

## Improved Word Representation for Arabic Sentiment Analyzing\*

Raid M. Saabni

School of Computer Science, The Academic College of Tel-Aviv Yafo, Israel

The Triangle Research and Development Center , Kafr Qarea, Israel

Correspondence should be addressed to: Raid M. Saabni; RaidSa@mta.ac.il

\* Presented at the 40<sup>th</sup> IBIMA International Conference, 23-24 November 2022, Seville, Spain

Copyright © 2022. Raid M. Saabni

### Abstract

In this article, we investigate a variety of documents to vector representation algorithms that may be used for text categorization and sentiment analysis. The document to vector representation technique known as doc2vec was the very first contender for such a job that we examined. This is followed by the TF and TF/IDF as basic approaches, and eventually the word2vec technique. Several techniques and combinations of these vectors, and other approaches to machine learning have been evaluated and assessed. Even when compared to deep pre-trained models such as the Bert model and others, the accuracy rates increase by 2% - 3% in average when the results of the created scheme in this study are compared to the state of-the-art results on various Arabic sentiment analysis benchmarks.

**Keywords:** Sentiment Analysis, Word2Vec, Doc2Vec, Pretrained Models.