

Jordan University of Science and Technology

Detecting Sentiment Embedded in Arabic Social Media ? A Lexicon-based Approach

Authors: Rehab M Duwairi, Nizar Mahyoub, Saleh Alrefai,

Abstract: Sentiment analysis aims at extracting sentiment embedded mainly in text reviews. The prevalence of semantic web technologies has encouraged users of the web to become authors as well as readers. People write on a wide range of topics. These writings embed valuable information for organizations and industries. This paper introduces a novel framework for sentiment detection in Arabic tweets. The heart of this framework is a sentiment lexicon. This lexicon was built by translating the SentiStrength English sentiment lexicon into Arabic and afterwards the lexicon was expanded using Arabic thesauri. To assess the viability of the suggested framework, the authors have collected and manually annotated a set of 4400 Arabic tweets. These tweets were classified according to their sentiment into positive or negative tweets using the proposed framework. The results reveal that lexicons are helpful for sentiment detection. The overall results are encouraging and open venues for future research.