

BAB 5

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang dilakukan dengan menerapkan algoritma SVM, *bagging*, *borderline-SMOTE* dan *grid search* untuk melakukan klasifikasi penyakit diabetes, maka dapat diambil kesimpulan sebagai berikut:

1. Penerapan algoritma SVM, *bagging*, *borderline-SMOTE* dan *grid search* untuk membentuk model klasifikasi penyakit diabetes dengan menggunakan *training data* yang berasal dari *dataset Pima Indians*. Skenario pembentukan model yang terjadi terdiri dari model *support vector machine*, model *support vector machine* dengan *bagging*, model *support vector machine* dan *bagging* dengan *training data* yang telah di-*oversampling* oleh *borderline-SMOTE* dan model *support vector machine* dengan *bagging* dan *borderline-SMOTE* yang telah dioptimasi oleh *grid search*. Kemudian model dari tiap skenario diuji dengan menggunakan *testing data* yang berasal dari *dataset Pima Indians* dan *dataset Frankfurt*.
2. Performa pengujian model pada skenario 4 yang terbentuk dari kolaborasi antara algoritma SVM, *bagging*, *borderline-SMOTE* dan *grid search* merupakan model klasifikasi yang memiliki generalisasi paling baik daripada model pada skenario lainnya. Hasil dari evaluasi terhadap *testing data* mendapat nilai akurasi sebesar 80,8%, nilai *precision* sebesar 81,96% untuk kelas Sehat dan 79,68% untuk kelas Diabetes, nilai *recall* sebesar 79,36% untuk kelas Sehat dan 82,25% untuk kelas Diabetes, dan nilai *f1-score* sebesar 80,64% untuk kelas Sehat dan 80,95% untuk kelas Diabetes. Hasil tersebut memang lebih rendah daripada yang didapat oleh model skenario 3, tetapi model skenario 3 mengalami penurunan performa pada saat melakukan klasifikasi *dataset Frankfurt*. Sedangkan hasil evaluasi model skenario 4 terhadap *dataset Frankfurt* mengalami kenaikan daripada hasil evaluasi terhadap *testing data*. Hasil evaluasi

model skenario 4 terhadap *dataset Frankfurt* mendapat akurasi sebesar 92,1%, nilai *precision* sebesar 95,51% untuk kelas Sehat dan 86,12% untuk kelas Diabetes, nilai *recall* sebesar 92,32% untuk kelas Sehat dan 91,66% untuk kelas Diabetes, dan nilai *f1-score* sebesar 93,39% untuk kelas Sehat dan 88,81% untuk kelas Diabetes.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat penulis berikan untuk pengembangan penelitian ini, yaitu:

1. Diharapkan pada penelitian selanjutnya dapat menggunakan metode optimasi *hyperparameter* selain *grid search*, karena optimasi dengan *grid search* memakan waktu yang cukup lama.
2. Penelitian selanjutnya diharapkan dapat melakukan penanganan terhadap data *outlier* yang dimiliki oleh *dataset* diabetes.