

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Curah hujan merupakan salah satu unsur cuaca yang memengaruhi ketersediaan air, debit aliran sungai, kegiatan pertanian, arus lalu lintas, dan aktivitas masyarakat lainnya. Curah hujan yang tinggi dapat meningkatkan risiko banjir, sedangkan curah hujan yang rendah dalam waktu lama dapat menyebabkan kekeringan. Prediksi curah hujan dibutuhkan untuk membantu pengelolaan sumber daya air dan mengurangi risiko bencana hidrometeorologi [1].

Curah hujan di Indonesia berubah mengikuti musim dan berbeda pada setiap wilayah. Beberapa wilayah dapat mengalami banjir ketika curah hujan tinggi, Seperti diawal tahun 2026 terjadi banjir di sejumlah titik di Kota Tangerang [2]. Banjir tersebut juga menyebabkan sebagian warga harus mengungsi [3]. Kondisi ini menunjukkan perlunya prediksi curah hujan pada setiap wilayah agar langkah penanganan dapat disiapkan lebih awal.

Besarnya curah hujan pada suatu wilayah dipengaruhi oleh berbagai kondisi cuaca seperti suhu udara, kelembapan, kecepatan angin, penyinaran matahari, dan tekanan udara [4]. Pengaruh setiap kondisi tersebut tidak selalu sama dan dapat berubah dari waktu ke waktu. Curah hujan pada suatu bulan juga dapat berkaitan dengan kondisi pada bulan-bulan sebelumnya. Prediksi curah hujan perlu memperhatikan hubungan antarvariabel cuaca dan pola waktu pada data.

Prediksi curah hujan telah dilakukan dengan berbagai metode mulai dari model statistik, *machine learning*, hingga *deep learning*. Beberapa model yang pernah digunakan antara lain ARIMA, KNN, SVR, *Random Forest*, RNN, GRU, dan LSTM. Perbandingan pada beberapa lokasi menunjukkan bahwa model *deep learning* menghasilkan kesalahan prediksi yang lebih rendah dibandingkan dengan beberapa model lainnya [1]. LSTM juga telah digunakan pada prediksi curah hujan di beberapa wilayah [5]. Kemampuan LSTM dalam menyimpan informasi dari periode sebelumnya membantu model mempelajari pola pada data curah hujan.

Hasil prediksi LSTM tetap dipengaruhi oleh pengaturan model saat pelatihan. Pengaturan tersebut meliputi jumlah *layer*, jumlah unit, *learning rate*, *batch size*, dan *dropout* [6]. Pemilihan *hyperparameter* yang kurang tepat dapat menurunkan hasil prediksi atau membuat model terlalu mengikuti data pelatihan.

Penentuan *hyperparameter* secara manual memerlukan banyak percobaan karena setiap parameter memiliki beberapa pilihan nilai [7].

Pencarian *hyperparameter* dapat dilakukan menggunakan metode optimasi. *Bayesian Optimization* menggunakan hasil evaluasi yang sudah diperoleh untuk menentukan konfigurasi berikutnya [8]. Metode ini telah diterapkan pada optimasi LSTM untuk beberapa kasus prediksi termasuk prediksi curah hujan bulanan [9]. Metode lain yang dapat digunakan adalah *Genetic Algorithm*.

Pada *Genetic Algorithm*, setiap konfigurasi mewakili satu kombinasi *hyperparameter*. Konfigurasi baru dibentuk melalui proses seleksi, *crossover*, dan mutasi [10]. Metode ini juga telah digunakan untuk menentukan *hyperparameter* LSTM [11]. Keberagaman konfigurasi perlu dijaga agar pencarian tidak berhenti terlalu cepat pada satu bagian ruang pencarian. Salah satu caranya adalah dengan mengubah laju mutasi selama proses optimasi.

Bayesian Optimization dan *Genetic Algorithm* memiliki cara kerja yang berbeda dalam menentukan *hyperparameter*. *Bayesian Optimization* menggunakan hasil evaluasi sebelumnya untuk menentukan konfigurasi berikutnya sedangkan *Genetic Algorithm* membentuk konfigurasi baru melalui seleksi, *crossover*, dan mutasi. Penelitian terdahulu menggunakan data, ruang *hyperparameter*, metrik evaluasi, dan jumlah percobaan yang berbeda. Hasil penelitian tersebut belum dapat digunakan untuk menyimpulkan metode mana yang lebih baik ketika diuji dalam kondisi yang sama. Penelitian ini membandingkan kedua metode menggunakan data, ruang *hyperparameter*, metrik evaluasi, dan jumlah percobaan yang sama.

Berdasarkan perbedaan cara pencarian tersebut, penelitian ini membandingkan *Bayesian Optimization* dan *Genetic Algorithm* dalam menentukan *hyperparameter* LSTM untuk prediksi curah hujan di Kota Tangerang. Kedua metode diuji menggunakan data, ruang pencarian, fungsi objektif, dan jumlah evaluasi yang sama. Model hasil optimasi kemudian dibandingkan dengan LSTM *baseline* menggunakan MAE, RMSE, dan R^2 , sedangkan proses optimasi dianalisis berdasarkan perkembangan RMSE validasi dan waktu pencarian.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan *Bayesian Optimization* dan *Genetic Algorithm* dalam proses optimasi *hyperparameter* model LSTM untuk prediksi curah hujan di

Kota Tangerang?

