Introduction of Computational Intelligence Laboratory

Yusuke Nojima and Naoki Masuyama

Osaka Metropolitan University

nojima@omu.ac.jp, masuyama@omu.ac.jp

https://ci-labo-omu.github.io

Laboratory Members

Yusuke Nojima
Professor

Naoki Masuyama
Associate Professor

Students

Doctor course: 4

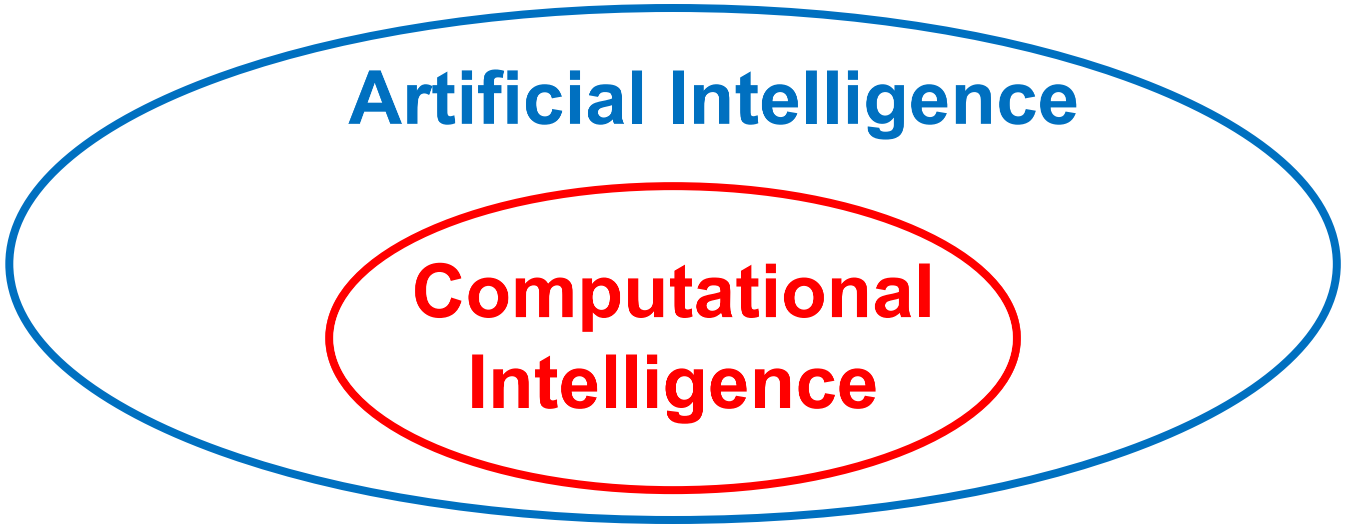
* One American student

Master course: 12

* One French student (Robin)
* One Chinese student

Bachelor course: 4

What is Computational Intelligence?



Artificial Intelligence（Broader Meaning）：

Artificially created intelligent techniques

Computational Intelligence:

Nature-inspired intelligent techniques

(Learning, adaptation, and evolution)

What is Computational Intelligence?

IEEE Computational Intelligence Society

https://cis.ieee.org/

What is Computational Intelligence?

What is Computational Intelligence?

Computational Intelligence (CI) is the theory, design, application and development of biologically and linguistically motivated computational paradigms. Traditionally the three main pillars of CI have been Neural Networks, Fuzzy Systems and Evolutionary Computation. However, in time many nature inspired computing paradigms have evolved. Thus CI is an evolving field and at present in addition to the three main constituents, it encompasses computing paradigms like ambient intelligence, artificial life, cultural learning, artificial endocrine networks, social reasoning, and artificial hormone networks. CI plays a major role in developing successful intelligent systems, including games and cognitive developmental systems. Over the last few years there has been an explosion of research on Deep Learning, in particular deep convolutional neural networks. Nowadays, deep learning has become the core method for artificial intelligence. In fact, some of the most successful AI systems are based on CI.

Core Topics in Our Laboratory

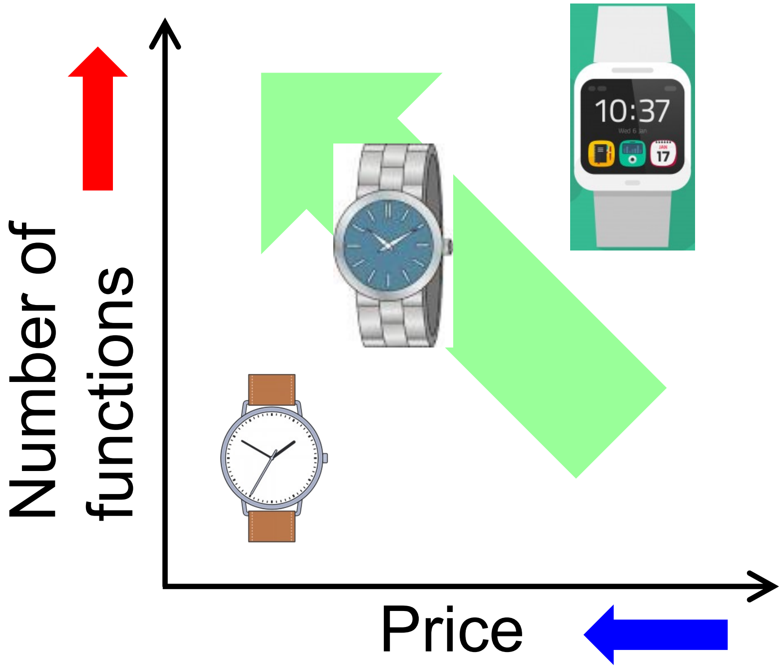
Algorithm Developments of Optimization and Machine Learning based on Computational Intelligence Methodologies

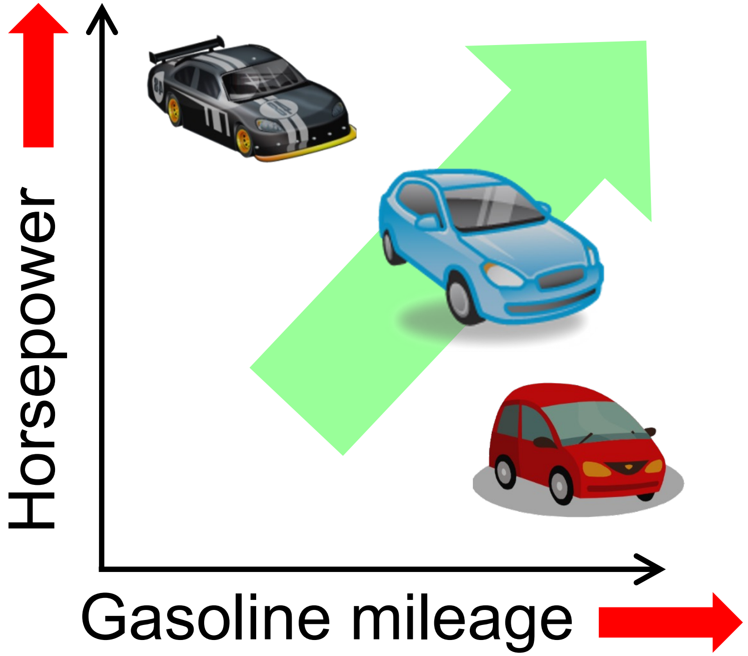
1. Evolutionary Multiobjective Optimization
2. Explainable Machine Learning
3. Continual Machine Learning

1. Evolutionary Multiobjective Optimization

Multiobjective Optimization Problems

- have **multiple conflicting objective functions** to be optimized simultaneously.
  - Minimize ***Price*** and maximize ***Number of functions***
  - Maximize ***Gasoline mileage*** and maximize ***Horsepower***

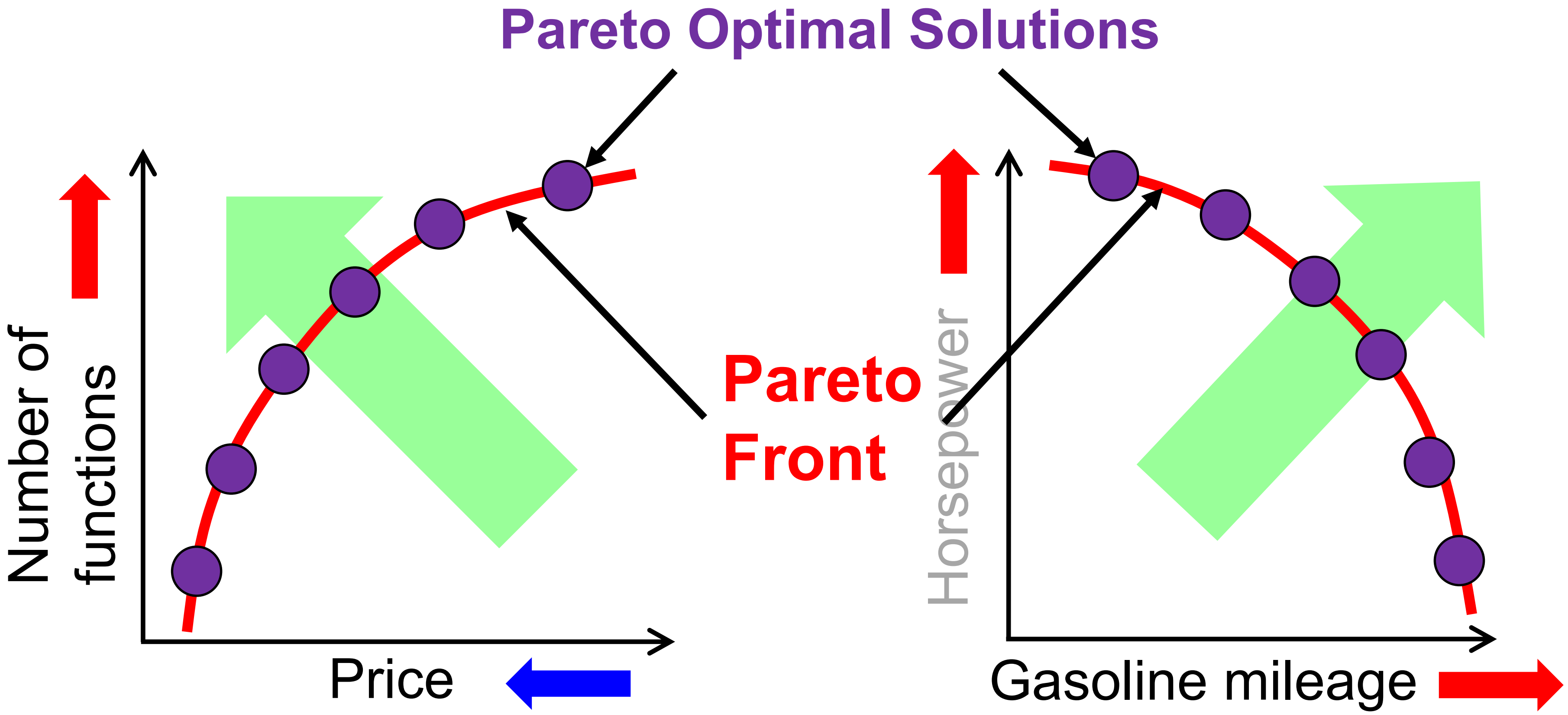




1. Evolutionary Multiobjective Optimization

Multiobjective Optimization Problems

- have **multiple conflicting objective functions** to be optimized simultaneously.
  - Minimize ***Price*** and maximize ***Number of functions***
  - Maximize ***Gasoline mileage*** and maximize ***Horsepower***

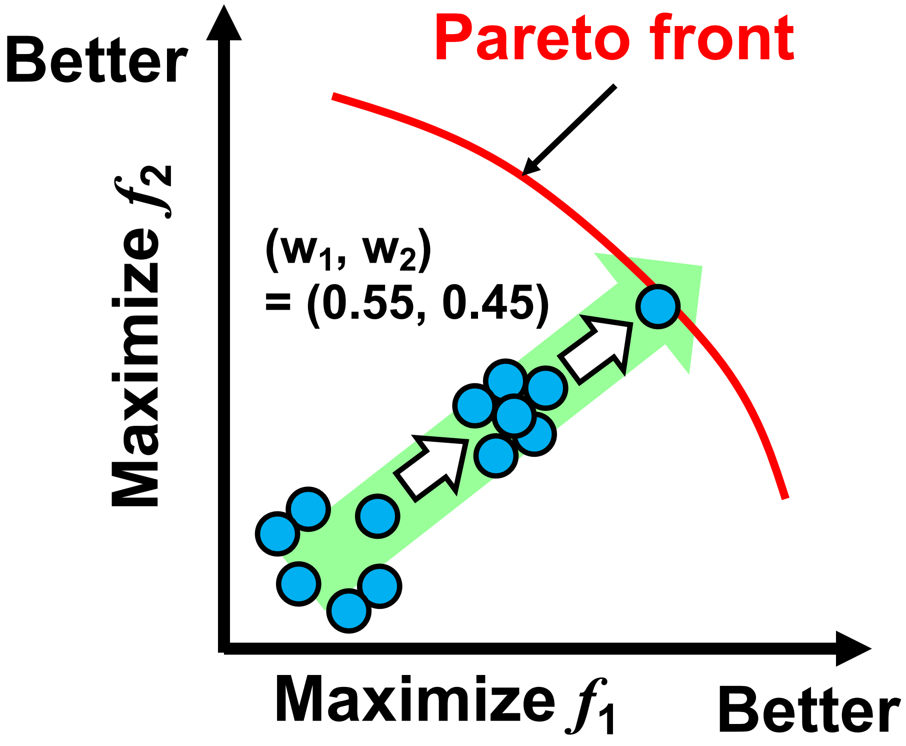


1. Evolutionary Multiobjective Optimization

Traditional Evolutionary Computation Approach

- uses a weighted sum fitness function F for parent selection and generation update.
- can find a single solution near to the Pareto front.

