

Feature extraction for sentiment analysis on twitter data with spanish language

Victor Muniz^{1*}

1. Centro de investigación en Matemáticas. Monterrey, México.

*Contact author: victor.m@cimat.mx

Keywords: Twitter, kernel methods, feature extraction, normalization.

It is well known that social networks such as Twitter, has become an important source of information about tendencies and opinions of people related to a wide range of topics. The analysis of such data is challenging due to the characteristics of texts (tweets): short text, informal writing, misspelling, among many others. In this work, we show the advances of an ongoing project about polarity classification on twitter data for spanish language, focusing on the preprocessing and feature extraction, where we used kernel based methods in order to normalize tweets and obtain a similarity measure between them. We will show some results obtained by applying different classification methods.