

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

1. Berdasarkan hasil penelitian mengenai penerapan *Random Forest Regression* untuk memprediksi konsumsi listrik rata-rata pelanggan di Kabupaten Labuhanbatu, dapat disimpulkan bahwa algoritma tersebut dapat diterapkan secara sistematis pada data pelanggan yang bersifat cross-sectional dan nonhistoris. Penelitian menggunakan 6.510 data pelanggan listrik prabayar dengan lima variabel prediktor, yaitu MERK_KWH, TARIF, DAYA_VA, NAMA_KECAMATAN, dan NAMA_KELURAHAN, serta KONSUMSI_RATA_RATA_TARGET sebagai variabel target. Data dibagi menjadi 5.208 data training dan 1.302 data testing. Empat variabel kategorikal diproses menggunakan one-hot encoding melalui pipeline, sedangkan DAYA_VA dipertahankan sebagai variabel numerik. Penyetelan hyperparameter dilakukan dengan *GridSearchCV* lima fold hanya pada data training sehingga data testing tetap berfungsi sebagai evaluasi akhir yang independen.
2. Hasil pengujian menunjukkan bahwa model Random Forest Regression menghasilkan MAE sebesar 57,06, MSE sebesar 5.823,30, RMSE sebesar 76,31, R^2 sebesar 0,0172, dan sMAPE sebesar 64,41%. Nilai tersebut menunjukkan bahwa model hanya mampu menjelaskan sekitar 1,72% variasi konsumsi listrik rata-rata pada data testing, sehingga

kemampuan prediksinya masih rendah. Dibandingkan baseline, Random Forest memberikan hasil sedikit lebih baik karena menghasilkan MAE, MSE, RMSE, dan sMAPE yang lebih rendah serta R^2 yang lebih tinggi. Namun, perbedaannya kecil sehingga keunggulan tersebut belum menunjukkan daya prediksi yang kuat.

3. Hasil permutation importance menunjukkan NAMA_KECAMATAN sebagai prediktor dengan kontribusi relatif tertinggi, diikuti NAMA_KELURAHAN, MERK_KWH, TARIF, dan DAYA_VA. Temuan ini tidak menunjukkan hubungan sebab-akibat. Secara keseluruhan, keterbatasan model lebih berkaitan dengan terbatasnya informasi prediktor daripada kegagalan teknis algoritma. Oleh karena itu, hasil penelitian harus dipahami sebagai estimasi konsumsi rata-rata antarpelanggan berdasarkan karakteristik yang tersedia, bukan sebagai peramalan konsumsi listrik pada periode mendatang.
4. Model juga menghasilkan rentang prediksi 69,11 sampai 118,56, sedangkan nilai aktual testing berada pada rentang 2,50 sampai 523,20. Kondisi ini menunjukkan model belum cukup sensitif mengenali pelanggan dengan konsumsi sangat rendah maupun sangat tinggi. Selain itu, ketidakseimbangan jumlah pelanggan antarwilayah membatasi kekuatan generalisasi hasil pada seluruh kecamatan penelitian yang tercakup dalam dataset penelitian.

5.2. Saran

1. Berdasarkan hasil penelitian, pengembangan berikutnya perlu memprioritaskan peningkatan kualitas dan kelengkapan data sebelum memperluas kompleksitas algoritma. Penelitian selanjutnya disarankan menambahkan variabel yang lebih berkaitan langsung dengan perilaku konsumsi listrik, seperti riwayat konsumsi satu sampai tiga bulan, periode bulan dan tahun, jenis pelanggan, karakteristik bangunan, jumlah penghuni, jenis kegiatan usaha, peralatan listrik, luas bangunan, cuaca atau suhu, serta riwayat gangguan. Jika riwayat konsumsi digunakan, struktur data perlu diubah menjadi data pelanggan per periode dan pembagian training serta testing dilakukan berdasarkan waktu agar tidak terjadi kebocoran informasi.
2. Cakupan data juga perlu diperluas dan diseimbangkan antar wilayah karena jumlah pelanggan pada beberapa kecamatan sangat sedikit dibandingkan wilayah dominan, sehingga pola kelompok minoritas sulit dipelajari secara stabil. Model *Random Forest* pada penelitian ini belum disarankan sebagai dasar tunggal untuk memprediksi konsumsi listrik masa depan karena data yang digunakan bersifat cross-sectional dan kemampuan prediksinya masih rendah. Setelah data lebih representatif dan informatif, penelitian dapat membandingkan *Random Forest* dengan algoritma regresi lain menggunakan prosedur evaluasi yang sama. Setiap pengembangan tetap perlu mempertahankan pemisahan data testing, pencegahan data leakage, pelaporan hasil apa adanya, serta pembatasan interpretasi pada hubungan prediktif, bukan sebab-akibat.