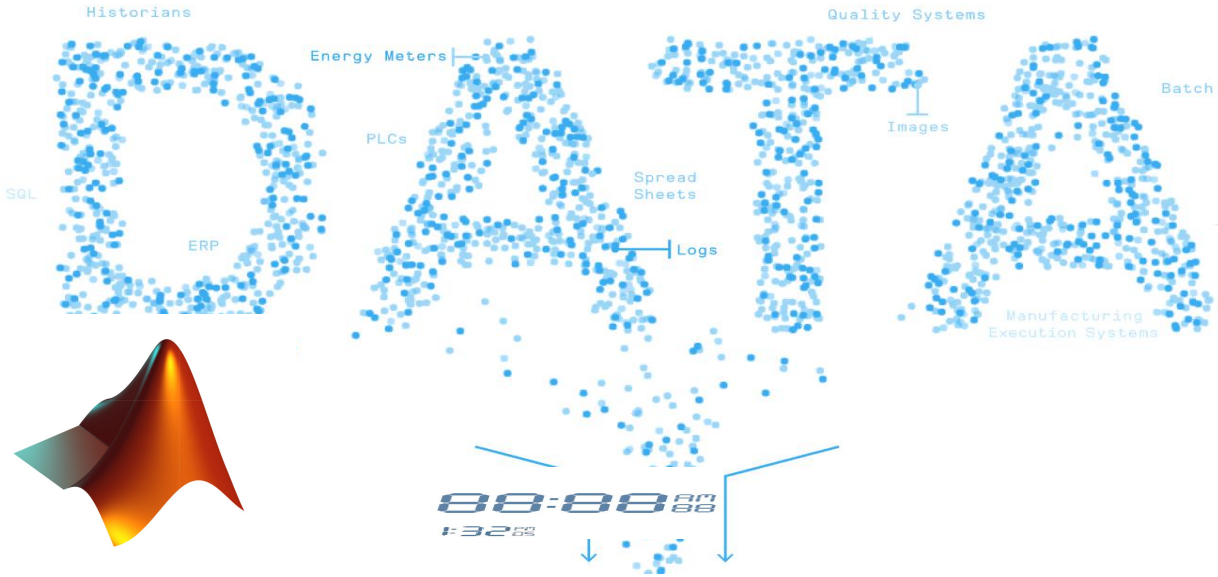
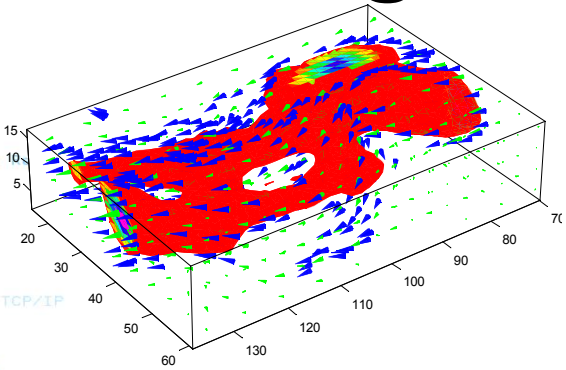
Introduction to Artificial Intelligence
- 05-01-01 Unsupervised Learning

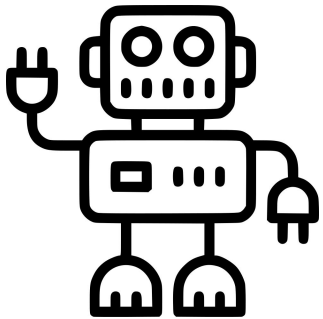




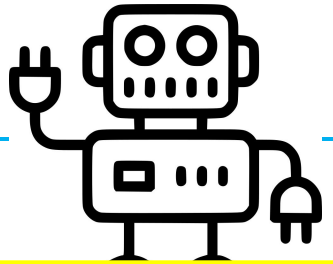
Dr Leo Chen
leo.chen@ieee.org
09/Mar/2022

