

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan pada hasil penelitian yang telah diselesaikan mengenai optimasi algoritma Random Forest dalam prediksi tentang kualitas air bersih menggunakan metode Random Search dan Grid Search, maka telah disimpulkan beberapa hal berikut :

1. Model Random Forest tanpa optimasi (*default*) berperan sebagai pembanding dasar dalam penelitian ini. Hasil pengujian menunjukkan bahwa model ini menghasilkan akurasi sebesar 84.12%, presisi 83.07%, recall 83.78%, dan F1-score 83.42%. Keunggulan utama dari model ini terletak pada efisiensi waktu komputasi, yakni hanya membutuhkan 3 detik. Meskipun demikian, performanya masih lebih rendah dibanding model yang telah melalui proses optimasi hyperparameter.
2. Optimasi algoritma Random Forest menggunakan metode Random Search mampu meningkatkan performa model *default*. Proses optimasi ini menghasilkan model dengan akurasi 84.38%, presisi 84.04%, recall 82.94%, dan F1-score 83.49%. Hal ini menunjukkan bahwa pencarian kombinasi hyperparameter secara acak tetap dapat memberikan peningkatan kinerja, meskipun proses pencariannya tidak menyeluruh. Waktu komputasi yang dibutuhkan adalah sekitar 2 jam 2 menit 54 detik.
3. Optimasi algoritma Random Forest menggunakan metode Grid Search memberikan hasil performa optimal dalam penelitian. Model ini mendapat akurasi tertinggi pada 85.62%, presisi 85.56%, recall 83.99%, dan F1-score 84.77%. Hal ini menunjukkan bahwa pencarian kombinasi hyperparameter secara sistematis melalui Grid Search mampu menemukan parameter yang optimal. Namun, metode ini juga memerlukan waktu komputasi yang paling tinggi, yaitu sekitar 5 jam 12 menit.
4. Berdasarkan hasil evaluasi secara keseluruhan, dapat disimpulkan bahwa proses optimasi hyperparameter dapat berpengaruh terhadap peningkatan

performa model klasifikasi Random Forest. Model dengan optimasi Grid Search menunjukkan performa terbaik dari sisi akurasi, diikuti oleh optimasi Random Search, dan terakhir model *default*. Namun, peningkatan performa tersebut diikuti oleh peningkatan waktu komputasi secara signifikan.

5.2 Saran

Berdasar dari hasil penelitian yang telah selesai dilakukan, terdapat beberapa saran yang disampaikan guna memberi masukan untuk pengembangan penelitian selanjutnya, yaitu :

1. Disarankan penggunaan dataset dengan jumlah lebih besar dan lebih banyak fitur lain agar model yang dikembangkan memiliki daya klasifikasi yang lebih baik terhadap data baru. Dataset dengan variasi yang lebih luas baik dari segi parameter kimia air dan lokasi geografis pengambilan sampel dapat membuat model menjadi lebih relevan terhadap kondisi di lapangan.
2. Mengingat distribusi kelas pada dataset yang digunakan bersifat tidak seimbang, maka apabila proses penyeimbangan kelas tidak diterapkan pada tahap prapemrosesan, disarankan untuk menggunakan metrik evaluasi tambahan yang lebih representatif, seperti ROC-AUC guna memperoleh gambaran kinerja model yang lebih akurat dan merata terhadap seluruh kelas.
3. Jika keterbatasan waktu komputasi menjadi pertimbangan utama, maka disarankan agar penelitian selanjutnya menerapkan metode optimasi hyperparameter yang lebih efisien, mengingat Grid Search memerlukan waktu komputasi yang relatif lama. Penggunaan model algoritma *default* maupun model algoritma optimasi lain dapat dipertimbangkan untuk memperoleh performa model yang optimal dengan waktu komputasi yang lebih efisien.