

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pembahasan yang telah diuraikan di atas maka dapat dibuat kesimpulan sebagai berikut:

1. Algoritma *Support Vector Machine* dapat diimplementasikan untuk mengklasifikasikan data Twitter berbahasa Indonesia.
2. Penelitian ini telah menguji kemampuan algoritma *Support Vector Machine* untuk klasifikasi data Tweet berbahasa Indonesia dengan total data 900 dan diperoleh rata – rata akurasi sebesar 71%.

5.2 Saran

Beberapa saran yang penulis berikan terkait penelitian yang dilakukan adalah sebagai berikut:

1. Melakukan survei untuk pelabelan data terhadap sumber yang dapat dipertanggung jawabkan dan terpercaya. Sehingga data yang didapatkan dapat dijadikan data latih yang baik.
2. Proses pembersihan tweet harus benar – benar bersih, hal ini bertujuan untuk proses pembuatan model dan prediksi lebih tepat.
3. Pada saat melakukan percobaan dengan tweet yang diunduh secara realtime terkadang model klasifikasi tidak dapat melakukan klasifikasi, hal ini dikarenakan media social seperti twitter memiliki banyak data yang masuk, dan data tersebut mungkin sama persis, hal ini membuat model klasifikasi tidak optimal.

4. Proses pengambilan data yang terkadang terlalu singkat membuat fitur pembobotan dan svm tidak optimal, untuk kedepannya bisa membuat parameter minimal panjang data tweet.
5. Melakukan perbandingan dengan algoritma lain seperti *K-NN* dan *Neural Network*.
6. Melakukan kombinasi fitur pembobotan seperti *Cosine Similarity*.

