

BAB V

KESIMPULAN

5.1 Kesimpulan

Penelitian ini melakukan pengujian model menggunakan *Support Vector Machine* (SVM). Pengujian dilakukan dengan data komentar yang telah dikumpulkan melalui proses pengunduhan komentar pada server Instagram. Setelah data terkumpul, untuk data *training* akan dilakukan proses pelabelan *spam* dan *ham* (*not spam*). Proses pelabelan ini dilakukan secara manual oleh peneliti. Data training ini nantinya akan diolah terlebih dahulu melalui serangkaian proses. Terdapat langkah *Preprocessing*, perhitungan TF-IDF, serta pelatihan. Adapun tahapan proses *Preprocessing* yang dilakukan yaitu komentar *cleansing*, *Case folding*, dan *Tokenizing*.

Setelah dilakukan *preprocessing*, akan dilakukan pembobotan TF-IDF. Setelah mendapatkan hasil TF-IDF bobot ini akan diklasifikasi dengan algoritma SVM kernel RBF. Akan dilakukan proses evaluasi terhadap hasil pelatihan maupun prediksi yang dilakukan oleh algoritma SVM. Adapun proses evaluasi dan pengujian yang dilakukan dalam penelitian ini didapatkan hasil dari ketiga skenario pengujian dengan ratio 70:30 dari 5000 data komentar akan dipecah menjadi data *training* dan data *testing*, data *training* sebanyak 3500 komentar dan data *testing* sebanyak 1500 komentar. *Accuracy score* yang diperoleh lebih tinggi yaitu 0.900%. Dengan demikian disimpulkan bahwa algoritma *Support Vector Machine* cocok digunakan dalam hal pengklasifikasian komentar *spam* pada Instagram.

5.2 Saran

Pada penelitian ini, terdapat beberapa keterbatas dan kekurangan. Kekurangan dan keterbatasan ini bisa dijadikan acuan dan pertimbangan untuk penelitian selanjutnya. Adapun saran yang dihasilkan dalam penelitian ini yaitu sebagai berikut:

1. Jumlah data *training* yang dipakai dalam proses pelatihan diusahakan bersih dari komentar duplikat atau komentar yang sama.
2. Sistem ini masih melakukan pengklasifikasian secara *offline*. Dengan kemudahan dari sisi pemrograman pada masa yang akan datang, diharapkan dapat mampu melakukan pengklasifikasian terhadap komentar pada waktu nyata (*realtime*).
3. Sistem pengklasifikasian ini diharapkan dapat dikembangkan lagi sehingga dapat dijalankan *cross-platform*.
4. Diharapkan menggunakan algoritma lain supaya hasil akan dapat dibandingkan untuk mendapatkan hasil yang terbaik.