

BAB V PENUTUP

5.1. Kesimpulan

Berdasarkan serangkaian eksperimen dan analisis yang telah dilakukan, beberapa kesimpulan utama yang dapat diambil adalah sebagai berikut:

1. Optimalisasi Model CatBoost dengan Optuna

Optimalisasi model CatBoost untuk memprediksi terapi regimen pengobatan pada pasien tuberkulosis menunjukkan peningkatan kinerja yang signifikan setelah diintegrasikan dengan Optuna. Optimalisasi model dilakukan melalui 300 *trials* menggunakan *Optuna* dengan pendekatan *Tree-Structured Parzen Estimator* (TPE). Proses ini difokuskan pada *tuning* beberapa *hyperparameter* kunci yang memiliki pengaruh signifikan terhadap performa model.

2. Hasil Pengujian Model

- Pembagian data terbaik dalam penelitian ini diperoleh melalui metode *train-test split* dengan rasio 90:10, yang menghasilkan performa optimal pada model setelah dilakukan proses optimasi. Pada dataset yang telah diseimbangkan, model menunjukkan peningkatan akurasi yang signifikan, di mana akurasi tertinggi mencapai 96% pasca optimasi, dibandingkan dengan hanya 90% sebelum dilakukan penyetelan parameter. Peningkatan ini menegaskan efektivitas strategi optimasi dalam memperkuat kapabilitas model dalam mengklasifikasikan data secara lebih tepat dan konsisten.
- Penerapan metode *K-Fold Cross Validation* dengan nilai *K* sebesar 5 menunjukkan bahwa model CatBoost yang telah melalui proses optimasi menggunakan *Optuna* memberikan performa yang cukup baik. Evaluasi dilakukan menggunakan skema pembagian data sebesar 90:10, di mana 90% digunakan untuk pelatihan dan 10% untuk pengujian pada setiap *fold*. Hasil yang diperoleh mencerminkan kinerja model yang konsisten dan andal, dengan nilai rata-rata *F1 Score* sebesar 0.9339, *accuracy* mencapai 0.9340, serta *precision* dan *recall* yang masing-masing berada di angka 0.9353 dan 0.9340. Temuan ini mengindikasikan bahwa *tuning* parameter menggunakan pendekatan berbasis *Optuna* mampu meningkatkan presisi dan sensitivitas model secara seimbang, serta memperkuat kemampuan generalisasi terhadap data baru.

3. Pengaruh Dataset Seimbang

Penyeimbangan data menggunakan SMOTE terbukti meningkatkan performa model prediksi regimen obat anti-tuberkulosis. Sebelumnya, data didominasi oleh kelas $2(HRZ)/4(HR)$ dengan 303 sampel, sementara dua kelas lainnya hanya 233 dan 84 sampel, menghasilkan *imbalance ratio* hingga 3,6. Setelah SMOTE, seluruh kelas memiliki 303 sampel, menciptakan distribusi yang seimbang. Hasilnya, model bekerja lebih akurat dan konsisten. Temuan ini menegaskan pentingnya penanganan ketidakseimbangan kelas untuk meningkatkan keandalan prediksi klinis dalam pemilihan terapi tuberkulosis.

5.2. Saran Pengembangan

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya adalah:

1. Menambahkan paduan pengobatan yang diklasifikasikan agar tidak terbatas pada tiga regimen pengobatan saja, namun juga mencakup regimen pengobatan lain yang biasa digunakan pada pasien tuberkulosis.
2. Menambah jumlah data dan variasi fitur yang digunakan, karena dalam penelitian ini data hanya terbatas pada tahun 2020–2024 dengan 15 fitur, sehingga data yang lebih luas dan fitur yang lebih beragam dapat meningkatkan performa model prediksi.
3. Membandingkan performa CatBoost yang dioptimasi dengan Optuna dengan metode optimasi lainnya seperti *Particle Swarm Optimization* (PSO), *Genetic Algorithm* (GA), atau *Bayesian Optimization* untuk melihat efektivitas masing-masing metode terhadap hasil prediksi komplikasi.
4. Menambahkan ruang lingkup hyperparameter CatBoost yang dioptimasi, karena dalam penelitian ini hanya mencakup *depth*, *random_strength*, *l2_leaf_reg*, *border_count*, *iterations*, *bagging_temperature*, dan *learning_rate*. Penambahan hyperparameter lain seperti *rsm* (*random subspace method*), *leaf_estimation_iterations*, *grow_policy*, *min_data_in_leaf*, serta *boosting_type* dapat dieksplorasi untuk meningkatkan akurasi model secara lebih maksimal.