Semantic similarity plays a significant role in the areas of Web mining, Information Retrieval, NLP and Text mining. Even though it is exploited in various applications accurately measuring semantic similarity still remains a challenging task. In this paper a method is proposed to measure semantic similarity between words using web as information source and by combining two existing approaches to measure semantic similarity they are: Semantic Word Distance (SWD) and Snippets. The SWD measure finds the semantic similarity by determining the frequency of occurrences of the words in web pages (corpus). The semantic relation between words are also obtained through lexical patterns which are extracted from text snippets. A
robust method is used to integrate these similarity scores using support vector machine. For the experimental purpose 100 word pairs are used to train the support vector machine and it classifies the word pair as either synonymous or non synonymous with higher accuracy.

References


Index Terms

Computer Science  
Current Trends In Advanced Computing