Cluster Interpretation of the Self-Organising Map

Jan Feyereisl
- email: jqf@cs.nott.ac.uk
- supervisor: Prof. Uwe Aickelin
- IMA Seminar, 10th February 2009

Part I
Talk Outline

- Clustering within Machine Learning
- Existing Clustering Approaches
- Current Issues with Clustering
- Examples
- Why Self-Organising Map?
- Cluster Interpretations of SOM
- Future Work
Types of Learning

- Supervised Learning
  - Classification
  - Regression

- Unsupervised Learning
  - Clustering
  - Density Estimation
  - Visualisation (Data Projection)
Existing Clustering Approaches

- **Hierarchical** (Demo)
  - Single/Average/Complete - Linkage
- **Partitional**
  - *K*-means
- **Density Estimation**
  - *Expectation Maximisation* (Demo)
- **Support Vector**
  - *Support Vector Clustering*
Issues with Existing Clustering Algorithms

- Clusters are everywhere
- There is no "best" algorithm [Kleinberg 2003]
- Data collection, normalisation and cluster validity as important as clustering technique!
- Limited theoretical knowledge
  - *Classification* – proofs in Statistical Learning Theory (Vapnik et.al)
  - *Clustering* – no proofs exist
Elementary Example

- **Data:**
  - 2 Dimensional
  - Gaussian Noise
  - 3 Clusters

- Simple for both:
  - Classification
  - Clustering
Elementary Example

- **Data:**
  - 2 Dimensional
  - Gaussian Noise
  - 3 Clusters

- Simple for both:
  - Classification
  - Clustering
Elementary Example

- **Data:**
  - 2 Dimensional
  - Gaussian Noise
  - 3 Clusters

- Simple for both:
  - Classification
  - Clustering
- **Data:**
  - 2 Dimensional
  - Gaussian Noise
  - 3 Clusters

- Simple for both:
  - Classification
  - Clustering
Elementary Example

- **Data:**
  - 2 Dimensional
  - Gaussian Noise
  - 3 Clusters

- Simple for both:
  - Classification
  - Clustering
A Less Elementary Example

- **IRIS dataset**
  - 4 Dimensional
  - 150 items
  - 3 Classes – labeled

<table>
<thead>
<tr>
<th>S.Length</th>
<th>S.Width</th>
<th>P.Length</th>
<th>P.Width</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>3.5</td>
<td>1.4</td>
<td>0.2</td>
<td>setosa</td>
</tr>
<tr>
<td>4.8</td>
<td>3.4</td>
<td>1.6</td>
<td>0.2</td>
<td>setosa</td>
</tr>
<tr>
<td>6.7</td>
<td>3.1</td>
<td>4.4</td>
<td>1.4</td>
<td>versicolor</td>
</tr>
<tr>
<td>6.1</td>
<td>2.8</td>
<td>4.7</td>
<td>1.2</td>
<td>versicolor</td>
</tr>
<tr>
<td>7.7</td>
<td>3.8</td>
<td>6.7</td>
<td>2.2</td>
<td>virginica</td>
</tr>
<tr>
<td>6.0</td>
<td>2.2</td>
<td>5.0</td>
<td>1.5</td>
<td>virginica</td>
</tr>
</tbody>
</table>
A Less Elementary Example

- Problems:
  - Visualisation
  - Understanding

- Plot two Features
- **No** labels
- **No** obvious groups
A Less Elementary Example

- Problems:
  - Visualisation
  - Understanding

- Two Attributes
  - With labels
  - Groups apparent, but...
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - PCA – Principal Component Analysis
  - MDS – Multidimensional Scaling
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - PCA – Principal Component Analysis
  - MDS – Multidimensional Scaling
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - **PCA** – Principal Component Analysis
  - **MDS** – Multidimensional Scaling
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - PCA – Principal Component Analysis
  - MDS – Multidimensional Scaling
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - PCA – Principal Component Analysis
  - MDS – Multidimensional Scaling

![Sammon's Mapping - No Labels](image)
A Less Elementary Example

- Graphical Techniques
  - Scatter Plot Matrix

- Projection Methods
  - PCA – Principal Component Analysis
  - MDS – Multidimensional Scaling
A Less Elementary Example

- Clustering:
  - K-means
  - EM
  - SVC
A Less Elementary Example

K-Means Clustering - First Two Features

IRIS Dataset - First Two Features with Labels
A Less Elementary Example

- Clustering:
  - K-means
  - EM
  - SVC

![K-Means Clustering - PCA](image)
A Less Elementary Example

K-Means Clustering - PCA

Principal Component Analysis
A Less Elementary Example

- Clustering:
  - K-means
  - EM
  - SVC
A Less Elementary Example

Classification

Principal Component Analysis
A Less Elementary Example

- Clustering:
  - K-means
  - EM
  - SVC
A Less Elementary Example

- Self-Organising Map:
  - Hybrid
  - Visualisation
  - Multidimensional Scaling
  - Data Reduction
  - Maintains Topography

- Complexity:
  - $O(n^m)$ – Map size
  - $O(n)$ – Dataset size
SOM Visualisation

(ref: Wikipedia)

10/02/2009
A Less Elementary Example

- Self-Organising Map
  - U-Matrix
Why Self-Organising Maps?

- There is no one ideal algorithm for clustering for majority of problems (such as SVM in classification)
- Clustering algorithms are very dataset/problem dependent!

Next time:
- Bio-Inspired Extension of SOM for Automatic Cluster Interpretation
  - Exploits existing cluster interpretation techniques for better cluster boundary recognition
  - Extends cluster interpretation beyond numerical analysis