

ABSTRACT

Al-Qur'an is a source and guidance for Muslims. It is a document that is 15 centuries ago and well written in the Arabic language. Many Muslims have to learn Arabic in addition to knowing the meaning of it. One of the most critical aspects of the Arabic language is morphology and identifying the word's morphological description. It is called by morphological analysis task. This task is essential because, from the morphological aspect of a word, it can know the different forms of a word, and from that, it can know its meaning. The Gonzales paper has successfully created a model to identify morphological features (MSD) of Arabic word verb only. This study will focus on adding some other Arabic type of word, which is a noun. Trying to use the current state-of-the-art approach method is neural-based with the recurrent neural network (RNN). RNN can capture more information about the sequence of sub-word like prefix, infix, root, and suffix to make a better msd identifier. The input is a single Arabic word Going through pattern extraction, subword vectorizing, verb form identification, pronoun and type of word identification, and finally MSD identification process to see the result. This study helps to improve the last gap (morphological analyzer on POS noun) and enhances what can be identified from the previous approach to perform better. This model successfully identifies MSD with 99% accuracy and 97% F1 - score. This method also has been compared with Jabalin Gonzales system using this system data testing and get better accuracy with 99% and Jabalin got 39% accuracy.

Keywords: morphological analyzer, morphosyntactic description, recurrent neural network, Arabic word classification , sub-word vectorizing