

Arabic Text Classification: Review Study

Authors : M. Hijazi, A. Zeki, A. Ismail

Abstract : An enormous amount of valuable human knowledge is preserved in documents. The rapid growth in the number of machine-readable documents for public or private access requires the use of automatic text classification. Text classification can be defined as assigning or structuring documents into a defined set of classes known in advance. Arabic text classification methods have emerged as a natural result of the existence of a massive amount of varied textual information written in the Arabic language on the web. This paper presents a review on the published researches of Arabic Text Classification using classical data representation, Bag of words (BoW), and using conceptual data representation based on semantic resources such as Arabic WordNet and Wikipedia.

Keywords : Arabic text classification, Arabic WordNet, bag of words, conceptual representation, semantic relations

Conference Title : ICCSE 2016 : International Conference on Computer Science and Engineering

Conference Location : Kuala Lumpur, Malaysia

Conference Dates : February 11-12, 2016