Module Contents

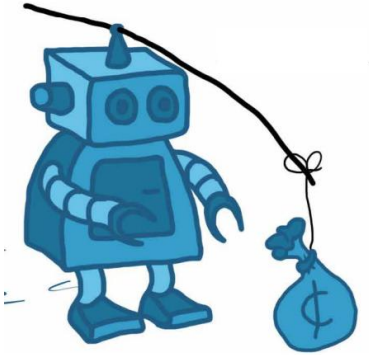
1. Introduction
2. Evolutionary Computation
3. Artificial Neural Network
4. Fuzzy Logic and Fuzzy Systems
5. More AI Subsets
6. AI and Industry 4.0
7. AI Applications
8. Labs
9. Courseworks

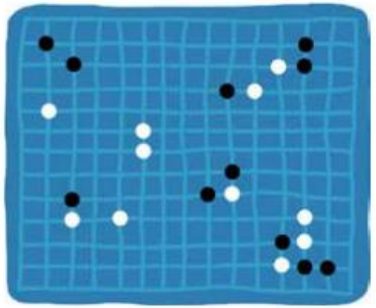


Chapter Contents



1.Deep Learning
2.Machine Learning
3.Swarm Intelligence
4.Heredity Algorithm
5.Quantum Computing
6.DNA Computing
7.Neuromorphic Computing



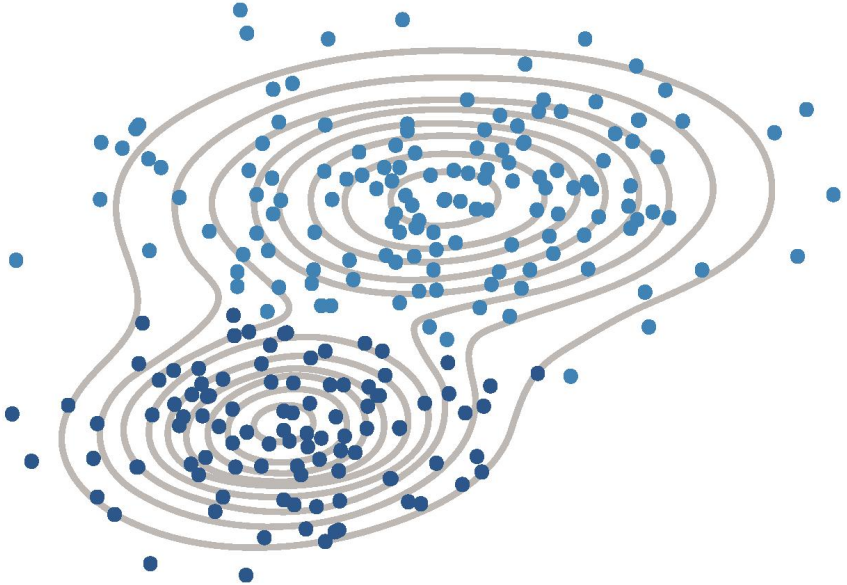


Unsupervised Learning (UL)

1. What is Unsupervised Learning
2. Why Use Unsupervised Learning
3. Fundamentals of Unsupervised Learning
4. Applications

