Summary 4: Hierarchical Dirichlet Processes

Susan Cherry

This paper introduces Hierarchical Dirichlet Processes, which a s nonparametric approach to modeling groups of data. Each group has a mixture model and we allow mixture components to be shared among groups. The paper goes on to describe three different representations of the Dirichlet Process: The stick breaking representation, the Chinese restaurant franchise, and an infinite limit of finite mixture models. The idea of hierarchical Dirichlet process that this paper discusses is an example of a dependency model for multiple Dirichlet processes, specifically aimed at the problem of sharing clusters among groups of related data.

HDP is necessary because if we simply use a Dirichlet Process, this simple hierarchical model will not solve the modeling problem. Clusters within each group arise through the discreteness of draws from a DP, the atoms associated with different group are different and there is no sharing between groups. To overcome this problem, the Hierarchical Dirichlet Process is introduced, which allows atoms to be shared among multiple Dirichlet Processes and results in the desired sharing of atoms among groups.

The paper then describes three related MCMC sampling schemes from the hierarchical Dirichlet process mixture model. The first is the Gibbs sampler, the second is an augmented representation based on both the Chinese restaurant franchise and the posterior for G_0, and the third is a variation on the second with streamlined bookkeeping.