$$F = w_1 f_1 + w_2 f_2$$

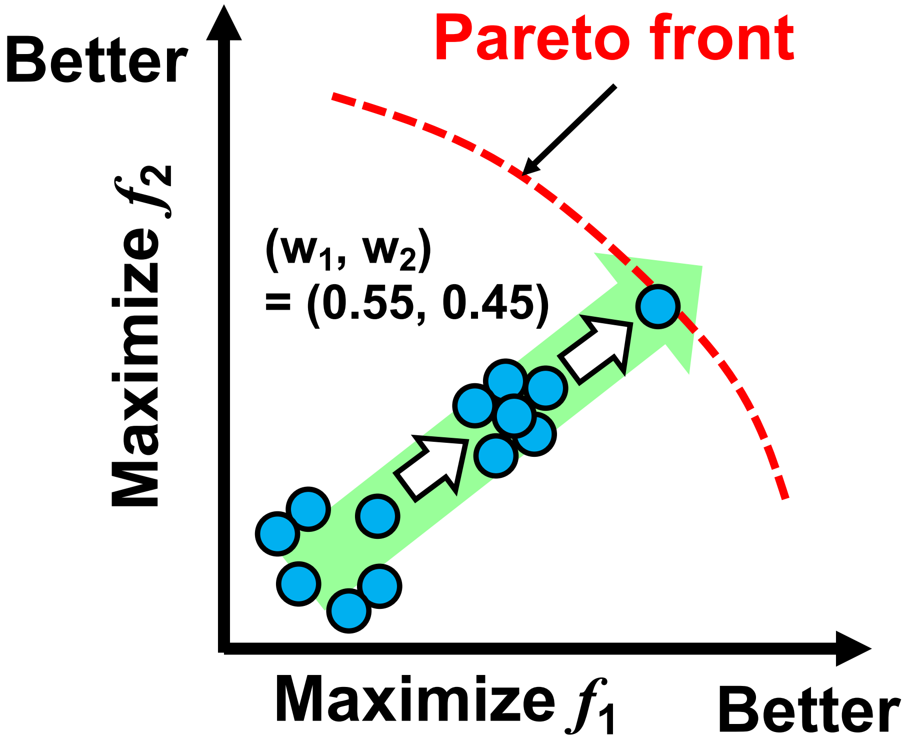


1. Evolutionary Multiobjective Optimization

Traditional Evolutionary Computation Approach

- uses a weighted sum fitness function F for parent selection and generation update.
- can find a single solution near to the Pareto front.
- hardly specifies a weight vector because of unknown Pareto front shape.

$$F = w_1 f_1 + w_2 f_2$$

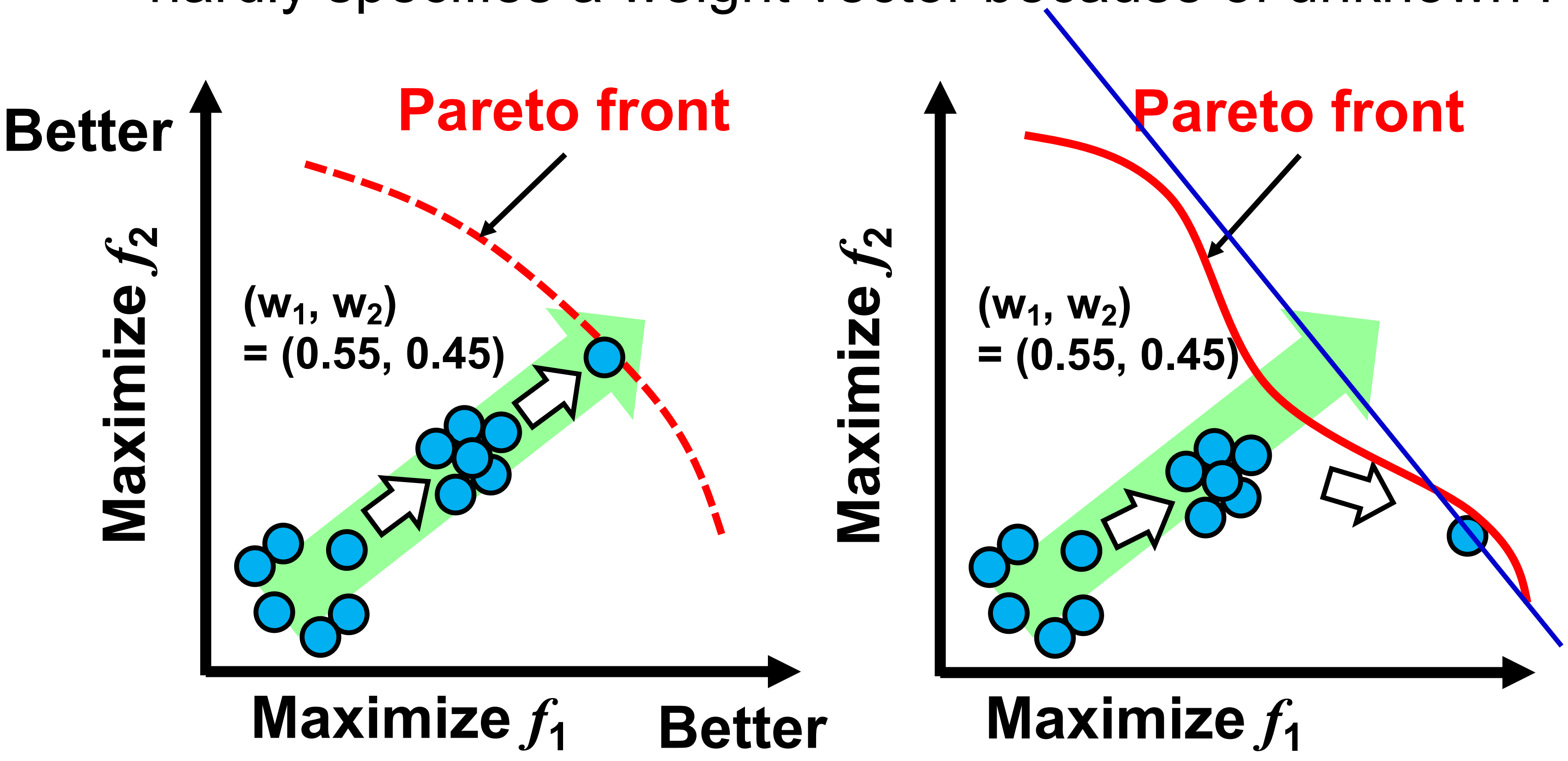


1. Evolutionary Multiobjective Optimization

Traditional Evolutionary Computation Approach

- uses a weighted sum fitness function F for parent selection and generation update.
- can find a single solution near to the Pareto front.
- hardly specifies a weight vector because of unknown Pareto front shape.

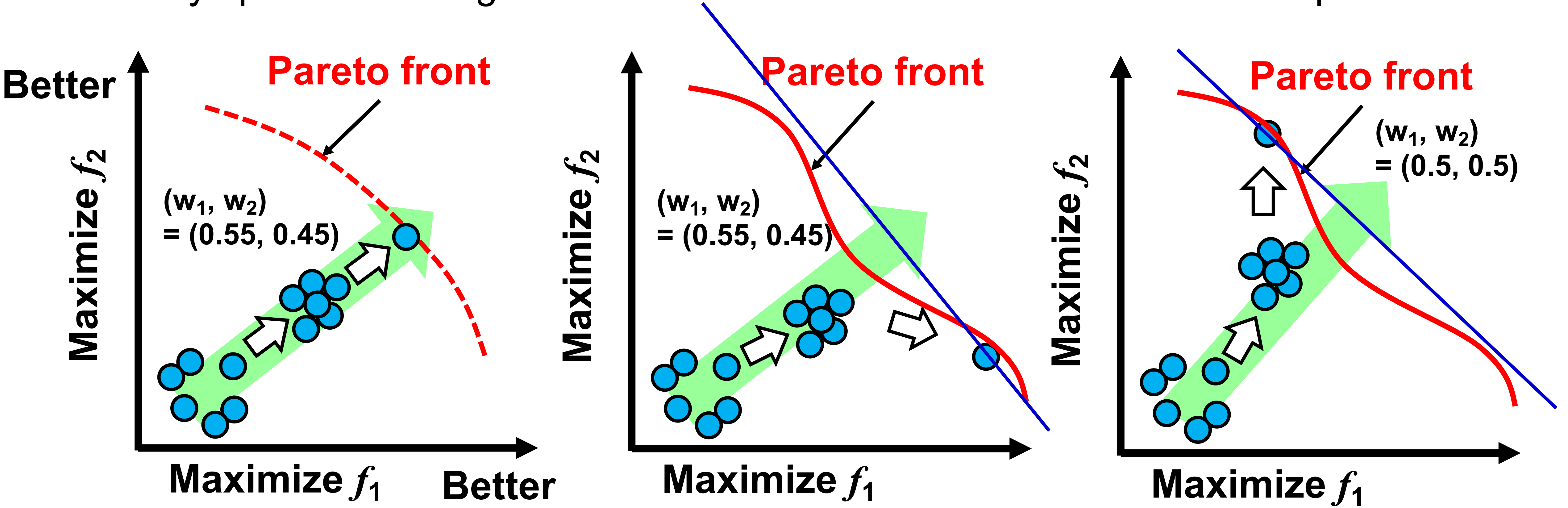
$$F = w_1 f_1 + w_2 f_2$$



1. Evolutionary Multiobjective Optimization
Traditional Evolutionary Computation Approach

- uses a weighted sum fitness function F for parent selection and generation update.
- can find a single solution near to the Pareto front.
- hardly specifies a weight vector because of unknown Pareto front shape.

$$F = w_1 f_1 + w_2 f_2$$



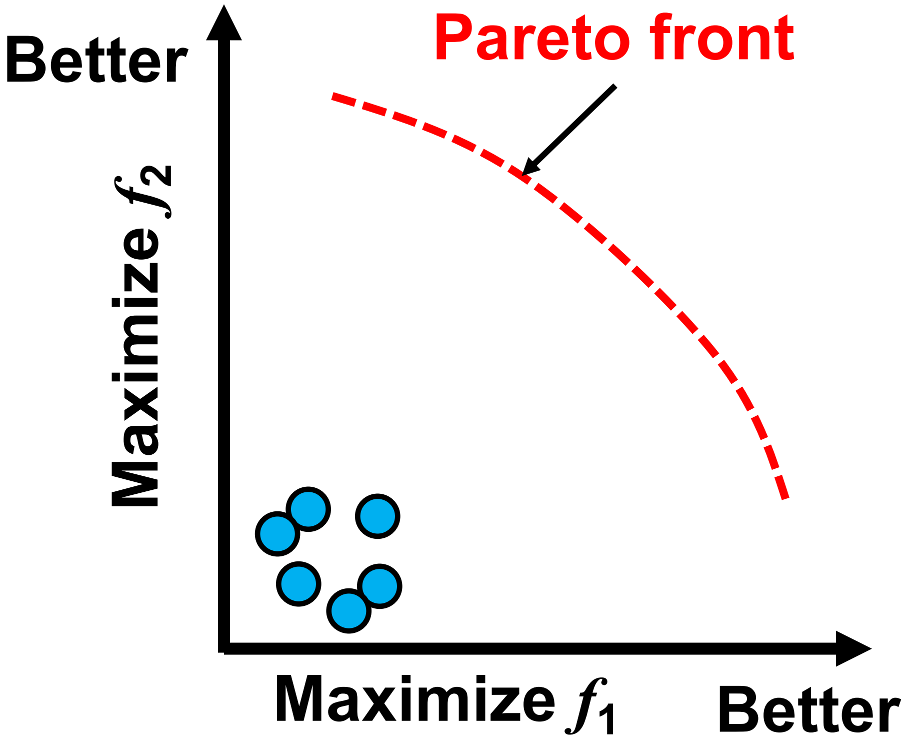
1. Evolutionary Multiobjective Optimization

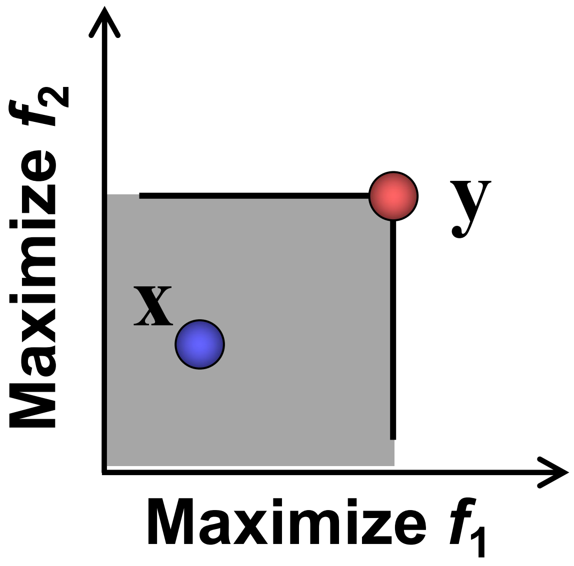
Evolutionary Multiobjective Optimization (EMO)

- uses the **dominance relation** for parent selection and generation update.

