

17.Hierarchy-Cutting Model Based Association Semantic For Analyzing Domain Topic On The Web

Abstract –

Traditional text mining techniques for semantic association transform free text into association semantic and Points of Interests (POI) for knowledge discovery. However, it does not consider time and precision while performing linear inseparability from textual web content. Effective relation association semantic technology using machine learning analysis not only can reduce time for knowledge discovery from textual web content, but also can improve the searching accuracy of the related information system. How to realize the semantic relation mining by the machine learning analysis is an important research topic. In this paper, a three-step procedure to mine associations of semantic relations for textual web document content, called, Polynomial Kernelized Maximum Entropy and Support Vector Machine (PKME-SVM) framework is presented. First, Associative Polynomial Kernel-based Maximum Entropy representing semantic relations are extracted from raw text web contents using Polynomial Kernel. The semantic relation extraction process also creates a sentence grammar tree in the form of reduced sentence. Then, Probabilistic Term (i.e. word) Taxonomy (PTT) Framework is applied to discover the probabilistic term on corresponding Web Content Domain. Finally, for pruning the semantic relation from textual web content, the Generalized Association Support Vector Mining algorithm adopts the notion of dual characteristic function for systematic overgeneralization reduction. The objective of PKME-SVM is to obtain accurate result from textual web content using semantic relation mining operations and satisfy the web user specific needs through polynomial kernel. The efficacy of our framework is demonstrated through empirical experiments conducted on Freebase Data Dump. Experiment is conducted on the factors such as personalized information search retrieval rate, computation time and precision ratio.