

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang sudah dilakukan, didapatkan hasil sebagai berikut:

1. Tidak terdapat peningkatan yang signifikan pada akurasi kedua algoritma. Akurasi Random Forest sebelum dilakukan optimasi sebesar 88%-90%, sesudah dilakukan optimasi tetap sama yaitu sebesar 88%-90%. Sedangkan untuk XGBoost sebelum dilakukan optimasi sebesar 88%-89%, sesudah dilakukan optimasi terjadi penyesuaian sebesar 87%-90%.
2. Nilai presisi, recall, dan f1-score untuk kelas 0 tetap tinggi untuk kedua algoritma. Dengan besaran nilai presisi berkisar 92%-96%, recall sebesar 90%-97%, dan f1-score sebesar 92%-94% baik sebelum maupun sesudah dilakukan optimasi.
3. Sama halnya dengan kelas 0, nilai presisi, recall, dan f1-score tidak terjadi peningkatan. Dengan besaran nilai untuk presisi berkisar 57%-74%, recall berkisar 51%-78%, dan f1-score berkisar 60%-67% untuk kedua algoritma sebelum maupun sesudah dilakukan optimasi.
4. Peningkatan recall pada kelas 1 terjadi saat menggunakan SMOTE. Teknik *feature selection* RFECV dan *embedded method* juga tidak memberikan perbedaan signifikan pada akurasi model.
5. Secara keseluruhan, Random Forest lebih unggul dibandingkan XGBoost untuk dataset ini, terkhusus pada prediksi kelas 1 dengan nilai presisi dan f1-score yang lebih tinggi pada semua skenario.

Dari penjabaran di atas, dapat disimpulkan bahwa penggunaan teknik optimasi *hyperparameter RandomizedSearchCV* pada Random Forest dan XGBoost tidak memberikan peningkatan performa yang signifikan. Hipotesis awal bahwa optimasi akan meningkatkan performa model tidak terbukti untuk penelitian ini. Hal ini menunjukkan bahwa kedua algoritma sudah cukup baik dalam melakukan klasifikasi meski diterapkan pada *imbalanced dataset*.

5.2 Saran

Perbedaan nilai metrik yang dihasilkan setelah optimasi pada penelitian ini sangat kecil, sehingga untuk penelitian selanjutnya lebih disarankan untuk mencari pendekatan lain untuk menangani ketidakseimbangan data yang ada pada dataset ini. Pertimbangkan juga penggunaan deep learning maupun model transformer, guna mencari metode yang paling bagus untuk diaplikasikan pada dataset ini.

