SGN-13000/SGN-13006 Introduction to Pattern Recognition and Machine Learning (5 cr)

Unsupervised Learning

Joni-Kristian Kämäräinen

October 2018

Department of Signal Processing
Tampere University of Technology

Material

- Lecturer's slides and blackboard notes
- A.R. Webb and K.D. Copsey. *Statistical Pattern Recognition*. 3rd. Wiley, 2011: Chapter 10 and 11

Introduction

Unsupervised learning

Wikipedia:
In machine learning, the problem of unsupervised learning is that of trying to find hidden structure in *unlabeled data*.

Since the examples given to the learner are unlabeled, there is no error or reward signal to evaluate a potential solution. This distinguishes unsupervised learning from supervised learning.

Clustering approaches

Clustering approaches

Hierarchical clustering

Hierarchical clustering

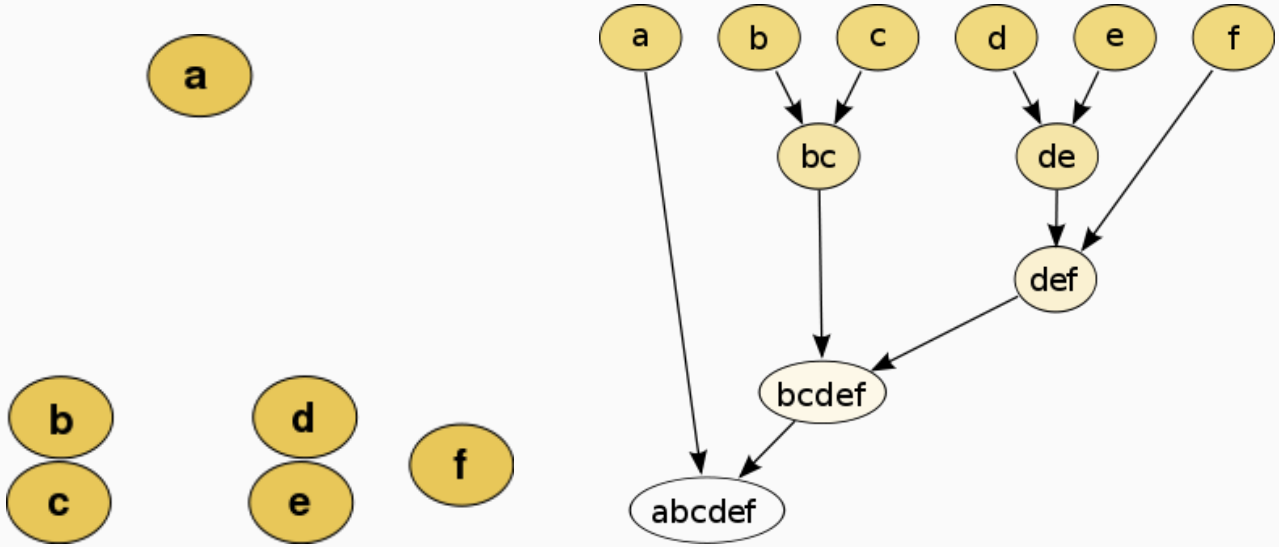


Figure 1: wikipedia.org

Single-link (minimum) distance

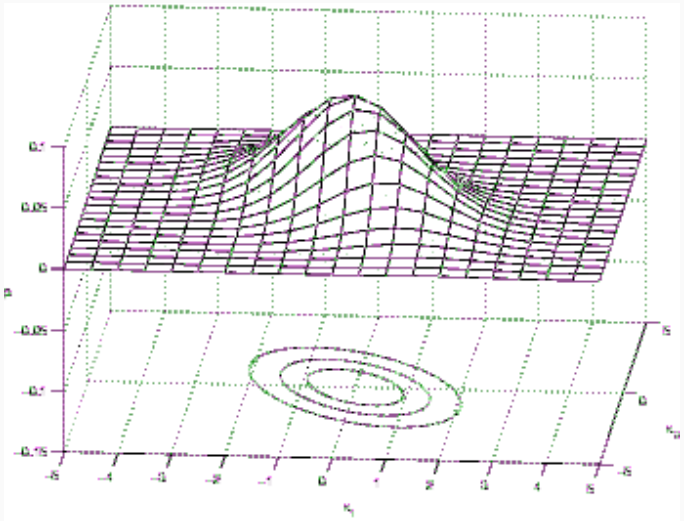
Run single-link distance based hierarchical clustering:

