

An Introduction to Data Clustering



Dr. Risman Adnan Mattotorang
Telkom University

Outline – Data Clustering

1. Definition of Clustering
2. Features Selection and Extraction
3. Clustering Models and Algorithms
4. Clustering Validation Measures
5. Other Frameworks and Libraries

Clustering Problems

•Non-Formal Definitions:

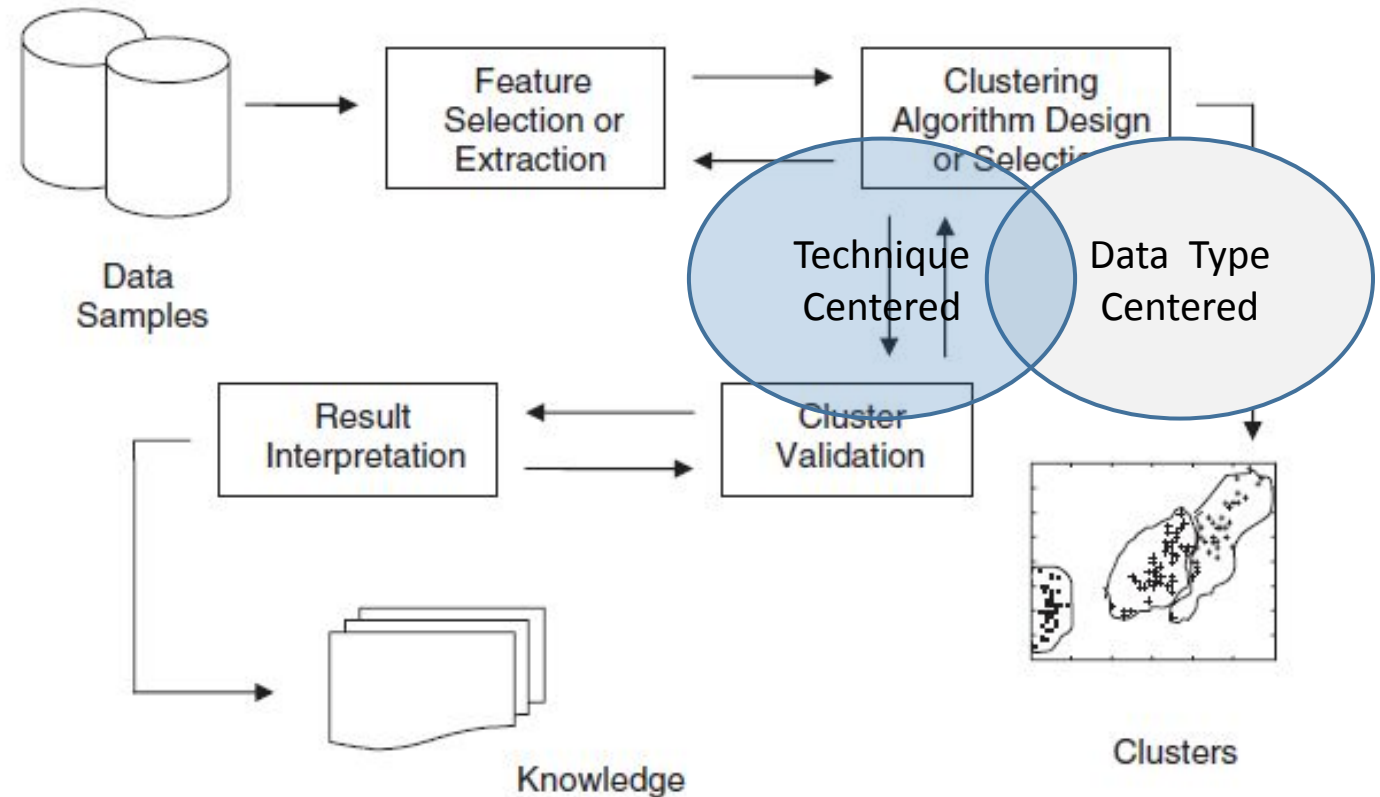
- **Similarity View:** Given a dataset, partition them into a set of groups which are as similar as possible.
- **Distance View:** Given a dataset, find an aggregate of points in the test space such that the distance between any two points in the cluster is less than the distance between any point in the cluster and any point not in it.
- **Probability View:** Given a dataset, find a continuous regions of data space (dimensional feature space) containing a relatively high density of points, separated from other such regions by regions containing a relatively low density of points.

Attention: *Formal definition is not only difficult but may even be misplaced.*

Clustering in EDA Process

We use clustering to understand the data without label:

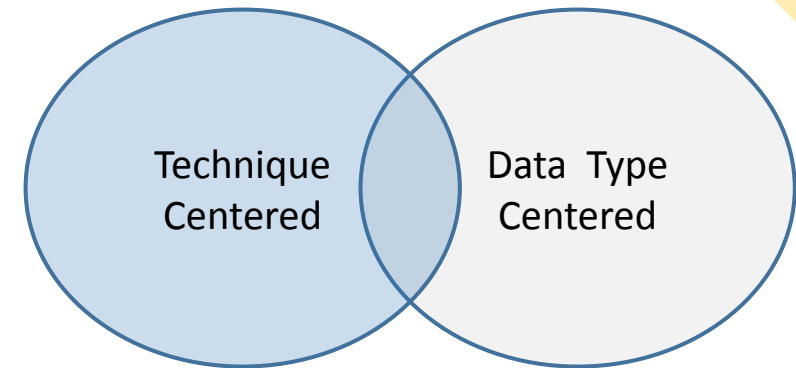
1. Data Samples
2. Feature Selection
3. **Cluster Visualization**
4. **Cluster Validation**
5. **Result Interpretation**
6. Knowledge/Insights
7. ML Model Selection



Application Domain of Clustering

- **Common Use Cases of Clustering:**

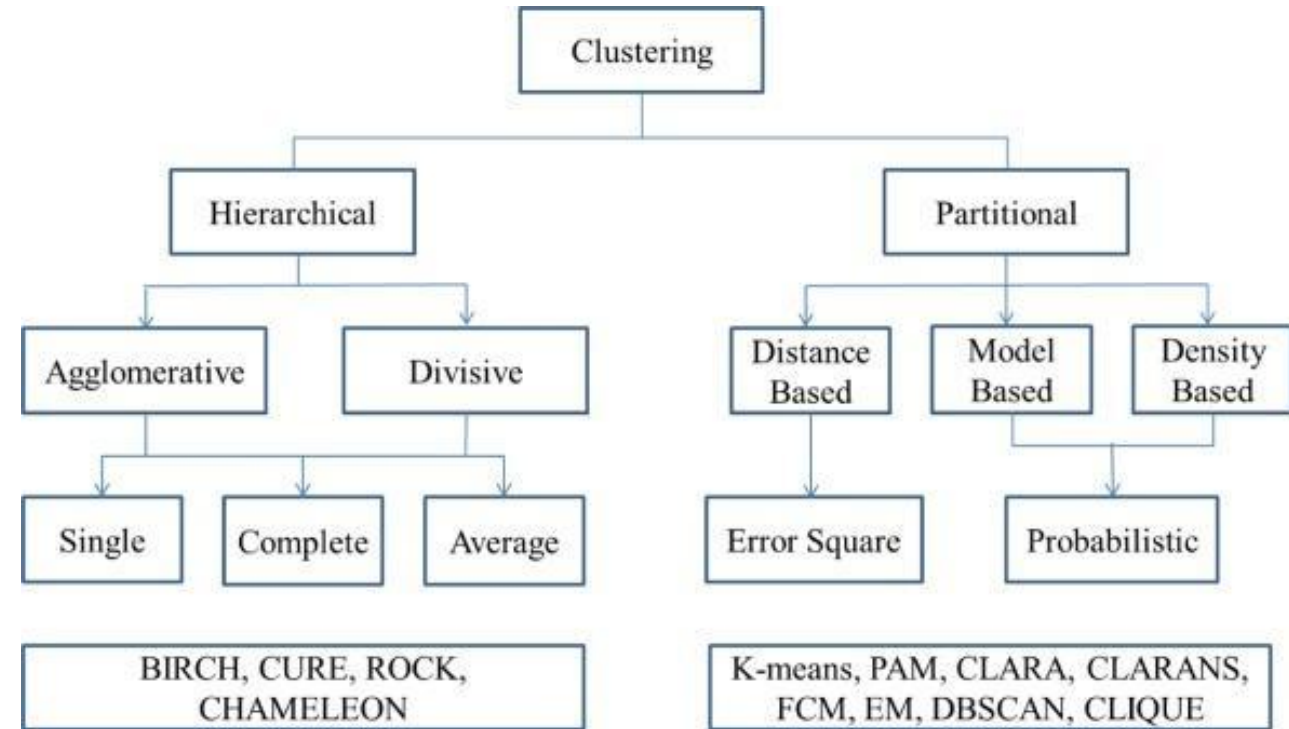
- **Exploratory and Intermediate Phase in ML Modeling**
- Summarization Phase of Collaborative Filtering
- Automatic Customer Segmentation
- Dimensionality Reduction and Data Summarization
- Dynamic Detection for Streaming Data
- Multimedia Data Analysis Eg. Image, Audio, Videos
- Various Use Cases in Biological Data Analysis
- Various Use Cases in Social Network Analysis
- Many more



