

ABSTRACT

Arabic language - the original language of Al-Quran - is quite rich in vocabulary. On the other hand, Indonesian language is quite simple and has very few vocabulary. While the Indonesian Translation of Al-Quran (ITQ) is done by expert scholars in Arabic and Indonesian languages, it is difficult for search engines to accurately search ITQ to provide precise results. Word sense disambiguation (WSD) is the ability to identify the meaning of words in context in a computational manner. The ambiguity in searching the Indonesian translation of Al-Quran comes from Synonyms, hypernyms, antonyms, hyponyms, homonymy. Based on these factors, therefore, that need special attention when searching ITQ. The main objective of this research is to develop searching system using WSD on concept for ITQ in order to improve the performance of searching and provide information needed by the user. The following process are selection of word senses, preprocessing text, use of external knowledge sources, set of features, feature vector, and apply of classification methods. This research has been applied WSD classification methods; supervised and Unsupervised method. Experiment result is selection of word senses, context interpretation, system srchitecture, and system protoype. The research will provide a benefit for both researchers and practitioners. Hence, the findings will provide theoretical, methodological, and practical contribution for fields of Information & Communication Technology (ICT) and Islam. The experimental results could be used to apply Indonesian machine translation, question and answering system, speech synthesis. The system development is also expected be an effective tool in facilitating people to learn about Al-Quran content in an easy and efficient way.

Keyword : word sense disambiguation, Search, Concept, unsupervised, supervised, Natural Language Processing