

ABSTRACT

Twitter is a microblog that is currently enthused by many people in many countries. The users can post everything whatever they want. Whether it's a breaking news, celeb gossip, and so on.

Everyday not less than 500 million tweets posted from users around the world. From those tweets, Twitter will list which topic that is currently being discussed in their Trending Topic feature. However, those trend topics actually refer to what words, or phrases that is frequently tweeted by users from various location about big event that is currently happening.

This research aims to classify tweets sent from Jakarta, Bandung, and Makassar into several topics, then observing the level of the popularity of those topics in various time.

This research started with data collection from twitter.com, followed with preprocessing data. Continued to forming a classifier using training data. Last step is to classify tweets into specific topics using Multinomial Naïve Bayes Classifier method. This research resulted an average *f1-score* up to 76,81%. Higher than Random Forest, SVM, and Nearest Neighbor Classifier (75,18%, 68,82%, and 50,22%).

Keywords: Topic classification, Twitter, Jakarta, Bandung, Makassar, Multinomial Naïve Bayes Classifier