

BAB 5

KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa dalam proses pengumpulan ulasan aplikasi Halodoc dari Google Play Store menggunakan teknik web scraping, berhasil diperoleh data sebanyak 7.9945 pada 23 April 2024. Setelah melalui proses preprocessing, jumlah data yang telah dibersihkan sebanyak 7.433 dengan data positif sebanyak 5.474 dan data negatif 1.958, data yang telah disiapkan kemudian diboboti menggunakan metode TF-IDF sebelum masuk ke tahap klasifikasi. Hasil analisis menunjukkan bahwa model NB secara konsisten menunjukkan performa yang lebih baik daripada KNN dalam beberapa metrik evaluasi. Spesifikasinya adalah sebagai berikut:

1. Akurasi model NB mencapai tingkat akurasi sebesar 90%, sedangkan KNN hanya mencapai 88%.
2. *F1-score* hasil dari rata-rata dari *precision* dan *recall*. Pada penelitian ini NB memiliki nilai *f1-score* sebesar 0.78. Nilai 0.78 berada di antara 0 dan 1, yang berarti model NB memiliki performa yang lebih baik tetapi tidak sempurna. Ini menunjukkan bahwa model NB berhasil mengidentifikasi sebagai besar kasus positif dengan tingkat kesalahan yang relatif rendah, sedangkan nilai *F1-score* 0.73 pada KNN berarti model KNN memiliki keseimbangan yang cukup baik antara *precision* dan *recall*, tetapi performanya sedikit lebih rendah dibandingkan NB. Hal ini menunjukkan bahwa metode NB lebih baik dalam menjaga keseimbangan antara *precision* dan *recall* dibandingkan metode KNN.
3. *Recall* melalui *k-fold cross validation*, NB menunjukkan konsistensi dalam kinerjanya dengan tingkat *recall* yang stabil di sekitar 90%, sementara KNN memerlukan penyesuaian yang cermat terhadap nilai *k* untuk mencapai hasil yang optimal, dengan nilai *recall* sekitar 87%.

5.2 SARAN

Untuk meningkatkan kinerja model klasifikasi analisis sentimen, beberapa Langkah strategi perlu dipertimbangkan. Berikut adalah beberapa saran yang dapat membantu mencapai hasil yang lebih optimal :

1. Melakukan eksplorasi terhadap metode klasifikasi alternatif selain NB dan KNN.
2. Jika KNN tetap dipilih, makan penyesuaian parameter k perlu dilakukan dengan cermat untuk mencapai hasil yang optimal.
3. Analisis lebih lanjut untuk dapat memahami faktor-faktor yang mempengaruhi performa model, seperti jenis data dan pemilihan parameter.