

## BAB VI

### SIMPULAN DAN SARAN

#### 6.1 Simpulan

Kesimpulan yang diperoleh dari hasil analisis, pengambilan data, implementasi, dan pengujian aplikasi ini adalah:

1. Menggunakan metode *Naïve Bayes Classifier* dengan dasar nilai kemunculan kata tertentu, aplikasi dapat mendeteksi pekerjaan dari penulis *tweet*.
2. Isi dari *tweet* yang terdapat di dalam Twitter tidak memiliki struktur dan bahasa yang baku. (Lihat pada Bab V bagian 5.4 dan 5.10).
3. Pembentukan model masing-masing pekerjaan secara *Naive Bayes Multinomial* lebih stabil dibandingkan dengan pembentukan model dengan *Naïve Bayes Binomial*.
4. Metode pemecahan kata secara n-gram yang paling cocok digunakan untuk *tweet* pada media sosial Twitter adalah unigram. (Lihat pada Bab V bagian 5.8)
5. Kata-kata yang sering muncul di dalam *tweet* merupakan *stopwords*, sehingga jika proses pembentukan model menggunakan fitur pencarian dan penggantian *stopwords* dengan huruf 's', kata yang paling sering muncul adalah huruf 's' untuk masing-masing model pekerjaan (Lihat pada Bab V bagian 5.4).
6. *Naive Bayes Binomial* memberikan hasil pengujian kesalahan tipe satu lebih baik dibandingkan dengan metode *Naïve Bayes Multinomial*.

#### 6.2 Saran

Berikut ini adalah saran-saran yang bisa dipakai untuk mengembangkan aplikasi:

1. Lingkup pekerjaan yang diuji diperbanyak.
2. Mencari solusi untuk menghilangkan semua *emoticon*
3. Menambah *list stopwords*, seperti kemungkinan singkatan-singkatan (Bab 5.10)
4. Menggunakan fitur *stemming* untuk mencari kata-kata yang bermakna sama walaupun bentuk katanya berbeda (Bab 5.10)
5. Pembuatan model dibentuk secara *incremental*