

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan sebelumnya, dapat ditarik beberapa kesimpulan, antara lain:

1. Agar akun Twitter pengguna dapat diklasifikasi dan diprediksi jenis kepribadiannya, diperlukan tahapan – tahapan untuk mengolah data berupa teks dari akun tersebut. Tahapan itu ialah *tweet scraping* guna mengunduh data, *tweet preprocessing* guna mengolah data agar dapat diklasifikasi oleh sistem, dan prediksi oleh algoritma Naïve Bayes Classifier.
2. Pada pengujian menggunakan dataset yang tidak *balance*, klasifikasi menggunakan algoritma Naïve Bayes Classifier untuk menentukan jenis kepribadian berdasarkan metode *The Four Temperaments* mendapatkan nilai akurasi sebesar 64.4%, sedangkan untuk dataset yang *balance* menghasilkan nilai akurasi sebesar 61.1%
3. Pengujian dengan metode *K-Fold Cross Validation* pada dataset tidak *balance* mendapatkan nilai rata – rata akurasi sebesar 52.77%, sedangkan untuk dataset yang *balance* mendapatkan rata – rata akurasi 39.29%.

5.2 Saran

Peneliti menyadari masih banyak keterbatasan dan kekurangan dalam penelitian ini. Keterbatasan dan kekurangan ini dapat menjadi acuan bagi

penelitian berikutnya. Berikut adalah saran untuk penelitian berikutnya berdasarkan penelitian yang telah dilakukan:

1. Dalam penelitian ini, data yang diolah dan dianalisis oleh algoritma masih terbatas pada data berupa teks kalimat saja. Oleh karena itu dapat dikembangkan lagi dengan cara menambahkan parameter lain, seperti: *emoticon*, tanda baca, dan gambar.
2. Pada proses normalisasi kata atau penggantian kata *slang* menjadi kata baku, masih banyak kata – kata yang belum tercantum dalam kamus bahasa *slang* yang peneliti buat.
3. Dataset yang peneliti dapatkan pada saat pengumpulan data tidak seimbang antara kepribadian satu dengan yang lain. Sampel data dengan kepribadian koleris dan sanguinis masih terlalu sedikit dibandingkan dengan kepribadian melankolis dan phlegmatis.
4. Penelitian ini hanya menggunakan algoritma Naïve Bayes Classifier untuk melakukan klasifikasi data, sehingga masih terdapat peluang untuk meningkatkan akurasi dengan algoritma lainnya.