Categories of Clustering Techniques

Hundreds Methods and Techniques in These Categories:

1. Features Selection (Partitioning) Methods
2. Probabilistic and Generative Models
3. Distance-Based Algorithms (Flat vs Hierarchical)
4. Density and Grid-Based Methods
5. Leveraging Dimensionality Reduction Methods
6. High-Dimensionality Scenarios
7. Scalable Techniques for Clustering Analysis

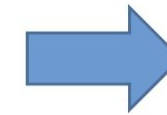
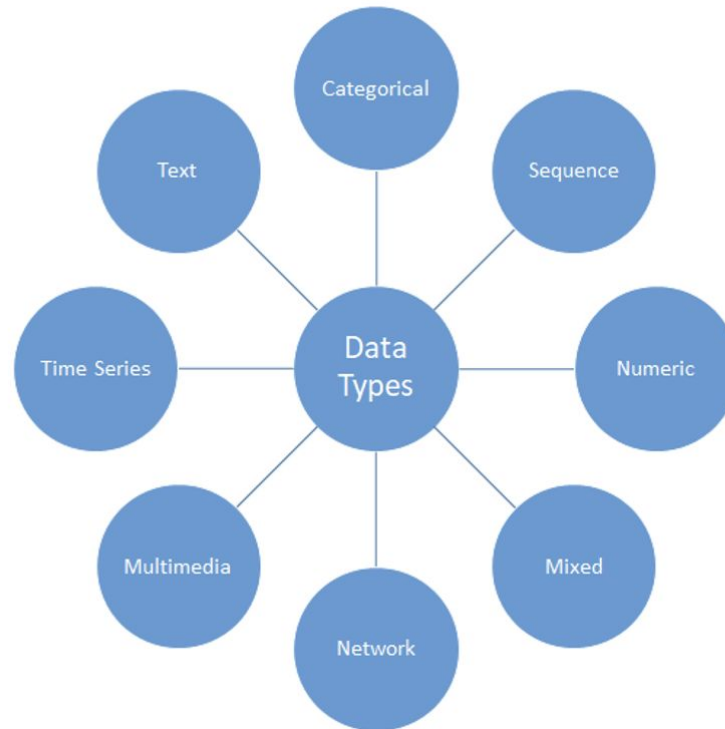


Attention: *The clustering problems can be addressed using a wide variation of methods..*

Categories of Clustering Data Types

Many Methods and Techniques :

1. Categorical
2. Text Data
3. Multimedia
4. Time-Series
5. Discrete Sequence
6. Network Data
7. Uncertain Data

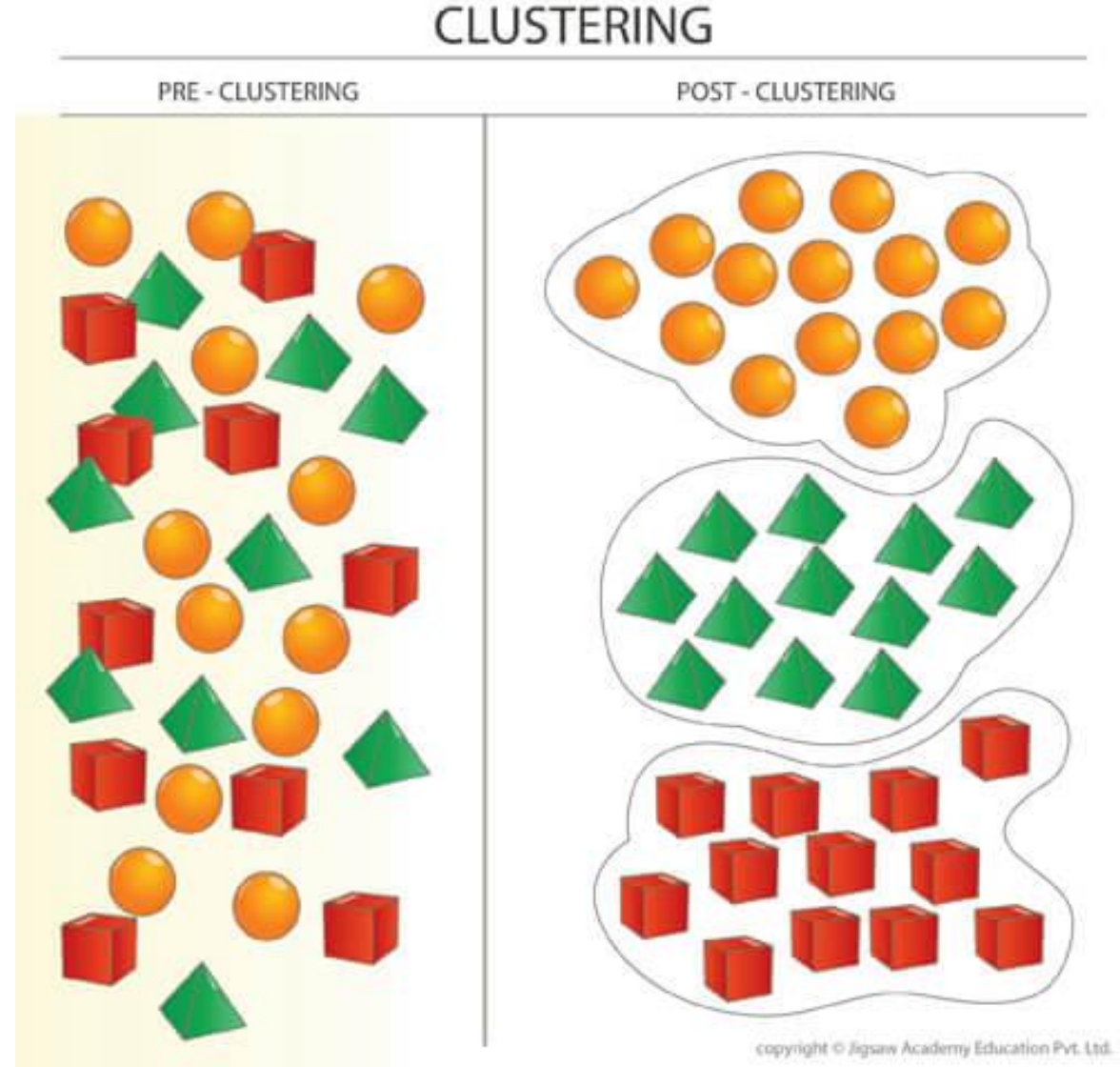


**Clustering Algorithm
Design and Selection**

Interpretation of Clustering

How do we know that a particular clustering is good or that it solves the needs of the application?

