

BAB V

KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, berikut adalah kesimpulan dari klasifikasi multikelas faktor risiko obesitas pada individu menggunakan algoritma *Machine Learning* :

1. Penelitian ini berhasil mengembangkan model klasifikasi multikelas faktor risiko obesitas pada individu menggunakan tiga algoritma *Machine Learning*, yaitu *K-Nearest Neighbor* (KNN), *Random Forest*, dan C4.5. Dataset yang digunakan terdiri dari 2.086 data dengan 15 atribut dan 1 target, yang telah melalui proses *preprocessing* sebelumnya seperti normalisasi dan encoding untuk memastikan kualitas data.
2. Hasil evaluasi menunjukkan bahwa algoritma *Random Forest* memiliki performa terbaik dengan akurasi sebesar 95,45%, diikuti oleh C4.5 dengan akurasi sebesar 94,02%, dan *K-Nearest Neighbor* (KNN) dengan akurasi sebesar 80,86%. Algoritma *Random Forest* dan C4.5 lebih disarankan karena menunjukkan akurasi dan stabilitas yang lebih tinggi dibandingkan dengan KNN.
3. *Random Forest* adalah pilihan terbaik untuk dataset ini karena kemampuannya menangani data yang kompleks, ketahanannya terhadap *noise*, dan kemampuan generalisasi yang tinggi. Sementara itu, KNN memiliki keterbatasan dalam menangani dimensi data yang tinggi dan

sensitivitas terhadap *outlier*, sedangkan C4.5 rentan terhadap *overfitting* dan kurang optimal dalam menangani data numerik. Oleh karena itu, *Random Forest* direkomendasikan sebagai algoritma utama untuk klasifikasi faktor risiko obesitas pada dataset ini.

4. Hasil penelitian menunjukkan bahwa faktor-faktor yang berpengaruh signifikan terhadap tingkat obesitas meliputi berat badan, tinggi badan, frekuensi konsumsi sayur, frekuensi aktivitas fisik, dan konsumsi air adalah faktor dominan dalam meningkatkan faktor risiko obesitas.
5. Penelitian ini memberikan kontribusi penting dalam pengembangan model prediksi faktor risiko obesitas pada individu. Model yang dihasilkan dapat dimanfaatkan dalam aplikasi atau sistem prediktif untuk mendukung upaya pencegahan dan penanganan obesitas secara lebih efektif.

5.2 SARAN

Penulis menyadari bahwa penelitian ini memiliki beberapa keterbatasan dalam proses klasifikasi multikelas faktor risiko obesitas pada individu menggunakan algoritma *Machine Learning* (KNN, *Random Forest*, dan C4.5). oleh karena itu, saran untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan representatif, mencakup populasi dengan berbagai latar belakang demografis, gaya hidup, dan kondisi kesehatan untuk meningkatkan generalisasi model.
2. Penelitian di masa depan disarankan untuk menguji algoritma *Machine*

Learning lain, seperti *Support Vector Machine (SVM)*, *Gradient Boosting*, atau *XGBoost*, untuk dibandingkan dengan algoritma yang telah digunakan dalam penelitian ini, yaitu *K-Nearest Neighbor (KNN)*, *Random Forest*, dan *C4.5*, guna mengetahui algoritma yang paling efektif dalam klasifikasi tingkat obesitas.

3. Penelitian mendatang dapat memperluas ruang lingkup dengan menambahkan variabel tambahan seperti pola tidur, tingkat stres, atau faktor psikologis yang berpengaruh terhadap obesitas, untuk menghasilkan analisis yang lebih komprehensif.
4. Penelitian berikutnya dapat mencakup pengembangan metode pengolahan data yang lebih kompleks, seperti penerapan teknik *feature engineering* atau penggunaan *deep learning* untuk mendeteksi hubungan non-linear yang lebih mendalam antar variable.
5. Lembaga kesehatan atau organisasi terkait disarankan untuk memanfaatkan model yang dikembangkan dalam penelitian ini dengan mengintegrasikannya ke dalam aplikasi atau sistem prediksi risiko obesitas berbasis teknologi, untuk mendukung pencegahan dan penanganan obesitas secara lebih efektif dan efisien.