

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian, peramalan curah hujan harian di Kota Medan menggunakan metode XGBoost yang memanfaatkan data historis BMKG dalam rentang tahun 2013-2022 didapatkan kesimpulan sebagai berikut.

1. Faktor-faktor meteorologis seperti temperatur maksimum (Tmax), temperatur minimum (Tmin), temperatur rata-rata (Trata-rata), kelembapan rata-rata (Rhrata-rata), lama penyinaran matahari (LPM), serta hasil *feature engineering* seperti *lag*, *rolling mean*, *differencing*, dan fitur siklik, terbukti mampu membantu model XGBoost dalam mengenali pola data *time series* curah hujan.
2. Metode XGBoost diimplementasi menggunakan bahasa pemrograman Python dengan tahapan yang meliputi *preprocessing* data, *feature engineering*, standardisasi data dan pemodelan. Proses ini dilakukan dengan mengatur *hyperparameter*, meliputi $n_estimator = 800$, $learning\ rate = 0,01$, $max_depth = 3$, $colsample_bytree = 0,8$, $subsample = 0,8$, $gamma = 1$, $reg_lamda = 1$ untuk mencegah *overfitting* dan *underfitting*.
3. Evaluasi kinerja model pada tiga skenario pembagian data (70:20, 80:20, 90:10). Pada skenario pembagian 70:30 model menghasilkan nilai MAE sebesar 1,05 dan RMSE sebesar 1,93. Selanjutnya, pada skenario 80:20 model memperoleh nilai MAE sebesar 0,97 dan RMSE sebesar 1,80. Sedangkan pada skenario 90:10 model menghasilkan nilai MAE sebesar 1,05 dan RMSE sebesar 1,91. Hasil evaluasi dari ketiga skenario, didapatkan bahwa skenario pembagian data 80:20 memberikan performa terbaik, dengan nilai kesalahan prediksi (MAE dan RMSE) paling rendah dibandingkan dua skenario lainnya. Visualisasi hasil peramalan menunjukkan bahwa model dapat mengikuti pola curah hujan harian, meskipun terdapat sedikit perbedaan pada puncak curah hujan ekstrem.
4. Akurasi model ditingkatkan melalui *feature importance*, yang mengidentifikasi fitur-fitur paling berpengaruh seperti CH_diff7 (0,449824), CH_diff1 (0,186684), dan Rhrata-rata (0,067113). Dengan menerapkan *feature*

importance untuk menyeleksi fitur dengan menggunakan *threshold* 0,01695, model dengan 9 fitur utama menghasilkan performa terbaik dengan MAE sebesar 0,93 mm dan RMSE sebesar 1,71 mm. Hal ini menunjukkan bahwa *feature importance* efektif dalam mengurangi kompleksitas model dan meningkatkan akurasi peramalan.

5.2 Saran

Penelitian ini menunjukkan bahwa fitur hasil *feature engineering* seperti *lag*, *rolling mean window*, *differencing*, dan fitur siklik cukup berpengaruh terhadap hasil peramalan berdasarkan *feature importance*. Untuk penelitian selanjutnya, model dapat dikembangkan dengan mencoba berbagai rentang waktu pada fitur *lag* dan *rolling mean*, serta menambahkan rekayasa fitur lain untuk melihat pengaruh pola temporal terhadap curah hujan. Selain itu, dapat pula dieksplorasi faktor lain yang berpotensi memengaruhi curah hujan guna meningkatkan performa model. Metode XGBoost sendiri cocok digunakan pada data berukuran menengah hingga besar seperti penelitian ini yang menggunakan 3.652 data harian, karena mampu menangani hubungan non-linear antar variabel dan data dengan nilai ekstrem seperti curah hujan yang tinggi.