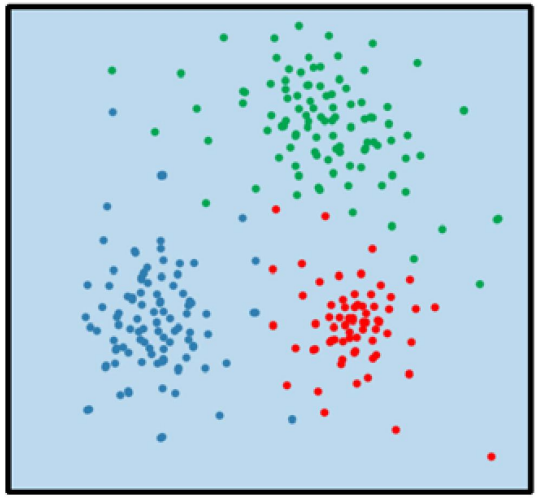
FAQ

Reference

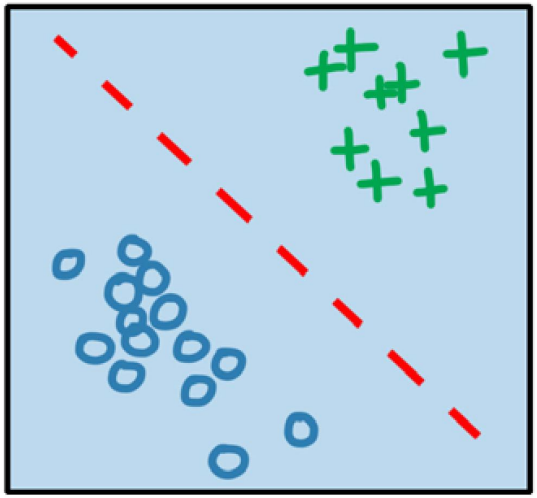


machine learning

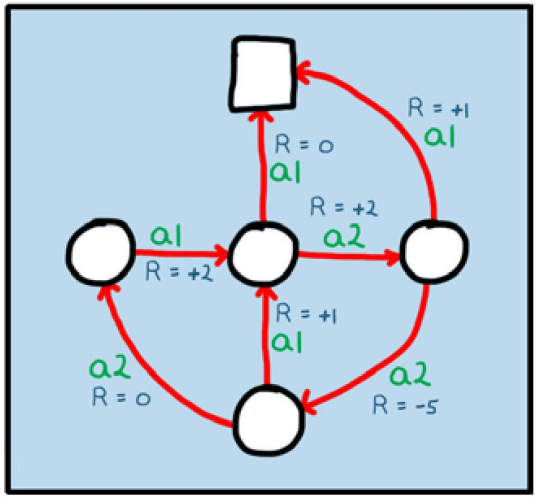
unsupervised learning



supervised learning



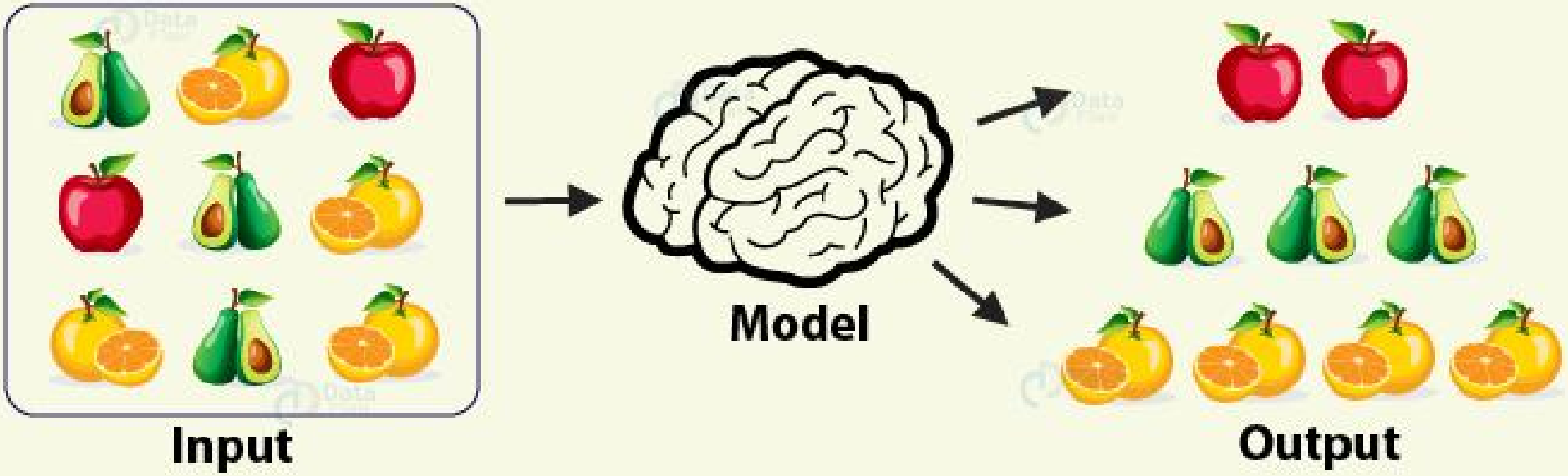
reinforcement learning



1 What is Unsupervised Learning

- **Unsupervised Learning** is a group of **ML** algorithms that work with "no-ground-truth" data **- unlabeled data.**
- **Unsupervised learning** is to **find** hidden **patterns** or intrinsic structures in input data. Unsupervised learning can be separated into **two** types of problems:
- **Clustering**: A clustering problem is where you want to discover the **inherent** groupings in the data.
- **Association**: An association rule learning problem is where you want to **discover rules** that describe large portions of your data.

