

BAB I

PENDAHULUAN

1.1 Latar Belakang

Peningkatan suhu udara menjadi salah satu fenomena iklim global yang semakin terasa dalam beberapa dekade terakhir. Suhu tertinggi dalam satu dekade terakhir tercatat sepanjang tahun 2025, ditandai dengan meningkatnya intensitas dan durasi gelombang panas di berbagai wilayah dunia. Indonesia menempati peringkat kedua sebagai negara dengan jumlah penduduk terbanyak yang terpapar panas ekstrem selama lebih dari 30 hari dalam setahun berdasarkan laporan *Climate Central* (World Bank, 2025). Fenomena tersebut mencerminkan peningkatan risiko iklim yang berdampak langsung terhadap kehidupan manusia dalam berbagai sektor, seperti kesehatan masyarakat, pertanian, serta infrastruktur.

Pernyataan tersebut sejalan dengan hasil observasi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) bahwa suhu udara rata-rata di Indonesia mengalami kenaikan sebesar $1,02^{\circ}\text{C}$ dalam periode 44 tahun (1981–2024), atau sekitar $0,23^{\circ}\text{C}$ per dekade (Chindi, 2026). Direktorat Perubahan Iklim BMKG (2025) juga menyebutkan dalam proyeksi perubahan suhu rata-rata tahunannya bahwa seluruh daerah Jawa Tengah diprediksi akan mengalami perubahan sebesar $3,1\text{--}3,4^{\circ}\text{C}$. Kota Semarang sebagai ibu kota Provinsi Jawa Tengah merupakan salah satu wilayah perkotaan yang dikenal memiliki suhu udara relatif tinggi. Kota ini memiliki iklim tropis basah dengan karakteristik beragam sepanjang tahun. Keunikan karakteristik geografis kota ini terletak pada perbedaan yang cukup signifikan antara wilayah utara yang merupakan daerah pesisir Laut Jawa dan

wilayah selatan yang merupakan daerah perbukitan (Yuwono *et al.*, 2017). Selain itu, Kota Semarang juga memiliki topografi yang bervariasi serta curah hujan yang tinggi. Berbagai karakteristik tersebut menjadikan pergerakan suhu udara harian di Kota Semarang cenderung lebih ekstrem dan berdampak langsung terhadap kenyamanan termal, kesehatan masyarakat, serta perencanaan tata kota. Dalam kondisi yang demikian, suhu udara tidak hanya menjadi fenomena yang diamati secara historis, tetapi juga menimbulkan kebutuhan akan kemampuan untuk mengantisipasi kondisi suhu di masa mendatang.

Prediksi atau peramalan suhu udara adalah upaya untuk mengetahui kondisi suhu udara di masa mendatang berdasarkan data historis yang telah dikumpulkan sebelumnya. Informasi prediksi suhu udara harian dapat membantu berbagai sektor dalam pengambilan keputusan, seperti mitigasi risiko kesehatan, perencanaan energi, dan adaptasi terhadap perubahan iklim (Limehuwey *et al.*, 2023). Data suhu udara di Indonesia, terutama di Kota Semarang, pada kenyataannya memiliki karakteristik yang cukup kompleks, berupa deret waktu, ketergantungan temporal, pola nonlinier, serta fluktuasi pada musim tertentu yang menjadi tantangan tersendiri dalam melakukan prediksi.

Pendekatan statistik konvensional sering kali memiliki keterbatasan dalam menangkap kompleksitas pola data suhu. Limehuwey *et al.* (2023) dalam penelitiannya tentang prediksi suhu di Kota Ambon menggunakan analisis regresi linier yang berhasil memodelkan data suhu dengan variabel tunggal dan menghasilkan tingkat kesalahan prediksi sekitar 1,3 satuan akibat adanya pola fluktuatif data. Hal ini menunjukkan bahwa suhu udara tidak hanya dipengaruhi oleh satu faktor, tetapi merupakan hasil interaksi berbagai variabel meteorologis

yang memiliki pola nonlinier dan ketergantungan temporal yang kuat, sehingga pendekatan menggunakan model statistika konvensional dianggap kurang optimal.

Seiring berkembangnya ilmu pengetahuan, berbagai pendekatan pemodelan yang lebih adaptif mulai dikembangkan untuk menangkap kompleksitas data suhu udara. Salah satunya adalah *machine learning* seperti *Light Gradient Boosting Machine* (LightGBM) yang memiliki keunggulan dalam efisiensi komputasi serta kemampuan menangani data yang kompleks dan memiliki hubungan yang non linier (Duan *et al.*, 2021). Kemampuan tersebut diperoleh algoritma ini melalui mekanisme pertumbuhan pohon *leaf wise* dan *histogram-based learning*, menjadikan LightGBM sebagai metode alternatif yang akurat dan efisien untuk melakukan prediksi (Sutaryani *et al.*, 2024).

