

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian melakukan analisis sentimen terhadap kebijakan PPKM menggunakan algoritma *support vector machine* dan seleksi fitur *particle swarm optimization* dapat disimpulkan sebagai berikut.

1. Pada penelitian ini dalam mengklasifikasi sentimen positif dan negatif terhadap kebijakan PPKM menggunakan algoritma *support vector machine* dan seleksi fitur *particle swarm optimization*. Data yang digunakan, data *tweet* pengguna media sosial twitter sebanyak 886 data yang dikategori menjadi 2 kelas yakni kelas positif 418 data dan kelas negatif 468 data. Dalam membangun model dilakukan beberapa tahapan yakni pelabelan secara manual dengan bantuan 3 *annotator*, setelah data di berikan label kelas selanjutnya data dilakukan pembersihan *noise* melalui tahapan pra proses, lalu data yang telah bersih dari *noise* dilakukan pembobotan kata atau pemberian nilai pada setiap kata dengan metode TF-IDF, selanjutnya data hasil TF-IDF akan dilakukan seleksi fitur kemudian data dilakukan pembagian data menjadi 80% data latih dan 20% data uji menggunakan *traint_split*. Lalu diklasifikasi menggunakan *support vector machine*.
2. Hasil evaluasi klasifikasi terhadap kebijakan PPKM menggunakan algoritma SVM tanpa seleksi fitur PSO memperoleh akurasi 0.7977, *recall* 0.6904, dan *precision* 0.8529. Sedangkan pada algoritma menggunakan seleksi fitur PSO nilai akurasi, *recall*, dan *precision* meningkat masing masing 0.0731, 0.0779, dan 0.0874 sehingga hasilnya nilai menjadi akurasi 0.8708, *recall* 0.7683, dan *precision* 0.9403 Namun, waktu pengujian menggunakan seleksi fitur PSO lebih lama dibandingkan pengujian tanpa seleksi fitur PSO.

3. Berdasarkan hasil evaluasi klasifikasi dapat disimpulkan bahwa menggunakan seleksi fitur *particle swarm optimization* dapat meningkatkan nilai akurasi, *recall* dan *precision* terhadap model yang dibangun.

5.2 Saran

Pada penelitian ini dapat dilakukan beberapa saran untuk evaluasi penelitian selanjutnya, yaitu:

1. Menggunakan data dengan jumlah data yang lebih banyak.
2. Dalam pelabelan menggunakan para ahli atau pakar Bahasa Indonesia supaya data yang dilabelkan lebih tepat.
3. Menggunakan seleksi fitur lain seperti *Chi Square*, *Information Gain* atau yang lainnya untuk membandingkan hasil akurasi, *recall*, dan *precision*.