SCV 2008: December 1-3, Taipei, Taiwan (Abstract)

Decision tree modeling for ranking data

Philip L.H. Yu
The University of Hong Kong, Hong Kong, China.

Abstract

Ranking data arises from many applications in marketing, psychology and politics. We establish new decision tree models for the analysis of ranking data by adopting the concept of classification and regression tree. We modify the existing splitting criteria, Gini and entropy, which can precisely measure the impurity of a set of ranking data. Two types of impurity measures for ranking data are introduced, namely n-wise and top-k measures. Minimal cost-complexity pruning is used to find the optimum-sized tree. In model assessment, the area under the ROC curve (AUC) is applied to evaluate the tree performance. The proposed methodology is implemented to analyze a partial ranking dataset of Inglehart's items collected in the 1993 International Social Science Programme survey. Change in importance of item values with country, age and level of education are identified. (Joint work with W.M. Wan and Paul H. Lee)