For maximization problems,

$$\boxed{f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \quad \text{for } \forall i \quad \text{and} \quad f_j(\mathbf{x}) < f_j(\mathbf{y}) \quad \text{for } \exists j}$$





$\mathbf{y}$ dominates $\mathbf{x}$.

$\mathbf{x}$ is dominated by $\mathbf{y}$.

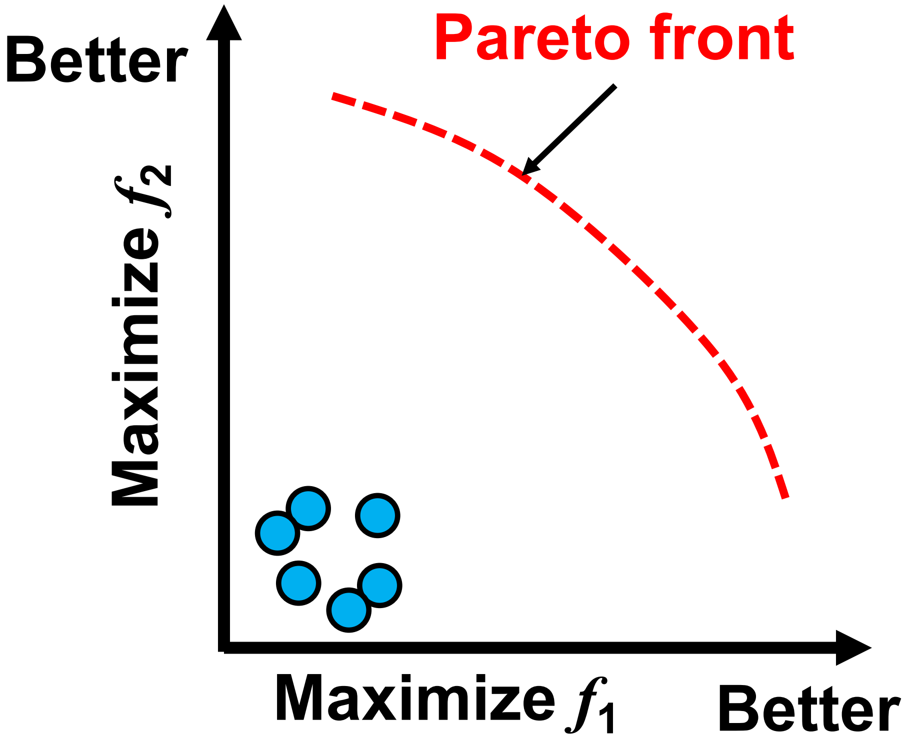
Dominance solutions have a high priority to be selected as parents and for the next population.

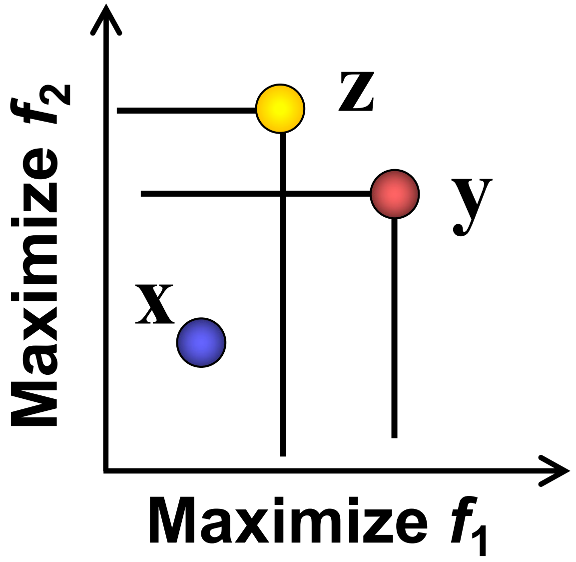
Evolutionary Multiobjective Optimization (EMO)

- uses the **dominance relation** for parent selection and generation update.

For maximization problems,

$$\boxed{f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \quad \text{for } \forall i \quad \text{and} \quad f_j(\mathbf{x}) < f_j(\mathbf{y}) \quad \text{for } \exists j}$$





y dominates **x**.

x is dominated by **y**.

z dominates **x**.

x is dominated by **z**.

y and **z** are non-dominated each other.

Dominance solutions have a high priority to be selected as parents and for the next population.

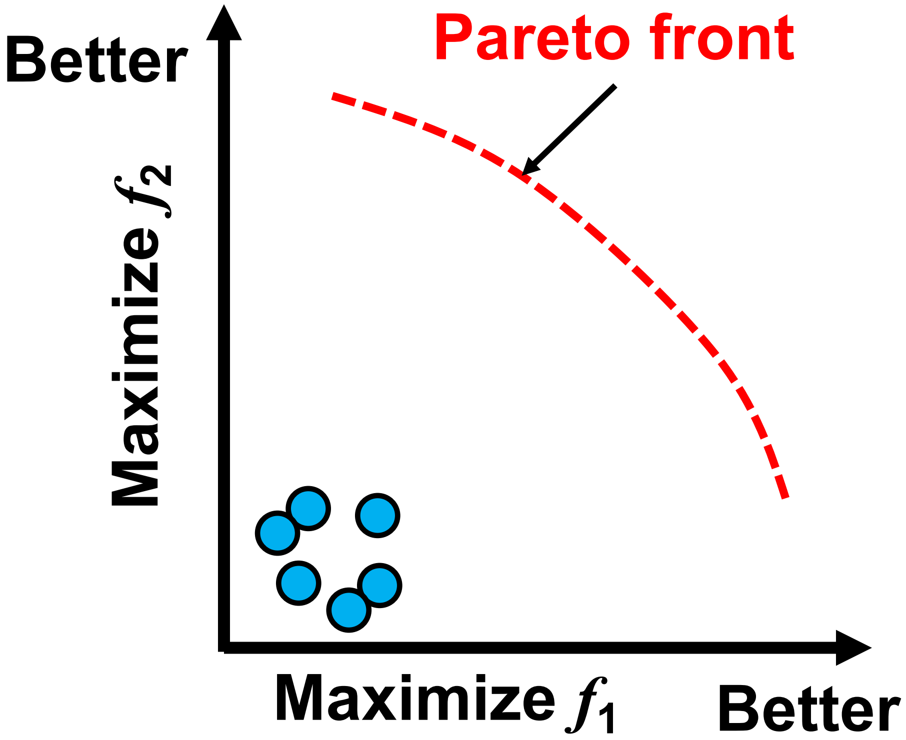
1. Evolutionary Multiobjective Optimization

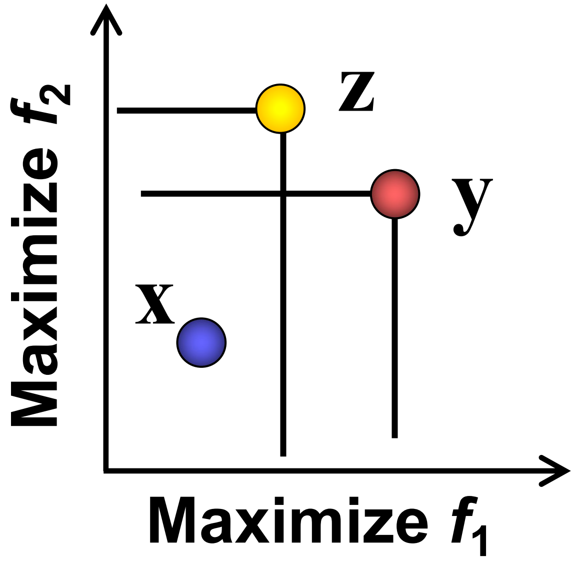
Evolutionary Multiobjective Optimization (EMO)

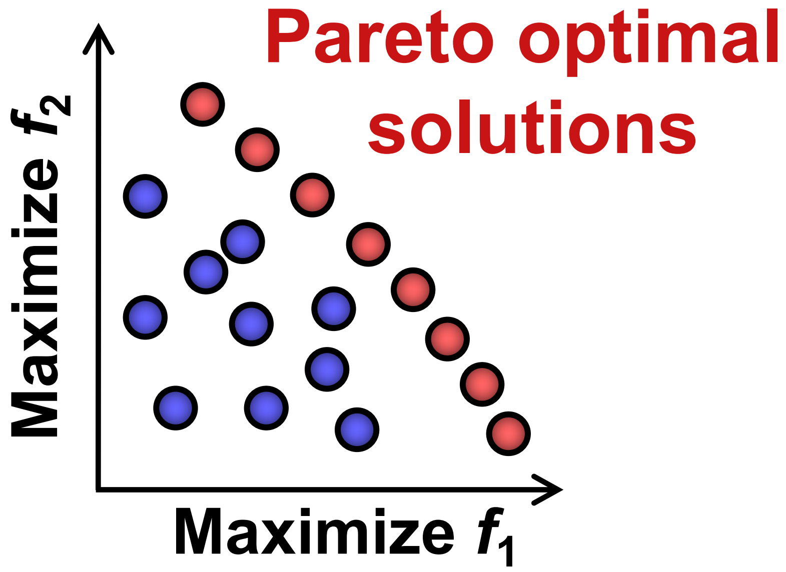
- uses the **dominance relation** for parent selection and generation update.

For maximization problems,

$$\boxed{f_i(\mathbf{x}) \le f_i(\mathbf{y}) \ \text{ for } \forall i \quad \text{and} \quad f_j(\mathbf{x}) < f_j(\mathbf{y}) \ \text{ for } \exists j}$$



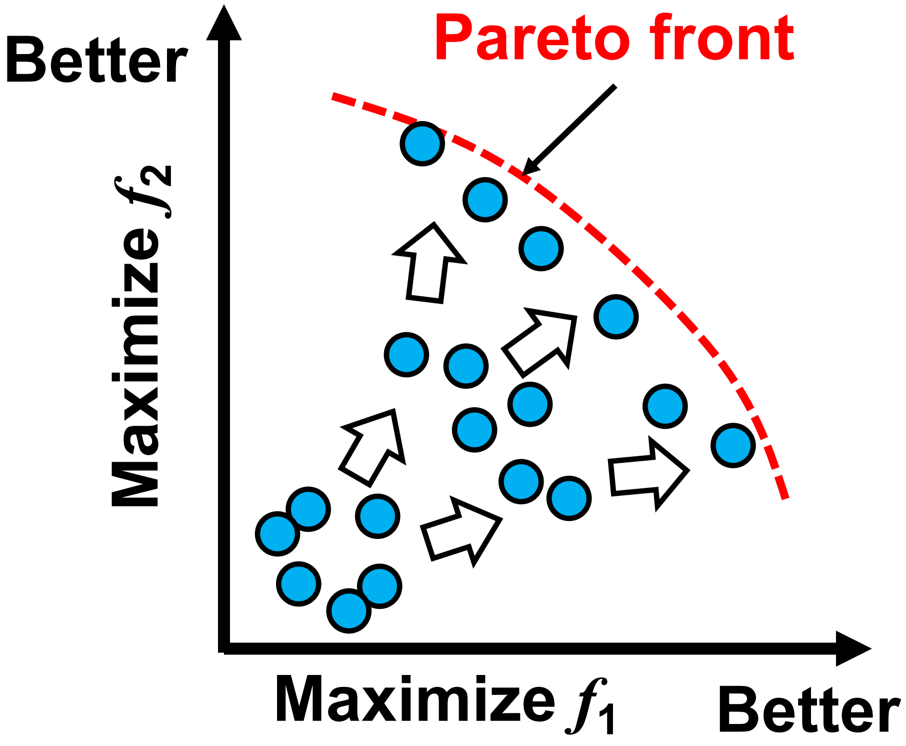




If a solution is not dominated by others, the solution is Pareto optimal.

