Deep Boltzmann Machines (Salakhutdinov, Hinton)

John Min

March 30, 2014

New learning algorithm for Boltzmann machines containing many layers of latent variables. Data-dependent expectations are estimated using a variational approximation that tends to focus on a single mode. Data-independent expectations are approximated using persistent Markov chains. These quite different techniques are used to estimate the two types of expectations that constitute the gradient of the log-likelihood.

Originally, the learning algos for Boltzmann machines (Hinton and Sejnowski, 1983) required randomly initialized Markov chains to approach their