Unsupervised Learning



2 Why Use Unsupervised Learning

- Unsupervised Learning finds all kinds of **unknown patterns** in data through **pattern recognition**.
- It takes place in **real-time**, so all the input data to be **analyzed & labeled** in learners' presence.
- Unsupervised methods help to search for **features** that can be useful for categorisation.

3 Fundamentals of Unsupervised Learning

In cluster analysis, **data** is partitioned into groups based on some **measure of similarity** or **shared characteristic**. Clusters are formed so that objects in the same cluster are very similar and objects in different clusters are very distinct.

Clustering algorithms fall into two broad groups:

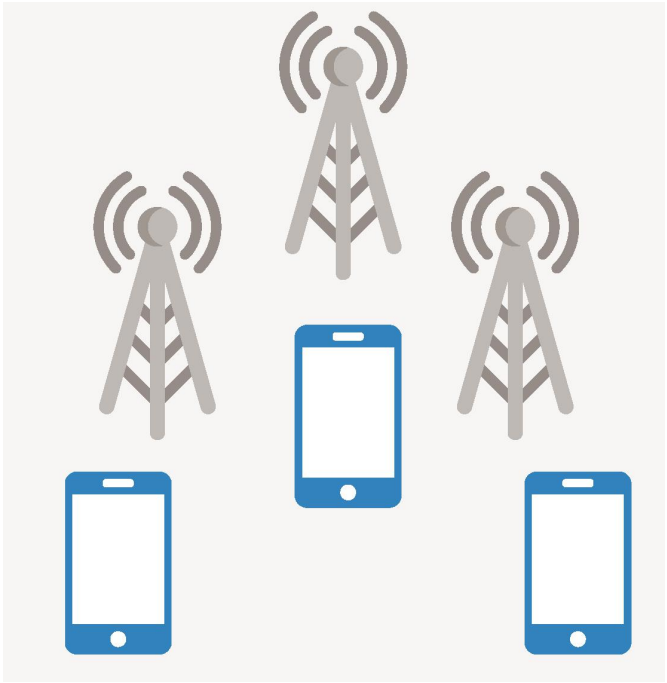
- **Hard** clustering, where each data point belongs to only one cluster
- **Soft** clustering, where each data point can belong to more than one cluster

If you don’t yet know how the data might be grouped:

- Use **self-organising** feature maps or **hierarchical** clustering to look for possible structures in the data.
- Use cluster evaluation to look for the “**best**” number of groups for a given clustering algorithm.

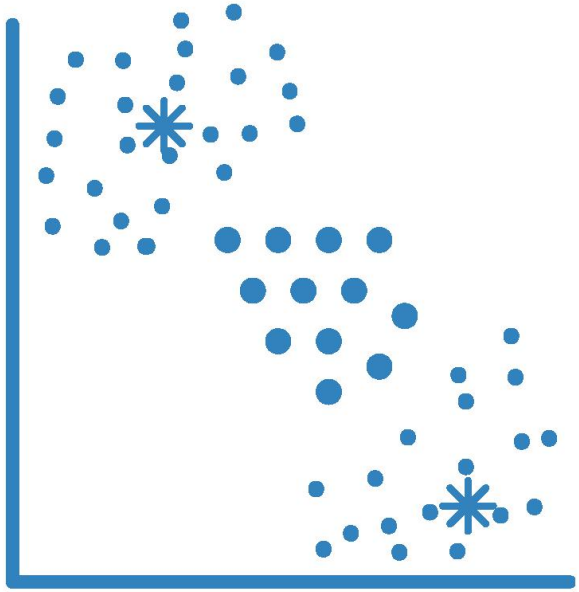
Hard Clustering Algorithms

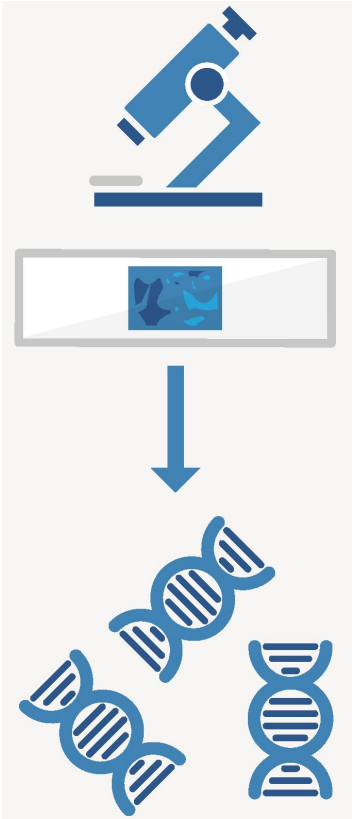
- K-Means
- K-Medoids
- Fuzzy C-Means
- Hierarchical
- Gaussian Mixture
- Neural Networks
- Hidden Markov Model
- Self-Organizing Map



