

Alian M. Awajan A. 2020. Paraphrasing Identification Techniques in English and Arabic Texts. The 11th International Conference on Information & Communication Systems April 7– 9 2020, Irbid, Jordan. Pages 155-160

Abstract

Rewriting sentences using different words leading to the same meaning of the original sentence is called paraphrasing while paraphrasing identification task is the task of detecting the sentence that is paraphrased from another. This research provides a literature survey of researches that have studied and proposed methods for paraphrasing identification of English and Arabic texts. Accordingly, the impact of paraphrasing on NLP applications and paraphrasing benchmarks are also studied. The comparative study on the available paraphrasing identification techniques shows that the best precision is provided with WordNet based techniques while the best accuracy is provided by deep learning with statistical features of English texts. In the case of Arabic, the best precision is provided by the distributed word vector representations with Convolutional Neural Network.