- Visual Insights (HD-Eye, IPCLUS)
- Supervised Insights
- Multiview and Ensemble-Based Insights
- Validation-Based Insights



Outline – Data Clustering

1. Definition of Clustering
2. Features Selection and Extraction
3. Clustering Models and Algorithms
4. Clustering Validation Measures
5. Other Frameworks and Libraries

Features Selection for Clustering

Facts:

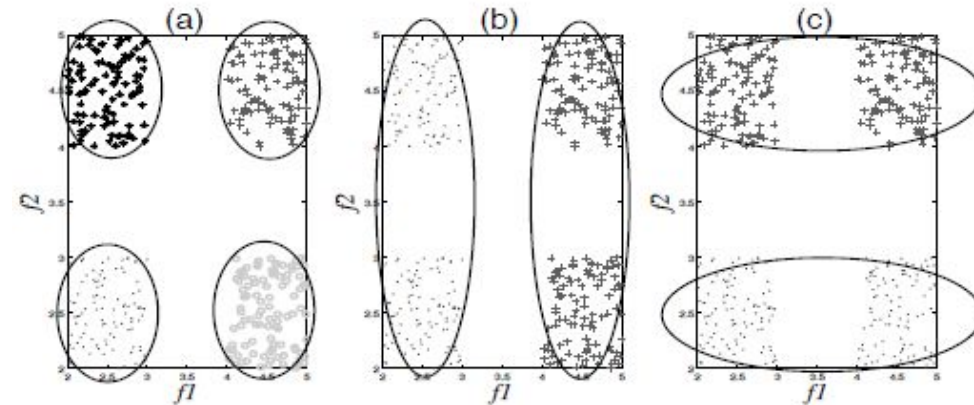
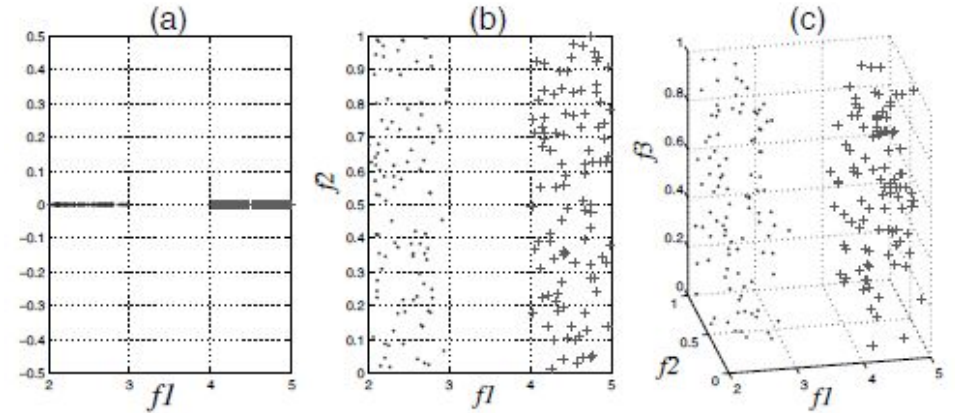
- Removing features don't affect clustering accuracy
- Different features produce different clusters

Features Selection for Clustering:

- Filter
- Wrapper
- Hybrid

Data Types & Algorithms:

- Generic, Text, Streaming, Linked data



Filter Model

Characteristics:

- Do not utilize any clustering algorithms.
- Evaluate score of each feature according to certain criteria.
- Select features with highest scores.
- Features evaluation could be either univariate and multivariate.
- Univariate = feature is evaluated independently of the feature space.
- Multivariate can handle redundant features.

Example of Methods:

- Spectral Feature Selection (SPEC)
- Feature Dependency
- Entropy-Based Distance
- Laplacian Score

D	Dataset
n	Sample size
m	Number of features
x_j	j^{th} sample
f_i	i^{th} feature
F	Selected feature set
l	Number of selected features
K	Number of clusters
C_k	k^{th} cluster

Wrapper Model

Characteristics:

- Utilize clustering algorithms to evaluate features.
- Heuristic search strategy to reduce the search space.
- **How it works:** Finding subset of features, evaluate, repeat.
- Computationally expensive compared to filter model.
- Produce better clustering but bias toward clustering methods.

Example of Methods:

- MLE-based search criteria.
- K-means search.
- Mixture of Gaussian search strategy.

Hybrid Model

Characteristics:

- Overcome drawbacks of wrapper and filter methods.
- Take efficiency from filtering and quality from wrapper methods.
- **How it works:**
 - Utilizes filtering criteria to select different candidate subsets.
 - evaluates the quality of clustering of each candidate subsets
 - Select the subset with the highest clustering quality is selected.

Example of Methods:

- All methods from filtering and wrapper

Common Feature Selection Algorithms

- **Algorithms for Generic Data**

- Spectral Feature Selection (SPEC)
- Laplacian Score (LS)
- Feature Selection for Sparse Clustering
- Localized Feature Selection Based on Scatter Separability (LFSBSS)
- Multicluster Feature Selection (MCFS)
- Feature Weighting k -Means

- **Algorithms for Streaming Data**

- Text Stream Clustering Based on Adaptive Feature Selection (TSC-AFS)

- **Algorithms for Text Data**

- Term Frequency (TF)
- Inverse Document Frequency (IDF)
- Term Frequency-Inverse Document Frequency (TF-IDF)
- Chi Square Statistic
- Frequent Term-Based Text Clustering
- Frequent Term Sequence

