Abstract

Clustering is a useful method that categorizes a large quantity of unordered text documents into a small number of meaningful and coherent collections, thereby providing a basis for instinctive and informative navigation and browsing mechanisms. Different type of distance functions and similarity measures have been used for clustering, such as squared, cosine similarity, Euclidean distance and relative entropy. This paper presents text document space dimension reduction in text document retrieval by agglomerative clustering and Hebbian-type neural network. Hebbian-type neural network reduce document space to two dimensions so each document is represented as a point in the reduced document space. Furthermore, the clusters are formed in compact document space.

References

- Unsupervised Learning of Semantic Relations for Molecular Biology Ontologies,
An Approach for Document Clustering using Agglomerative Clustering and Hebbian-type Neural Network


Index Terms

Computer Science  
Artificial Intelligence
Keywords
Agglomerative and Oja Learning Rule of hebbian-type neural network F-measure