Feature Weighting and Instance Selection for Collaborative Filtering: An Information-Theoretic Approach

Kai Yu[1,2], Xiaowei Xu[1,3], Martin Ester[2] and Hans-Peter Kriegel[2]

[1]Corporate Technology, Siemens AG, Munich, Germany
[2]Institute for Computer Science, University of Munich, Munich, Germany
[3]Information Science Department, University of Arkansas at Little Rock, Little Rock, AR, USA

Abstract. Collaborative filtering (CF) employing a consumer preference database to make personal product recommendations is achieving widespread success in E-Commerce. However, it does not scale well to the ever-growing number of consumers. The quality of the recommendation also needs to be improved in order to gain more trusts from consumers. This paper attempts to improve the accuracy and the effciency of collaborative filtering. We present a unified information-theoretic approach to measure the relevance of features and instances. Feature weighting and instance selection methods are proposed for collaborative filtering. The proposed methods are evaluated on the well-known EachMovie data set and the experimental results demonstrate a significant improvement in accuracy and efficiency.

Keywords: Collaborative filtering; Recommender systems; Feature weighting; Instance selection; Instance-based learning; Data mining