

An Application Preference Text Classification Method Based on TextRank

Abstract

This invention provides an application preference text classification method based on TextRank, including the steps as follows: generate keywords of each App according to the TextRank algorithm to form a first keywords stock; indicate a seed keyword for each sub-category according to the plurality of sub-categories; get the Apps including the seek keywords from the first keywords stock by fuzzy searching according to the seed keywords and indicate such Apps with sub-categories; conduct full calculation for the seek keywords of all Apps under the sub-categories by the TextRank algorithm and generate the second keywords stock under a plurality of sub-categories; traverse the list of Apps again and compare the contents of each keyword with the second keywords stock in the similarity of character strings; if the similarity is lower than the preset threshold, delete the association between the Apps and the current sub-categories. This invention can study by itself and gradually remove the unconcerned keywords according to the effect of core keyword generation to improve the accuracy.

FIG. 1