

BAB I

PENDAHULUAN

1. 1 Latar Belakang

Pencemaran udara telah berevolusi menjadi salah satu krisis lingkungan terbesar. Sejak era revolusi industri, aktivitas manusia menjadi pemicu berbagai polutan berbahaya ke atmosfer, yang menyebabkan penurunan kualitas lingkungan. Berdasarkan Laporan akhir kegiatan pemantauan kualitas udara Provinsi DKI Jakarta, (Dinas Lingkungan Hidup DKI Jakarta, 2022), Menurut WHO pencemaran udara telah berkontribusi sebesar 7,6% kematian di seluruh dunia dan tidak kurang dari 7 juta kematian dini setiap tahun. Bahkan 99% penduduk dunia diperkirakan telah terpapar pencemaran udara yang menyebabkan resiko penyakit, seperti stroke, jantung, gangguan paru-paru dan kanker. Kondisi penurunan kualitas udara terakumulasi secara signifikan di Kawasan perkotaan dengan laju urbanisasi dan pertumbuhan ekonomi yang pesat. Sebagaimana, dijelaskan oleh (Hameed et al., 2023), aktivitas perkotaan terutama dari sektor transportasi dan industri berkontribusi secara signifikan terhadap pencemaran lingkungan yang berdampak buruk pada kesehatan masyarakat secara luas.

Di antara berbagai jenis polutan, perhatian utama dunia kesehatan kini tertuju pada Particulate Matter 2.5 (PM_{2.5}) karena partikel ini memiliki diameter aerodinamis kurang dari 2,5 mikrometer. Ukurannya yang mikroskopis memungkinkan partikel ini menembus penghalang sistem pernapasan terdalam hingga masuk ke aliran darah. Urgensi penanganan polutan ini sangat dirasakan di DKI Jakarta. Dalam Laporan (Kementerian Koordinator Bidang Infrastruktur & Pembangunan Kewilayahan, 2024), mencatat bahwa rata-rata tahunan konsentrasi PM_{2.5} di ibu kota mencapai 34,78 ug/m³, angka yang jauh melampaui ambang batas panduan WHO sebesar 5ug/m³ dan menempatkan Jakarta dalam status kedaruratan kualitas udara.

Tingginya konsentrasi polusi di Jakarta tidak terlepas dari posisinya sebagai inti dari wilayah Jabodetabek (Jakarta, Bogor, Depok, Tangerang,

Bekasi), yang merupakan salah satu kawasan perkotaan terbesar di Indonesia. Kawasan ini dihuni oleh sekitar 30 juta jiwa penduduk yang terus berkembang seiring pertumbuhan penduduk dan perkembangan ekonominya. Peningkatan aktivitas transportasi, industri, permukiman, dan komersial semakin memberikan tekanan terhadap kualitas udara di wilayah Jabodetabek. Sebagaimana dalam Laporan (Kementerian Koordinator Bidang Infrastruktur dan Pembangunan Kewilayahan, 2024), berbagai sumber emisi yang terkonsentrasi di wilayah Jabodetabek, ditambah dengan tingginya intensitas mobilitas dan kegiatan ekonomi, berkontribusi terhadap memperburuknya kondisi udara ambien.

Selain faktor emisi, dinamika $PM_{2.5}$ juga sangat ditentukan oleh mekanisme pembersihan alamiah melalui faktor meteorologi dan pola waktu. (Gencarelli et al., 2025) menjelaskan bahwa curah hujan (*total precipitation*) memiliki peran krusial dalam proses deposisi basah (*wet deposition*) yang secara efektif "mencuci" polutan dari udara, sementara kecepatan angin mempengaruhi dispersi partikel. Hal ini sejalan dengan (Song et al., 2023) yang mencatat bahwa faktor iklim memodulasi variabilitas spasial dan temporal polusi. Di sisi lain, dimensi waktu juga menjadi determinan penting, (Gencarelli et al., 2025) menyoroti bahwa variabel waktu (*day*) merupakan salah satu faktor terpenting dalam pemodelan karena merefleksikan siklus aktivitas manusia (kerja vs libur) dan variasi musiman yang mempengaruhi akumulasi polutan.