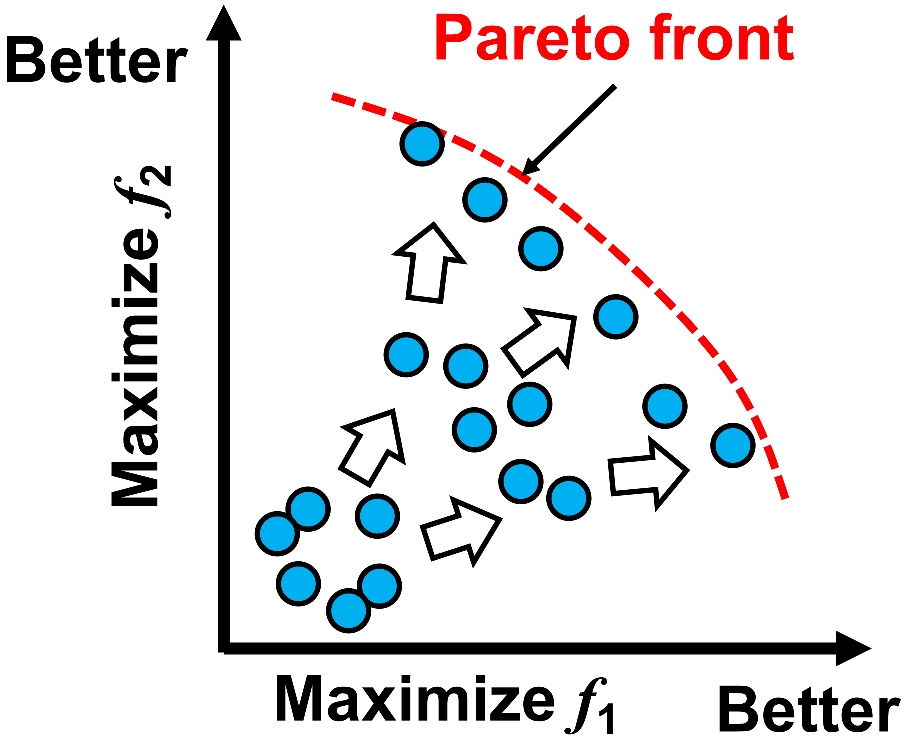
Evolutionary Multiobjective Optimization (EMO)

- uses the **dominance relation** for parent selection and generation update.
- **highly selects non-dominated solutions and gives a higher priority to solutions in a sparse area to improve the convergence toward the Pareto front and the diversity over the Pareto front.**



Evolutionary Multiobjective Optimization (EMO)

- uses the **dominance relation** for parent selection and generation update.
- **highly selects non-dominated solutions and gives a higher priority to solutions in a sparse area to improve the convergence toward the Pareto front and the diversity over the Pareto front.**



Advantages of EMO

- **The shape of the Pareto front can be approximate.**
- **Various options can be shown to the user.**
- **Population diversity improves the search performance.**

Evolutionary Multiobjective Optimization (EMO)

The current research interests on EMO

1. **Evolutionary many-objective optimization**
2. **Constrained multiobjective problems**
3. **Multimodal multiobjective problems**
4. **Analysis on real-world optimization problems**

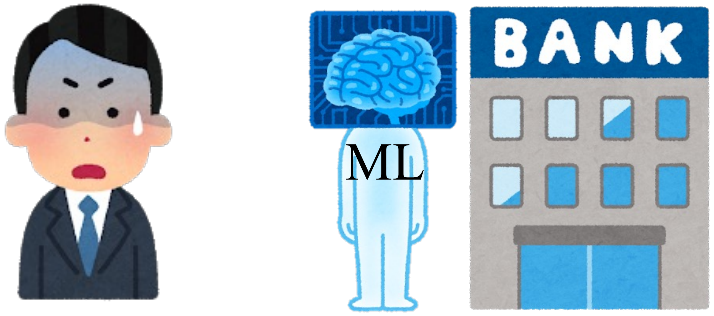
2. Explainable Machine Learning
Explainability in Machine Learning

- **is necessary to make users convince ML-based systems/models for real-world applications.**

Can I apply for a mortgage?

No. Your score is too low.

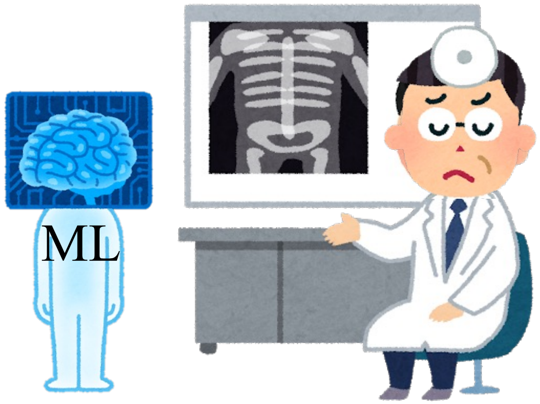
Why?



Why did this autonomous car hit a bike?

This patient has a lung cancer.

This patient has a lung cancer.



Really!? This shadow is not a cancer from my experience.

Could you explain?

2. Explainable Machine Learning
Explainability in Machine Learning

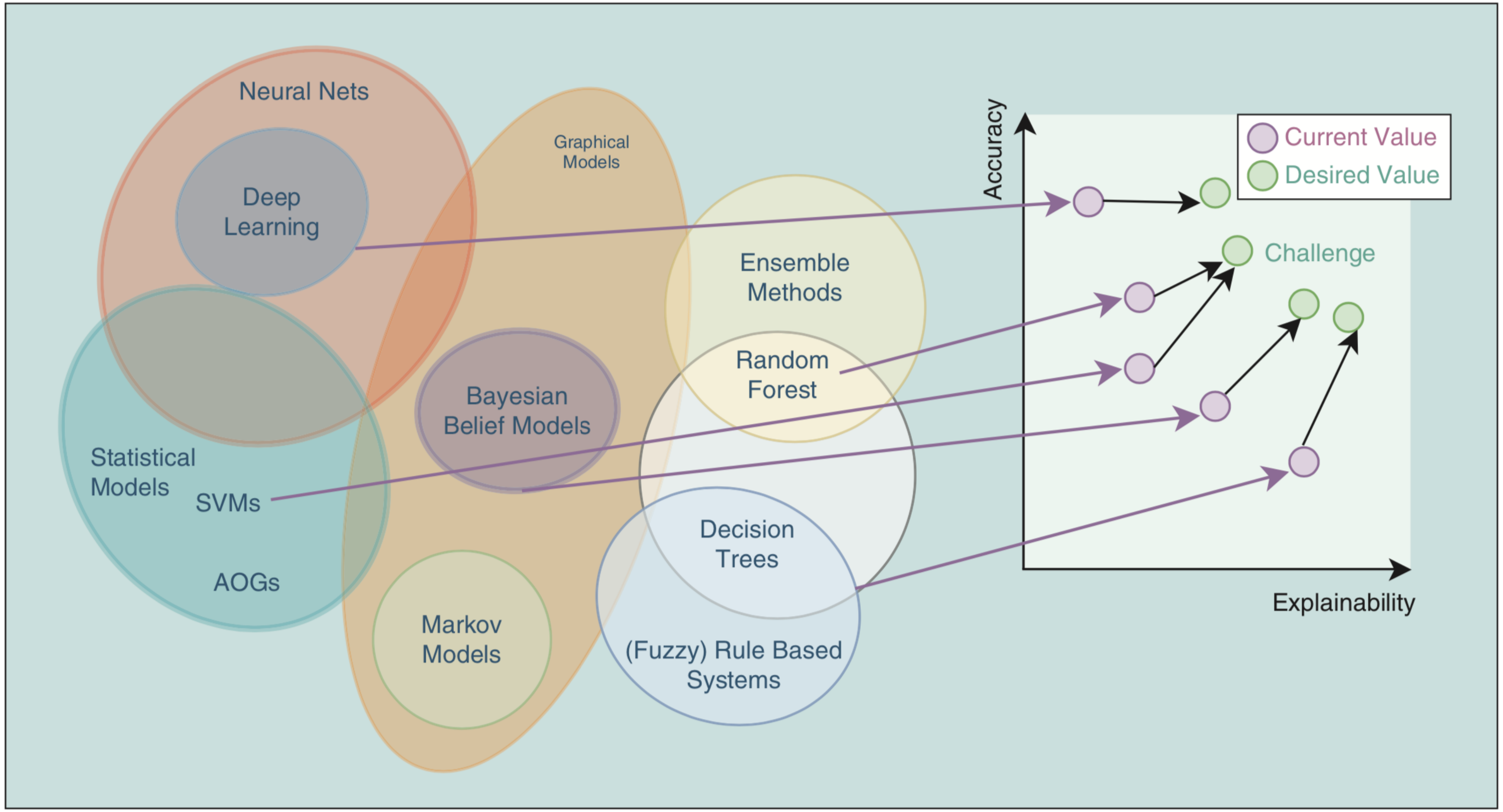


FIGURE 3 Accuracy vs. explainability: Venn diagram for several ML models (partially inspired from [74]).

A. Fernández, F. Herrera, O. Cordon, M. J. del Jesus, and F. Marcelloni, "Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?," *IEEE Computational Intelligence Magazine*, 14 (1), 69–81, Jan. 2019.

Why are Fuzzy Systems Promising for an Explainable Issue?

Example: Classification of body shape balance

Class 1 : **Balanced Body Shape**

Rule 1: If Height is *short* and Weight is *light* then Class 1

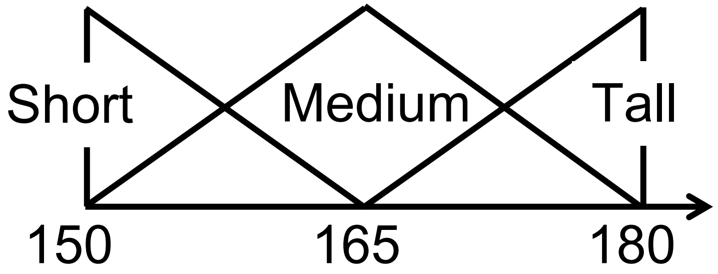
Rule 2: If Height is *medium* and Weight is *medium* then Class 1

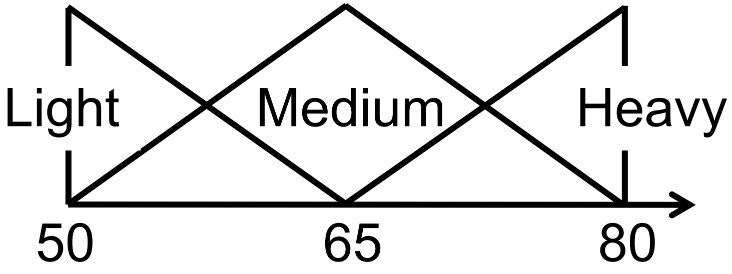
Rule 3: If Height is *tall* and Weight is *heavy* then Class 1

Class 2 : **Unbalanced Body Shape**

Rule 4: If Height is *short* and Weight is *heavy* then Class 2

Rule 5: If Height is *tall* and Weight is *light* then Class 2



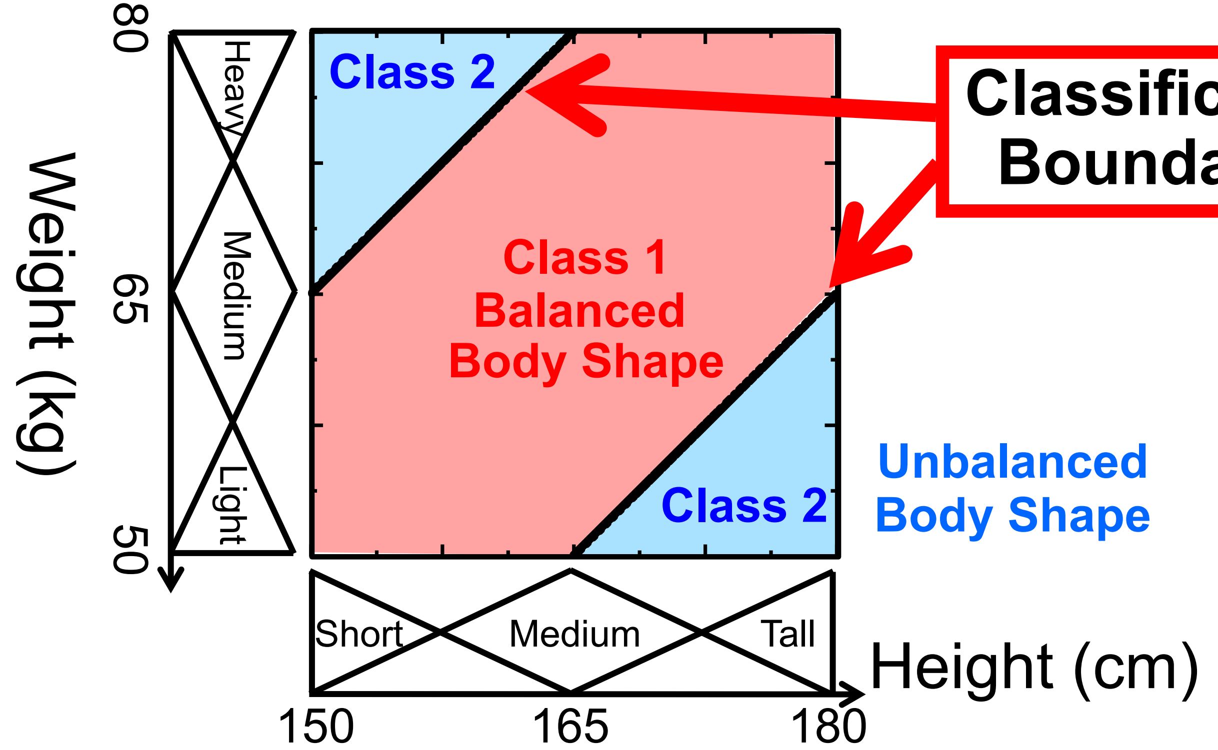


Our knowledge can be easily represented by linguistic values.

2. Explainable Machine Learning

Why are Fuzzy Systems Promising for an Explainable Issue?

