Information Visualization and Dimensionality Reduction

Scuola di Dottorato in Ingegneria Industriale

Post-graduate course (July 4-8, 2011)

Contact info

Teacher: Francesco Corona
Aalto University School of Science (Finland)
Dept. of Information and Computer Science
Email: francesco.corona@aalto.fi
Web: users.ics.tkk.fi/fcorona

Schedule and location

Monday July 4 to Friday July 8, 2011. Lectures start at 9:10 AM (give 5 take 5 minutes) and will finish around 1 PM.
Facoltà di Ingegneria, Dipartimento di Ingegneria Chimica e Materiali, Aula DICM.

Overview

Information visualization is the practice of displaying measured quantities or data by means of the combined use of points, lines, a coordinate system, numbers, symbols, words, shading and color. At their best, data visualizations are instruments for reasoning about quantitative information and, of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful.

- The first part of the course reviews the graphical practice of the last two centuries and provides a language for discussing graphics and a practical theory of data visualization.

Methods of dimensionality reduction are innovative and important tools in the fields of data analysis and machine learning, for they provide a way to understand and visualize the structure of complex data sets. Since the late nineties, many new methods have been developed and dimensionality reduction is now a hot topic.

- The second part of the course discusses properties and features of some advanced methods to reduce the dimensionality of numerical databases.

Syllabus

Information visualization:

- Graphical Practice: Graphical Excellence (Data-maps, Time-series, Space-time narrative and Relational graphics), and Graphical Integrity (Distortion, Design and Data variation, Visual area and Numerical measure).
- Theory of Data graphics: Data-ink, Graphical redesign, Chartjunk and Data-ink maximization.

Dimensionality reduction:

- Anatomy of a dimensionality reduction method: Purpose, Expected functionalities (Manifold learning and Latent variable extraction), Internal characteristics (Models, Algorithms and Criteria) and a walk-through example (Principal Component Analysis).
- Distance preservation: Multidimensional Scaling, Sammon's Mapping, Curvilinear Component Analysis, Isomap, and Geodesic Sammon's Mapping.
- Topology preservation: The Self-Organizing Map and Locally Linear Embedding.

Passing the course

To pass the course, students must actively participate to the lectures, and hand-in a coursework project in the format of conference poster.
Further information and the course material will be given during the lectures.