

Growing Self Organising Map based Exploratory Analysis of Text Data

Authors : Sumith Matharage and Damminda Alahakoon

Abstract : Textual data plays an important role in the modern world. The possibilities of applying data mining techniques to uncover hidden information present in large volumes of text collections is immense. The Growing Self Organizing Map (GSOM) is a highly successful member of the Self Organising Map family and has been used as a clustering and visualisation tool across wide range of disciplines to discover hidden patterns present in the data. A comprehensive analysis of the GSOM's capabilities as a text clustering and visualisation tool has so far not been published. These functionalities, namely map visualisation capabilities, automatic cluster identification and hierarchical clustering capabilities are presented in this paper and are further demonstrated with experiments on a benchmark text corpus.

Keywords : Text Clustering, Growing Self Organizing Map, Automatic Cluster Identification, Hierarchical Clustering.

Conference Title : ICEP 2014 : International Conference on Electronic Publications

Conference Location : journal city, WASET

Conference Dates : November 23-23, 2014