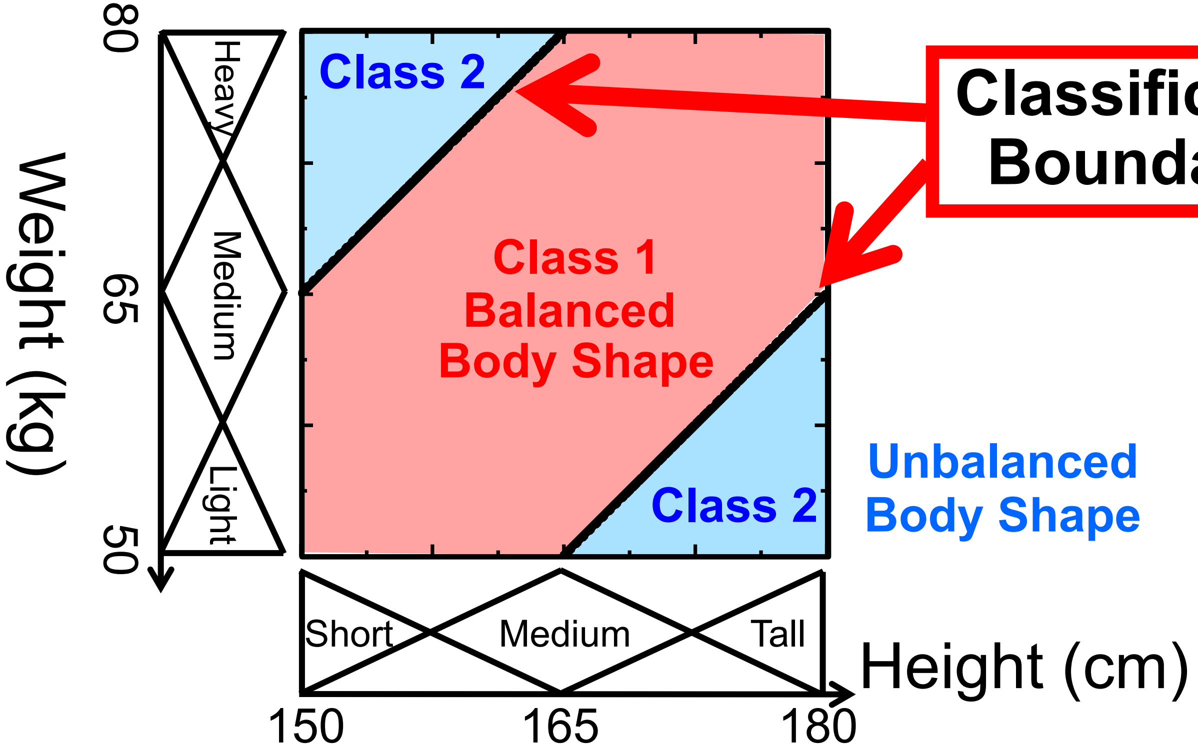
Rule 1: If Height is *short* and Weight is *light* then Class 1
Rule 2: If Height is *medium* and Weight is *medium* then Class 1
Rule 3: If Height is *tall* and Weight is *heavy* then Class 1
Rule 4: If Height is *short* and Weight is *heavy* then Class 2
Rule 5: If Height is *tall* and Weight is *light* then Class 2



2. Explainable Machine Learning

Why are Fuzzy Systems Promising for an Explainable Issue?

Rule 1: If Height is *short* and Weight is *light* then Class 1
Rule 2: If Height is *medium* and Weight is *medium* then Class 1
Rule 3: If Height is *tall* and Weight is *heavy* then Class 1
Rule 4: If Height is *short* and Weight is *heavy* then Class 2
Rule 5: If Height is *tall* and Weight is *light* then Class 2



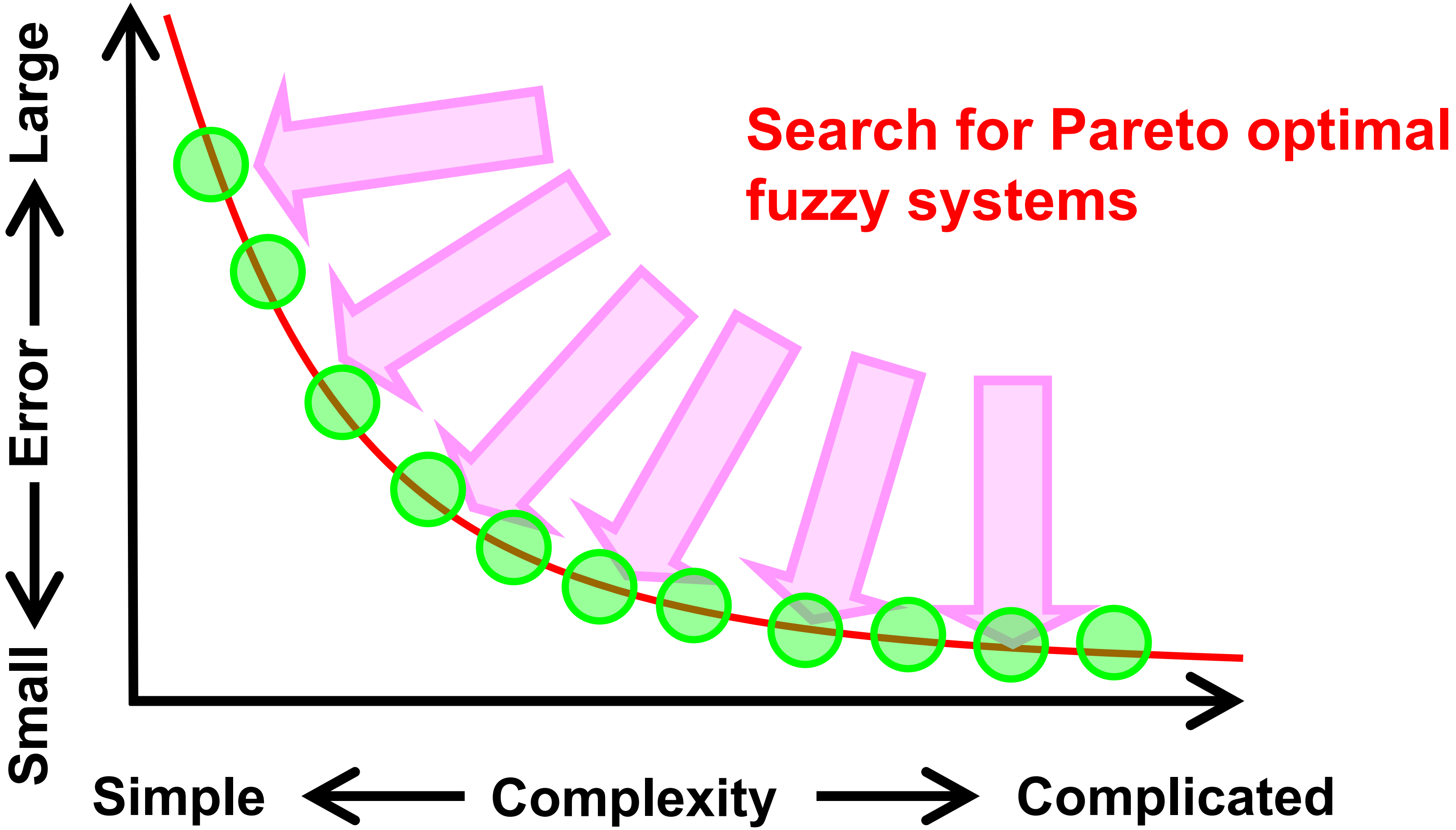
All fuzzy systems are not necessarily explainable.

2. Explainable Machine Learning

Application of EMO Algorithms to Fuzzy System Design

Multi-objective optimization

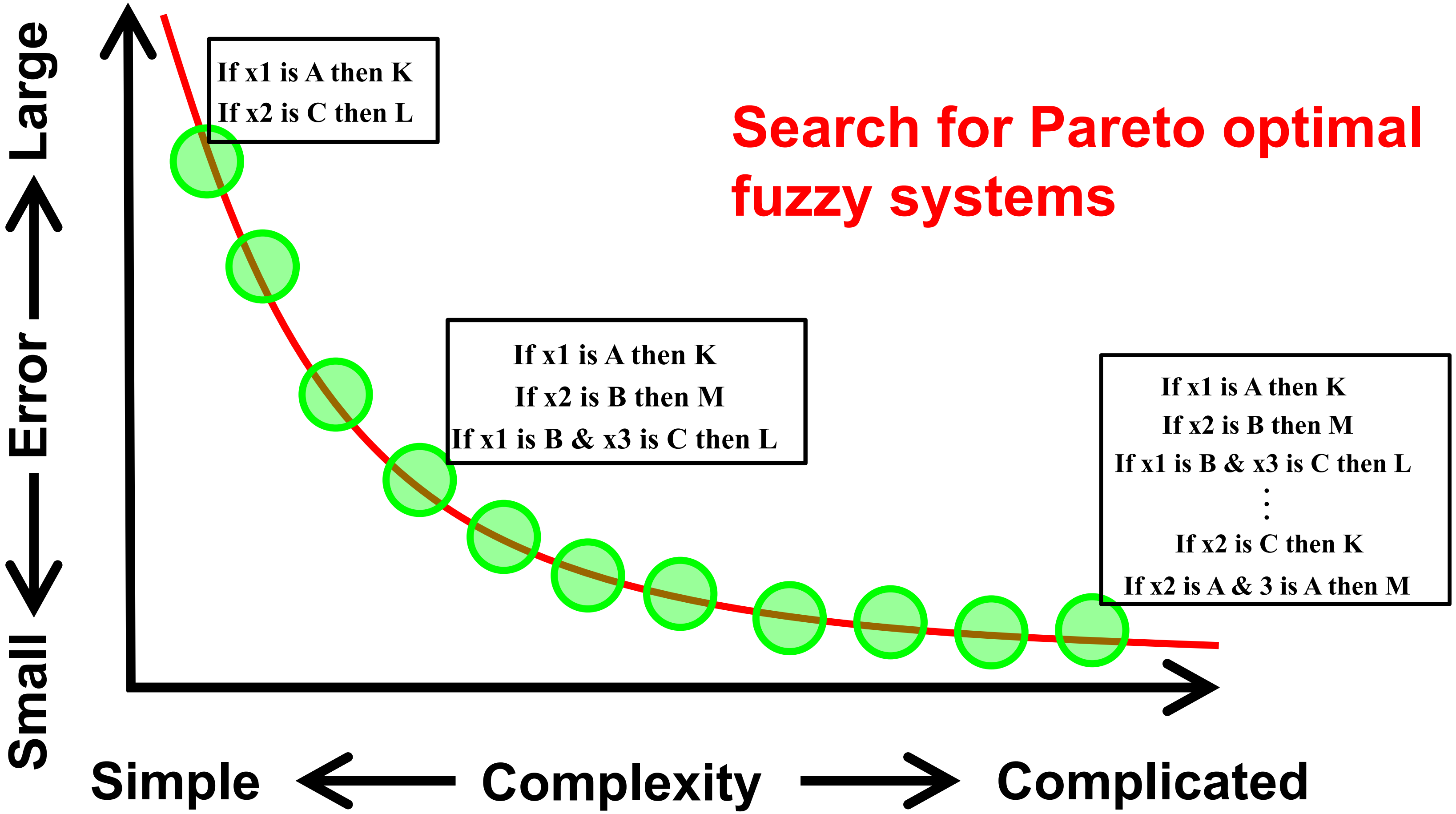
Minimize *Error* and minimize *Complexity*