Soft Clustering Algorithms

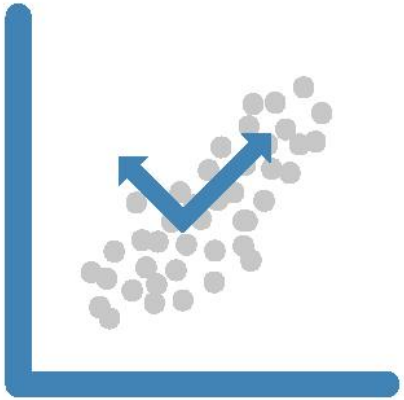
- Fuzzy c-Means
- Gaussian Mixture Model

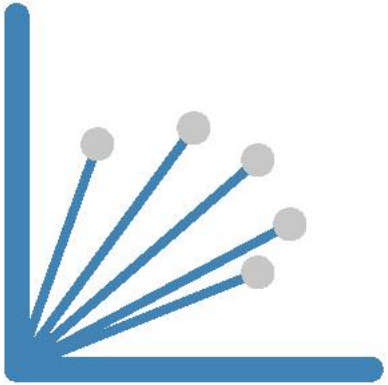


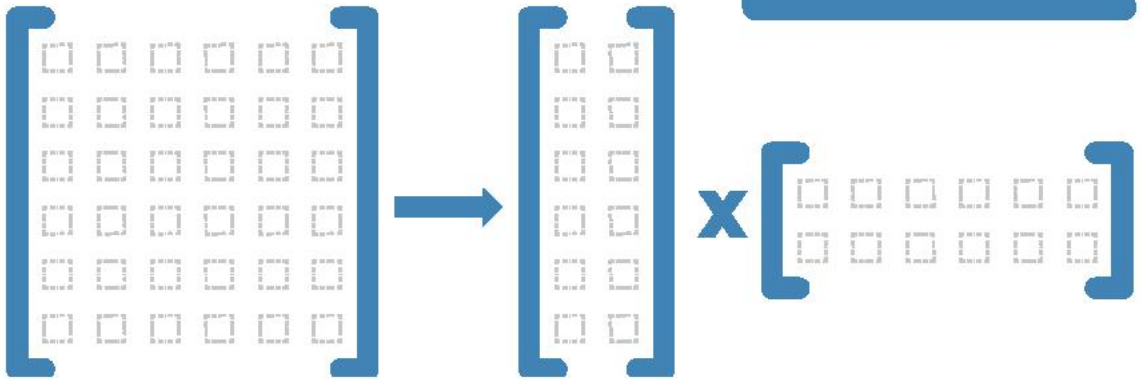


Dimensionality Reduction

- Principal Components Analysis
- Non-Negative Matrix Factorisation
- Factor analysis





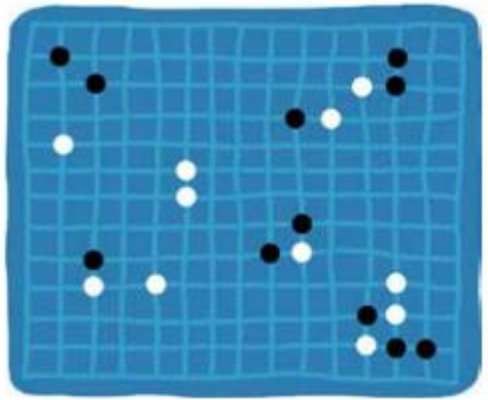


4 Applications[1]

- **Clustering** automatically devide the dataset into groups based on their **similarities**.
- **Anomaly detection** can discover unusual text or data points in your dataset. It is useful for **finding fraudulent** transactions.
- **Association** mining identifies sets of items that often occur together in your datapoints/dataset.
- **Latent variable** models are broadly used for data preprocessing. Like **reducing** the amount of **features** in a dataset or decomposing the dataset into multiple components

