**Introduction**

Multi-relational data—capture relations among multiple entities of interest.

**Bayesian Multi-relational Clustering**

Hanhui Shan and Arindam Banerjee

Department of Computer Science and Engineering

University of Minnesota, Twin Cities

**Bayesian Multi-relational Clustering**

• Given a data matrix denoting the relationship between row and column entities, obtaining row and column clusters simultaneously

• A Bayesian model allowing mixed membership on both sides

**Mixed membership naïve Bayes models**

• One-way Bayesian clustering

  • A naïve Bayes model with Dirichlet priors to allow mixed membership

  • Micro-precision

    | TFMMB | NB
    |------|------|
    | 0.634 | 0.629 |

  • Perplexity

    | TFMMB | NB
    |------|------|
    | 0.015 | 0.014 |

**Discriminative latent Dirichlet allocation**

• Introducing labels to latent Dirichlet allocation for classification

• Allow the number of classes to be larger than the number of topics

#topics increases from c to c+100. A higher accuracy is usually observed with a larger #topics.

DLDA is better than or comparable to other classification algorithms, including SVM.

**Fast mixed membership naïve Bayes & Fast latent Dirichlet allocation**

By using a fast variational inference algorithm, Fast MMNB and Fast LDA are about 10 times faster than standard MMNB and LDA, with a similar perplexity.

**Future work—Bayesian multi-relational clustering**

**Book Model**—One entity (User) is connected with multiple other entities (Actor, Movie, Other Users) through corresponding relations.

**Tensor Model**—Multiple entities (User, Movie, Word) are connected through one relation.

**General case**—A combination of Book Model and Tensor Model. Arbitrary forms of relationships