- **Algorithms for Linked Data**

- LUFFS: An Unsupervised Feature Selection Framework for Linked Data

Outline – Data Clustering

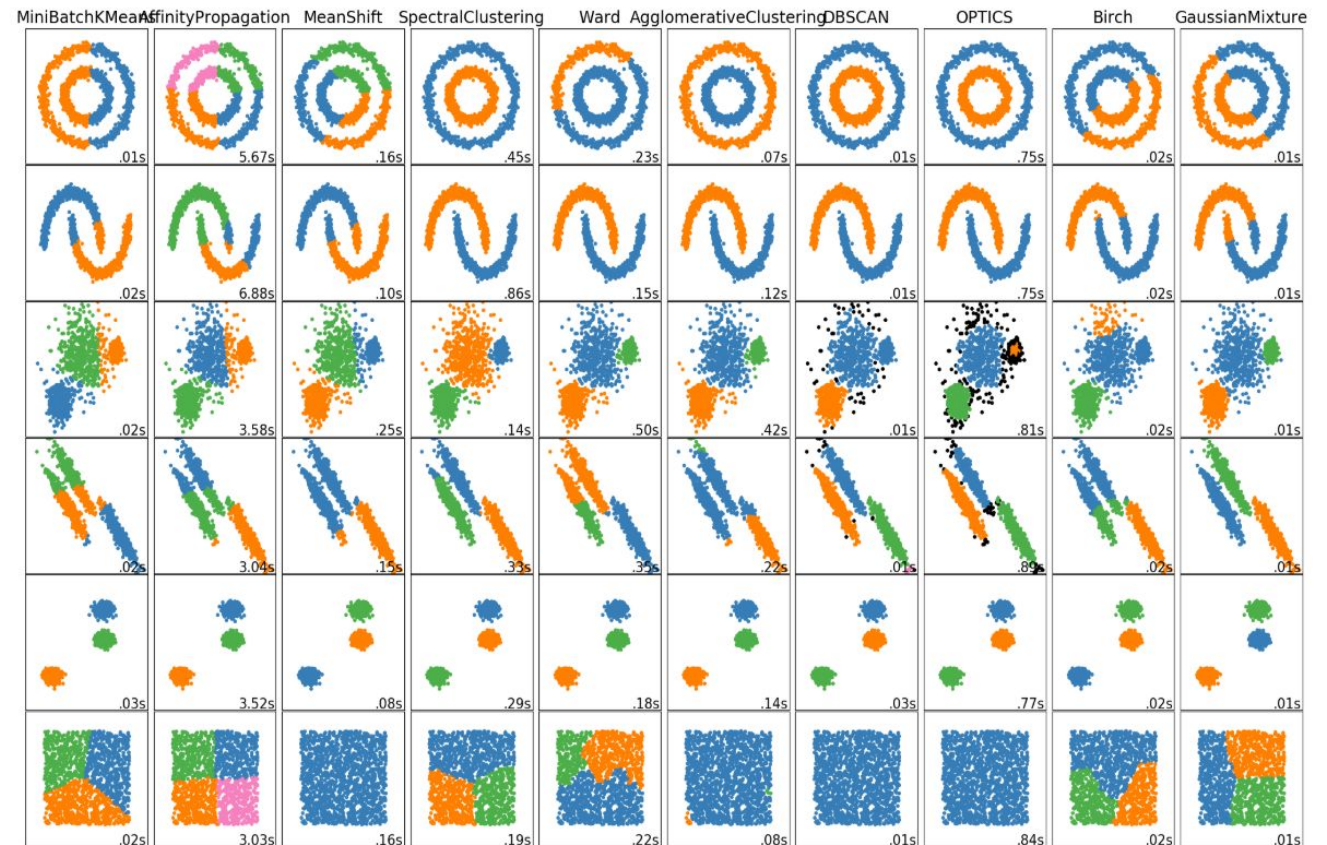
1. Definition of Clustering
2. Features Selection and Extraction
3. Clustering Models and Algorithms
4. Clustering Validation Measures
5. Other Frameworks and Libraries

Various Models, Methods and Algorithms

Common Approaches:

- Partitioning Methods
- Probabilistic Models
- Generative Models
- Distance-Based Algorithms
- Density-Based Methods
- Grid-Based Methods

Need more Ph.Ds to understand it all.... don't be a wikipedia.



Model Comparison

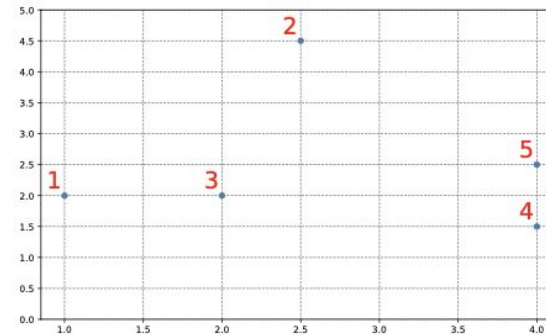
Method name	Parameters	Scalability	Usecase	Geometry (metric used)
K-Means	number of clusters	Very large <code>n_samples</code> , medium <code>n_clusters</code> with MiniBatch code	General-purpose, even cluster size, flat geometry, not too many clusters	Distances between points
Affinity propagation	damping, sample preference	Not scalable with <code>n_samples</code>	Many clusters, uneven cluster size, non-flat geometry	Graph distance (e.g. nearest-neighbor graph)
Mean-shift	bandwidth	Not scalable with <code>n_samples</code>	Many clusters, uneven cluster size, non-flat geometry	Distances between points
Spectral clustering	number of clusters	Medium <code>n_samples</code> , small <code>n_clusters</code>	Few clusters, even cluster size, non-flat geometry	Graph distance (e.g. nearest-neighbor graph)
Ward hierarchical clustering	number of clusters or distance threshold	Large <code>n_samples</code> and <code>n_clusters</code>	Many clusters, possibly connectivity constraints	Distances between points
Agglomerative clustering	number of clusters or distance threshold, linkage type, distance	Large <code>n_samples</code> and <code>n_clusters</code>	Many clusters, possibly connectivity constraints, non Euclidean distances	Any pairwise distance
DBSCAN	neighborhood size	Very large <code>n_samples</code> , medium <code>n_clusters</code>	Non-flat geometry, uneven cluster sizes	Distances between nearest points
OPTICS	minimum cluster membership	Very large <code>n_samples</code> , large <code>n_clusters</code>	Non-flat geometry, uneven cluster sizes, variable cluster density	Distances between points
Gaussian mixtures	many	Not scalable	Flat geometry, good for density estimation	Mahalanobis distances to centers
Birch	branching factor, threshold, optional global clusterer.	Large <code>n_clusters</code> and <code>n_samples</code>	Large dataset, outlier removal, data reduction.	Euclidean distance between points

Hierarchical Agglomerative Clustering

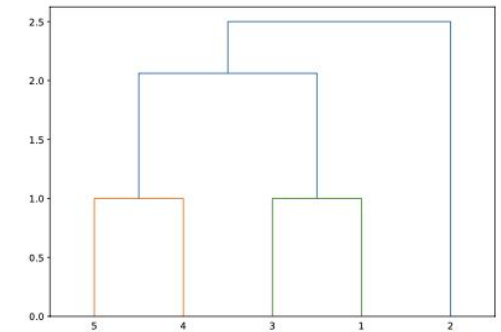
- The input to the algorithm is an $N \times N$ dissimilarity matrix $D_{ij} \geq 0$, and the output is a tree structure in which groups i and j with small dissimilarity are grouped together in a hierarchical fashion. For example **city block distance** between the points to define the dissimilarity, i.e.,

$$d_{ij} = \sum_{k=1}^2 |x_{ik} - x_{jk}|$$

- Start with a tree with N leaves, each corresponding to a cluster with a single data point. Next we compute the pair of points that are closest, and merge them.



(a)



(b)

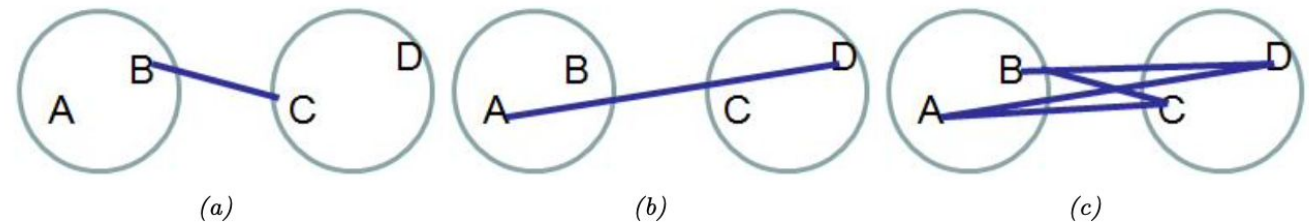


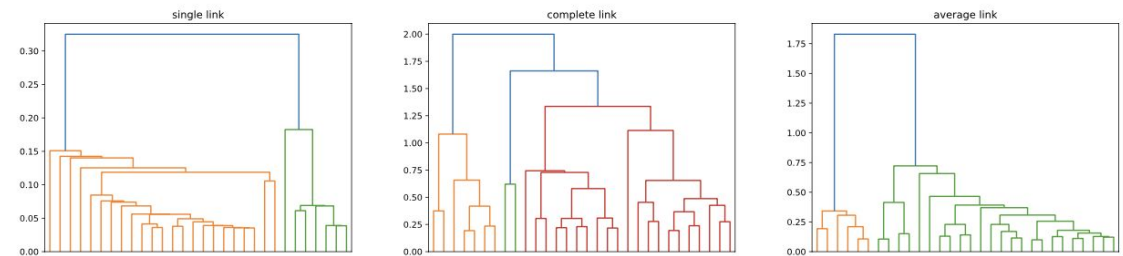
Figure 21.3: Illustration of (a) Single linkage. (b) Complete linkage. (c) Average linkage.