2. Bagaimana hasil dan perbandingan model LSTM *baseline*, LSTM dengan *Bayesian Optimization* dan LSTM dengan *Genetic Algorithm* berdasarkan nilai MAE, RMSE, dan R^2 ?
3. Bagaimana perbandingan *Bayesian Optimization* dan *Genetic Algorithm* berdasarkan konfigurasi *hyperparameter* yang dihasilkan, perubahan nilai RMSE selama proses optimasi, dan waktu pencarian?

1.3 Batasan Masalah

Agar pembahasan penelitian tetap sesuai dengan tujuan yang telah ditetapkan, batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan data historis meteorologi yang diperoleh melalui *Open Meteo Historical Weather*. Lokasi pengambilan data berada di Kota Tangerang pada koordinat $-6,1511^\circ$ lintang dan $106,6165^\circ$ bujur dengan elevasi 18 meter.
2. Data meteorologi yang digunakan mencakup periode Januari 1981 sampai Juni 2026. Data harian diolah menjadi data bulanan sebelum digunakan dalam proses pemodelan.
3. Prediksi dilakukan untuk memperkirakan curah hujan pada bulan berikutnya berdasarkan data meteorologi pada beberapa bulan sebelumnya.
4. Model prediksi yang digunakan adalah *Long Short-Term Memory* (LSTM). Model yang dibandingkan terdiri dari LSTM *baseline*, LSTM dengan *Bayesian Optimization*, dan LSTM dengan *Genetic Algorithm*.
5. Optimasi dibatasi pada jumlah *layer*, jumlah unit pada setiap *layer*, *dropout*, *learning rate*, dan *batch size*. Parameter lain pada model ditetapkan sama selama proses optimasi.
6. *Bayesian Optimization* dan *Genetic Algorithm* menggunakan ruang pencarian, data validasi, metrik evaluasi, dan jumlah percobaan yang sama.
7. Evaluasi model dilakukan menggunakan *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan koefisien determinasi (R^2).

8. Penelitian ini tidak membahas sistem prediksi secara *real-time*, sistem peringatan dini, pengembangan aplikasi, maupun perbandingan dengan model prediksi selain LSTM.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan *Bayesian Optimization* dan *Genetic Algorithm* untuk mengoptimasi *hyperparameter* model LSTM pada prediksi curah hujan di Kota Tangerang.
2. Membandingkan hasil prediksi LSTM *baseline*, LSTM dengan *Bayesian Optimization*, dan LSTM dengan *Genetic Algorithm* berdasarkan nilai MAE, RMSE, dan R^2 .
3. Menganalisis perbedaan *Bayesian Optimization* dan *Genetic Algorithm* berdasarkan konfigurasi *hyperparameter* yang dihasilkan, perubahan nilai RMSE selama proses optimasi, dan waktu pencarian.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah referensi mengenai penggunaan LSTM untuk prediksi curah hujan berdasarkan data meteorologi.
2. Memberikan gambaran mengenai penerapan *Bayesian Optimization* dan *Genetic Algorithm* untuk menentukan *hyperparameter* model LSTM.
3. Memberikan informasi mengenai perbedaan hasil prediksi dan proses optimasi *Bayesian Optimization* dan *Genetic Algorithm* sehingga dapat menjadi bahan pertimbangan pada penelitian sejenis.

1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini disusun dalam lima bab sebagai berikut.

1. **Bab 1 PENDAHULUAN**

Bab ini menjelaskan latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan laporan.

2. **Bab 2 LANDASAN TEORI**

Bab ini membahas teori dan penelitian terdahulu yang menjadi dasar penelitian, meliputi prediksi curah hujan, data deret waktu, model *Long Short-Term Memory*, *Bayesian Optimization*, *Genetic Algorithm*, serta metrik evaluasi yang digunakan.

3. **Bab 3 METODOLOGI PENELITIAN**

Bab ini membahas metodologi penelitian, mulai dari pengumpulan dan persiapan data, agregasi data bulanan, pembentukan target, pembagian data, *Exploratory Data Analysis*, penetapan fitur, transformasi target, penentuan nilai *window size*, pembangunan model LSTM, proses optimasi *hyperparameter*, hingga evaluasi model.

4. **Bab 4 HASIL DAN DISKUSI**

Bab ini membahas hasil penelitian dan pembahasannya, meliputi hasil data *preprocessing*, analisis data, penetapan konfigurasi, pelatihan model LSTM *baseline*, hasil optimasi *Bayesian Optimization* dan *Genetic Algorithm*, serta perbandingan kinerja ketiga model.

5. **Bab 5 SIMPULAN DAN SARAN**

Bab ini membahas simpulan berdasarkan tujuan dan hasil penelitian, serta saran untuk penelitian selanjutnya.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A