# Bab 5 Kesimpulan dan Saran

## 5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan maka kesimpulan yang dapat ditarik adalah sebagai berikut:

1. Teknik pengumpulan bukti digital pada Jejaring Sosial Twitter dilakukan dengan cara, melakukan registrasi aplikasi pada link yang telah disediakan oleh Twitter untuk mendapatkan *Access Token*. Setelah akses token didapatkan kemudian buat *script* sebagai media *crawling* antara akun *Twitter* dan *API Search Twitter*, *Access Token* yang telah didapatkan di input pada *script* agar mendapat akses untuk melakukan *crawling* data, pembuatan *script* bisa menggunakan bahasa pemrograman apa saja seperti java, PHP dan lain-lain, namun untuk penelitian ini peneliti menggunakan Python dari Anaconda 4.0 untuk membuat *script* tersebut, selanjutnya melakukan *Crawling* data untuk mendapatkan *Log Twitter* dan di simpan dalam bentuk file *Json.*
2. Klasifikasi bukti digital *Cyberbullying* pada Jejaring Sosial Twitter menggunakan *Naïve Bayes Classifier* dilakukan dengan cara: pertama adalah mengumpulkan *Log tweet* dari database Twitter. Kedua, lakukan pembersihan data atau preprocessing data baik secara manual maupun otomatis dengan *Machine Learning* yang sesuai, dalam hal ini peneliti menggunakan WEKA. Ketiga, lakukan pembobotan term Frequency dan Invers Document Frequency (TF-IDF) untuk menghitung bobot masing-masing teks, kemudian lakukan validasi data, penelitian ini menggunakan *10 fold cross validation* dan terakhir adalah lakukan klasifikasi menggunakan Naïve Bayes. Adapun Hasil klasifikasi bukti digital *log tweet* cyberbullying adalah untuk teks yang positif bullying sebanyak 86.97%, negatif bullying 13.02%. untuk Jenis cyberbullying yang digunakan related *psychology* sebanyak 61.63%, related animals sebanyak 3.08%, general bullying sebanyak 19.57% dan sexuality sebanyak 3.08% dan negatif bullying 12.67%. hasil tersebut diatas untuk data periode November sampai desember 2016 dan dapat disimpulkan bahwa pelaku cyberbullying di Indonesia untuk periode tersebut cukup tinggi.

## 5.2 Saran

Berdasarkan hasil pengujian tersebut maka penulis memberikan saran sebagai berikut:

1. Hasil *crawling* data tidak semuanya berbahasa Indonesia, terdapat gabungan bahasa daerah dan bahasa asing, maka perlu pengembangan pada *script* agar dapat melakukan *filter* untuk *crawl* bahasa Indonesia saja.
2. Proses *preprocessing* menjadi lambat karena *Library stemming* yang digunakan masih dilakukan perbaikan pada hasil *stemming*, sehingga pada data yang lebih banyak butuh waktu yang lebih lama untuk mengubah data tersebut menjadi terstruktur, untuk penelitian selanjutnya Perlu dikembangkan kamus aplikasi untuk melakukan *stemming* dan *stopwords* khusus untuk analisis *cyberbullying*, *harassment* dan *hate speech* agar akurasi untuk klasifikasi lebih baik dan lebih cepat. Selain itu data yang digunakan diperbanyak.