

Course 8	Unsupervised Pattern Recognition
Programme	<p>0. Introduction to clustering</p> <p>1. Data Exploration and preparation</p> <p>1.1 Types of features</p> <p>1.2 Feature extraction</p> <p>1.3 Graphical examination</p> <p>1.4 Missing Data and outlier removal</p> <p>1.5 Principal component analysis</p> <p>1.6 Kernel functions</p> <p>1.7 Data reduction</p> <p>1.8 Distance measures</p> <p>2. Prototype-based clustering</p> <p>2.1. K-Means</p> <p>2.2. Fuzzy K-means</p> <p>2.3. Partitioning Around Medoids (PAM)</p> <p>2.4. Mixture models (EM algorithm)</p> <p>2.5. Self-Organizing Maps (SOM)</p> <p>2.6. Other prototype-based algorithms</p> <p>3. Density-based clustering</p> <p>3.1. Density Based Spatial Clustering</p> <p>3.2. Grid Clustering</p> <p>3.3. DENCLUE (DENsity CLUstEring)</p> <p>3.4. Other density-based clustering</p> <p>5. Graph-based clustering</p> <p>5.1. Hierarchical clustering: Introduction</p> <p>5.2. Hierarchical clustering Locally optimal algorithm</p> <p>5.3. Hierarchical clustering Linking comparison</p> <p>5.4. Chameleon</p> <p>5.5. Hybrid Graph-Density based clustering: SNN-DBSCAN</p> <p>5.6. Other graph-based clustering</p> <p>6. Cluster evaluation</p> <p>6.1. Clustering tendency</p> <p>6.2. Unsupervised cluster evaluation</p> <p>6.3. Supervised cluster evaluation</p> <p>6.4. Criteria to determine the number of clusters</p> <p>7. Miscellanea</p> <p>7.1 Dynamic clustering</p> <p>7.2 Ensemble/Consensus clustering</p> <p>7.3 Subspace clustering</p> <p>7.4 Semisupervised clustering</p> <p>7.5 Clustering with obstacles</p> <p>7.6 Turning a supervised classification into a clustering algorithm</p> <p>8. Conclusions and final advise</p>
Bibliography	<p>- Alboukadel Kassambara. Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning. ISBN: 1542462703</p> <p>- Rui Xu, Don Wunsch, 2009. Clustering. IEEE Press Series on Computational Intelligence. ISBN: 9780470276808.</p> <p>- Guojun Gan, Chaoqun Ma, Jianhong WuData, 2007. Clustering: Theory, Algorithms, and Applications (ASA-SIAM Series on Statistics and Applied Probability). ISBN: 9780898716238</p> <p>- Slawomir Wierzchoń , Mieczyslaw Kłopotek. Modern Algorithms of Cluster Analysis. ISBN 978-3-319-69307-1</p>
Prerequisites	<p>Basic knowledge of programming is desirable, but not essential, to follow the course. Students must bring their own laptop with R installed. http://www.r-project.org/. It is recommended to also install the RStudio IDE: https://www.rstudio.com</p> <p>The last day of the course the student may work on his/her own dataset; if this were not possible, there will be several standard data sets to choose from.</p>