

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kualitas udara di DKI Jakarta terus mengalami akibat meningkatnya aktivitas transportasi dan industri. Kadar partikulat polutan dan gas sering kali melampaui ambang batas anjuran World Health Organization (WHO), yaitu $5 \mu\text{g}/\text{m}^3$ untuk $\text{PM}_{2.5}$ per tahun. WHO (2023) melaporkan bahwa lebih dari tujuh juta kematian per tahun secara global disebabkan oleh paparan polusi udara, dengan kontribusi signifikan dari partikel halus yang memicu penyakit pernapasan dan kardiovaskular.

Pemantauan kualitas udara di Jakarta saat ini masih terbatas pada pengukuran langsung oleh Dinas Lingkungan Hidup DKI Jakarta, yang bersifat deskriptif dan belum mampu memberikan prediksi secara akurat. Model statistik konvensional seperti regresi linier hanya memanfaatkan hubungan sederhana antarvariabel dan tidak cukup kuat untuk menangkap hubungan non-linear serta dinamis antara polutan dan faktor meteorologis. Kondisi ini menimbulkan kebutuhan akan model prediksi yang lebih adaptif terhadap kompleksitas data serta mampu mengenali dinamika spasial serta temporal dari perubahan kualitas udara.

Metode Gradient Boosting, khususnya CatBoost, menjadi salah satu algoritma yang menjanjikan. CatBoost dibangun berdasarkan pendekatan *ensemble sequential learning*, setiap model baru berfungsi memperbaiki kesalahan dari model sebelumnya. Pendekatan ini memungkinkan peningkatan performa prediksi secara bertahap serta mampu mengurangi *bias-variance tradeoff* melalui mekanisme *boosting iterations*. Seperti penelitian Toharudin dkk. (2023) yang membandingkan beberapa metode *gradient boosting* untuk memprediksi indeks kualitas udara, seperti AdaBoost, LightGBM, XGBoost, dan CatBoost.

Namun pendekatan berbasis *boosting* saja belum sepenuhnya mampu menangkap pola spasial-temporal. Distribusi polutan di wilayah Jakarta tidak hanya bergantung pada waktu, tetapi juga pada lokasi geografis dan interaksi antarwilayah. Sebagai contoh, konsentrasi PM_{2.5} di Jakarta Timur dapat dipengaruhi oleh arah angin serta aktivitas industri di Jakarta Utara. Oleh karena itu, integrasi antara dimensi ruang (spasial) dan waktu (temporal) menjadi penting agar model dapat mengenali pola penyebaran polutan secara lebih realistis dan dinamis.

Sejalan dengan hal tersebut, beberapa penelitian seperti Chen dan Zhao (2024) melalui model *Spatio-Temporal Attention Dynamic Graph Network* (STA-DGN) berhasil mempelajari keterkaitan antarstasiun pemantauan menggunakan graph attention serta menangkap pola waktu melalui *temporal attention*. Model ini menunjukkan penurunan error signifikan dibandingkan LSTM, Transformer, dan GCRNN. Demikian pula, penelitian Zhang dkk. (2024) yang mengembangkan model MGAtt-LSTM (*multi-graph attention*) untuk menangkap hubungan spasial dan temporal. Hasil penelitian tersebut membuktikan bahwa perhatian spasial dan temporal yang terintegrasi dapat meningkatkan kemampuan model dalam memprediksi difusi dan fluktuasi polutan udara.

Berdasarkan temuan-temuan tersebut, penelitian ini mengusulkan pengembangan model prediksi spasial-temporal Indeks Kualitas Udara di DKI Jakarta dengan menggabungkan algoritma CatBoost dan *Extreme Feature Attention Weighting*. Pendekatan ini dirancang untuk memberikan bobot perhatian (*attention weighting*) pada variabel polutan dan faktor meteorologis yang paling berpengaruh (menyentuh titik konsentrasi ekstrem) terhadap perubahan kualitas udara. Sistem yang dihasilkan diharapkan mampu memberikan prediksi yang lebih akurat dan adaptif terhadap variasi spasial dan temporal.

1.2 Tujuan Penelitian

Tujuan dari studi ini adalah :

1. Merancang model prediksi AQI spasial-temporal DKI Jakarta berbasis CatBoost Regressor dengan Attention Mechanism, yang memberikan pembobotan pada parameter AQI dan meteorologi ekstrem di setiap stasiun pemantauan dengan bobot maksimum 1.0 untuk kondisi ekstrem dan 0.5 untuk kondisi normal, di mana seluruh bobot dinormalisasi dalam rentang 0–1 sesuai prinsip *softmax attention*.
2. Menerapkan hasil prediksi dalam sistem informasi geografis (GIS) indeks kualitas udara berbasis dashboard Streamlit agar masyarakat DKI Jakarta agar dapat memantau wilayah rawan polusi dan turut berpartisipasi dalam reduksi emisi.

1.3 Manfaat Penelitian

Dalam jangka pendek, hasil penelitian ini diharapkan dapat berperan sebagai sistem informasi untuk monitoring kondisi kualitas udara di DKI Jakarta dengan memberikan informasi kepada masyarakat ketika konsentrasi polutan maupun unsur meteorologi (cuaca) diprediksi melebihi batas anjuran WHO ketika jam sibuk lalu lintas dan industri.

Dalam jangka panjang, sistem ini diharapkan dapat berkontribusi terhadap perencanaan dan pengelolaan lingkungan perkotaan DKI Jakarta, seperti :

1. Penetapan zona rendah emisi dengan pembatasan kendaraan bermotor di kawasan dengan tingkat polusi udara yang tinggi.
2. Pengembangan sistem transportasi ramah lingkungan, seperti perluasan jaringan transportasi publik dan penggunaan kendaraan berbasis listrik.
3. Optimalisasi pemanfaatan Ruang Terbuka Hijau (RTH) sebagai area penyerap polutan dan peningkat kualitas udara di wilayah perkotaan.