<table>
<thead>
<tr>
<th>Course 10</th>
<th><strong>Unsupervised Pattern Recognition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program</strong></td>
<td>0. Introduction to clustering</td>
</tr>
<tr>
<td></td>
<td>1. Data Exploration and preparation</td>
</tr>
<tr>
<td></td>
<td>1.1 Types of features</td>
</tr>
<tr>
<td></td>
<td>1.2 Feature extraction</td>
</tr>
<tr>
<td></td>
<td>1.3 Graphical examination</td>
</tr>
<tr>
<td></td>
<td>1.4 Missing Data and outlayer removal</td>
</tr>
<tr>
<td></td>
<td>1.5 Principal component analysis</td>
</tr>
<tr>
<td></td>
<td>1.6 Kernel functions</td>
</tr>
<tr>
<td></td>
<td>1.7 Data reduction</td>
</tr>
<tr>
<td></td>
<td>1.8 Distance measures</td>
</tr>
<tr>
<td></td>
<td>2. Prototype-based clustering</td>
</tr>
<tr>
<td></td>
<td>2.1. K-Means</td>
</tr>
<tr>
<td></td>
<td>2.2. ISODATA</td>
</tr>
<tr>
<td></td>
<td>2.3. Fuzzy K-means</td>
</tr>
<tr>
<td></td>
<td>2.4. Partitioning Around Medoids (PAM)</td>
</tr>
<tr>
<td></td>
<td>2.5. Mixture models (EM algorithm)</td>
</tr>
<tr>
<td></td>
<td>2.6. Self-Organizing Maps (SOM)</td>
</tr>
<tr>
<td></td>
<td>2.7. Other prototype-based algorithms</td>
</tr>
<tr>
<td></td>
<td>3. Density-based clustering</td>
</tr>
<tr>
<td></td>
<td>3.1. Density Based Spatial Clustering</td>
</tr>
<tr>
<td></td>
<td>3.2. Grid Clustering</td>
</tr>
<tr>
<td></td>
<td>3.3. DENCLUE (DENsity CLUstEring)</td>
</tr>
<tr>
<td></td>
<td>3.4. Other density-based clustering</td>
</tr>
<tr>
<td></td>
<td>5. Graph-based clustering</td>
</tr>
<tr>
<td></td>
<td>5.1. Hierarchical clustering: Introduction</td>
</tr>
<tr>
<td></td>
<td>5.2. Hierarchical clustering Locally optimal algorithm</td>
</tr>
<tr>
<td></td>
<td>5.3. Hierarchical clustering Linking comparison</td>
</tr>
<tr>
<td></td>
<td>5.4. Chameleon</td>
</tr>
<tr>
<td></td>
<td>5.5. Hybrid Graph-Density based clustering: SNN-DBSCAN</td>
</tr>
<tr>
<td></td>
<td>5.6. Other graph-based clustering</td>
</tr>
<tr>
<td></td>
<td>6. Cluster evaluation</td>
</tr>
<tr>
<td></td>
<td>6.1. Clustering tendency</td>
</tr>
<tr>
<td></td>
<td>6.2. Unsupervised cluster evaluation</td>
</tr>
<tr>
<td></td>
<td>6.3. Supervised cluster evaluation</td>
</tr>
<tr>
<td></td>
<td>6.4. Criteria to determine the number of clusters</td>
</tr>
<tr>
<td></td>
<td>7. Miscellanea</td>
</tr>
<tr>
<td></td>
<td>7.1 Subspace clustering</td>
</tr>
<tr>
<td></td>
<td>7.2 Ensemble/Consensus clustering</td>
</tr>
<tr>
<td></td>
<td>7.3 Semisupervised clustering</td>
</tr>
<tr>
<td></td>
<td>7.4 Clustering with obstacles</td>
</tr>
<tr>
<td></td>
<td>7.5 Biclustering, Co-clustering, Two-mode clustering</td>
</tr>
<tr>
<td></td>
<td>7.6 Turning a supervised classification algorithm into a clustering algorithm</td>
</tr>
<tr>
<td></td>
<td>8. Conclusions and final advise</td>
</tr>
</tbody>
</table>

**Bibliography**

**Prerequisites**
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/). Ideally, on the last day of the course the student should work on his/her own dataset; if this were not possible, there will be several standard data sets to choose from.

**Readings before the course**
The student will benefit more from the course if before attending she reads (these are not compulsory, only advisable):