HAC Implementation

- Agglomerative clustering starts with N groups, each initially containing one object, and then at each step it merges the two most similar groups until there is a single group, containing all the data.
- Since picking the two most similar clusters to merge takes $O(N^2)$ time, and there are $O(N)$ steps in the algorithm, the total running time is $O(N^3)$. However, by using a priority queue, this can be reduced to $O(N^2 \log N)$.
- Three types of HAC Implementation:
 - Single Link (Nearest Neighbor) Clustering
 - Complete Link (Further Neighbor) Clustering
 - Average Link Clustering

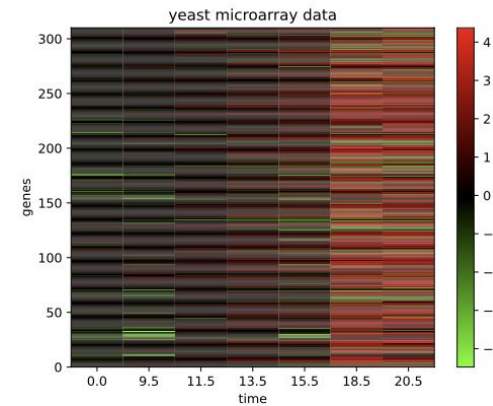
Algorithm 11: Agglomerative clustering

```
1 Initialize clusters as singletons: for  $i \leftarrow 1$  to  $n$  do  $C_i \leftarrow \{i\}$ ;  
2 ;  
3 Initialize set of clusters available for merging:  $S \leftarrow \{1, \dots, n\}$ ; repeat  
4   Pick 2 most similar clusters to merge:  $(j, k) \leftarrow \arg \min_{j, k \in S} d_{j, k}$ ;  
5   Create new cluster  $C_\ell \leftarrow C_j \cup C_k$ ;  
6   Mark  $j$  and  $k$  as unavailable:  $S \leftarrow S \setminus \{j, k\}$ ;  
7   if  $C_\ell \neq \{1, \dots, n\}$  then  
8     Mark  $\ell$  as available,  $S \leftarrow S \cup \{\ell\}$ ;  
9   foreach  $i \in S$  do  
10    Update dissimilarity matrix  $d(i, \ell)$ ;  
11 until no more clusters are available for merging;
```

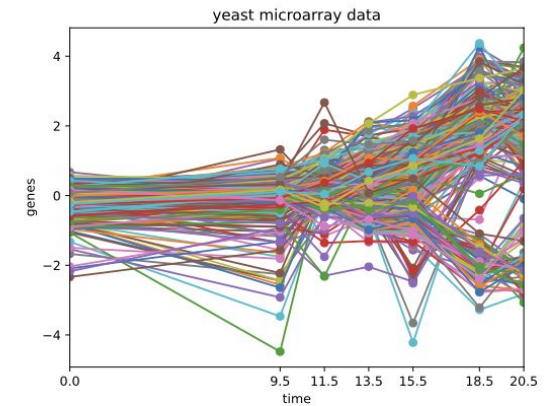


Example of HAC

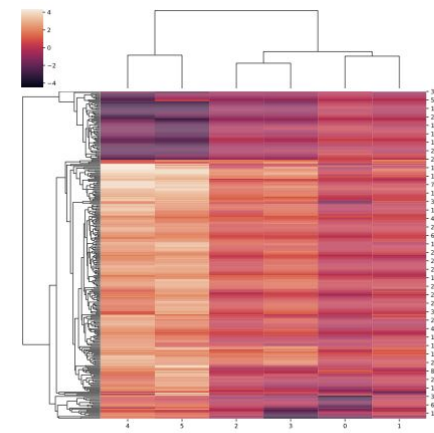
- Suppose we have a set of time series measurements of the expression levels for $N = 300$ genes at $T = 7$ points. We see that there are several kinds of genes, such as those whose expression level goes up monotonically over time (in response to a given stimulus), those whose expression level goes down monotonically, and those with more complex response patterns.
- Suppose we use Euclidean distance to compute a pairwise dissimilarity matrix, and apply HAC using average linkage. Hierarchical clustering applied to the yeast gene expression data. (a) The rows are permuted according to a hierarchical clustering scheme (average link agglomerative clustering), in order to bring similar rows close together. (b) 16 clusters induced by cutting the average linkage tree at a certain height.



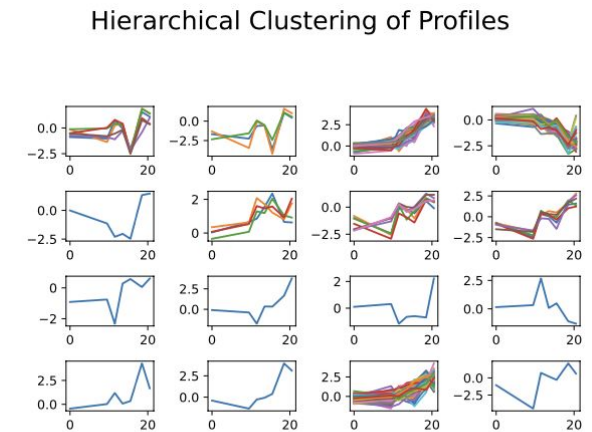
(a)



(b)



(a)



(b)

K-Means Clustering

- **Problems with HAC:**

- It takes $O(N^3)$ time (for the average link method), making it hard to apply to big datasets.
 - It assumes that a dissimilarity matrix has already been computed, whereas the notion of “similarity” is often unclear and needs to be learned.
 - It is just an algorithm, not a model, and so it is hard to evaluate how good it is. That is, there is no clear objective that it is optimizing.
- K-Means algorithm addresses these issues. In particular, it runs in $O(N)$ time, it computes similarity in terms of Euclidean distance to learned cluster centers $\mu_k \in \mathbb{R}^D$, and it optimizes a well-defined cost function.

Kmeans Algorithm

We assume there are K cluster centers $\mu_k \in \mathbb{R}^D$, so we can cluster the data by assigning each data point $\mathbf{x}_n \in \mathbb{R}^D$ to its closest center:

$$z_n^* = \arg \min_k \|\mathbf{x}_n - \mu_k\|_2^2$$

Of course, we don't know the cluster centers, but we can estimate them by computing the average value of all points assigned to them:

$$\mu_k = \frac{1}{N_k} \sum_{n: z_n=k} \mathbf{x}_n$$

We can then iterate these steps to convergence.

