Combining preference and social information in recommendation systems using nonparametric Bayesian mixtures

Abel Rodriguez*
University of California, Santa Cruz, USA – abel@soe.ucsc.edu

Radhakrishna Vuppala
University of California, Santa Cruz, USA – vrk@soe.ucsc.edu

Recommender systems, which attempt to provide users with suggestions about items (such as scientific papers, movies, songs, or news articles) on the basis of the item's content and/or previous user selections, have become an important area of research with numerous applications on e-commerce. This paper introduces a joint statistical model for user preferences, social relationships among users and item features that can serve as the basis for a recommendation system in online platforms. The model is constructed using ideas from the literature on Bayesian nonparametric mixture modeling. More specifically, user preferences are modeled using an infinite relational model (IRM) (e.g., see Kemp et al, 2006) in which both users and items are independently partitioned into homogeneous groups, social relationships are modeled using a stochastic blockmodel that also partitions individuals into groups with similar social roles, and item features are modeled using a topic model (e.g., see Blei, Ng and Jordan, 2003), which partitions items into groups with homogenous features. Information is shared across both components of the model through a common partition of items and users. The model can be interpreted as projecting the original data into a two-dimensional discrete latent space; the first dimension of this latent space captures similarities among users driven by their preferences for previously examined items and their social relationships, while the second dimension captures similarities among items driven by the similarities in item content and similarities on users preferences. Hence, the recommender system derived from our model is a hybrid system that combines ideas from collaborative, content-based and social filtering. The model is implemented using a stochastic gradient variational algorithm that allows us to apply the model to real-life applications.

Keywords: Recommendation System; Nonparametric Bayes; Stochastic Blockmodel; Variational Algorithms.