BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian terkait penerapan algoritma *Naïve Bayes* dan *KNN* dalam klasifikasi Status Gizi pada anak balita di Puskesmas Cimahi Selatan, dapat diambil beberapa kesimpulan sebagai berikut :

1. Penelitian ini memanfaatkan data Status Gizi anak balita yang didapat dari Puskesmas Cimahi Selatan yang terdiri dari 6 atribut utama yaitu, Jenis Kelamin, Usia, Berat Badan, Tinggi Badan, dan ZS BB/TB, untuk mengklasifikasikan Status Gizi. Sebelum proses klasifikasi kedua algoritma tersebut melalui beberapa tahapan proses yaitu, prapemrosesan data, transformasi data, pemilhan fitur, *balancing* data, pembagian data menjadi data latih dan data uji, serta pelatihan dan pengujian model sebagai tahap terakhir.
2. Hasil pengujian memperlihatkan bahwa algoritma *KNN* memiliki nilai akurasi yang lebih tinggi dibandingkan dengan algoritma *Naïve Bayes*. *KNN* mampu meraih nilai akurasi sebesar 99.70%, sedangkan *Naïve Bayes* hanya memperoleh nilai akurasi 96.18%. Meskipun hanya berbeda sedikit namun perbedaan nilai akurasi disebabkan oleh perbedaan cara kerja kedua algoritma. Algoritma *Naïve Bayes* bekerja dengan asumsi bahwa setiap fitur dalam dataset independen, sehingga kurang optimal dalam mengidentifikasi hubungan antar fitur yang mungkin saling berkaitan. Di sisi lain, algoritma *KNN* tidak bergantung pada asumsi tersebut, melainkan melakukan klasifikasi berdasarkan kemiripan antar data, sehingga lebih fleksibel dalam mengenali pola-pola kompleks yang ada di dalam data. Dengan pendekatan tersebut, *KNN* menunjukkan kinerja yang lebih baik dalam mengklasifikasikan Status Gizi pada anak balita di Puskesmas Cimahi Selatan.

5.2 Saran

Berdasarkan hasil penelitian ini, terdapat beberapa saran untuk penelitian selanjutnya.

1. Meskipun *KNN* memberikan hasil yang lebih baik, disarankan untuk mengeksplorasi algoritma lain seperti *SVM* atau *Random Forest* guna membandingkan kinerjanya.

2. Teknik prapemrosesan data seperti pada proses normalisasi dapat ditingkatkan untuk memperoleh hasil dari model yang dibuat agar lebih optimal. Selain itu, pengujian menggunakan dataset yang lebih besar dan bervariasi juga sangat penting untuk memastikan kemampuan generalisasi model untuk mendapatkan hasil yang baik.