

Jordan University of Science and Technology

Using Enhanced Lexicon-Based Approaches for the Determination of Aspect Categories and Their Polarities in Arabic Reviews

Authors: Mohammad Al Smadi, Islam Obaidat, Mahmoud Al-Ayyoub, Rami Mohawesh, Yaser Jararweh

Abstract: Sentiment Analysis (SA) is the process of determining the sentiment of a text written in a natural language to be positive, negative or neutral. It is one of the most interesting subfields of natural language processing (NLP) and Web mining due to its diverse applications and the challenges associated with applying it on the massive amounts of textual data available online (especially, on social networks). Most of the current work on SA focus on the English language and work on the sentence-level or the document-level. This work focuses on the less studied version of SA, which is aspect-based SA (ABSA) for the Arabic language. Specifically, this work considers two ABSA tasks: aspect category determination and aspect category polarity determination, and makes use of the publicly available human annotated Arabic dataset (HAAD) along with its baseline experiments conducted by HAAD providers. In this work, several lexicon-based approaches are presented for the two tasks at hand and show that some of the presented approaches significantly outperforms the best-known result on the given dataset. An enhancement of 9% and 46% were achieved in the tasks aspect category determination and aspect category polarity determination respectively