2. Explainable Machine Learning
Application of EMO Algorithms to Fuzzy System Design

Multi-objective optimization

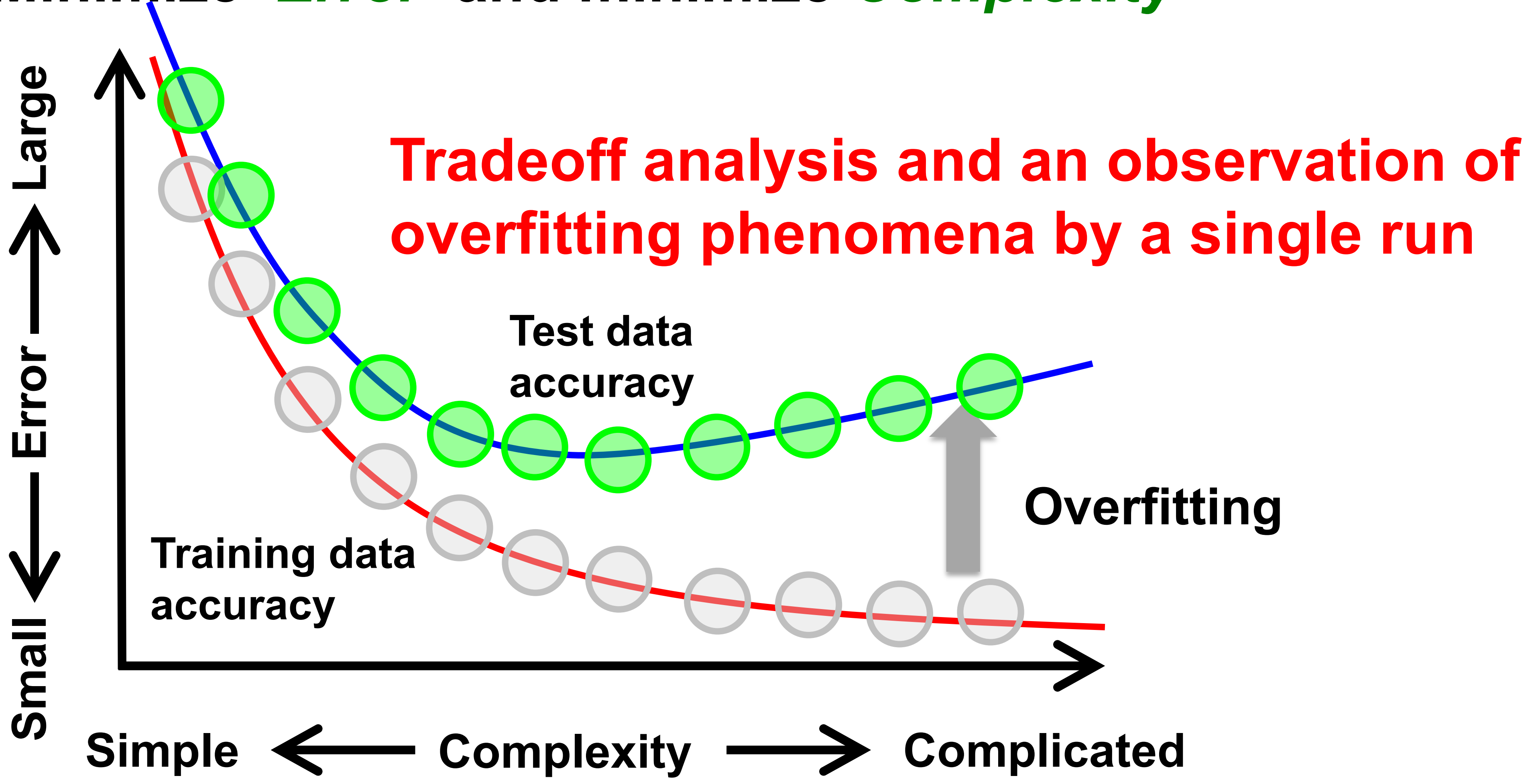
Minimize *Error* and minimize *Complexity*



Application of EMO Algorithms to Fuzzy System Design

Multi-objective optimization

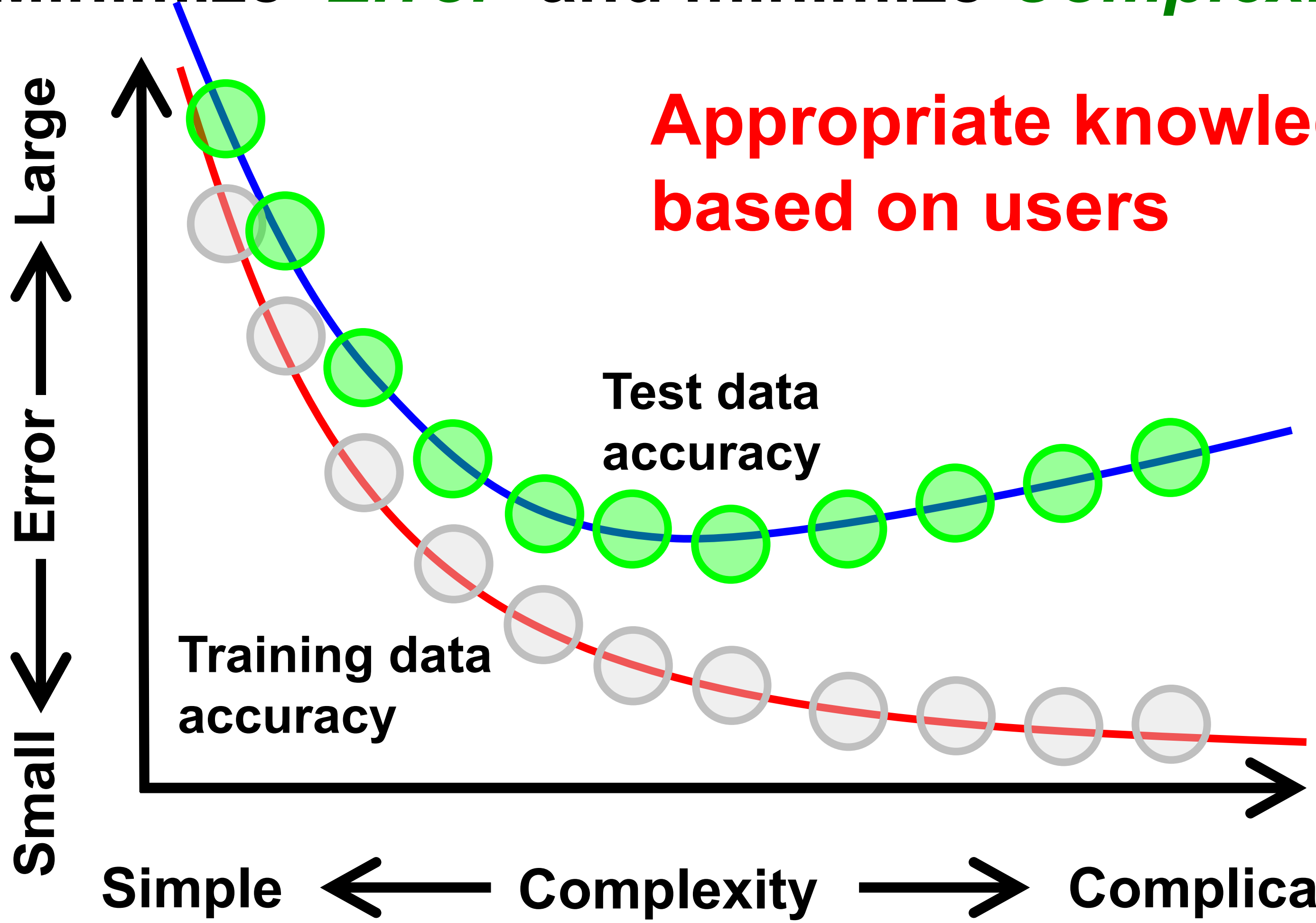
Minimize *Error* and minimize *Complexity*



2. Explainable Machine Learning
Application of EMO Algorithms to Fuzzy System Design

Multi-objective optimization

Minimize *Error* and minimize *Complexity*



Appropriate knowledge selection based on users

2. Explainable Machine Learning

Evolutionary Multiobjective Fuzzy Systems

The current research interests on Evolutionary Multiobjective Fuzzy Systems

1. **Accuracy Enhancement**
2. **Incorporation of Fairness Measures**
3. **Classifier Design with a Reject Option**
4. **Partial Interpretable Models**

Core Topics

Algorithm Developments of Optimization and Machine Learning based on Computational Intelligence Methodologies

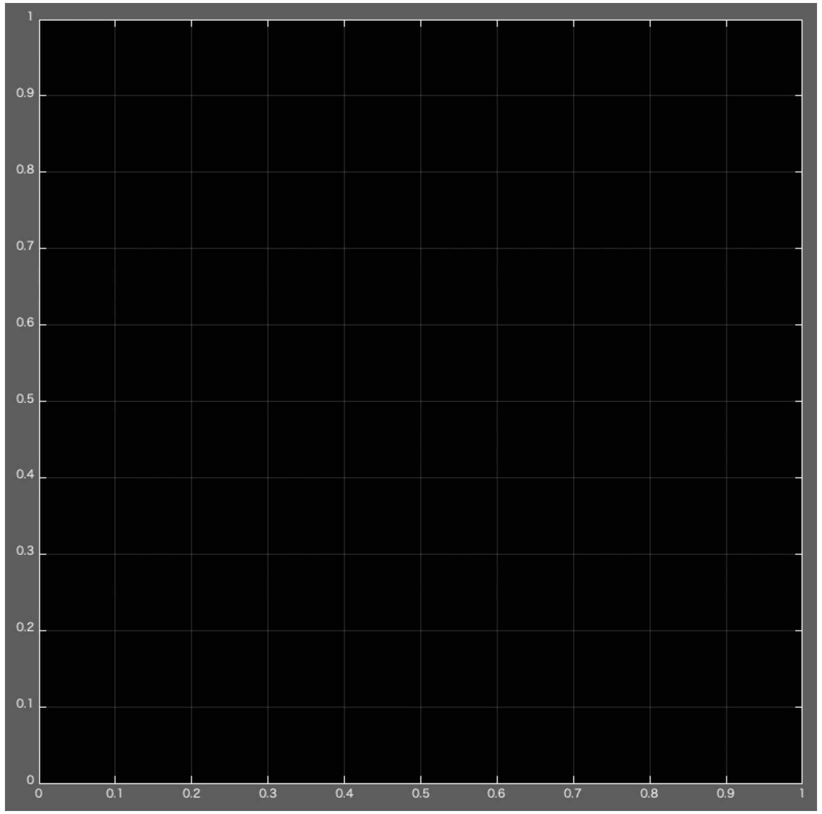
1. Evolutionary Multiobjective Optimization
2. Explainable Machine Learning
3. Continual Machine Learning