Terdapat beberapa penelitian yang menerapkan algoritma LightGBM untuk melakukan prediksi pada beberapa topik. Seperti Duan *et al.* (2021) yang berhasil membuktikan bahwa LightGBM mampu memodelkan data deret waktu lebih baik dibandingkan metode peramalan deret waktu konvensional, berkat efisiensi dan efektivitas algoritma *boosting* yang dimiliki oleh LightGBM. Lubis dan Syahputra (2022) juga melakukan peramalan dalam konteks bisnis untuk membandingkan kemampuan LightGBM dengan empat algoritma lainnya seperti *Linear Regression* (LR), *Ridge Regression* (RR), *XGBoost Classifier*, dan *Long Short-Term Memory* (LSTM). LightGBM memperoleh akurasi tertinggi pada data multivariat dalam penelitian tersebut.

Berdasarkan berbagai penelitian yang telah dilakukan sejauh ini, memaksimalkan performa LightGBM untuk menghasilkan prediksi yang akurat menjadi target utama. Fokus penelitian terdahulu pada umumnya masih terbatas

pada optimasi parameter internal model, sementara pengayaan informasi temporal dalam struktur data sering kali dikesampingkan. Pengayaan informasi temporal dilakukan dengan menambahkan fitur *time series* berupa fitur (variabel) *lag* dan *rolling mean* yang merupakan hasil dari *feature engineering* (rekayasa fitur) untuk menambahkan informasi temporal pada algoritma berbasis pohon ini.

Metode *feature engineering* serupa telah dilakukan oleh Cao *et al.* (2023) dan diperoleh kesimpulan bahwa penambahan fitur *time series* pada LightGBM mampu menekan nilai galat seperti MSE dan MAE. Perbedaan utama antara penelitian ini dan penelitian yang dilakukan oleh Cao *et al.* (2023) terletak pada objek penelitian serta metode *feature engineering* yang digunakan. Penelitian oleh Cao *et al.* (2023) berfokus pada prediksi suhu di dalam rumah kaca dengan resolusi data per menit selama periode 30 hari, sedangkan penelitian ini memprediksi suhu udara rata-rata harian menggunakan data selama dua tahun. Pada tahap rekayasa fitur, Cao *et al.* (2023) menerapkan fitur selisih tingkat pertama (*first-order difference*), sementara penelitian ini menggunakan fitur *lag* dan *rolling mean* yang dibentuk dari variabel target. Penelitian ini juga melibatkan sembilan variabel prediktor (X_1 – X_9) dan mengevaluasi kinerja model menggunakan metrik RMSE, MAE, dan MAPE, sedangkan penelitian acuan menggunakan metrik MSE, MAE, dan R^2 .

Penelitian ini tidak hanya mengandalkan penambahan fitur *time series*, melainkan juga mendalami mekanisme internal algoritma melalui pendekatan *leaf wise growth* dan *histogram-based learning* dalam mengenali pola data iklim yang nonlinier. Integrasi antara penambahan fitur *time series* serta keunggulan struktural LightGBM ini diharapkan mampu menghasilkan model peramalan yang lebih

akurat, sekaligus memberikan pemahaman metodologis yang mendalam mengenai peran informasi temporal dalam memitigasi risiko fluktuasi suhu udara harian.

1.2 Rumusan Masalah

Rumusan masalah yang ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Bagaimana mengimplementasikan algoritma *Light Gradient Boosting Machine* (LightGBM) dengan penambahan fitur *time series* untuk memprediksi suhu udara harian Kota Semarang?
2. Bagaimana pengaruh penambahan fitur *time series* terhadap performa prediksi model LightGBM?
3. Bagaimana hasil peramalan suhu udara harian Kota Semarang untuk tiga periode ke depan menggunakan model LightGBM terbaik?

1.3 Batasan Masalah

Batasan masalah yang ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data iklim harian Kota Semarang yang diperoleh dari hasil pengamatan Stasiun Meteorologi Jawa Tengah, BMKG, mencakup periode 13 Februari 2024 hingga 25 April 2026 dengan total 803 data.
2. Fitur *time series* yang ditambahkan dalam penelitian ini adalah fitur *lag* dan *rolling mean*.
3. Metode yang digunakan untuk memprediksi suhu udara harian rata-rata Kota Semarang adalah algoritma *Light Gradient Boosting Machine* (LightGBM).

4. Evaluasi performa model dilakukan menggunakan metrik *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan penelitian ini dilakukan untuk mencapai tujuan sebagai berikut:

1. Mengimplementasikan algoritma LightGBM dengan penambahan fitur *time series* untuk memprediksi suhu udara harian Kota Semarang.
2. Menganalisis pengaruh penambahan fitur *time series* terhadap performa prediksi model LightGBM.
3. Mendapatkan hasil peramalan suhu udara harian Kota Semarang untuk tiga periode ke depan menggunakan model LightGBM terbaik.