

BAB V

PENUTUP

5.1 KESIMPULAN

Berdasarkan pengujian yang telah dilakukan peneliti memperoleh kesimpulan bahwa:

1. Pada saat *preprocessing* data didapatkan kesimpulan sebagai berikut: *Feature Importances* berhasil merankingkan variabel prediktor yang berkorelasi tinggi dengan variabel target dari yang teratas yaitu *time*, *serum_creatinine*, *ejection fraction*, *age*, *creatinine_phosphokinase*, *platelets*, *serum_sodium*, *sex*, *anaemia*, *smoking*, diabetes, dan terakhir *high_blood_pressure*. Setelah dilakukan *Min Max Normalization*, nilai min dan max setiap fitur menjadi sama yaitu 0 dan 1. Setiap skema yang dilakukan ADASYN memiliki penambahan jumlah data yang berbeda-beda dikarenakan jumlah fitur yang digunakan pada setiap skema berbeda-beda.
2. Berdasarkan hasil pengujian, nilai akurasi data *testing* tertinggi pada penelitian ini sebesar 97.6% dengan nilai 95,8% untuk *precision*, 100% untuk *recall*, dan 97,9% untuk *f1-score*. Nilai ini diperoleh dari *fold* 10 pada skema 6 fitur dan skema 7 fitur. Serta nilai akurasi data *training* keduanya sebesar 100% dengan nilai 100% untuk *precision*, 100% untuk *recall*, dan 100% untuk *f1-score*. Hasil akurasi yang diperoleh menampilkan kompetensi *Random Forest* untuk mengklasifikasi pasien yang meninggal akibat gagal jantung dan selamat dari gagal jantung berdasarkan data *history* pasien gagal jantung.

5.2 SARAN

1. Peneliti dapat membandingkan *Min Max Normalization* dengan metode *feature Scaling* lainnya untuk mengetahui metode mana yang baik untuk mengurangi penyimpangan data.
2. Peneliti dapat membandingkan ADASYN dengan SMOTE untuk mengetahui metode mana yang baik untuk menyeimbangkan dataset.

3. Peneliti dapat menambah model lainnya selain *Random Forest* agar mendapatkan berbagai macam *insight* mengenai model mana yang menghasilkan nilai akurasi terbaik dari dataset *Heart failure prediction*.