Utilization of Text-based Emotions in social media for Depression Analysis.

C.K. Ranhoti Gedara* and C.R. Liyanage

Department of ICT, Faculty of Technology, University of Ruhuna
**Corresponding Author E-mail: chamodikavindaya@gmail.com, TP: +94752470481*

As estimated, millions of people around the world suffer from depression every year. It is a common and serious medical illness that negatively affects the way a person feels, thinks, and acts. Emotions can be used to identify depressed people through changes in their moods, expressions of thoughts, ideas, and opinions. With the expansion of online social media networks, many people share their thoughts and opinions as text-based posts and these are rich sources of human emotions. The current study has used text posts and comments which were published on public groups on Facebook related to depression as the labeled dataset for text analysis. Although previous studies have detected depression using different techniques, such as facial expressions analysis, behavioral analysis, and investigation of linguistic characteristics of written text, the results were not adequate for life-saving applications. Therefore, intending to solve with higher accuracy, this study investigated a novel emotional intensity-based approach to detect depressed people. During this study, depression is identified by analyzing emotional intensities in-text sources using a supervised learning approach. The existing NRC lexicon model which contains 8 basic emotions; anticipation, trust, joy, fear, sadness, surprise, anger, and disgust was used to get similar word lists with real-valued scores of weights for each word in eight basic emotions. After pre-processing the text, the techniques, such as developing a Bag of Words to collect all the words related to emotions, performing vectorization to extract words from posts that are similar to the Bag of Words, and finding intensities for each emotion using cosine similarity were conducted. Finally, the intensities of all 8 emotions of each post were fed into a Feed-Forward Neural Network to learn and predict the pattern of intensities to classify depression or non-depression according to the text posts. As the main output of the study, people with depression were identified, with 90% accuracy according to the patterns of emotional words in their posts on social media.

Keywords: Bag of Words; Cosine Similarity; Emotional Intensities; Feed Forward Neural Network; NRC Lexicon; Vectorization