URN etd-0407109-183709-639
Author Jong Yu Chow
Author's Email Address No Public.
Statistics This thesis had been viewed 42 times. Download 57 times.
Department
Year 2004
Semester 2
Degree Master
Type of Document Master's Thesis
Language zh-TW.Big5 Chinese
Title A Study of Feature Selection and Pre-Sorting Strategy for COBWEB Algorithms
Date of Defense 0000-00-00
Page Count 63
Keyword
hierarchical conceptual clustering
feature selection
search strategy
COBWEB
Gini index

Abstract
This thesis studies hierarchical conceptual clustering for feature selection and search strategy. Clustering is often used for discovering structure in data. Clustering systems differ in the objective function used to evaluate clustering quality and control strategy used to search the space of clustering. Ideally, the search strategy should consistently construct clusterings of high quality, but be computationally inexpensive as well.
Hierarchical conceptual clustering by a system known as COBWEB. COBWEB, a conceptual clustering system that store knowledge in conceptual hierarchical , and uses an heuristic evaluation function called category utility that measures the quality for a set of probabilistic categories.
In this thesis, we propose two strategies for COBWEB algorithm. Gini index is similar in form to the category utility, and based on standard squared-difference metric. It is used for feature selection and search strategy. More, Gini index and category utility are predicted class labels as probability. In particular, this thesis investigates the presorting strategy, it is more efficiency to reducing ordering effects on input data, furthermore, improve the clustering quality. Continue, being a computationally information gain to removes importantless feature subset, therefore, make lower time complexity in particular clustering. Ideally, these two strategies should consistently construct hierarchical conceptual clustering of high quality as well.
Advisory Committee Guang Yeh Tung - advisor

Files etd-0407109-183709-639-001.pdf

indicate access worldwide
Date of Submission 2005-07-29