

Abstract

Word sense disambiguation is the process of determining the proper meaning of a word according to its context. In this study, we represent the impact of word embedding on building Arabic sense inventory by an unsupervised approach. Three pre-trained embeddings are tested to investigate their effect on the resulting sense inventory and their efficiency in word sense disambiguation for Arabic context. Sense inventories are constructed using a fully unsupervised method based on graph-based word sense induction algorithm. The results show that Aravec-Twitter inventory achieves the best accuracy of 0.47 for 50-neighbors and a close accuracy to the Fasttext inventory for 200-neighbors.