BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan temuan dari penelitian yang sudah dilakukan tentang klasifikasi pengaduan masyarakat pada data Tanggap Karawang menggunakan algoritma *Random Forest* dan SVM, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penelitian ini berhasil membangun sistem klasifikasi aduan masyarakat berbasis teks dengan memanfaatkan algoritma *machine learning*. Proses dimulai dari tahapan preprocessing data teks, konversi fitur menggunakan TF-IDF, pelabelan data, hingga pelatihan dan pengujian model klasifikasi. Hal ini menjawab rumusan masalah pertama, yaitu bagaimana merancang model klasifikasi aduan berbasis teks untuk menentukan unit organisasi yang relevan.
2. Dua algoritma digunakan, yaitu *Random Forest* dan SVM. Keduanya diuji pada dataset yang telah dibagi menjadi data latih (80%) dan data uji (20%). Berdasarkan hasil evaluasi, model SVM menunjukkan performa memiliki keunggulan dibandingkan *Random Forest*. SVM mencapai akurasi sebesar 81,66%, *precision* sebesar 80,54%, *recall* sebesar 81,66%, dan *F1-score* sebesar 80,59%. Sedangkan *Random Forest* memperoleh akurasi sebesar 76,41%, *precision* sebesar 74,14%, *recall* sebesar 76,41%, dan *F1-score* sebesar 74,33%. Dengan demikian, rumusan masalah kedua yang berkaitan dengan performa algoritma telah terjawab, bahwa SVM lebih efektif pada kasus ini.
3. Dari hasil klasifikasi yang telah dilakukan, ditemukan bahwa unit organisasi yang paling sering dikenali oleh model adalah Dinas Pekerjaan Umum dan Penataan Ruang. Hal ini disebabkan oleh banyaknya laporan yang memuat kata-kata seperti "banjir", "jalan rusak", dan "saluran", yang secara semantik mengarah pada tugas pokok dari dinas tersebut. Temuan ini menjawab rumusan masalah ketiga, yaitu mengidentifikasi unit

organisasi yang paling sering menjadi tujuan aduan berdasarkan hasil prediksi model klasifikasi.

5.2 Saran

Penelitian ini masih memiliki ruang untuk dikembangkan lebih lanjut, terutama dalam aspek pemilihan algoritma dan pemrosesan data. Salah satu saran yang dapat diberikan adalah untuk melakukan eksplorasi terhadap algoritma lain seperti *XGBoost* dan *Naïve Bayes* pada penelitian selanjutnya. Algoritma *XGBoost* dikenal memiliki performa unggul dalam klasifikasi teks karena kemampuannya dalam menangani data yang bersifat sparse seperti hasil transformasi TF-IDF, serta fitur regularisasi internal yang mendukung kestabilan model. Sementara itu, algoritma *Naïve Bayes*, meskipun sederhana, seringkali memberikan hasil yang kompetitif dan efisien dalam waktu komputasi, serta sangat cocok digunakan sebagai *baseline* model dalam pemrosesan teks.

Selain itu, perlu juga diperhatikan masalah ketidakseimbangan kelas pada data aduan, di mana beberapa unit organisasi memiliki jumlah data yang secara signifikan lebih banyak dibandingkan yang lain. Hal ini dapat mengakibatkan model lebih condong ke kelas mayoritas. Untuk mengatasi hal tersebut, metode seperti SMOTE (*Synthetic Minority Over-sampling Technique*) atau penerapan *class weighting* dapat menjadi solusi agar performa model tetap adil dalam memprediksi kelas minoritas.

Pengembangan sistem klasifikasi ini ke dalam bentuk aplikasi atau antarmuka berbasis web juga menjadi prospek yang menjanjikan, karena dapat membantu pemerintah daerah dalam menangani laporan masyarakat secara cepat dan otomatis. Selain itu, penambahan fitur analisis semantik atau pemetaan berbasis topik akan memperkaya pemahaman model terhadap konteks isi aduan dan meningkatkan ketepatan klasifikasi.