Interactive Dimensionality Reduction for Visual Analytics

Ignacio Díaz¹, Abel A. Cuadrado¹, Daniel Pérez¹, Francisco J. García¹ and Michel Verleysen² *

1- Electrical Engineering Dept. University of Oviedo
Edif. Dept. 2, Campus de Viesque s/n 33204, Gijón, SPAIN
2- Univ. Catholique de Louvain - Machine Learning Group
ICTEAM/ELEN - Place du Levant, 3 1348 Louvain-la-Neuve, Belgium

Abstract: In this work, we present a novel approach for data visualization based on interactive dimensionality reduction (IDR). The main idea of the paper relies on considering the intermediate results of non-convex DR algorithms under changes on the metric of the input data space driven by the user. With an appropriate visualization interface, our approach allows the user to focus on the relationships among dynamically selected groups of variables, as well as to assess the impact of a single variable or groups of variables in the structure of the data.

The IDR approach

1. Analysis of time-varying input datasets

Analysis of a fixed set of samples, each one characterized by a set of measurements that evolve with time (e.g., analysis of a batch of fruits, analysis of the evolution of a set of patients on an epidemics, time evolution of social networks, etc.)

2. Changes in the metric of the input space

User-driven modification of the distance metrics allows for detection of correlations in groups of variables

Interactive exploration of correlations

In the case of parametric dependency of the type

\[ x_n = f_n(t), x_m = f_m(t), \ldots, x_{km} = f_{km}(t) \]

DR methods yield an easily recognizable "snake shape" figure

Weighting variables can be exploited to select subsets of variables to "test" each type of dependency, more general than linear correlation

Example

Sample user-driven weight variation

3. Interactive incorporation of class knowledge

Extended data points using class info

Application demo: fault analysis of AC motor

Application interface with IDR user-driven modification of the input metric space

Analysis of vibration signals \(x_1, x_2, x_3, x_4\) and two phase currents \(y_1, y_2\) of a 4kW 2 pole-pair asynchronous motor. (http://visural.bitbucket.org/idr-benchmark/Motor/

IDR approach using the SNE algorithm


*The authors would like to thank financial support from the Spanish Ministry of Economy (MINECO) and FEDER funds from the EU.