Namun, upaya untuk memetakan hubungan kausalitas antara variabel-variabel tersebut sering kali terbentur pada karakteristik data lingkungan yang kompleks dan stokastik. Hubungan antara emisi transportasi dan konsentrasi $PM_{2.5}$ di wilayah tropis seperti Jakarta sering kali bersifat non-linier, di mana fluktuasi polusi tidak selalu berbanding lurus dengan jumlah kendaraan akibat adanya pengaruh dominan dari dinamika atmosfer yang berubah-ubah. Fenomena ini menyebabkan analisis yang hanya mengandalkan asumsi hubungan linier berpotensi menyederhanakan masalah dan menghasilkan

kesimpulan yang bias atau tidak akurat dalam menjelaskan variabilitas polusi udara yang sesungguhnya.

Untuk menganalisis hubungan kompleks antara variabel-variabel tersebut, pemilihan metode pemodelan yang tepat menjadi sangat krusial. Penelitian terdahulu menunjukkan bahwa metode statistik konvensional seperti *Ordinary Least Squares* (OLS) sering kali memiliki keterbatasan dalam akurasi. (Gencarelli et al., 2025) mencatat bahwa model regresi linier sederhana umumnya menghasilkan nilai R^2 yang rendah (0,10–0,65) karena ketidakmampuannya menangkap hubungan non-linier pada data lingkungan. Sebaliknya, pendekatan Machine Learning, seperti *Random Forest* dan *Extreme Gradient Boosting* (XGBoost) terbukti lebih unggul. Studi terapan oleh (Kaya Keleş et al., 2025) dan (Gencarelli et al., 2025) membuktikan bahwa kedua algoritma ini mampu menangkap pola non-linier dengan jauh lebih baik, menghasilkan akurasi prediksi yang tinggi dan nilai error yang lebih rendah dibandingkan model tradisional.

Meskipun berbagai studi terdahulu telah berhasil menerapkan metode Machine Learning untuk prediksi kualitas udara, masih terdapat kesenjangan penelitian (research gap) yang signifikan. Studi seperti Tang et al. (2025) dan Song et al. (2023) mayoritas berfokus pada wilayah subtropis dan menggunakan data statis atau penginderaan jauh (*remote sensing*), sehingga kurang menangkap dinamika mobilitas harian masyarakat kota secara riil. Di sisi lain, studi lokal di Jakarta seperti yang dilakukan oleh Ilahi et al. (2024) cenderung masih menggunakan pendekatan statistik konvensional dan mengabaikan variabel transportasi, padahal sektor ini menjadi salah satu penyumbang emisi terbesar. Ketiadaan integrasi data mobilitas riil dalam model prediksi menyebabkan evaluasi kebijakan menjadi kurang presisi.

Berangkat dari permasalahan tersebut, penelitian ini bertujuan untuk menganalisis secara komprehensif efektivitas transportasi publik sebagai instrumen pengendalian pencemaran udara di DKI Jakarta. Penelitian ini menerapkan pendekatan komparatif dengan menggunakan algoritma Machine

Learning (Random Forest dan XGBoost) untuk menangkap pola non-linier. Dengan mengintegrasikan data volume penumpang transportasi umum riil dan parameter meteorologi, penelitian ini tidak hanya berfokus pada akurasi prediksi, tetapi juga diperdalam dengan simulasi skenario kebijakan. Melalui pendekatan ini, diharapkan dapat dihasilkan estimasi yang presisi mengenai batas efektivitas sektor mobilitas dalam menurunkan beban pencemaran udara di tengah pengaruh faktor meteorologi yang dinamis.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, rumusan masalah yang akan dijawab dalam penelitian ini, antara lain:

1. Apakah terdapat pola temporal (Bulan dan Hari) yang signifikan dalam fluktuasi konsentrasi $PM_{2.5}$ di Jakarta?
2. Bagaimana pengaruh mobilitas masyarakat pengguna transportasi umum dan faktor meteorologi terhadap konsentrasi $PM_{2.5}$ di Jakarta?
3. Manakah model yang memberikan akurasi prediksi terbaik antara Random Forest, dan XGBoost?
4. Seberapa besar efektivitas peningkatan jumlah pengguna transportasi umum dalam penurunan konsentrasi $PM_{2.5}$ berdasarkan simulasi skenario model?

