Abstract

In recent years, data mining applications have been found quite extendible in the area of social science like mass communication and religion studies. In traditional approach used for such work, hidden semantics between documents were not considered well. In this study, we have shown that text mining can be applied to classify social figures like politician, religious leaders.
Such classification is based on text mining of speeches delivered by social figures. These social figures are famed personalities and their speeches are collected from their official websites. Our text classification is based on tf.idf followed by cosine and Jaccard Similarity. To improve the results on discerning features, we have designed a hash graph modeling technique Knowledge Discovery System for Social Figures (KDSSF) based on synonym words dictionary. In the comparative analysis of speeches made by social figures, we did not focus on the provision of the optimal matches but overall classification of the social figures in any domain of interests. Preliminary experiments have illustrated that inclusion of hash based graph modeling can significantly improve the results of classification.

Reference

377-386.

**Index Terms**

Computer Science

Pattern Recognition

**Key words**

Graph Modeling

Term Frequency

Inverse Document Frequency

Social Sciences

Synonyms