   Growing Self-Organizing Clustering

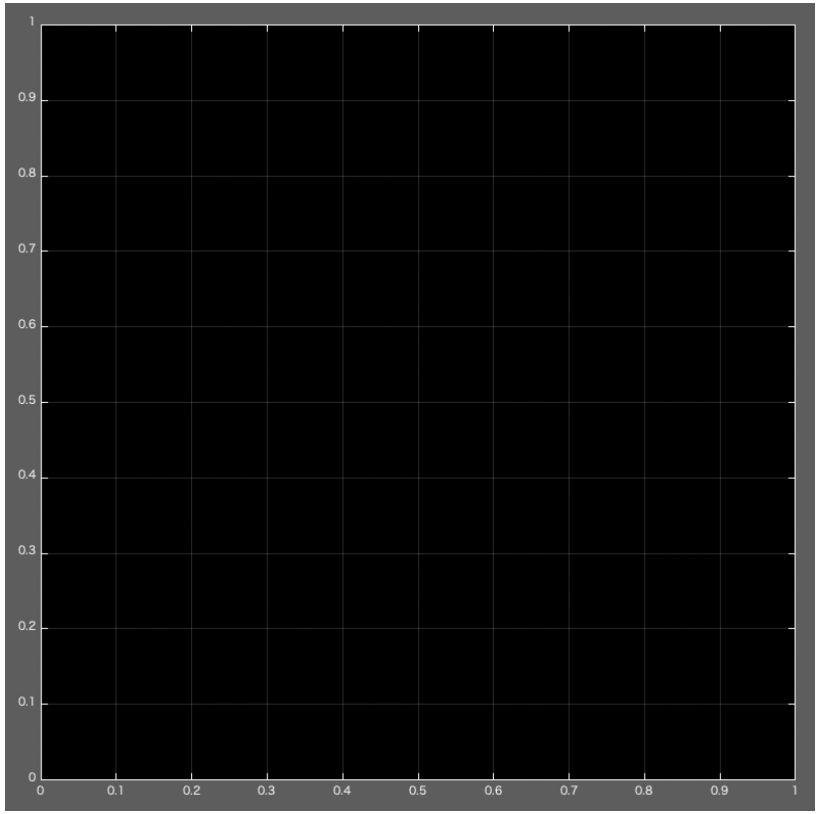
Continual Machine Learning

Growing Self-organizing Clustering

- Adaptively and continually extract information from given data



Input

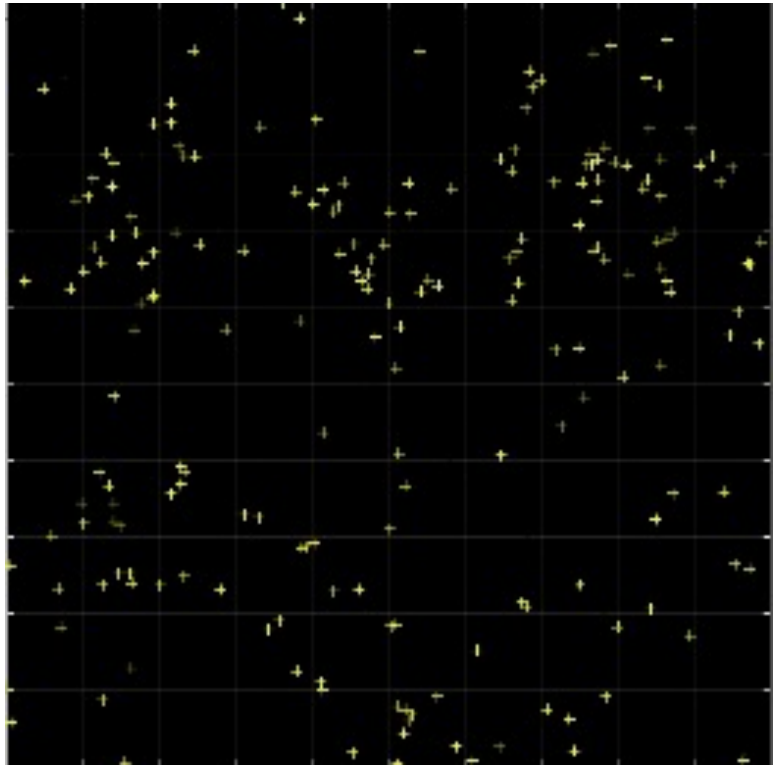


Extract information by nodes and construct their relationships by edges

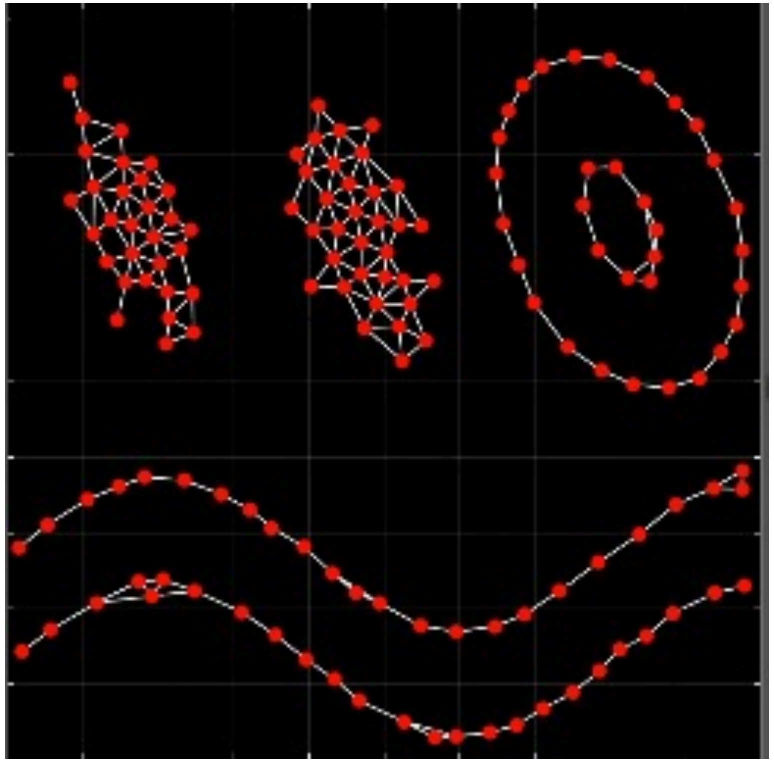
Continual Machine Learning

Growing Self-organizing Clustering

- Adaptively and continually extract information from given data



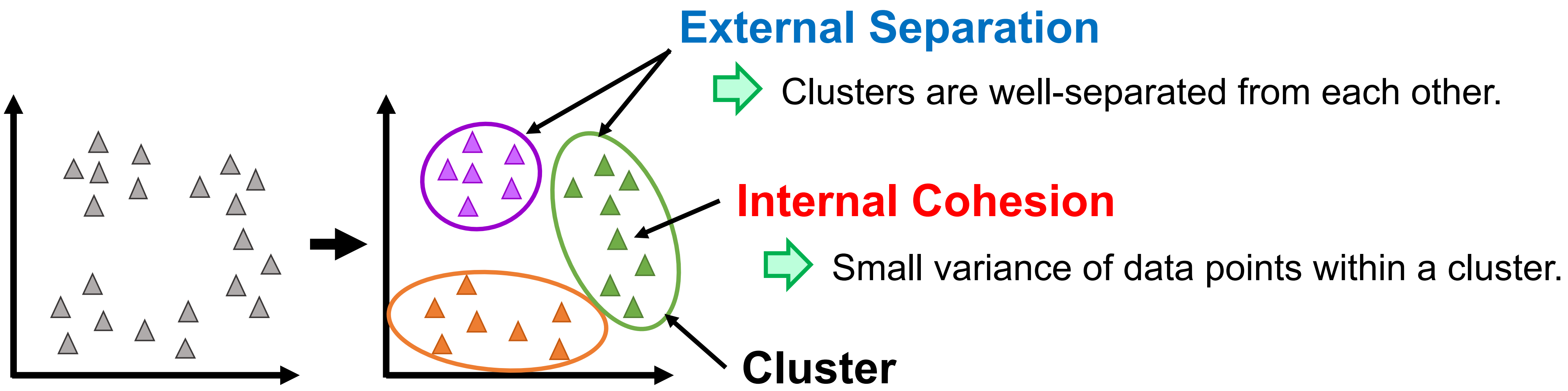
Input



Extract information by nodes and construct their relationships by edges

Clustering
Overview

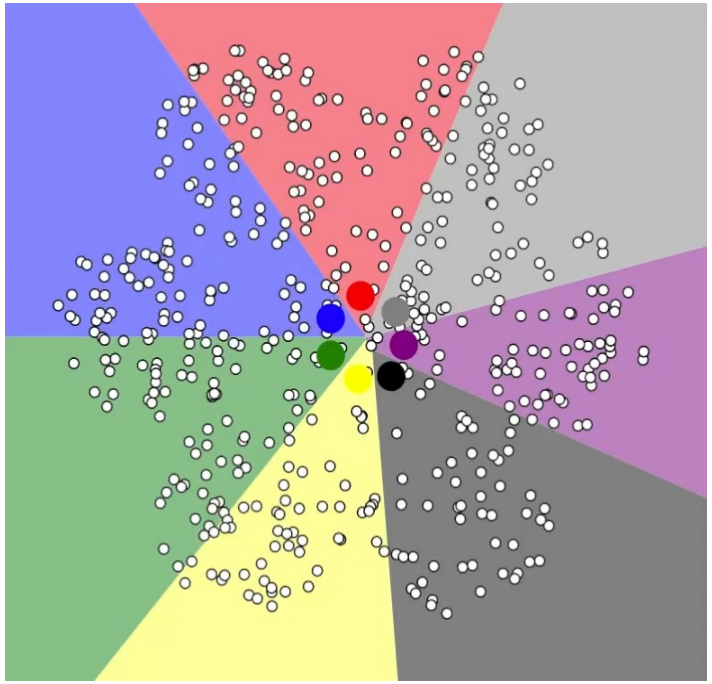
- Clustering aims to divide observed data into subsets that achieve internal cohesion and external separation.



- Hard Clustering
  - Each data point belongs to only one cluster. e.g.) k-means
- Soft Clustering
  - Each data point can belong to multiple clusters. e.g.) fuzzy c-means

Clustering
k-means

- The objective of the k-means is to minimize the sum of distances between the data points and their respective centroid (node).



https://www.naftaliharris.com/blog/visualizing-k-means-clustering/

➢ Objective Function

$$\mathcal{L} = \sum_{k=1}^{K} \sum_{i \in C_k} \|\boldsymbol{x}_i - \boldsymbol{\mu}_k\|^2$$

where $$\boldsymbol{\mu}_k = \frac{1}{|C_k|} \sum_{i \in C_k} \boldsymbol{x}_i$$

$\mathcal{L}$: Loss

K : Number of Centroids

C_k : k-th Centroids

x_i : Data Point

μ_k : Position of k-th Centroids

- The number of centroids (nodes) needs to be specified in advance.
- **Not suitable for complex, non-spherical, or dynamically changing data distributions.**

Clustering
Effects of Node Organization

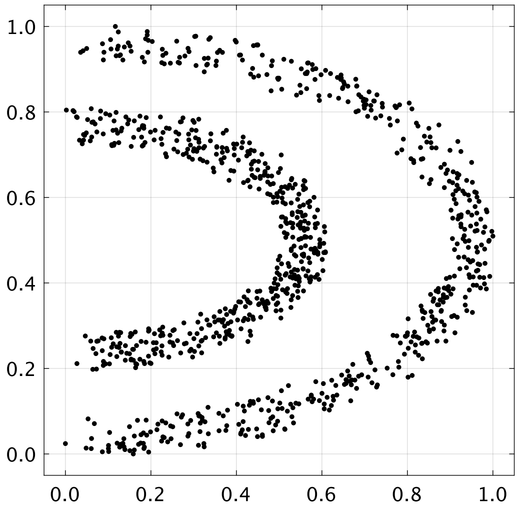
ARI : Adjusted Rand Index

AMI : Adjusted Mutual Information

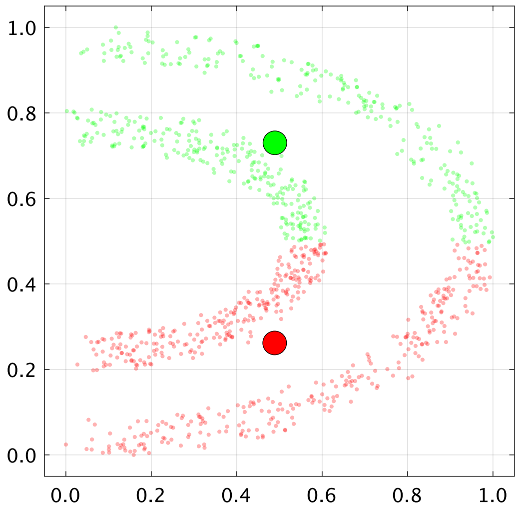
The higher, the better.

- Depending on the data distribution, the number of nodes should NOT be the same as the number of true clusters.

dataset (2 clusters)



2 nodes (k-means)



ARI = 0.00 , AMI = 0.00

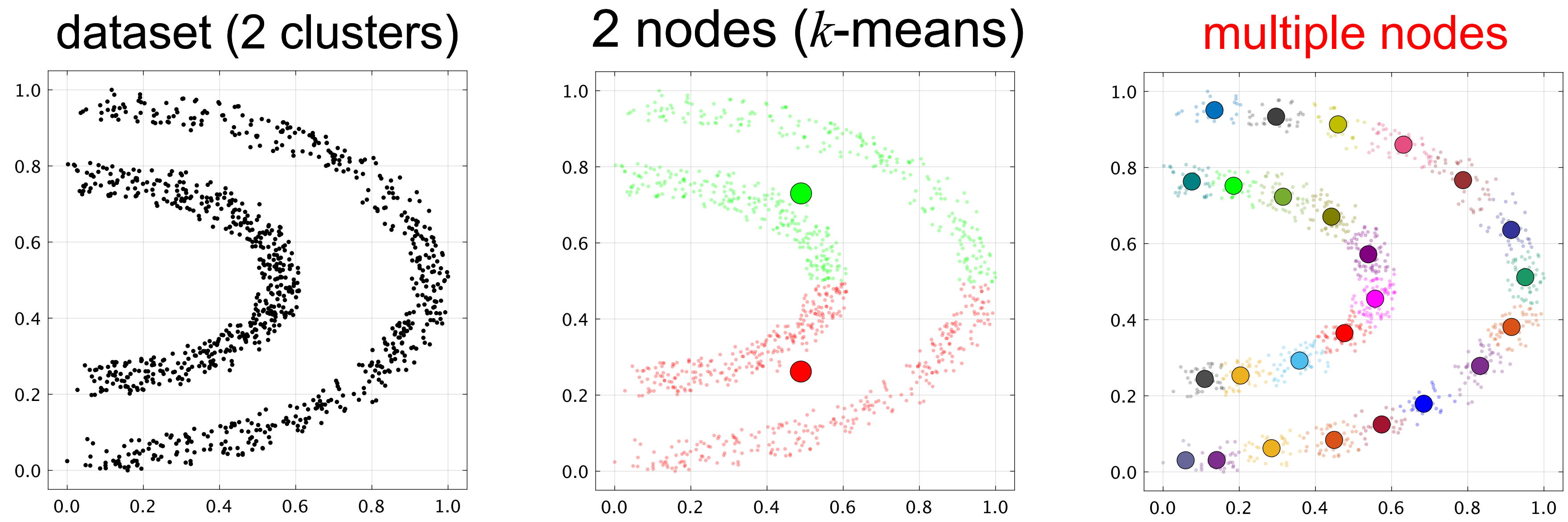
Clustering
Effects of Node Organization

ARI : Adjusted Rand Index

AMI : Adjusted Mutual Information

The higher, the better.

- Depending on the data distribution, the number of nodes should NOT be the same as the number of true clusters.
- The simple solution is to use multiple nodes.
  - ⇨ Using more nodes than true clusters may deteriorate clustering performance.



ARI = 0.00 , AMI = 0.00

ARI = 0.22 , AMI = 0.09

Clustering
Effects of Node Organization

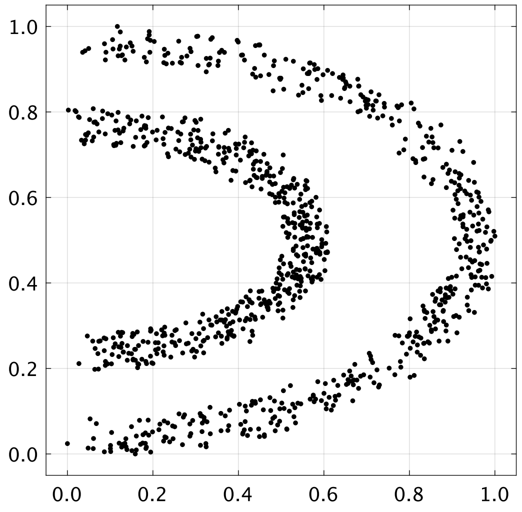
ARI : Adjusted Rand Index

AMI : Adjusted Mutual Information

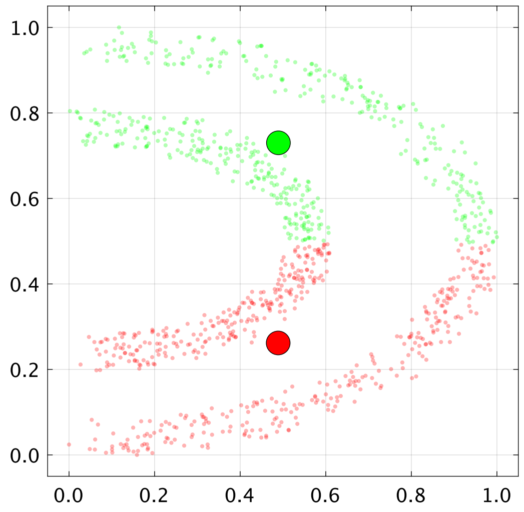
The higher, the better.

