

BAB V

PENUTUP

5.1 Kesimpulan

Dari penelitian yang telah dilakukan. Berikut beberapa kesimpulan yang dapat diambil dari penelitian tersebut:

- a. Untuk menerapkan teknik pembobotan kata dengan model *Naive Bayes* dalam analisis sentimen, perlu dilakukan tahapan – tahapan sehingga data *tweet* dapat digunakan dalam pembobotan kata, dan klasifikasi. Tahapan tersebut adalah pengunduhan data, pemberian label, dan *preprocessing*. *Tweet* hasil *preprocessing* diolah agar dapat digunakan untuk pembobotan kata dan klasifikasi model *Naive Bayes Classifier* sehingga mampu menghasilkan nilai akurasi.
- b. Dari 4 skema yang dilakukan, skema TF-IDF dan TF-RF dengan *Bernoulli Naive Bayes* dari pengujian *Confusion Matrix* menghasilkan akurasi terbaik yaitu 61%, sedangkan rata - rata nilai akurasi dari validasi *5-fold* sebesar 60%.
- c. Dan hasil akurasi terendah terdapat pada kombinasi TF-IDF dan *Multinomial Naive Bayes* 58% dari hasil pengujian menggunakan *Confusion Matrix*, dan 59% dari hasil akurasi rata –rata validasi *5-fold*.

5.2 Saran

Peneliti menyadari masih terdapat kekurangan dan keterbatasan dalam penelitian ini. Kekurangan dan keterbatasan ini dapat menjadi acuan untuk penelitian berikutnya. Berikut saran untuk penelitian berikutnya berdasarkan penelitian yang telah dilakukan:

- a. Dalam penelitian ini, data yang digunakan peneliti masih berupa teks kalimat. Oleh karena itu dapat dikembangkan dengan menambahkan parameter lain, seperti: tanda baca, *emoticon*, dan gambar.
- b. Dataset yang peneliti dapatkan ketika pengumpulan data tidak seimbang antara yang berlabel positif, netral, dan negatif. Data yang berlabel negatif masih terlalu sedikit.
- c. Dalam penelitian ini tidak ada proses normalisasi atau penggantian kata slang menjadi kata baku, sehingga perlu dilakukan pengembangan.
- d. Dataset pada penelitian ini menggunakan dataset yang berhubungan dengan corona saja, sehingga dapat menggunakan data yang bersifat lebih umum, sehingga prediksi dan komparasi tidak hanya dapat digunakan 1 objek.
- e. Penelitian ini hanya membandingkan model algoritma jika menggunakan pembobotan kata. Untuk penelitian selanjutnya dapat membandingkan hasil akurasi model algoritma tanpa menggunakan pembobotan kata.