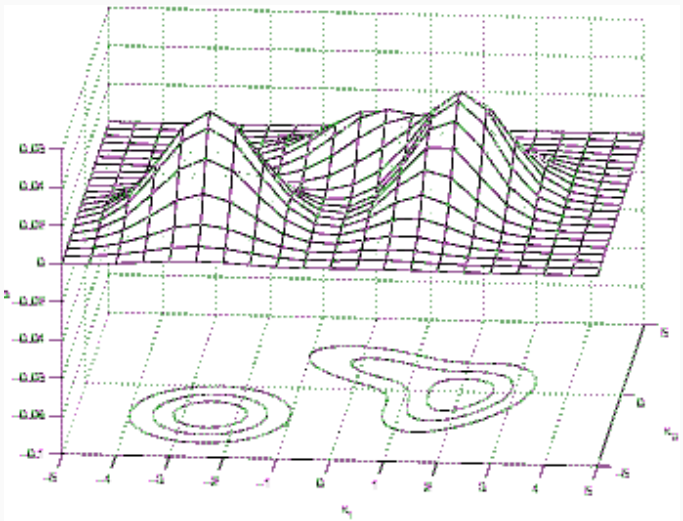
	1	2	3	4	5	6
1	0	4	13	24	12	8
2		0	10	22	11	10
3			0	7	3	9
4				0	6	18
5					0	8.5
6						0

Clustering approaches

Mixture models

Mixture models clustering





Clustering approaches

Sum-of-squares methods

Sum-of-Squares Methods

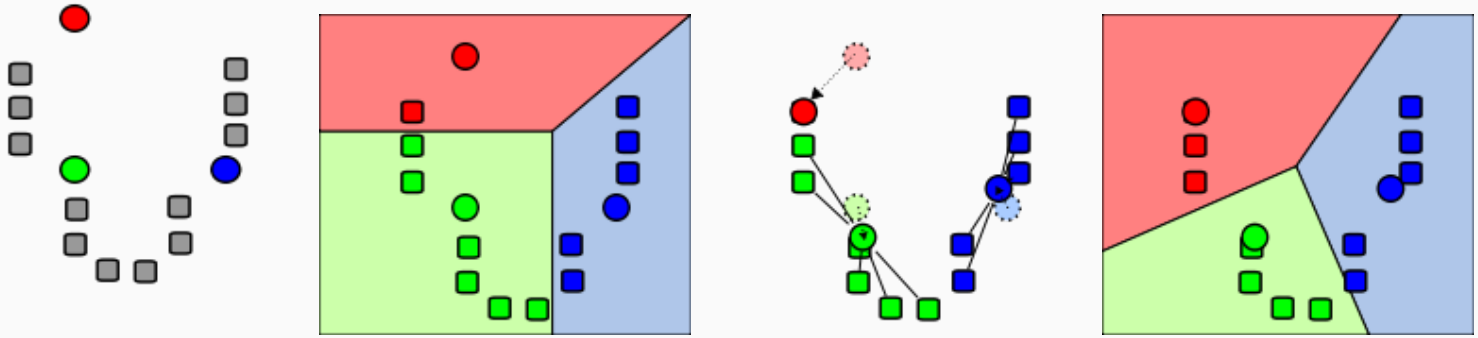


Figure 2: K-means algorithm (Wikipedia).

Clustering approaches

Spectral clustering

Spectral clustering

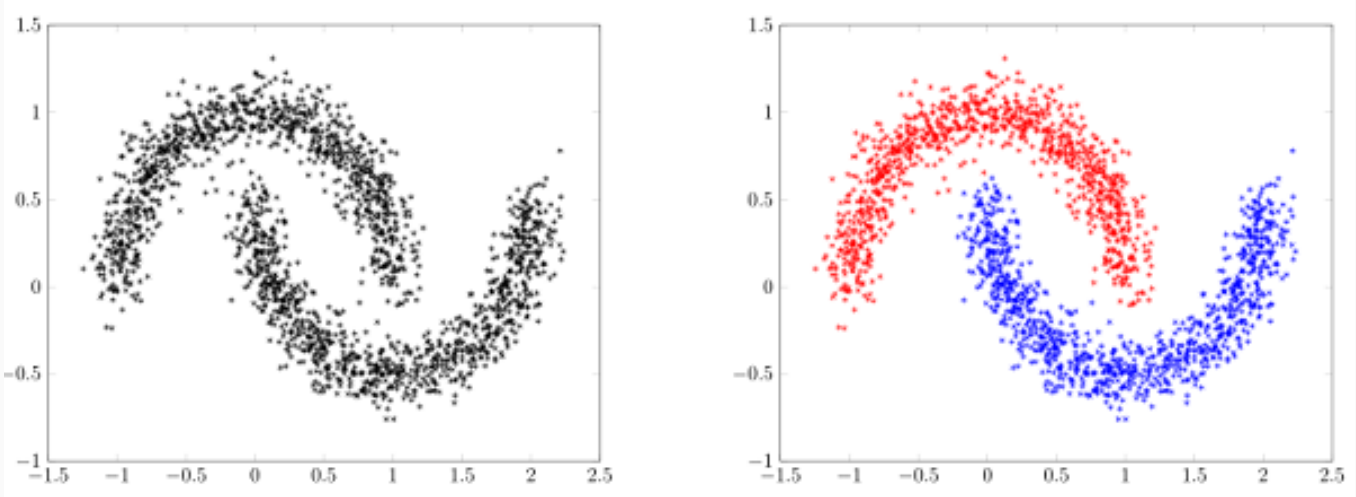


Figure 3: Spectral clustering example (Mathworks).