



كلية الحاسبات والذكاء الاصطناعي
Faculty of Computers and Artificial Intelligence
كلية معتمدة



بيانات البحث رقم (٢):

Arabic semantic similarity approach for Farmers' complaints	عنوان البحث باللغة الإنجليزية
Rehab A.Farouk, Mohamed H. Khafagy, Mostafa Ali, Kamran Munir, Rasha M.Badry	أسماء المؤلفين
International Journal of Advanced Computer Science and Applications	اسم المجلة
October 2021	تاريخ النشر
International Journal	التصنيف
An Approach for Measuring Arabic Text Semantic Similarity	البحث مشتق من رسالة ماجستير
2156-5570	ISSN
10.14569/IJACSA.2021.0121038	DOI

ملخص البحث باللغة الإنجليزية :

Semantic similarity is applied for many areas in Natural Language Processing, such as information retrieval, text classification, plagiarism detection, and others. Many researchers used semantic similarity for English texts, but few used for Arabic due to the ambiguity of Arabic concepts in both sense and morphology. Therefore, the first contribution in this paper is developing a semantic similarity approach between Arabic sentences. Nowadays, the world faces a global problem of coronavirus disease. In light of these circumstances and distancing's imposition, it is difficult for farmers to physically communicate with agricultural experts to provide advice and find suitable solutions for their agricultural complaints. In addition, traditional practices still are used by most farmers. Thus, our second contribution is helping the farmers solve their Arabic agricultural complaints using our proposed approach. The Latent Semantic Analysis approach is applied to retrieve the most problem-related semantic to a farmer's complaint and find the related solution for the farmer. Two methods are used in this approach as a weighting schema for data representation are Term Frequency and Term Frequency-Inverse Document Frequency. The proposed model has also classified the big agricultural dataset and the submitted farmer complaint according to the crop type using MapReduce Support Vector Machine to improve the performance of semantic similarity results. The proposed approach's performance with Term Frequency-Inverse Document Frequency-based Latent Semantic Analysis achieved better than its counterparts with an F-measure of 86.7%.