

Summary of Clustering Methods

- Model based approaches (mixture model)
 - Uses an implicit similarity function (model → clustering bias)
 - Cluster structure is “built” into a generative model
 - Complex generative models can discover complex structures
 - Prior can be leveraged to further customize the clustering algorithm
 - However, no easy way to directly control the similarity measure
- Similarity-based approaches
 - Allows for direct and flexible specification of similarity
 - Objective function to be optimized is not always clear
- Both approaches can generate both term clusters and doc clusters