

## BAB V

### KESIMPULAN DAN SARAN

#### 4.4 Kesimpulan

Kesimpulan dari penelitian Sistem Pemantau *Influenza Like Illness* dan Visualisasinya Memanfaatkan Twitter sebagai berikut.

1. Data twitter dapat digunakan untuk mendeteksi adanya *influenza like illness* di Indonesia. Namun tweet tidak dapat digunakan secara langsung karena data tweet banyak mengandung *noise*. Diperlukan proses-proses sebelum data disajikan dalam bentuk visualisasi. Dengan cara membagi klasifikasi dalam dua tahap, hasil eksperimen dengan menggunakan *improved* kNN mendapatkan akurasi yang lebih tinggi jika dibandingkan dengan satu tahap klasifikasi. Untuk klasifikasi dua tahap pada tahap pertama untuk mengklasifikasikan tweet sakit dan bukan sakit menghasilkan akurasi sebesar 92,706% dengan  $k=1$ . Hasil eksperimen tahap kedua untuk tweet flu dan bukan flu menghasilkan akurasi sebesar 91,847% dengan  $k=1$ . Sedangkan jika hanya menggunakan satu tahap klasifikasi, nilai akurasi sebesar 78,809% dengan  $k=10$  hasilnya sama dengan penelitian sebelumnya oleh Cullota (2010).
2. Visualisasi yang disajikan mendapatkan respon yang baik dari pengguna. Antar muka yang disajikan baik dari sisi grafis maupun pemilihan warna dinilai membantu pengguna untuk memantau penyebaran ILI. Beberapa responden menginginkan informasi dapat dipersempit dari setiap provinsi menjadi setiap kota dengan informasi tweet secara *realtime*. Secara keseluruhan, visualisasi dinilai baik dan dapat membantu untuk pemantauan ILI meskipun perlu dilakukan beberapa perbaikan sebelum visualisasi dapat digunakan secara luas.

#### 4.5 Saran

Saran-saran yang diberikan pada penelitian ini untuk pengembangan lebih lanjut adalah sebagai berikut.

1. Perlu dilakukan praproses teks yang lebih baik agar data yang digunakan lebih berkualitas dan meminimalkan *noise*, cara-cara yang dapat dilakukan seperti membuat kamus kata baku, membuat proses untuk mengatasi kesalahan penulisan dan menambah daftar *stopword* dan sinonim.
2. Pada tahap pelabelan tweet dapat ditambahkan analisis dari sisi tata bahasa dan kultur bahasa.
3. Perlu dilakukan penelitian untuk kasus yang sama dengan menggunakan algoritma yang berbeda, lalu membandingkan hasil akurasi dengan algoritma kNN.
4. Diperlukan data tweet yang lebih banyak dengan rentang waktu yang lama. Pencarian data lokasi *user* sangat sulit di Twitter, maka dapat dikombinasikan dengan menggunakan *social media* lain seperti *foursquare*.
5. Berdasarkan beberapa saran dari responden, visualisasi dapat disajikan secara *realtime* dan tidak hanya visualisasi sebatas provinsi namun juga setiap kota. Disarankan juga untuk menambah jenis visualisasi seperti penambahan grafik perbandingan ketika provinsi di klik.