

Abstract

Right now, the circulation of information happened so fast in the social media space. However, not all information on the social media is true and valid. Twitter as one of the biggest social media also had the same problem as stated before. A lot of hashtags on twitter are coming and spreading from buzzer. In this paper, we are going to make a classification system that can differentiate between a hashtag that come from buzzer and a hashtag that come naturally on twitter. We will use random forest algorithm with community feature from previous study as the foundation. We also going to see what is the most influential attribute that affecting the classification process. In this study, the system that we created can distinguished the hashtag that appear from buzzer with 98% accuracy and 98% f-measure based on the dataset that we gather. We also get the most influential attribute in this classification are unique_user_pertweet and q1_rasio

Keywords: buzzer, random forest, hashtag, trending topic, twitter