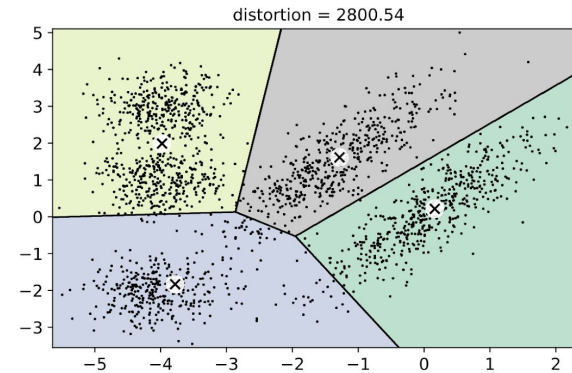
More formally, we can view this as finding a local minimum of the following cost function, known as the **distortion**:

$$J(\mathbf{M}, \mathbf{Z}) = \sum_{n=1}^N \|\mathbf{x}_n - \mu_{z_n}\|^2 = \|\mathbf{X} - \mathbf{Z}\mathbf{M}^T\|_F^2$$

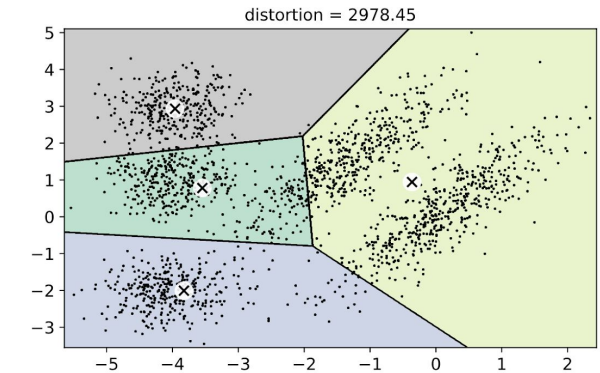
where $\mathbf{X} \in \mathbb{R}^{N \times D}$, $\mathbf{Z} \in [0, 1]^{N \times K}$, and $\mathbf{M} \in \mathbb{R}^{D \times K}$ contains the cluster centers μ_k in its columns. K-means optimizes this using alternating minimization.

Example of K-Means

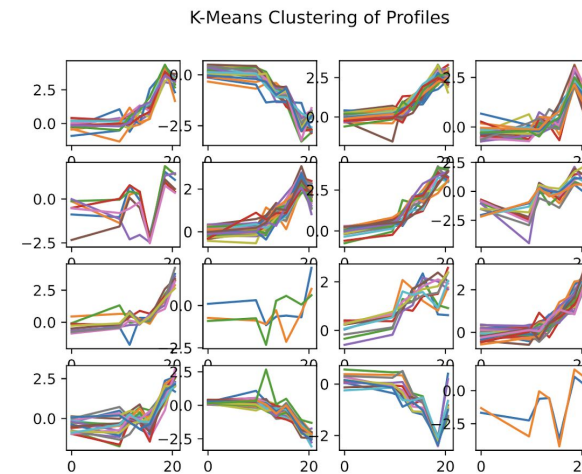
- Illustration of K-means clustering in 2d. We show the result of using two different random seeds.
- Clustering the yeast data using K-means clustering with $K = 16$. (a) Visualizing all the time series assigned to each cluster. (b) Visualizing the 16 cluster centers as prototypical time series.



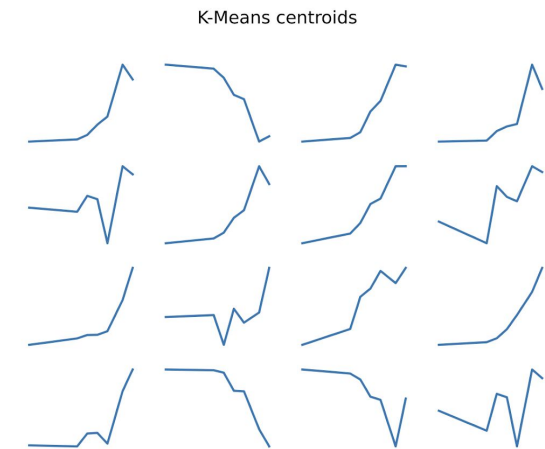
(a)



(b)



(a)



(b)

Introduction to EduClust

<https://educlust.dbvis.de>

EduClust

A Visual Education Platform
for Teaching Clustering Algorithms

Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, Matthias Miller, Daniel Keim

k-means

Three Not Equal Circles

Custom Data

Parameters

k:

Distance
Measure:

Euclidian

$$d_2(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

Navigation

Iteration #

Step #

Current Animation Speed



Current Step

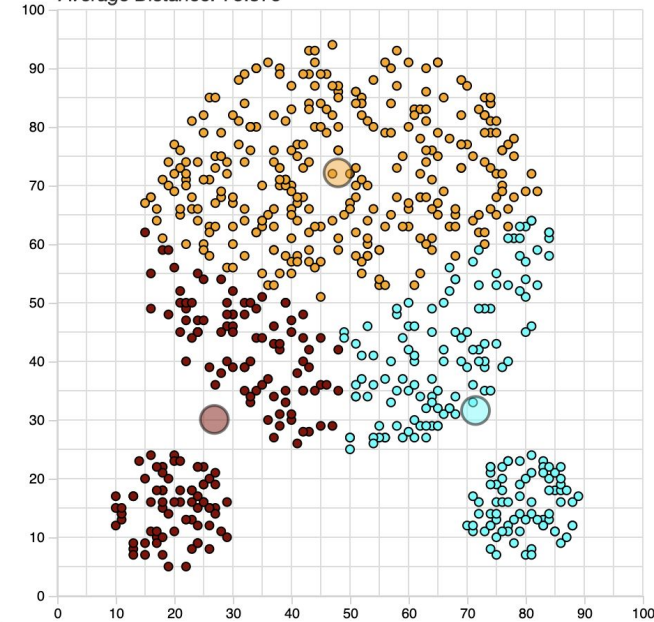
You have feedback?

Imprint

Data Privacy

Visualization

Average Distance: 75.375



Pseudocode: k-means

Complexity Range: $O(k \times n \times t)$

Input: k clusters

Output: k clusters

Pseudocode:

1. Choose k objects as initial **cluster centers**.
2. Assign each data point to the cluster which has the closest **mean point (centroid)** under chosen distance metric.
3. When all data points have been assigned, recalculate the positions of k **centroids (mean points)**.
4. Repeat steps 2 and 3 until the **centroids do not change any more**. All data points remain in their most recently assigned cluster.

Best Clustering Found

Universität
Konstanz



Outline – Data Clustering

1. Definition of Clustering
2. Features Selection and Extraction
3. Clustering Models and Algorithms
4. Clustering Validation Measures
5. Other Frameworks and Libraries

Clustering Validation Measures

- **Clustering Validation:** measure of goodness of clustering results.
- **Facts:** No consistent and conclusive solution to cluster validation. The best suitable measures to use in practice remain unknown.
- **Major Issues:**
 - Importance of normalizing validation measures.
 - Relationship between different validation measures.
 - Selection of validation measures.
 - Validation for specific algorithms.
- **Category of techniques:**
 - External: Use external information not presenting in the data
 - Internal: Use only internal information in the data
 - Both are crucial to apply and both have limitations

External Validation Measures

- **The Contingency Matrix.** Given a data set D with n objects, assume that there is a partition of D , and K is the number of clusters. If the “true” class labels for the data are given, another partition can be generated on D , called C , then the information on the overlap between the two partitions can be written in the form of a contingency matrix.