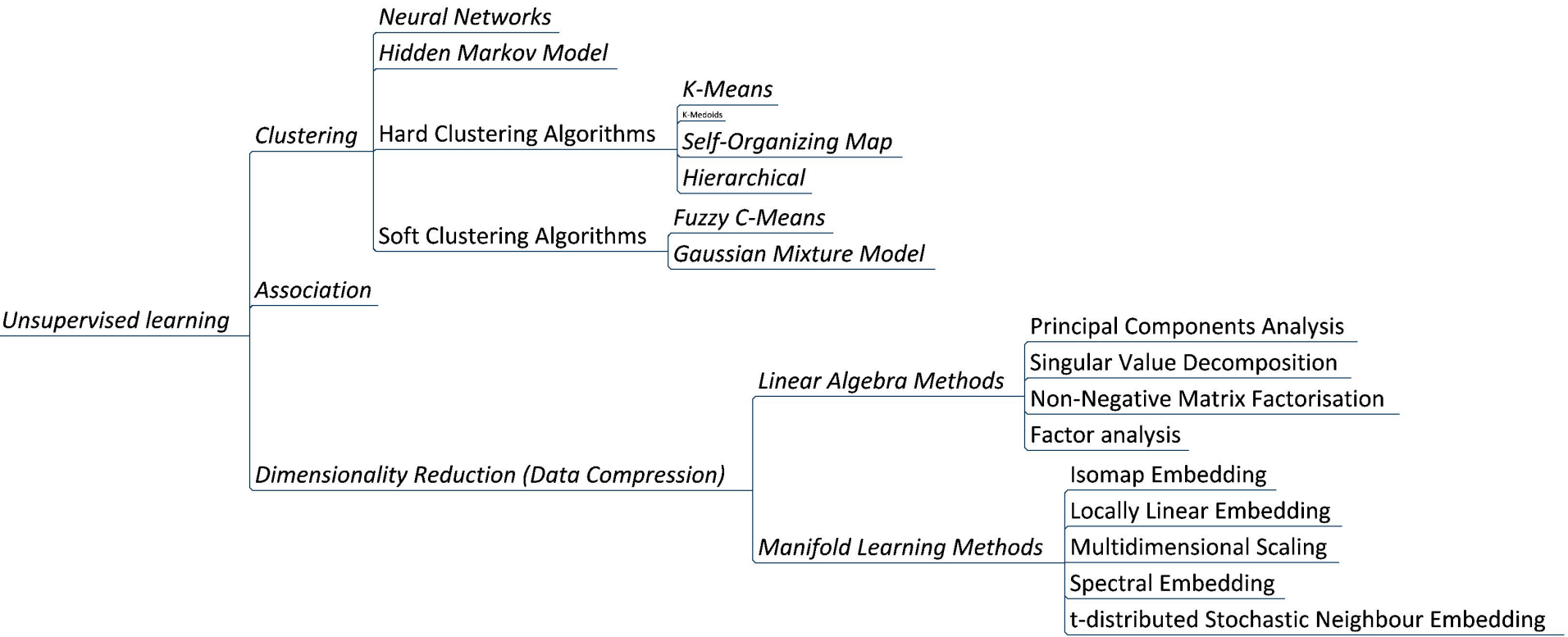
FAQ

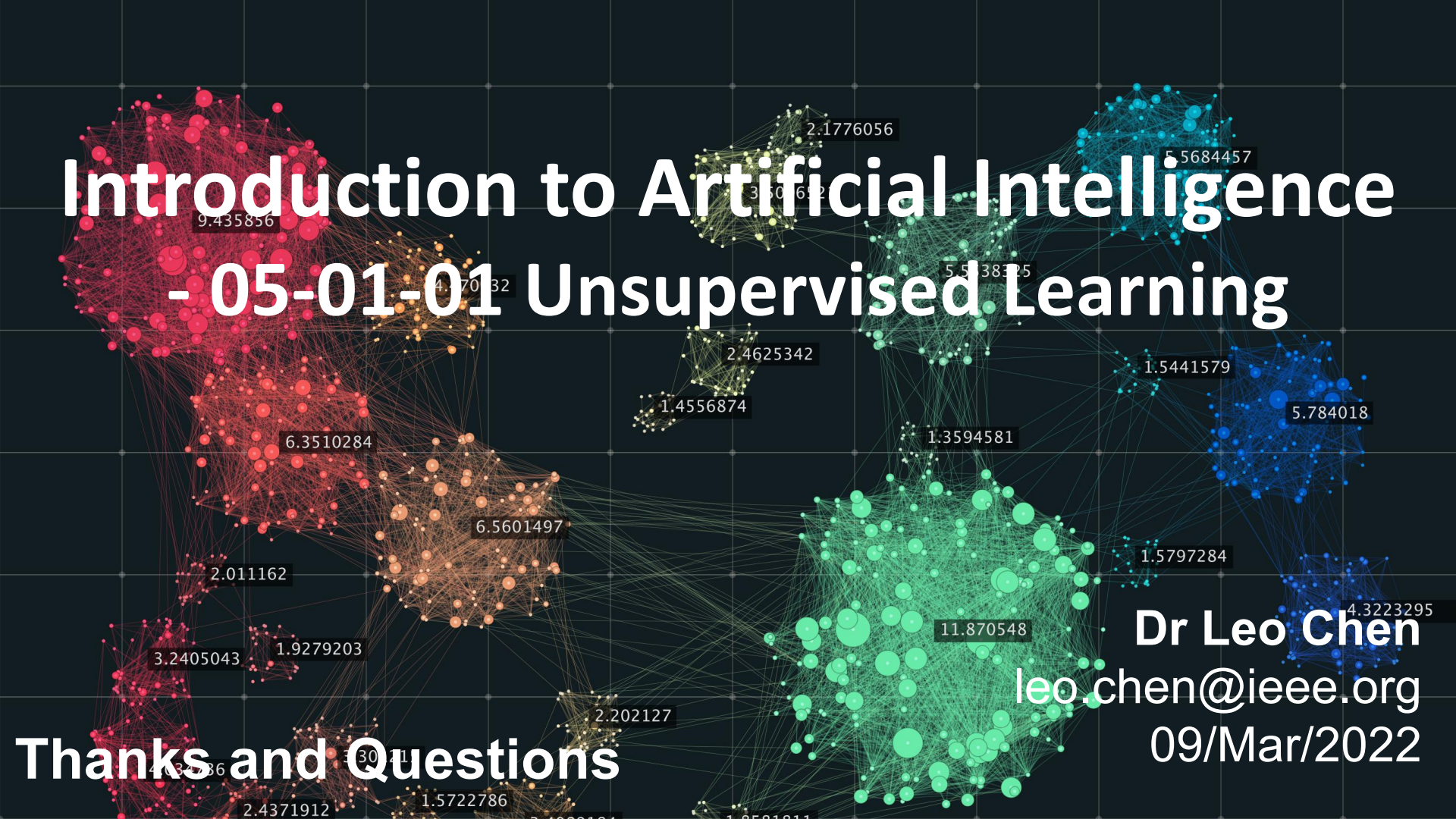
FAQ 01 Why Unsupervised Learning cannot be applied to a regression?



FAQ 01 Why Unsupervised Learning cannot be applied to a regression?

Since it is unknown what the output values/results could be, making it impossible to train the algorithm how you normally would.


Unsupervised learning
Clustering
Neural Networks
Hidden Markov Model
Hard Clustering Algorithms
K-Means
K-Medoids
Self-Organizing Map
Hierarchical
Soft Clustering Algorithms
Fuzzy C-Means
Gaussian Mixture Model
Association
Dimensionality Reduction (Data Compression)
Linear Algebra Methods
Principal Components Analysis
Singular Value Decomposition
Non-Negative Matrix Factorisation
Factor analysis
Manifold Learning Methods
Isomap Embedding
Locally Linear Embedding
Multidimensional Scaling
Spectral Embedding
t-distributed Stochastic Neighbour Embedding



Introduction to Artificial Intelligence - 05-01-01 Unsupervised Learning

Dr Leo Chen
leo.chen@ieee.org
09/Mar/2022

Thanks and Questions