Abstract

Most of the text mining methods use term-based mining. All those methods are affected by common problems such as synonymy and polysemy. Mining of patterns have more advantage than other term based methods. Pattern Taxonomy Mining can be used to increase the effectiveness in the discovery of useful patterns. In addition to solving the common problems in term based mining, this paper tries to address the low occurring problems as well. Algorithms to deploy patterns and to evolve inner pattern are used to improve the effectiveness of pattern discovery. RCV1 text collection is used for experiments in this paper. Performance and execution of text categorization have significantly enhanced without any lose in the accuracy rate.

References


12. S. t. Wu, Y. Li and Y. Xu, "Deploying Approaches for Pattern Refinement in Text Mining," Sixth International Conference on Data Mining (ICDM’06), Hong Kong, 2006, pp. 1157-1161.


Index Terms

Computer Science

Data Mining
Keywords

Pattern Mining, Text Categorization, Inner Patterns, Pattern Taxonomy, Useful Information.