



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian yang dilakukan dalam penanganan *imbalance data* untuk klasifikasi penyakit jantung dengan menggunakan metode SMOTE pada algoritma *naïve bayes*. Penerapan klasifikasi pada *data mining* didasarkan pada penggunaan *knowledge discovery in database* atau KDD sebagai framework penelitian. Tahapan pada KDD dimulai dari *data selection*, *data pre-processing*, *data transformation*, *data mining*, dan *evaluation/interpretation*. Dimana, penerapannya adalah untuk melihat hasil kinerja algoritma *naïve bayes* menggunakan metode optimasi SMOTE dalam menangani data yang tidak seimbang, sehingga hasil penelitian menunjukkan bahwa SMOTE berhasil dalam memberikan optimasi pada kinerja *naïve bayes*.

Dalam melakukan pembentukan model klasifikasi penyakit jantung, proses *split data* untuk *data training* dan *data testing*, dibandingkan 3 proporsi rasio *split data* mulai dari 70:30, 60:40, dan 50:50, sebelum dan sesudah menggunakan SMOTE. Ketiga rasio kemudian dibandingkan kinerjanya dengan melihat peningkatan yang terjadi pada kinerja evaluasi metrik, dimana rasio 50:50 menunjukkan nilai yang konsisten dan paling optimal dalam peningkatan kinerja metrik presisi, *recall*, dan AUC baik sebelum penggunaan SMOTE maupun sesudah.

Berdasarkan hasil penerapannya, nilai kinerja rasio *split data* 50:50 pada *data training* sebelum SMOTE pada rasio 50:50 memiliki nilai akurasi 86.42%, presisi 29.93%, *recall* 43.66%, dan AUC 0.820. Setelah menggunakan metode SMOTE, akurasi menurun menjadi 73.05%, namun presisi meningkat menjadi 77.72%, *recall* meningkat menjadi 64.61%, dan AUC meningkat menjadi 0.831. Dalam pengujiannya *data testing* memiliki nilai yang hampir sama dan tidak terdapat perbedaan yang signifikan antara *data training* dan *data testing*, sehingga hasil pengujian ini masuk kedalam kategori *good fit*. Dalam konteks ini, peningkatan pada metrik spesifik ini bisa dianggap sebagai peningkatan performa yang

diinginkan, sehingga SMOTE terbilang berhasil dalam penanganan *imbalance data*.

5.2 Saran

Dari kesimpulan yang telah didapatkan atas hasil penelitian, maka beberapa saran untuk penelitian selanjutnya, diantaranya:

1. Pembuatan sistem informasi atas model klasifikasi penyakit jantung akan memudahkan para tenaga medis, yang kemudian ide ini sangat baik untuk dikembangkan.
2. Penelitian selanjutnya diharapkan dapat menggunakan metode untuk penanganan *imbalance data* lain, seperti *adaptive booster* atau *random over sampling* yang nantinya dapat dilakukan sebagai pembanding metode optimasi yang lebih baik.

