ABSTRAK

Perubahan iklim global menyebabkan pola curah hujan menjadi semakin tidak menentu sehingga meningkatkan risiko banjir, kekeringan, dan gangguan sektor pertanian. Kompleksitas geografis Indonesia memperumit proses prediksi curah hujan karena dipengaruhi berbagai faktor atmosfer lokal dan regional. Perkembangan deep learning memungkinkan pemodelan deret waktu yang lebih baik melalui arsitektur *Long Short-Term Memory* (LSTM) dan *Bidirectional Long Short-Term Memory* (BiLSTM). Namun, belum diketahui secara pasti apakah BiLSTM selalu lebih unggul dibandingkan LSTM serta bagaimana pengaruh variasi *hyperparameter* dan panjang urutan data (*timesteps*) terhadap performa prediksi.

Penelitian ini bertujuan menganalisis perbandingan kinerja LSTM dan BiLSTM dalam memprediksi curah hujan harian multivariat serta mengevaluasi pengaruh variasi *timesteps* dan *hyperparameter* terhadap akurasi model. Data diperoleh dari Stasiun Meteorologi Citeko Kabupaten Bogor dengan beberapa parameter cuaca harian sebagai variabel input dan curah hujan sebagai target. Model diuji menggunakan *timesteps* 7, 14, dan 30 serta dioptimasi melalui eksplorasi *units* LSTM, dense *units*, *dropout rate*, *batch size*, dan *learning rate*. Evaluasi dilakukan menggunakan *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE).

Hasil penelitian menunjukkan bahwa kedua model mencapai performa terbaik pada *timesteps* 7. LSTM menghasilkan RMSE 13,0448 dan MAE 7,5624, sedangkan BiLSTM menghasilkan RMSE 13,0883 dan MAE 7,5676. Peningkatan panjang *timesteps* justru meningkatkan *error* karena membawa informasi yang kurang relevan dan *noise*. Kompleksitas tambahan pada BiLSTM tidak memberikan peningkatan akurasi yang signifikan karena karakteristik curah hujan bersifat kausal satu arah. Faktor yang paling memengaruhi performa model adalah interaksi antara *timesteps*, jumlah *units*, dan *dropout rate*, sementara *learning rate* relatif konstan pada nilai optimal 0,003.

Penelitian ini menyimpulkan bahwa model yang lebih kompleks tidak selalu menghasilkan prediksi yang lebih baik. LSTM dengan konfigurasi *hyperparameter* optimal dan *timesteps* pendek memberikan keseimbangan terbaik antara akurasi dan stabilitas, sehingga lebih cocok untuk prediksi curah hujan harian multivariat. Hasil penelitian diharapkan menjadi referensi dalam pengembangan model prediksi curah hujan yang akurat, efisien, dan aplikatif untuk mendukung pengambilan keputusan di berbagai sektor.

Kata Kunci : Curah Hujan, Multivariat, Prediksi, *Hyperparameter*, *Timesteps*

ABSTRACT

Global climate change is causing increasingly erratic rainfall patterns, increasing the risk of flooding, drought, and disruption to the agricultural sector. Indonesia's geographical complexity complicates rainfall prediction due to the influence of various local and regional atmospheric factors. Advances in deep learning enable improved time series modeling through Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM) architectures. However, it is unclear whether BiLSTM is consistently superior to LSTM, and how varying hyperparameters and timesteps affect prediction performance.

This study aims to analyze the comparative performance of LSTM and BiLSTM in predicting multivariate daily rainfall and evaluate the effect of varying timesteps and hyperparameters on model accuracy. Data were obtained from the Citeko Meteorological Station in Bogor Regency, with several daily weather parameters as input variables and rainfall as the target. The model was tested using 7, 14, and 30 timesteps and optimized through exploration of LSTM units, dense units, dropout rate, batch size, and learning rate. The evaluation was conducted using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

The results showed that both models achieved the best performance at 7 timesteps. LSTM produced an RMSE of 13.0448 and an MAE of 7.5624, while BiLSTM produced an RMSE of 13.0883 and an MAE of 7.5676. Increasing the timestep length actually increased the error due to the inclusion of less relevant information and noise. The added complexity of BiLSTM did not significantly improve accuracy because rainfall characteristics are one-way causal. The factors that most influenced model performance were the interaction between timesteps, the number of units, and the dropout rate, while the learning rate remained relatively constant at the optimal value of 0.003.

This study concluded that more complex models do not always produce better predictions. LSTM with optimal hyperparameter configurations and short timesteps provided the best balance between accuracy and stability, making it more efficient for multivariate daily rainfall prediction. The research results are expected to serve as a reference in developing accurate, efficient, and applicable rainfall prediction models to support decision-making in various sectors.

Keywords: Rainfall, Multivariate, Prediction, Hyperparameter, Timesteps