

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Adapun kesimpulan akhir dari penelitian *Event Detection* pada *Microblogging* Twitter dengan Algoritma DBSCAN:

1. Algoritma DBSCAN membantu mendeteksi adanya *event* dengan cara mengelompokkan sebuah data berdasarkan tingkat kepadatannya sehingga *cluster* akan dipisahkan dengan data-data dengan tingkat kepadatan yang rendah. DBSCAN membutuhkan dua parameter input untuk melakukan clustering, yaitu *eps* dan *MinPts*. Kedua parameter tersebut sangat berpengaruh terhadap hasil *cluster*. Sebuah *cluster* hasil dari DBSCAN ini merepresentasikan sebuah *event* yang ingin dideteksi sehingga satu *cluster* berarti satu *event*.
2. Pengambilan informasi *event* dapat dilakukan dengan teknik *Named Entity Recognition* (NER). Implementasi NER yang digunakan yaitu pemanfaatan *regular expression* (regex). Regex akan mengambil informasi dengan mencocokkan pola *string* tertentu dengan data hasil *cluster*. Jika terdapat kecocokan, maka dapat diambil informasi mengenai lokasi dan deskripsi dari *event* yang terdeteksi.

5.2. Saran

Untuk pengembangan lebih lanjut, saran-saran yang diberikan pada penelitian ini:

1. Perlu dilakukan adanya klasifikasi terlebih dahulu terhadap data *tweet* yang akan diproses pada tahap *clustering*. Klasifikasi bertujuan untuk memisahkan data yang memiliki makna banjir sebagai bencana dan makna banjir sebagai kiasan. Hal ini berguna untuk meminimalisasi *noise*.

2. Perlu dilakukan penelitian lebih lanjut dengan menggunakan algoritma lain dari metode *density-based clustering* dan membandingkan hasilnya dengan algoritma DBSCAN.
3. Perlu dilakukan penelitian lebih lanjut untuk mengambil informasi dari setiap *cluster* (ekstraksi informasi) sehingga data nama dan lokasi kejadian suatu *event* lebih presisi, misalnya dengan menggunakan teknik NER berdasarkan pembelajaran mesin (*machine learning*).
4. Perlu dilakukan penelitian lebih lanjut dalam penghitungan pembobotan TFIDF dan fungsi jarak selain menggunakan *Euclidean Distance*. Hal ini untuk mengantisipasi kelemahan DBSCAN yang tidak mampu melakukan *clustering* terhadap variasi perbedaan jarak yang besar antar objek data.

