

Deeper Emotional Capture and Extraction of Discussion Pathways in Twitter Data

K.V. Epaliyana, K.N.A. Jayatissa, A.J.H. Liyanage and K.G.L. Sandeepa

Department of Computer Science and Engineering, University of Moratuwa, Moratuwa, Sri Lanka

The rise of micro-blogging has resulted in people expressing their daily thoughts, often resulting in far more emotion-laden than might normally occur. Recently, it has become a trend to post about the places one has visited and experiences the person had. Finding the emotions expressed through these texts can be used in understanding community thinking and decision making, but it is quite complex since it requires thorough knowledge in psychology and linguistic. Furthermore, processing the microblog text is challenging, because they are informal and less consistent in terms of language. This paper presents a lexicon and rule based approach, which breaks the emotions behind a tweet into the eight basic categories a person is capable of expressing, as defined by Robert Plutchik. Hence, we broke the emotional tone behind a tweet into 8 basic emotions namely anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. In this, we selected a tweet corpus related to Sri Lankan tourism by querying a Twitter tool with suitable keywords often used by tweeters. We have augmented the accuracies of emotions capturing through a series of extensive text preprocessing steps fitting to twitter texts. The experimental results have shown that processing the informal tokens, hashtags and repeating character sequences can make a significant improvement in the emotion capturing accuracies. This research includes separating the tweets into cohesive discussion pathways using an existing incremental unsupervised machine learning algorithm, in order to come up with a rich decision making tool for the selected domain. No previous research has broken down emotions into these eight basic emotions and integrated the separation of discussion pathways into emotional capturing. We have contributed to the research world by succeeding in both these untapped research areas and developing a user friendly tool to use in constructive decision making.

Keywords: Sentiment Analysis, Emotion Detection, Topical Pathway Separation, Pre-processing, Twitter