

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan rangkaian analisis yang telah dilakukan, dapat disimpulkan bahwa distribusi keluhan pengguna aplikasi BCA Mobile tidak tersebar secara proporsional pada setiap kategori. Dari 3.534 ulasan yang telah melalui tahap pembersihan dan pelabelan, kategori Keamanan, Login, dan Transaksi tercatat sebagai kelompok dengan frekuensi keluhan tertinggi. Sebaliknya, kategori UI dan Lainnya memiliki jumlah data yang lebih sedikit. Ketimpangan jumlah tersebut menunjukkan adanya class imbalance yang turut memengaruhi performa model, khususnya dalam mengidentifikasi kategori dengan sampel terbatas. Kondisi ini menegaskan bahwa model cenderung bekerja lebih optimal pada kelas mayoritas dibandingkan pada kelas minoritas.

Dari aspek implementasi algoritma, kedua metode yang diterapkan mampu melakukan klasifikasi multi-kelas terhadap enam kategori keluhan dengan hasil yang tergolong baik. Namun demikian, hasil evaluasi menunjukkan bahwa Support Vector Machine memiliki performa yang lebih stabil dibandingkan Random Forest. SVM mencatat akurasi sebesar 79,35% dengan weighted average F1-score 0,79, sedangkan Random Forest memperoleh akurasi 75,11% dan weighted average F1-score 0,76. Selain unggul dari sisi nilai metrik, pola kesalahan pada confusion matrix juga memperlihatkan bahwa SVM memiliki pemisahan kelas yang lebih terpusat pada diagonal utama. Hal ini mengindikasikan bahwa SVM lebih adaptif dalam menangani representasi teks berbasis TF-IDF dengan dimensi fitur yang besar serta distribusi data yang tidak sepenuhnya seimbang.

Model SVM yang ditetapkan sebagai model terbaik kemudian direalisasikan dalam bentuk prototipe dashboard berbasis Streamlit. Implementasi ini membuktikan bahwa model tidak hanya berhenti pada tahap eksperimen, tetapi dapat diintegrasikan ke dalam sistem yang bersifat aplikatif. Dashboard tersebut menyajikan ringkasan distribusi keluhan, hasil evaluasi model, visualisasi tren berdasarkan periode waktu, serta fitur klasifikasi otomatis untuk keluhan baru.

Selain itu, sistem menyediakan rekomendasi tindak lanjut berbasis aturan sesuai kategori hasil prediksi. Dengan demikian, penelitian ini tidak hanya menghasilkan model dengan performa yang memadai, tetapi juga menunjukkan potensi pemanfaatannya sebagai alat bantu dalam memetakan pola keluhan dan menetapkan prioritas penanganan secara lebih terstruktur.

Secara keseluruhan, temuan penelitian ini menegaskan bahwa pendekatan klasifikasi berbasis *machine learning* efektif digunakan untuk menganalisis pola keluhan pengguna secara sistematis. Proses pengelompokan membantu mengidentifikasi permasalahan yang dominan sekaligus menyediakan dasar yang lebih objektif dalam menentukan prioritas respons. Walaupun masih terdapat keterbatasan, hasil penelitian menunjukkan bahwa penerapan klasifikasi teks pada layanan perbankan digital memiliki peluang pengembangan yang signifikan, terutama dalam mendukung pengambilan keputusan berbasis data dan peningkatan kualitas layanan secara berkelanjutan.

5.2 Saran

Berdasarkan temuan penelitian serta sejumlah keterbatasan yang telah dijelaskan sebelumnya, terdapat beberapa rekomendasi yang dapat menjadi arah pengembangan pada studi berikutnya. Salah satu hal yang perlu mendapat perhatian adalah persoalan ketidakseimbangan distribusi data pada tiap kategori. Perbedaan jumlah sampel antar kelas dalam penelitian ini terbukti berdampak pada performa model, terutama pada kategori dengan jumlah data yang relatif sedikit. Oleh sebab itu, penelitian selanjutnya dapat mempertimbangkan strategi yang lebih efektif dalam mengatasi kondisi tersebut, baik melalui penambahan jumlah data, penerapan teknik penyeimbangan dataset, maupun pendekatan lain yang memungkinkan model mengenali kelas minoritas secara lebih optimal.

Di samping itu, teknik representasi teks yang diterapkan masih berpotensi untuk disempurnakan. Meskipun metode TF-IDF menunjukkan hasil yang cukup baik, pendekatan ini belum sepenuhnya mampu menangkap makna kontekstual secara mendalam. Penelitian lanjutan dapat menguji metode alternatif seperti word embedding atau model berbasis *transformer* yang memiliki kemampuan lebih baik dalam memahami relasi antar kata dalam suatu kalimat. Selain itu, proses optimasi

parameter model juga dapat dilakukan secara lebih komprehensif guna memperoleh konfigurasi yang benar-benar sesuai dengan karakteristik dataset yang digunakan.

Dari perspektif implementasi, sistem yang dikembangkan dalam penelitian ini masih bersifat *prototype* dan belum terintegrasi secara langsung dengan sistem pengelolaan keluhan yang digunakan secara nyata. Pengembangan berikutnya dapat diarahkan pada perancangan sistem pemantauan yang berjalan secara real-time dan terhubung dengan basis data institusi. Selain itu, sumber data dapat diperluas dengan memasukkan ulasan dari *platform* lain seperti App Store, media sosial, maupun kanal layanan pelanggan *internal*. Dengan cakupan data yang lebih beragam serta integrasi sistem yang lebih komprehensif, hasil penelitian diharapkan mampu memberikan gambaran yang lebih luas dan meningkatkan kemampuan model dalam menangani variasi keluhan pengguna secara lebih akurat dan adaptif.

