

## BAB V PENUTUP

### 5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan tentang analisis klasifikasi dari dataset diabetes menggunakan metode K-Nearest Neighbor, dapat dikemukakan kesimpulan yaitu untuk menentukan tingkat akurasi tertinggi dari metode K-Nearest Neighbor berdasarkan dengan N-neighbor yang digunakan yaitu 1,3,5,7 dan 9 dimana sumber dari dataset diabetes yang berjumlah 768 baris dan 9 kolom dimana variabel independen (x) yaitu pada kolom *Pregnancies*, *Glucose*, *BMI* dan *Age* dengan variabel dependen (y) adalah kolom *Outcome*. Dataset dataset dibagi ke dalam training dataset dan test dataset dengan perbandingan 70:30 yang memiliki arti bahwa 70% digunakan untuk training dan 30% untuk proses testing. Kinerja dari metode yang digunakan yaitu KNN dievaluasi dengan confusion matrix, berikut kesimpulan yang didapat :

1. Hasil pengujian metode K-Nearest Neighbor dengan N-neighbor yang digunakan yaitu 3,5,7 dan 9 pada dataset diabetes diperoleh nilai performa terbaik pada  $n=3$  dengan tingkat akurasi sebesar 75%. Hal ini menjelaskan bahwa keakuratan model dalam mengklasifikasikan dengan benar sebesar 75%.
2. Hasil uji performa KNN yang diukur meliputi akurasi, presisi, recall dan f-1 score pada class 1 mengalami kenaikan cukup signifikan setelah dilakukan balancing data.
3. Nilai akurasi dari setiap pengujian mengalami kenaikan, tetapi kenaikan kurang signifikan ketika dataset diabetes dilakukan balancing data.

## 5.2 Saran

Dari kesimpulan yang telah didapatkan oleh penulis dari hasil percobaan yang dilakukan saat penelitian ini, penulis memiliki beberapa saran untuk penelitian lanjutan dengan topik yang sama, yaitu sebagai berikut:

- a) Diharapkan untuk penelitian selanjutnya perlu dilakukan perbandingan terhadap algoritma lain, untuk menguji sejauh mana akurasi algoritma KNN dapat diandalkan.
- b) Mencari jumlah dataset yang lebih besar atau menggunakan data yang sudah seimbang pada penelitian selanjutnya agar dapat menghitung nilai performa yang baru.

