On Integrating Information Visualization techniques into Data Mining: A Review

Keqian Li
Problem synopsis

- Data Mining: the extraction of implicit, previously unknown, and potentially useful information from data.
- Information Visualization: allow users to see, explore, and understand large amounts of information at once
- This paper: explore how Information Visualization can be applied to complement and improve the different stages of data mining
Overview

- Visualization for Preliminary Data Analysis
- Visualization for Model Construction
- Visualization for Model Evaluation
Visualization for Preliminary Data Analysis

- Scalar value data visualization
- Multi-dimensional data Visualization
- Network Data
- Text Data
Multi-dimensional data Visualization

- Glyph based vector visualization
- Similarly encoding each dimension
- 3D Visual Data Mining
- Bipartite flow visualization
Network Data

- Node-link diagrams
- Adjacency matrix view
Text Data

- Analysis on raw text and word frequency
- Analysis with dimension reduction
Visualization for Model Construction

- Progressive construction
- Iterative prototyping
- Interactive pipeline construction
Visualization for Model Evaluation

- Classification
- Unsupervised Learning
- Neuron Networks
Classification

- Test oriented visualization
- Instance space based visualization
- Factor analysis
Unsupervised Learning

- Clustering
- Association Rules
- hidden structure in Graphical Model
Neuron Networks

- SOM
- Deep Learning
- Convolutional Networks