

BAB V

KESIMPULAN DAN REKOMENDASI

5.1 Kesimpulan

Eksperimen mengenai analisis prediksi DBD menggunakan metode *Machine Learning* telah dilakukan, maka penulis mendapatkan beberapa kesimpulan antara lain sebagai berikut:

1. Penelitian ini berhasil mengimplementasikan empat metode *Machine Learning* yaitu *Multiple Linear Regression* (MLR), *Random Forest* (RF), *Support Vector Regression* (SVR), dan *XGBoost* (XGB) untuk Prediksi Jumlah Kasus Demam Berdarah Dengue (DBD) dengan baik dan menggunakan *hyperparameter tuning GridSearch-CV* untuk mencari parameter optimal.
2. Perbandingan nilai *error* evaluasi metrik *Mean Absolute Error* (MAE) untuk kedua kota adalah sebagai berikut:
 - a) Kota San Juan: *Multiple Linear Regression* menghasilkan MAE 23.947; *Random Forest* menghasilkan MAE 20.513; *Support Vector Regression* menghasilkan MAE 18.349; *XGBoost* menghasilkan MAE 17.592.
 - b) Kota Iquitos: *Multiple Linear Regression* menghasilkan MAE 8.230; *Random Forest* menghasilkan MAE 7.087, *Support Vector Regression* menghasilkan MAE 6.320, *XGBoost* menghasilkan MAE 6.565.
3. Dengan data-data yang digunakan pada penelitian ini, model *Machine Learning* yang paling baik pada data dua kota yang berbeda ini adalah *XGBoost* dengan nilai *error* evaluasi metrik *Mean Absolute Error* (MAE) paling rendah. Pada Kota San Juan menghasilkan MAE 17.592 dan pada Kota Iquitos menghasilkan MAE 6.565.
4. Berdasarkan *feature importance* yang dihasilkan oleh model *XGBoost* pada kedua kota ini, terdapat variabel yang sama di keduanya. *reanalysis_specific_humidity_g_per_kg* (kelembaban spesifik rata-rata), *year* (tahun) dan *weekofyear* (minggu dalam setahun) termasuk *feature* atau

variabel yang paling berpengaruh terhadap data total kasus DBD di Kota San Juan dan Iquitos. Sehingga dapat disimpulkan bahwa variabel waktu dan variabel perubahan iklim sangat berpengaruh terhadap jumlah kasus DBD pada kedua kota tersebut.

5.2 Rekomendasi

Eksperimen pada penelitian ini sudah berhasil dilakukan, namun masih terdapat kekurangan yang mampu disempurnakan pada penelitian berikutnya, untuk itu penulis mengusulkan beberapa saran, antara lain:

- 1) Berdasarkan hasil pada penelitian ini meskipun salah satu dari empat metode yang digunakan menghasilkan MAE paling rendah, namun semua yang digunakan ini tidak dapat memprediksi jumlah kasus DBD yang melonjak secara tiba-tiba, model-model tersebut tidak dapat memprediksi pola perubahan data aktual jumlah kasus DBD yang lonjakannya sangat signifikan. Penulis menyarankan untuk mengefisienkan pada konfigurasi praproses untuk semua algoritma.
- 2) Penelitian ini menghasilkan model XGBoost yang paling baik untuk memprediksi jumlah kasus DBD berdasarkan data pengukuran iklim, untuk meningkatkan hasil prediksi penulis merekomendasikan penelitian berikutnya untuk mengefisienkan model training pada XGBoost.
- 3) Berdasarkan perhitungan *feature importance*, variabel yang paling berpengaruh terhadap jumlah kasus DBD di kedua kota adalah *reanalysis_specific_humidity_g_per_kg* (kelembaban spesifik rata-rata), *year* (tahun) dan *weekofyear* (minggu dalam setahun), sehingga pada penelitian selanjutnya disarankan agar variabel tersebut tetap digunakan.
- 4) Untuk meningkatkan hasil prediksi, penulis menyarankan untuk menggunakan set data latih yang lebih banyak dan terbaru dan akan lebih baik jika data yang digunakan adalah data harian.