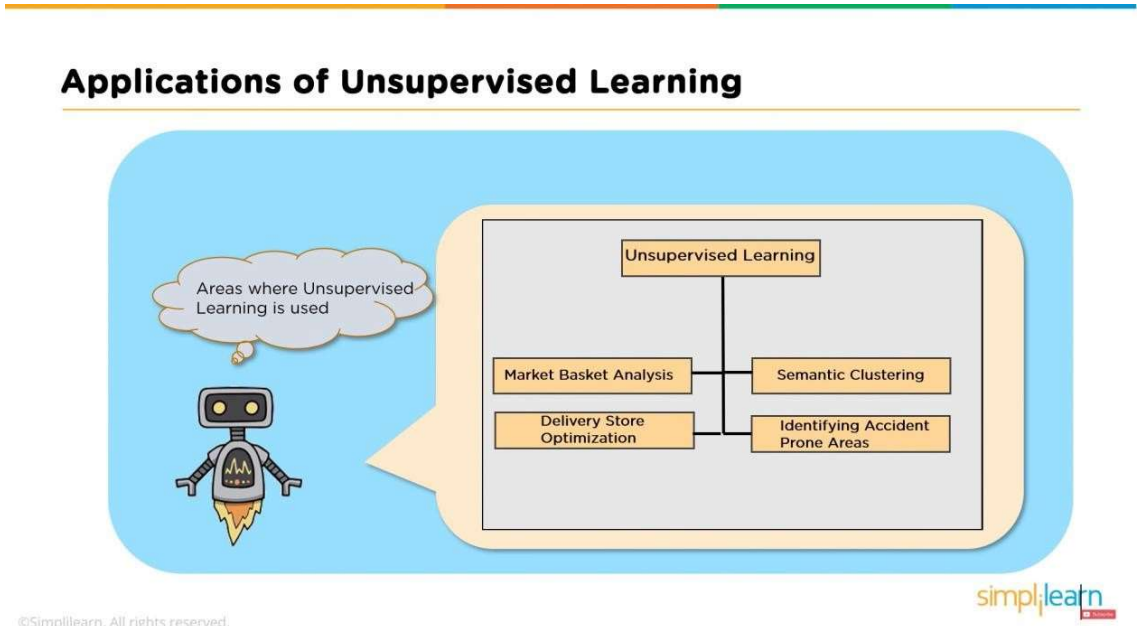
Unsupervised learning is defined as using unlabeled data to train the model.

Types of unsupervised learning:

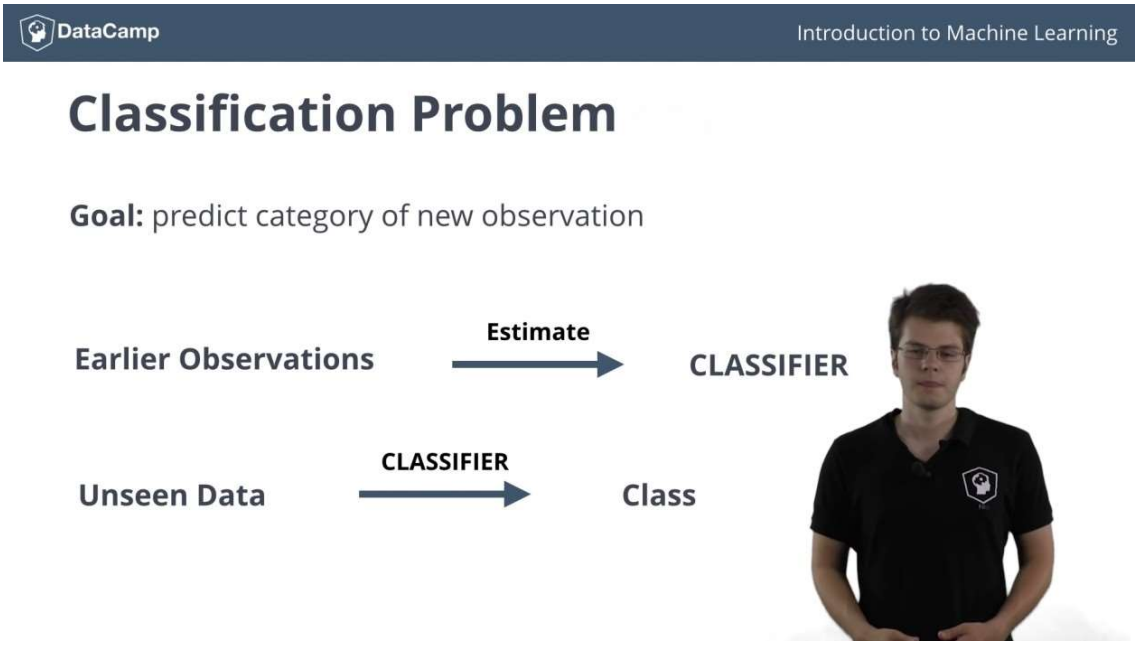
a) Clustering
b) Association



Applications of unsupervised learning:



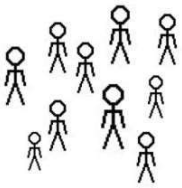
Classification:



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)
```

DataCamp Introduction to Machine Learning

Regression Applications

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

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

Clustering:

DataCamp Introduction to Machine Learning

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