	Measure	Definition	Range
1	Entropy (E)	$-\sum_i p_i (\sum_j \frac{p_{ij}}{p_i} \log \frac{p_{ij}}{p_i})$	$[0, \log K']$
2	Purity (P)	$\sum_i p_i (\max_j \frac{p_{ij}}{p_i})$	$(0, 1]$
3	F-measure (F)	$\sum_j p_j \max_i [2 \frac{p_{ij}}{p_i} \frac{p_{ij}}{p_j} / (\frac{p_{ij}}{p_i} + \frac{p_{ij}}{p_j})]$	$(0, 1]$
4	Variation of Information (VI)	$-\sum_i p_i \log p_i - \sum_j p_j \log p_j - 2 \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i p_j}$	$[0, 2 \log \max(K, K')]$
5	Mutual Information (MI)	$\sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i p_j}$	$(0, \log K']$
6	Rand statistic (R)	$[(\binom{n}{2} - \sum_i \binom{n_i}{2}) - \sum_j \binom{n_j}{2} + 2 \sum_{ij} \binom{n_{ij}}{2}] / \binom{n}{2}$	$(0, 1]$
7	Jaccard coefficient (J)	$\sum_{ij} \binom{n_{ij}}{2} / [\sum_i \binom{n_i}{2} + \sum_j \binom{n_j}{2} - \sum_{ij} \binom{n_{ij}}{2}]$	$[0, 1]$
8	Fowlkes & Mallows index (FM)	$\sum_{ij} \binom{n_{ij}}{2} / \sqrt{\sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2}}$	$[0, 1]$
9	Hubert Γ statistic I (Γ)	$\frac{(\sum_i \sum_j \binom{n_{ij}}{2}) - \sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2}}{\sqrt{\sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2}}}$	$(-1, 1]$
10	Hubert Γ statistic II (Γ')	$[(\binom{n}{2} - 2 \sum_i \binom{n_i}{2}) - 2 \sum_j \binom{n_j}{2} + 4 \sum_{ij} \binom{n_{ij}}{2}] / \binom{n}{2}$	$[0, 1]$
11	Minkowski score (MS)	$\sqrt{\sum_i \binom{n_i}{2} + \sum_j \binom{n_j}{2} - 2 \sum_{ij} \binom{n_{ij}}{2}} / \sqrt{\sum_j \binom{n_j}{2}}$	$[0, +\infty)$
12	classification error (ϵ)	$1 - \frac{1}{n} \max_{\sigma} \sum_j n_{\sigma(j), j}$	$[0, 1]$
13	van Dongen criterion (VD)	$(2n - \sum_i \max_j n_{ij} - \sum_j \max_i n_{ij}) / 2n$	$[0, 1]$
14	micro-average precision (MAP)	$\sum_i p_i (\max_j \frac{p_{ij}}{p_i})$	$(0, 1]$
15	Goodman-Kruskal coeff (GK)	$\sum_i p_i (1 - \max_j \frac{p_{ij}}{p_i})$	$[0, 1]$
16	Mirkin metric (M)	$\sum_i n_i^2 + \sum_j n_j^2 - 2 \sum_i \sum_j n_{ij}^2$	$[0, 2 \binom{n}{2})$

Note: $p_{ij} = n_{ij}/n$, $p_i = n_i/n$, $p_j = n_j/n$.

		Partition C				
		C_1	C_2	\cdots	$C_{K'}$	Σ
Partition P	P_1	n_{11}	n_{12}	\cdots	$n_{1K'}$	$n_{1\cdot}$
	P_2	n_{21}	n_{22}	\cdots	$n_{2K'}$	$n_{2\cdot}$
	\vdots	\vdots	\vdots	\cdots	\vdots	\vdots
	P_K	n_{K1}	n_{K2}	\cdots	$n_{KK'}$	$n_{K\cdot}$
	Σ	$n_{\cdot 1}$	$n_{\cdot 2}$	\cdots	$n_{\cdot K'}$	n

Internal Validation Measures

- As the goal of clustering is to make objects within the same cluster similar and objects in different clusters distinct, internal validation measures are often based on two criteria: **Compactness** and **Separation**.

- Understanding the impacts of:**

- Monotonicity
- Noise
- Density
- Subclusters
- Skewed Distributions
- Arbitrary Shapes

	Measure	Definition
1	$RMSSTD^1$	$\{\sum_i \sum_{x \in C_i} \ x - c_i\ ^2 / [P \sum_i (n_i - 1)]\}^{\frac{1}{2}}$
2	R-squared (RS)	$(\sum_{x \in D} \ x - c\ ^2 - \sum_i \sum_{x \in C_i} \ x - c_i\ ^2) / \sum_{x \in D} \ x - c\ ^2$
3	Modified Hubert Γ statistic (Γ)	$\frac{2}{n(n-1)} \sum_{x \in D} \sum_{y \in D} d(x, y) d_{x \in C_i, y \in C_j}(c_i, c_j)$
4	Calinski-Harabasz index (CH)	$\frac{\sum_i n_i d^2(c_i, c) / (NC - 1)}{\sum_i \sum_{x \in C_i} d^2(x, c_i) / (n - NC)}$
5	I index (I)	$(\frac{1}{NC} \cdot \frac{\sum_{x \in D} d(x, c)}{\sum_i \sum_{x \in C_i} d(x, c_i)} \cdot \max_{i,j} d(c_i, c_j))^p$
6	Dunn's indices (D)	$\min_i \{ \min_j (\frac{\min_{x \in C_i, y \in C_j} d(x, y)}{\max_k \{ \max_{x, y \in C_k} d(x, y) \}}) \}$
7	Silhouette index (S)	$\frac{1}{NC} \sum_i \{ \frac{1}{n_i} \sum_{x \in C_i} \frac{b(x) - a(x)}{\max\{b(x), a(x)\}} \}$ $a(x) = \frac{1}{n_i - 1} \sum_{y \in C_i, y \neq x} d(x, y), b(x) = \min_{j, j \neq i} [\frac{1}{n_j} \sum_{y \in C_j} d(x, y)]$
8	Davies-Bouldin index (DB)	$\frac{1}{NC} \sum_i \max_{j, j \neq i} \{ [\frac{1}{n_i} \sum_{x \in C_i} d(x, c_i) + \frac{1}{n_j} \sum_{x \in C_j} d(x, c_j)] / d(c_i, c_j) \}$
9	Xie-Beni index (XB)	$[\sum_i \sum_{x \in C_i} d^2(x, c_i)] / [n \cdot \min_{i, j \neq i} d^2(c_i, c_j)]$
10	SD validity index (SD)	$Dis(NC_{max}) Scat(NC) + Dis(NC)$ $Scat(NC) = \frac{1}{NC} \sum_i \ \sigma(C_i) \ / \ \sigma(D) \ $ $Dis(NC) = \frac{\max_{i,j} d(c_i, c_j)}{\min_{i,j} d(c_i, c_j)} \sum_i (\sum_j d(c_i, c_j))^{-1}$
11	S-Dbw validity index (S_Dbw)	$Scat(NC) + Dens_bw(NC)$ $Dens_bw(NC) = \frac{1}{NC(NC-1)} \sum_i [\sum_{j, j \neq i} \frac{\sum_{x \in C_i \cup C_j} f(x, u_{ij})}{\max\{ \sum_{x \in C_i} f(x, c_i), \sum_{x \in C_j} f(x, c_j) \}}]$
12	$CVNN^2$ index	$Sep(NC, k) / \max_{NC} Sep(NC, k) + Com(NC) / \max_{NC} Com(NC)$ $Com(NC) = \sum_i [\frac{2}{n_i \cdot (n_i - 1)} \sum_{x, y \in C_i} d(x, y)]$ $Sep(NC, k) = \max_i (\frac{1}{n_i} \sum_{j=1, 2, \dots, n_i} \frac{q_j}{k})$

