

Jordan University of Science and Technology

Sentiment Analysis for Dialectical Arabic

Authors: Rehab Duwairi

Abstract: This article investigates sentiment analysis in Arabic tweets with the presence of dialectical words. Sentiment analysis deals with extracting opinionated phrases from reviews, comments or tweets. i.e. to decide whether a given review or comment is positive, negative or neutral. Sentiment analysis has many applications and is very vital for many organizations. In this article, we utilize machine learning techniques to determine the polarity of tweets written in Arabic with the presence of dialects. Dialectical Arabic is abundantly present in social media and micro blogging channels. Dialectical Arabic presents challenges for topical classifications and for sentiment analysis. One example of such challenges is that stemming algorithms do not perform well with dialectical words. Another example is that dialectical Arabic uses an extended set of stopwords. In this research we introduce a framework that is capable of performing sentiment analysis on tweets written using either Modern Standard Arabic or Jordanian dialectical Arabic. The core of this framework is a dialect lexicon which maps dialectical words into their corresponding Modern Standard Arabic words. The experimentation reveals that the dialect lexicon improves the accuracies of the classifiers.