- Depending on the data distribution, the number of nodes should NOT be the same as the number of true clusters.
- The simple solution is to use multiple nodes.
- The topology structure clarifies the distribution of data.

dataset (2 clusters)

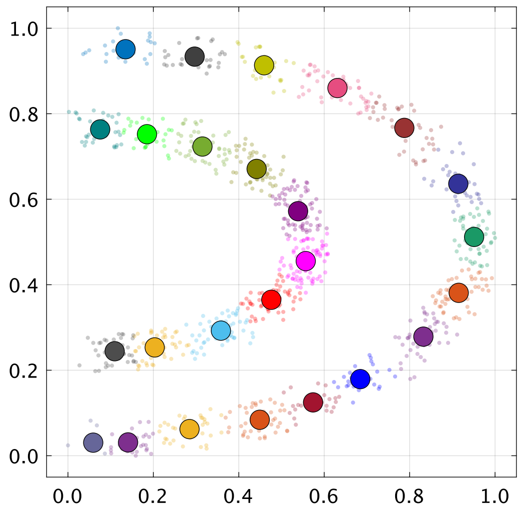


2 nodes (k-means)



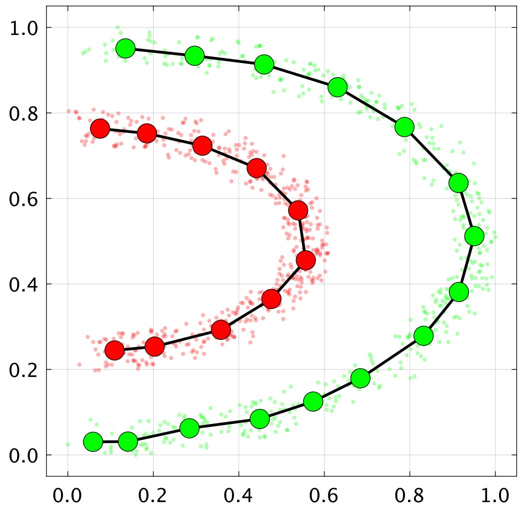
ARI = 0.00 , AMI = 0.00

multiple nodes



ARI = 0.22 , AMI = 0.09

multiple nodes
connected by edges



ARI = 1.00 , AMI = 1.00

Clustering

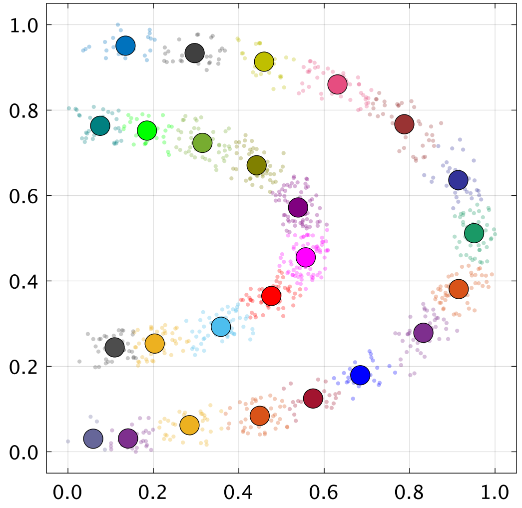
Usefulness of Generating Multiple Nodes

- The simple solution is to use multiple nodes.
  - ⇨ Using more nodes than true clusters may deteriorate clustering performance.

⇩

Useful for Functional Enhancements

multiple nodes



ARI = 0.22, AMI = 0.09

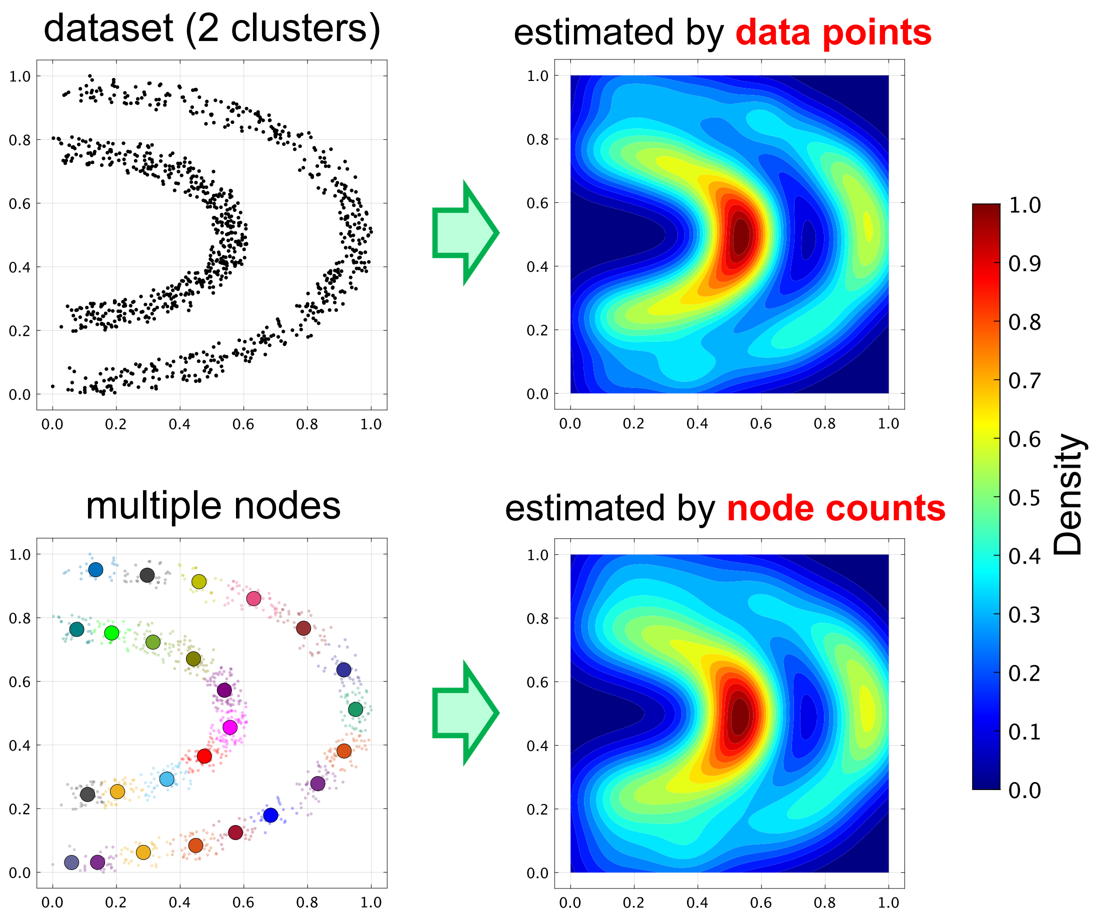
Clustering

Usefulness of Generating Multiple Nodes

- The simple solution is to use multiple nodes.
  - Using more nodes than true clusters may deteriorate clustering performance.

Useful for Functional Enhancements

- Density estimation
- Using nodes as a classifier (if labels exist)
- Instance selection while preserving the features of a distribution

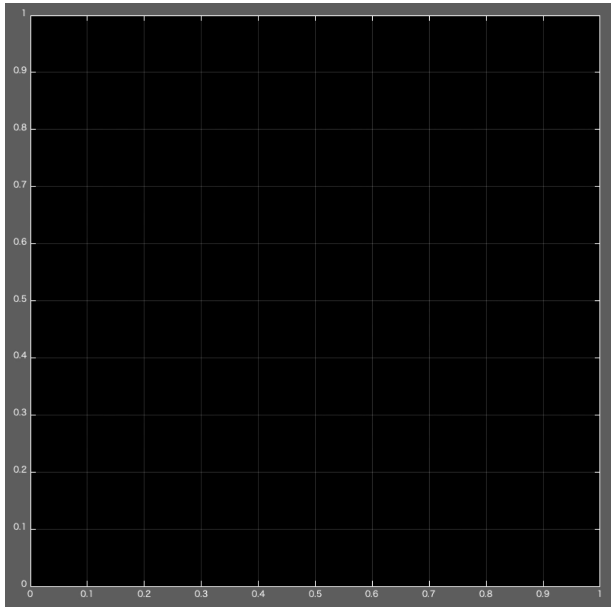


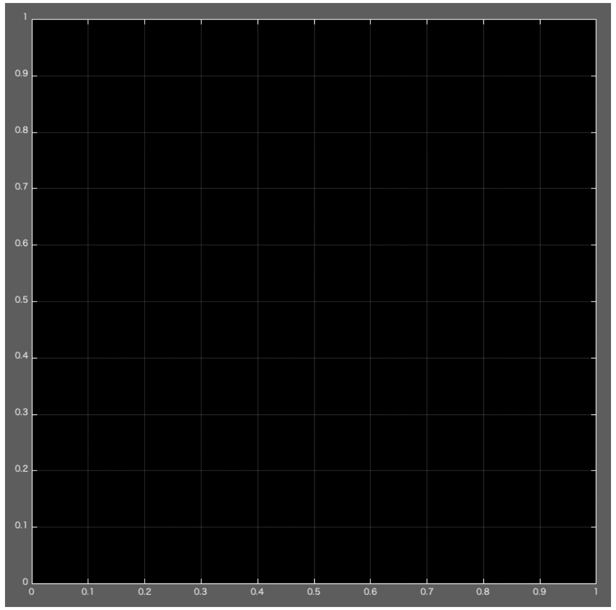
Clustering

Motivation for Current Research

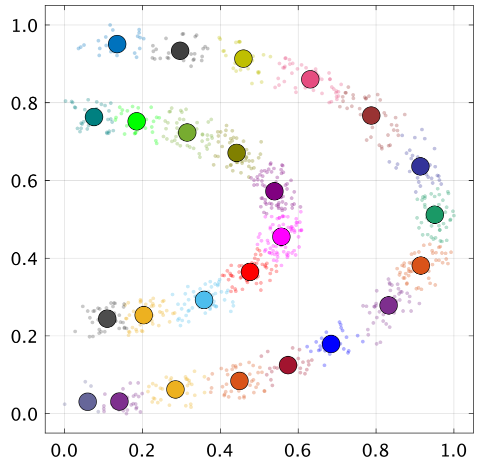
- ✓ How to adaptively adjust network structure in response to evolving data?
- ✓ What mechanisms allow for the continual extraction of meaningful information and relationships between the extracted information?

Growing Self-organizing Clustering



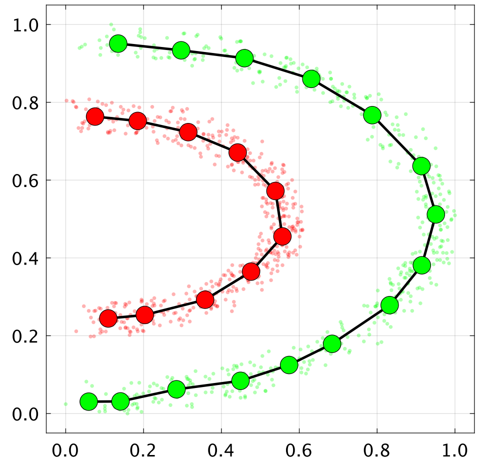


multiple nodes



multiple nodes

connected by edges



Recent Research Topics

Other Extensions by using Clustering

□ Applied Studies

Multi-Label Classifier

N. Masuyama, Y. Nojima, C. K. Loo, and H. Ishibuchi, "Multi-label classification via adaptive resonance theory-based clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 7, pp. 8696-8712, July 2023.

Federated Clustering

N. Masuyama, Y. Nojima, Y. Toda, C. K. Loo, H. Ishibuchi, and N. Kubota, "Privacy-preserving continual federated clustering via adaptive resonance theory,", *arXiv*, arXiv.2309.03487, 2023.

□ Applying Clustering Techniques to Evolutionary Computation

Y. Liu, H. Ishibuchi, N. Masuyama, and Y. Nojima, “Adapting reference vectors and scalarizing functions by growing neural gas to handle irregular Pareto fronts,” *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 3, pp. 439-453, June 2020.

T. Kinoshita, N. Masuyama, Y. Liu, Y. Nojima, and H. Ishibuchi, "Reference vector adaptation and mating selection strategy via adaptive resonance theory-based clustering for many-objective optimization," *IEEE Access*, vol. 11, pp. 126066-126086, November 2023.