

Jordan University of Science and Technology

Sentiment Analysis for Arabizi Text

Authors: Rehab Duwairi, Mosab Alfaqeh

Abstract: This paper has used supervised learning to assign sentiment or polarity labels to tweets written in Arabizi. Arabizi is a form of writing Arabic text which relies on using Latin letters rather than Arabic letters. This form of writing is common with the Arab youth. A rule-based converter was designed and applied on the tweets to convert them from Arabizi to Arabic. Subsequently, the resultant tweets were annotated with their respective sentiment labels using crowdsourcing. This ArabiziDataset consists of 3206 tweets.