Abstract

Automated Text Categorization (ATC) is a useful technology to build such software tool that can categorize a document to one of many predefined categories. Unfortunately there is no as such classifier for Holy Quran, one of the most important documents of the universe. This is because Holy Quran is written in Arabic and Arabic language processing is yet not that mature that language processing can be done. This paper aims on building a software tool that can categorize any verse of Holy Quran to one of predefined categories and if it is not falling in one of predefined categories it can automatically define a new category. New category will be added to database for future categorization process. Moreover, this categorization is not just based upon word count rather it is based upon word as well as meaning of the words in the verse and there correspondence in semantic network. The semantic network for this task is created and used in this classification. This tool will help those people who want to know the theme of a verse of Holy Quran.

Index Terms

Computer Science Artificial Intelligence

Keywords
Automated Text Categorization Semantic Network Semantic Search