

ABSTRAK

Twitter merupakan sebuah situs *microblogging* yang populer dibandingkan dengan situs *microblogging* lainnya. Twitter mampu mengirimkan pesan pendek 140 karakter. Isi dari pesan yang dikirim atau dinamakan *tweet* umumnya berisi laporan mengenai kejadian sehari-hari. Pada skripsi ini, penulis memfokuskan penelitian untuk mendeteksi adanya banjir melalui Twitter. Teknik yang dapat digunakan untuk memanfaatkan hal tersebut salah satunya dengan teknik *clustering*. *Clustering* dapat mengelompokkan *tweet* yang memiliki kemiripan isi ke dalam kelompok-kelompok. Metode pengelompokan *tweet* yang digunakan yaitu metode *Density-based Clustering* dengan Algoritma DBSCAN. Metode *Density-based Clustering* melakukan pengelompokan berdasarkan tingkat kepadatan dari suatu *tweet*. *Cluster* akan dipisahkan berdasarkan area yang memiliki kepadatan tinggi dengan area yang memiliki kepadatan rendah. Setiap *cluster* merepresentasikan satu *event*. Dari setiap *event* yang terdeteksi diambil informasi mengenai lokasi banjir dan deskripsi banjir tersebut menggunakan *Named Entity Recognition* (NER). Deteksi *event* menggunakan algoritma DBSCAN memberikan hasil yang baik, terbukti dengan nilai evaluasi *cluster* yang besar (0.86) dari data sebanyak 5354 *tweet* dengan jumlah *event* yang terdeteksi sebanyak 24 *event*.

Kata Kunci : *Event Detection, Clustering, DBSCAN, Twitter, Named Entity Recognition*

**EVENT DETECTION IN TWITTER MICROBLOGGING USING DBSCAN
ALGORITHM
(CASE STUDY: FLOOD)**

ABSTRACT

Twitter is the popular microblogging site's than others. Twitter allows user to send short messages contains 140 characters. The Twitter post, called tweets, typically contains about real-life events. In this paper, the authors focus on the study to detect flooding through Twitter. The technique can be used is clustering. Clustering can group tweets that have similar contents into groups. The authors uses Density-based Clustering with DBSCAN algorithm to group tweets. Density-based clustering methods perform clustering based on the density of a tweet. Clusters are separated by areas that have a high density area with a low density. Each cluster represents a single event. Each of detected event, extract information about the location and description of the flood using Named Entity Recognition (NER). Event detection using DBSCAN gives good results as evidenced by the large cluster evaluation value (0.86) of 5354 tweets with 24 number of events detected.

Keywords: *Event Detection, Clustering, DBSCAN, Twitter, Named Entity Recognition*