Contribution ID: **27** Type: **not specified**

Fast multimodal clustering: searching for optimal patterns

Saturday 20 February 2016 10:40 (40 minutes)

In Machine Learning, we usually deal with object-attribute tables. However, underlying objects may have other modalities than attributes only. For instance, an object may have a certain attribute only under specific conditions. The real examples came from gene expression data, where a gene can be active (expressed) in particular situations at a certain moment of time, implying ternary relation with triples (g,s,t). One more example came from resource sharing systems like Flickr or Bibsonomy, i.e. a user u can assign a certain tag t to a resource r. One may ask how to find homogeneous patterns, groups of genes with similar properties or communities in such data.

This talk presents several definitions of "optimal patterns"in triadic data and results of experimental comparison of five triclustering algorithms on real-world and synthetic datasets. The evaluation is carried over such criteria as resource efficiency, noise tolerance and quality scores involving cardinality, density, coverage, and diversity of the patterns. An ideal triadic pattern is a totally dense maximal cuboid (formal triconcept). Relaxations of this notion under consideration are: OAC-triclusters; triclusters optimal with respect to the least-square criterion; and graph partitions obtained by using spectral clustering. We show that searching for an optimal tricluster cover is an NP-complete problem, whereas determining the number of such covers is #P-complete. Our extensive computational experiments lead us to a clear strategy for choosing a solution at a given dataset guided by the principle of Pareto-optimality according to the proposed criteria. In the end on the talk, we will outline future prospects of multimodal triclustering and its relationship with tensor factorisation.

Author: Dr IGNATOV, Dmitry (HSE)

Presenter: Dr IGNATOV, Dmitry (HSE)

Session Classification: Data Science Applications