Outline – Data Clustering

1. Definition of Clustering
2. Features Selection and Extraction
3. Clustering Models and Algorithms
4. Clustering Validation Measures
5. Other Frameworks and Libraries

OSS Frameworks and Libraries

General Purposes

- WEKA
- Spider
- Cluster
- ELKI

Specialized

- OpenSubspace (Subspace Algorithm)
- cross-bow (Text Data Clustering)
- MALLET (Topic Modeling NLP)
- CLUTO (High Dimensional Data)
- Gait-CAD (Time-Series)
- Cluster (Gene Expression)
- METIS (Graph Partitioning)
- CFinder (Nodes in Network)

■ Public Datasets

- KDNuggets
- STATOO
- David Dowe Links
- KDD Cup
- UCR Time-Series
- SNAP repository
- Reuters News
- 20NewsGroups
- TREC
- Cora
- Gene Expression Omnibus

■ Public Datasets

- DIP
- BioGRID
- STRING
- MIPS
- GenBank
- EMBL
- PIR
- UniProt
- ImageCLEF
- ImageNet
- Autonomous Car

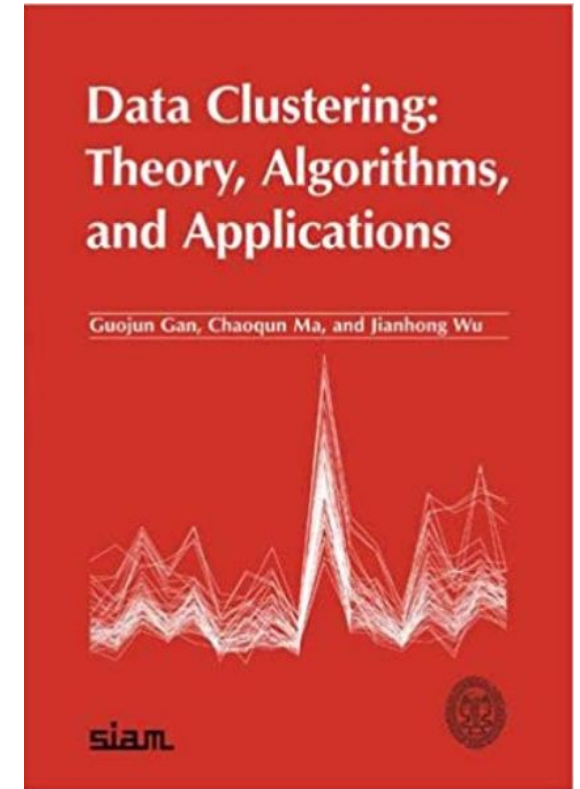
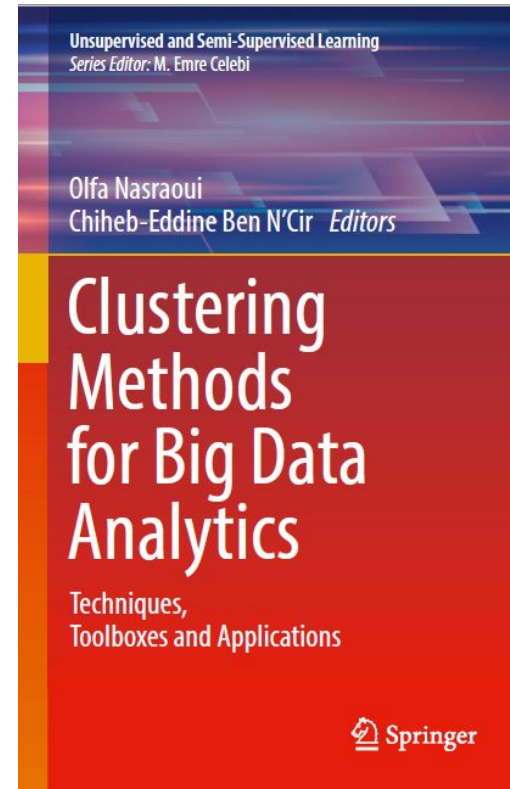
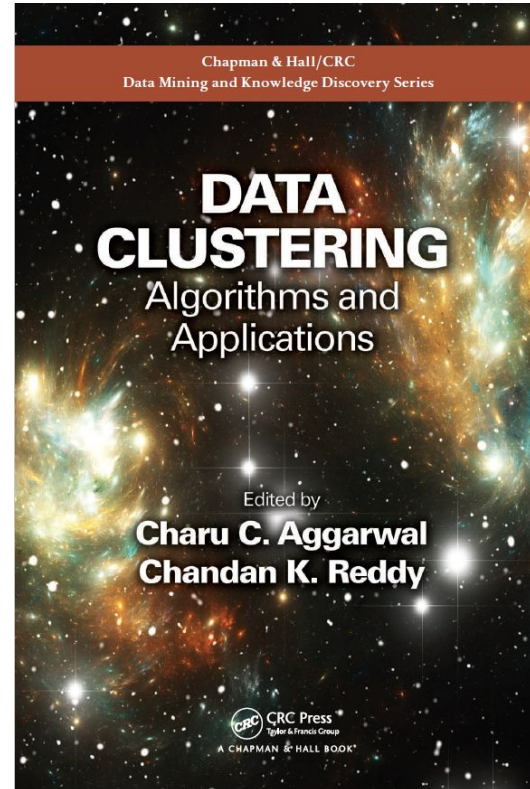
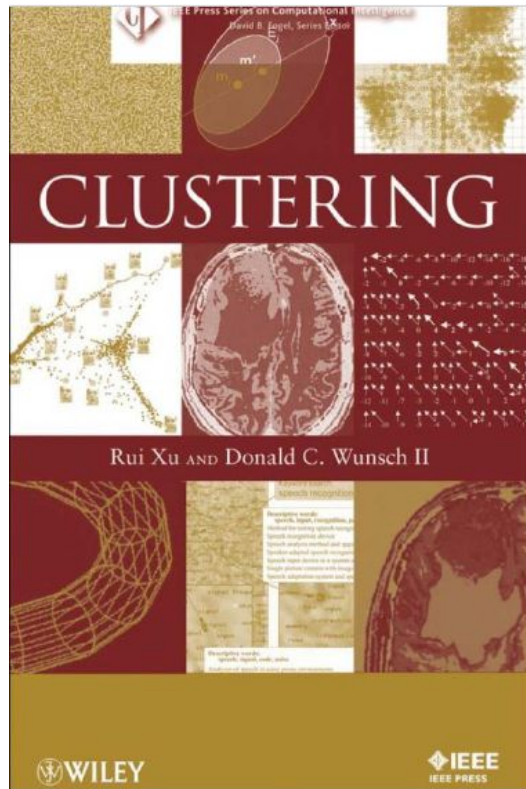
A pair of black-rimmed glasses with round lenses is resting on an open notebook. The notebook has a red ribbon bookmark. The background is blurred, showing more books and a wooden surface.

Homework

Please Reproduce the following link:

<https://scikit-learn.org/stable/modules/clustering.html>

To Learn More(when you need it)



A top-down view of several Christmas gifts on a dark wooden surface. The gifts are wrapped in various patterns: red with white hearts, red with white snowflakes, brown with red plaid, and brown with red geometric patterns. Most are tied with red ribbons, while one is tied with a green and white checkered ribbon and another with a red and white striped string. A pair of hands is visible in the bottom left, adjusting a red ribbon on a gift wrapped in red and white snowflake paper.

THANKS

Dr. Risman Adnan Mattotorang