Interactive Dimensionality Reduction for $V_{i}^{\{\sigma'_i\}^{(k)}}$ 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 Viesques 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 for visualization the intermediate results of nonconvex 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.





enables a quick feedback to the user and hence a much better user integration in the exploration process

I. 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.)



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

Interactive exploration of correlations

Weighting variables can be exploited to select

dependency, more general than linear correlation

subsets of variables to "test" such type of

sample user-driven weight variation

Example

In the case of parametric dependency of this type

 $x_{q_1} = f_{q_1}(t), x_{q_2} = f_{q_2}(t), \dots, x_{q_K} = f_{q_K}(t)$

DR methods yield an easily recognizable "snake shape" figure



Sensitivity analysis

Whenever the user increases a single weight " w_q " the input distance pattern d_{ij} becomes more sensitive to variable q. Points (samples) showing significantly large deviations in variable q, will move apart upon changes in w_q , thus revealing their dependency on variable q.

and ISOMAP methods, end is reflected as

rifed dimension poles Acta Ista france

effectependenaignetestworffactes



Hary Epopulities a long time.

Application demo: fault analysis of AC motor



Application interface with iDR user-driven modification of the input metric space Javascript application using processing (<u>http://processingjs.org</u>) Analysis of three vibration signals $a_x(t) a_y(t) a_z(t)$ and two phase currents $i_R(t)$, $i_S(t)$ of a 4kW 2 pole-pair asynchronous motor <u>http://isa.uniovi.es/~idiaz/demos/iDR-vibracionesMotor/</u>



an be completed on base deen fanced by the spanis the bart then dary nder grants DPI2009-13398 Co petuleelCS Cading the different degrees of electrical imbal every pr dsinta obtain^Pa p-LLE and ISOMAP, while lmb-var-**BELANZAVOL** ms to show a dependency or desident and set int SetC. Hand 1 $\mathbf{f}_{n}^{\text{th}}\mathbf{d}_{2}$ herenier en ein tere v ved water sate dreepter frea + Imb And Bwidthry Byindfor window ETRESE TAGETORS H E-dorly e exp the set of DS. Guny meenigs Cont And fer the baseling of the sector and the sector a Usien Getween the points to project new da ins transition between the largest conduction of a specific problem is to obtained for projection of a dependence of the big dimensional feature space. This yields a distribution of dimensional feature space. This yields a factors that define the actual but ion that reflects the factors on which ature space. the different condi-Stalization of dynamics is required such operating condition ariginatedection of owd but happen of sease actors in the different condi-operating condition ariginatedection of owd but happen of sease actors condition can be established. the process conditions depend. Best in the signal spectrum lead in the signal spectrum of Economy is hypothesis, the problem is to Tohonea with cars would chike nt low dimensional structures in we paie as f prefactly; DE that is and/or glyphs to represent features with a physical sense, this physical sense phy space. This yields a distribution and meetings asymmetries) PJ200946831084Cd2n0ens20nal MINISTERIO ature * roj<u>e</u>tions using a spatial distri-* In this work four state-of-the-art meth-DE CIENCIA The resulting projections can be effihat reflects the factors on which ods, namely, LTSA, LLE, L-Eig, and **E INNOVACIÓN** The lives while generations wan abrate the * 🛧 🖈 ess conditions depend. ISOMAP are tested, showing that all sieptlyds en sesented dusin an siened ization produce insightful results and lead to onethedatthatingovideanitasightfulsving aidatatentities using col- conceptually similar conclusions. effetbetehenzinge dynamicsropussestipg pheaenotantiaanyseproblems proposed ap-

pronebrian many problems in which wi

onditions -Inb-funnerImb-10

a 4th order Butterworth type filter in order to limit aliasing

nd latent low dimensional structures in ors and/or glyphs to represent features the feature space an precisely, DR_tech , a variables dvith a physical sense, this