Dependence Structure Mining, Statistical Approaches, and Bayesian Ying-Yang Learning

Lei Xu, Professor of Computer Science and Engineering,
IEEE Fellow, IAPR Fellow,
Member of European Academy of Sciences
Chinese University of Hong Kong, Shatin, NT, Hong Kong,
lxu@cse.cuhk.edu.hk phone 852 26098423, fax: 852 26035024

Mining certain dependence structures from data are important to various tasks of statistical learning, data mining, and pattern recognition. This tutorial consists of three parts. ***The first part*** introduces fundamentals. From a general statistical learning perspective, an overview is made not only on the tasks of mining major dependence structures but also on the tasks of perception oriented intelligent learning. Moreover, major efforts towards a key challenge, namely making learning on a finite size of samples with model selection ability, are summarized in two typical streams. Furthermore, the fundamentals of Bayesian Ying Yang (BYY) harmony learning are introduced as a unified statistical learning framework for various dependence structures and as a promising tool for solving this key challenge. ***The second part*** of this tutorial introduces major results of BYY harmony learning on seven categories of tasks of unsupervised learning, supervised learning, and temporal modeling. Namely, (1) optimal inversion such as Bayesian classifier and Kalman filter; (2) independence learning that consists of independent component analysis (ICA) , Gaussian factor analysis, binary factor analysis, nonGaussian factor analysis, and least mean square error reconstruction (LMSER); (3) Gaussian mixture based learning that covers MSE clustering, elliptic clustering, subspace clustering, and density estimation; (4) Independence finite mixture that combines the features of (2) and (3), e.g., local ICA, local NFA, local binary factor analysis, and local LMSER, etc; (5) Supervised learning on three layer nets, mixture-of-experts (ME) models, alternative ME models, radial basis function (RBF) nets, and kernel regressions; (6) Temporal extensions of (2)&(3)&(4) to hidden markov model (HMM) and variants, temporal ICA, temporal factor analysis, and temporal LMSER; and (7) Various BYY learning based topological maps in help of two typical strategies for topology recovering. These results are featured by not only new regularization techniques for parameter learning but also a new mechanism that implements model selection, on the number of clusters /Gaussians and the number of factors/states, either automatically during parameter learning or via a new class of model selection criteria used after parameter learning. ***The third part*** demonstrates several applications on mining dependence structures, pattern recognition, and financial engineering.