



BAB V

Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat diambil beberapa kesimpulan sebagai berikut:

1. Metode K-Nearest Neighbors (KNN) dapat diterapkan untuk mengklasifikasikan penyakit pasien rawat jalan berdasarkan data gejala yang dimasukkan ke dalam sistem. Proses klasifikasi dilakukan melalui tahapan reprocessing data, pembentukan fitur menggunakan TF-IDF, perhitungan kemiripan menggunakan cosine distance, hingga penentuan kelas berdasarkan tetangga terdekat. Hasil penelitian menunjukkan bahwa metode KNN mampu mengenali pola kemiripan gejala pasien dengan menghasilkan rekomendasi diagnosis penyakit secara otomatis.
2. Hasil pengujian menunjukkan bahwa model terbaik diperoleh pada nilai $K=7$ dengan nilai Accuracy sebesar 94.67%, Macro F1 sebesar 94.54% dan Top-5 Accuracy sebesar 98%. Nilai tersebut menunjukkan bahwa model mampu melakukan klasifikasi penyakit dengan cukup baik. Tingginya nilai Top-5 Accuracy menunjukkan bahwa diagnosis yang hamper selalu muncul dalam lima rekomendasi teratas yang diberikan sistem
3. Tahapan preprosesing seperti pembersihan data, normalisasi istilah gejala, transformasi data, dan pembentukan fitur lama sakit terbukti membantu

meningkatkan kualitas data yang digunakan dalam proses klasifikasi. Selain itu, penerapan teknik random undersampling mampu mengurangi dominasi kelas mayoritas yang terdapat pada dataset.

6.2 Saran

Berdasarkan hasil penelitian, penulis memberikan saran sebagai berikut:

1. Penelitian selanjutnya disarankan menggunakan dataset yang lebih besar dan lebih beragam agar model dapat mempelajari pola penyakit dengan lebih baik. Semakin banyak data yang digunakan, maka kemampuan model dalam melakukan klasifikasi diharapkan menjadi lebih akurat.
2. Penelitian ini hanya menggunakan fitur gejala pasien lama sakit sebagai fitur utama. Pada penelitian berikutnya dapat ditambahkan atribut lain seperti usia, suhu tubuh, tekanan darah, riwayat penyakit, hasil pemeriksaan laboratorium, dan faktor kesehatan lainnya untuk meningkatkan performa model.

