

BAB V

KESIMPULAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi sentimen ulasan pengguna menggunakan metode *support vector machine* pada aplikasi bibit dan bareksa, berikut kesimpulan dari penelitian ini:

1. Penerapan preprocessing pada dataset yang meliputi *case folding*, *cleansing data*, *normalization*, *stemming*, *stopword removal*, dan *tokenization*, pembobotan *tfidf*, dan seleksi fitur *chi-square* diterapkan sebelum masuk ke dalam proses pemodelan *support vector machine*. Dalam pencarian nilai parameter cost terbaik pada proses *support vector machine* digunakan *GridSearchCV*. Dengan menggunakan kernel linear, nilai cost terbaik yang dihasilkan adalah 0.5. Proses-proses tersebut dilakukan untuk menghasilkan performa pemodelan *support vector machine* terbaik.
2. Hasil pemodelan *support vector machine* menghasilkan nilai akurasi, presisi, *recall*, dan *f1-score* untuk mengukur seberapa baik performa yang dihasilkan. Pada aplikasi bibit tanpa menggunakan seleksi fitur *chi-square* menghasilkan nilai akurasi sebesar 91%, presisi 94%, *recall* 88%, dan *f1-score* 91%. Sedangkan dengan seleksi fitur *chi-square* menghasilkan nilai akurasi 92%, presisi 94%, *recall* 89%, dan *f1-score* 91%. Pada aplikasi bareksa tanpa menggunakan seleksi fitur *chi-square* menghasilkan akurasi 84%, presisi 83%, *recall* 85%, dan *f1-score* 84%. Sedangkan dengan seleksi fitur *chi-square* menghasilkan akurasi sebesar 86%, presisi 85%, *recall* 87%, dan *f1-score* 86%.
3. Penerapan *chi-square* menggunakan nilai alpha 0.9 yang menghasilkan jumlah fitur 1366 untuk aplikasi bibit dan 1139 untuk aplikasi bareksa. Nilai alpha ini dipilih berdasarkan hasil akurasi terbaik yang dihasilkan.
4. Penelitian ini membandingkan hasil analisis sentimen dengan dan tanpa seleksi fitur *chi-square*. Penggunaan seleksi fitur *chi-square* berhasil meningkatkan nilai akurasi pada dataset bibit dan bareksa. Pada dataset bibit menggunakan seleksi fitur *chi-square* nilai akurasi meningkat sebesar 1%. Sedangkan pada dataset bareksa, menggunakan seleksi fitur berhasil meningkatkan nilai akurasi sebesar 2%.

5. Dihasilkan sentimen analyzer berbasis website sederhana untuk mengklasifikasikan input ulasan komentar aplikasi bibit dan bareksa.

5.2 Saran

Berikut adalah saran yang dapat diberikan untuk penelitian selanjutnya guna menghasilkan output yang jauh lebih maksimal dari penelitian ini adalah sebagai berikut:

1. Menggunakan lebih banyak data untuk hasil analisis yang lebih optimal.
2. Mempertimbangkan penggunaan kernel rbf dan polynomial untuk pencarian parameter optimal dengan GridSearchCV.
3. Membandingkan klasifikasi *support vector machine* (SVM) dengan model pengklasifikasian lain, untuk membandingkan performa yang dihasilkan.