

Biltawi M., Awajan A., Tedmori S. 2019. Evaluation of Question Classification. Proceedings of the 2nd International Conference on new Trends in Computing Science (ICTCS'19). 9-11 October 2019, Amman- Jordan.

Abstract

The goal of this paper is to study question classification for the Arabic language using machine learning approaches. Different experiments were conducted using two types of weighting schemes and three classifiers; Multinomial Naïve Bayes, Decision Trees, and Support Vector Machine. The dataset used in the experiments is an updated version of CLEF. The best results were obtained when the dataset was preprocessed by removing punctuation, diacritics, stop-words, and performing normalization and stemming and then using the TF-IDF weighting scheme; with SVM being the best classifier among the three with an F1-score of 81%.