Abstract

Social networks play an increasingly important role in online world as it enables individuals to easily share opinions, experiences and expertise. The capability to extract latent communities based on user interest is becoming vital for a wide variety of applications. However, existing literature on community extraction has largely focused on methods based on the link structure of a given social network. Such link-based methods ignore the content of social interactions, which may be crucial for accurate and meaningful community extraction. In this paper, we present a novel approach for community extraction which naturally incorporates the content published within the social network with its semantic features. Two layer generative Restricted Boltzmann Machines model is applied for community discovery. The model assumes that users within a community communicate based on topics of mutual interest. The proposed model naturally allows users to belong to multiple communities. Through extensive experiments on the Twitter data for scientific papers, we demonstrate that the model is able to extract well-connected and topically meaningful communities.


an application to information retrieval. Advances in neural information processing systems, 17, 1481-1488.


Index Terms

Computer Science

Information Systems

Keywords

Social network    Community discovery    Machine learning    Restricted Boltzmann Machines

modeling

Semantic similarity.