

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini mengevaluasi performa model prediksi curah hujan di Kota Surabaya, berdasarkan jumlah variabel yang dipilih, dengan menggunakan *Principal Component Analysis* (PCA) dan *Long Short-Term Memory* (LSTM). Dataset yang digunakan mencakup periode tahun 2018 hingga 2024. Total data yang digunakan sebanyak 2.525 data harian. Data tersebut terdiri dari sembilan parameter sebagai fitur dan satu parameter sebagai target.

Dari hasil pengujian, performa terbaik diperoleh saat menggunakan 6 variabel PCA, dengan MSE sebesar 0.0010, RMSE sebesar 0.0333, dan MAE sebesar 0.0110. Hasil ini menunjukkan bahwa pemilihan jumlah variabel yang optimal dapat meningkatkan akurasi prediksi secara signifikan. Meskipun penggunaan 7 hingga 9 variabel menunjukkan penurunan nilai MAE, nilai RMSE dan MSE cenderung meningkat dibandingkan dengan skenario 6 variabel, menunjukkan bahwa penambahan variabel tidak selalu meningkatkan performa model. Sebaliknya, penggunaan lebih sedikit variabel (2 hingga 5) menghasilkan error yang lebih besar.

Hasil penelitian ini juga menegaskan bahwa PCA berperan penting dalam menentukan jumlah variabel yang optimal untuk meningkatkan akurasi prediksi. Dengan mereduksi dimensi data, PCA membantu menghilangkan informasi yang kurang relevan, sehingga model dapat fokus pada fitur yang paling berkontribusi terhadap prediksi. Namun, hasil eksperimen menunjukkan bahwa menambahkan lebih banyak variabel tidak selalu menghasilkan model yang lebih akurat. Dalam beberapa skenario, peningkatan jumlah variabel justru menyebabkan error yang lebih tinggi, kemungkinan akibat adanya redundansi atau noise dalam data. Oleh karena itu, pendekatan yang seimbang dalam memilih jumlah variabel hasil PCA sangat penting untuk mencapai keseimbangan antara akurasi dan efisiensi model prediksi curah hujan.

5.2 Saran

Meskipun penelitian ini telah menunjukkan bahwa penggunaan PCA dapat mempengaruhi hasil prediksi curah hujan menggunakan LSTM, hasil yang diperoleh masih belum menunjukkan jarak yang sangat dekat antara nilai prediksi dan nilai aktual. Masih terdapat kesenjangan atau error yang dapat disebabkan oleh berbagai faktor, seperti pola curah hujan yang tidak menentu, jumlah data yang digunakan, serta pemilihan parameter dalam proses pelatihan model. Untuk meningkatkan akurasi prediksi pada penelitian selanjutnya, beberapa saran yang dapat dipertimbangkan adalah sebagai berikut:

1. Melakukan *tuning hyperparameter* secara lebih optimal. Dalam penelitian ini, penggunaan hyperparameter seperti jumlah neuron, jumlah epoch, dan batch size masih menggunakan pendekatan standar. Dengan melakukan tuning hyperparameter yang lebih optimal menggunakan teknik seperti Grid Search atau Bayesian Optimization, performa model dapat ditingkatkan lebih lanjut.
2. Menambah jumlah data pelatihan. Semakin banyak data historis yang digunakan, semakin baik model dalam menangkap pola curah hujan. Oleh karena itu, penggunaan dataset dengan rentang waktu yang lebih panjang atau resolusi data yang lebih tinggi (misalnya data harian dibandingkan data bulanan) dapat meningkatkan performa model.
3. Mencoba metode lain selain LSTM. Meskipun LSTM dikenal sebagai model yang baik dalam menangani data deret waktu, masih terdapat metode lain yang dapat dicoba, seperti Transformer, GRU (Gated Recurrent Unit), atau kombinasi CNN-LSTM. Model-model ini mungkin dapat memberikan hasil yang lebih akurat dalam menangkap pola curah hujan.

Dengan menerapkan berbagai saran di atas, diharapkan hasil prediksi intensitas curah hujan dapat lebih akurat dan mendekati nilai aktual, sehingga dapat memberikan manfaat yang lebih besar dalam bidang meteorologi dan perencanaan mitigasi bencana.