Content Driven Tweets Annotation during Natural Disasters

Si Si Mar Win
sisimarwin@gmail.com
University of Computer Studies, Mandalay, UCSM

Than Nwe Aung
mdytna@gmail.com
University of Computer Studies, Mandalay, UCSM

Nowadays, Twitter, Social Networking Site, becomes most popular microblogging service and people have started publishing data on the use of it in natural disasters. Twitter has also created the opportunities for first responders to know the critical information and work effective reactions to impacted communities. This paper presents the automated annotation system that can detect the tweets which contain critical information or not. Annotation is done at tweet level with three labels by using the publicly available annotated datasets. LibLinear classifier is used to build a model for automatic tweets annotation. The annotation system also creates disaster related corpus with new tweets collected from Twitter API and annotated on real time manner. The performance of this model is evaluated based on different disaster related datasets and new Myanmar_Earthquake_2016 dataset derived from Twitter. The experiments show a high agreement rate between the annotation of this system and the annotators.