BAB V

PENUTUP

5.1 KESIMPULAN

Berdasarkan hasil penelitian serta analisis yang telah dilakukan, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Penerapan algoritma *K-Nearest Neighbor* (KNN) berhasil pada klasifikasi tingkat obesitas dengan hasil akurasi tinggi. Melalui pengujian menggunakan metode *Grid Search* dengan *Cross Validation*, diperoleh nilai K terbaik = 1, dengan akurasi 100% pada data latih dan 97,33% pada data uji. Hal tersebut menunjukkan bahwa model mampu mengenali pola data dengan baik, tidak mengalami *overfitting* atau *underfitting*, serta efektif digunakan dalam klasifikasi obesitas.
2. Data yang diambil berdasarkan hasil dari pembersihan data yang berjumlah 3.182 data dengan 6 fitur utama (*gender, age, weight, height, SMOKE, MTRANS*) dan *Obesity_Level* sebagai label, memperoleh beberapa variabel yang merupakan faktor yang paling berpengaruh terhadap tingkat obesitas. Variabel tersebut memiliki korelasi yang kuat terhadap label obesitas sehinnga berperan penting dalam klasifikasi.
3. Implementasi KNN dilakukan melalui tahapan pra-pemrosesan data yang mencakup pembersihan data, *encoding*, normalisasi, seleksi fitur, serta penerapan SMOTE untuk mengatasi ketidakseimbangan data antar kelas. Hasil pengujian tersebut menunjukkan bahwa KNN dapat diimplementasikan

dengan baik dalam mengklasifikasikan tingkat obesitas berdasarkan data pasien dengan performa yang akurat serta seimbang pada semua kelas obesitas.

5.2 SARAN

Untuk pengembangan lebih lanjut, peneliti memberikan beberapa saran berikut:

1. Untuk meningkatkan akurasi model disarankan agar mencoba algoritma klasifikasi lain seperti *Random Forest*, *Support Vector Machine* (SVM), atau *Deep Learning* untuk mengetahui adanya metode lain yang mampu memberikan hasil lebih optimal dibandingkan dengan KNN.
2. Melakukan eksplorasi parameter lebih lanjut menggunakan beberapa variasi baik untuk nilai K ataupun dengan penggunaan metrik jarak lain (seperti *Manhattan* atau *Minkowski*) agar model lebih optimal.
3. Menambahkan fitur pendukung yang relevan seperti pola makan, Riwayat kesehatan keluarga, atau tingkat stress, yang bertujuan agar klalsifikasi obesitas lebih komprehensif.
4. Mengembangkan teknik *balancing data* selain SMOTE, misalnya ADASYN atau kombinasi *oversampling-undersampling*, dengan tujuan dapat mengurangi bias pada distribusi kelas.
5. Menerapkan sistem nyata yang dapat dikembangkan menjadi aplikasi berbasis web atau *mobile* sehingga dapat membantu pengguna memantau status obesitas serta gaya hidup sehari-hari secara langsung.