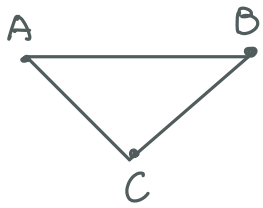
Data Representation :

1. Records (Name, age, balance)

2. Graphs (Nodes connected by edges) → Adjacency matrix, adjacency list



	A	B	C
A	0	1	1
B	1	0	1
C	1	1	0

(if they're connected)

→ matrix of a triangle

3. Images (Pixel of 0 or 1)

4. Text

5. Strings (DNA sequence → list of words)

6. Time Series (data at specific interval of time)

Types of learning

1. Supervise learning → create model and make predictions
   - Regression
   - classification

2. unsupervise learning
   - clustering

Data
- n data points, m features

Feature Space

Distance
- dissimilary function
- large value from distance → dissimilar
- easy to understand and graph

Minkowski Distance

euclidean distance (p=2)

$$L_p(x, y) = \left(\sum_{i=1}^{d} |x_i - y_i|^p \right)^{\frac{1}{p}}$$