

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Perkiraan curah hujan memainkan peran penting dalam berbagai kegiatan manusia, terutama di wilayah beriklim tropis yang memiliki pola cuaca yang dinamis. Guna meningkatkan keakuratan prediksi tersebut, studi ini membandingkan performa dua algoritma *machine learning*, yakni *Generalized Linear Model* dan *Linear Regression*. Data ini menggunakan data *Weather Forecast Data.csv* yang berasal dari platform Kaggle, kemudian melalui tahapan *preprocessing* seperti label *encoding*, normalisasi, serta *class imbalance*. Dataset selanjutnya dipisah menjadi data pelatihan dan pengujian sebelum diterapkan ke masing-masing algoritma. *Generalized Linear Model* merupakan model regresi yang dirancang untuk menangani distribusi *non-normal* pada variabel target dan menghubungkannya dengan variabel prediktor melalui fungsi tautan. Sementara itu, *Linear Regression* merupakan pendekatan sederhana yang menggambarkan hubungan linier antara variabel bebas dan terikat. Hasilnya, *Generalized Linear Model* memiliki nilai akurasi sebesar 90.17% lalu untuk RMSE sebesar 0.3949 dan MAE 0.3836, sedangkan *Linear Regression* memiliki nilai MAE sebesar 0.2656 dan RMSE 0.3218 lalu untuk akurasi sebesar 89.26%. Dengan demikian, hasil penelitian ini tidak hanya memberikan gambaran efektivitas kedua model dalam prediksi curah hujan, tetapi juga membuka peluang pengembangan model yang lebih kompleks dan adaptif di masa depan, sehingga dapat mendukung pengambilan keputusan yang lebih tepat dalam mitigasi risiko bencana atau kebutuhan lainnya.

5.2 Saran

Sebagai arah untuk pengembangan penelitian selanjutnya, disarankan untuk menelusuri pendekatan hybrid atau ensemble yang mampu mengombinasikan kelebihan *Generalized Linear Model* (GLM) dalam klasifikasi serta kekuatan *Linear Regression* dalam mengurangi galat prediksi. Penambahan variabel meteorologis lain atau perluasan rentang data historis yang dianalisis juga berpotensi meningkatkan kualitas serta kemampuan generalisasi model prediksi hujan. Eksplorasi metode alternatif seperti *Random Forest*, *Gradient Boosting*, atau

Neural Network dapat memberikan perspektif perbandingan yang lebih luas sekaligus membuka peluang peningkatan performa model. Di sisi lain, penyesuaian *hyperparameter* yang lebih optimal serta penerapan teknik preprocessing data lanjutan, seperti penanganan data tidak seimbang dan rekayasa fitur, diharapkan mampu menyempurnakan hasil prediksi. Terakhir, membangun sistem prediktif hujan berbasis data sensor cuaca secara langsung dan bekerja secara real-time akan sangat berguna untuk aplikasi nyata, khususnya dalam bidang meteorologi dan penanggulangan bencana.