1.3 Tujuan

Adapun tujuan dari penelitian yang dilakukan oleh penulis sebagai berikut:

1. Mengidentifikasi pola temporal (musiman dan harian) polusi udara di jakarta.
2. Menganalisis pengaruh mobilitas masyarakat pengguna transportasi dan faktor meteorologi terhadap konsentrasi $PM_{2.5}$ di Jakarta.
3. Mengevaluasi dan membandingkan kinerja model prediksi antara Random Forest, dan XGboost
4. Mensimulasikan dampak skenario peningkatan pengguna transportasi umum terhadap penurunan konsentrasi $PM_{2.5}$ untuk menguji efektivitas kebijakan pengurangan emisi.

1.4 Manfaat

Manfaat Akademis yang diharapkan oleh penulis, antara lain:

1. Penelitian ini memberikan wawasan mengenai perbandingan kinerja metode machine learning yang berbeda (Random Forest, dan XGBoost) dalam bidang lingkungan, khususnya parameter emisi PM_{2.5}.
2. Penelitian ini menyediakan studi kasus nyata tentang bagaimana data pengguna transportasi publik dapat diintegrasikan dalam model peramalan kualitas udara.
3. Sebagai persyaratan untuk mendapatkan gelar Serjana.

Manfaat Praktis yang diharapkan, antara lain:

1. Hasil penelitian dapat digunakan untuk mendukung pengambilan Keputusan berbasis data dalam merancang kebijakan mitigasi emisi PM_{2.5}.
2. Penelitian ini dapat meningkatkan kesadaran publik mengenai tren emisi PM_{2.5} dan dapat mendukung sistem peringatan dini terhadap peningkatan konsentrasi PM_{2.5},
3. Memberikan gambaran praktis mengenai efektivitas penggunaan transportasi publik dalam menekan emisi, sehingga dapat dijadikan bahan evaluasi program transportasi ramah lingkungan.
4. Hasil Penelitian dapat digunakan dasar untuk mengembangkan alat prediksi yang lebih akurat dan mudah digunakan dimasa depan untuk memprediksikan emisi PM_{2.5}

1.5 Batasan Masalah

Dalam melakukan penelitian sangat diperlukan adanya pembatasan dalam suatu permasalahan agar penelitian ini menjadi terarah dan memudahkan dalam pembatasan hingga tujuan utama dalam penelitian ini dapat sesuai dan tercapai. Ada beberapa Batasan masalah dalam penelitian ini, yaitu:

1. Penelitian ini hanya menggunakan data sekunder Januari 2023 – September 2025 yang berasal Satu Data Jakarta dengan detail Lembaga sebagai berikut, Dinas Lingkungan Hidup untuk data PM_{2.5}, Dinas Perhubungan untuk data

jumlah penumpang transportasi umum, dan Badan Meteorologi, Klimatologi, dan Geofisika untuk data Curah hujan dan kecepatan angin. Data dari sumber lain atau data dengan karakteristik yang berbeda tidak dipertimbangkan.

2. Penelitian ini hanya menggunakan model Random Forest, dan XGBoost.
3. Penelitian ini menggunakan metrik evaluasi tertentu, yaitu Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), untuk membandingkan kinerja model. Metrik evaluasi lain yang mungkin relevan tidak dipertimbangkan.
4. Penelitian ini hanya menggunakan beberapa faktor eksternal yang digunakan untuk menilai faktor-faktor yang mempengaruhi konsentrasi PM_{2.5}, diantaranya Faktor Jumlah Penumpang Transportasi umum, Curah Hujan, Kecepatan Angin, Faktor Hari, dan Faktor bulan. Faktor-Faktor selain factor tersebut tidak dipertimbangkan dalam penelitian ini.
5. Area penelitian hanya mencakup area DKI Jakarta.
6. Penelitian hanya menggunakan moda transportasi umum transjakarta berbahan bakar minyak dengan kapasitas 86 penumpang per bus.

1.6 Sistematika Penulisan

Untuk mempermudah melihat dan mengetahui pembahasan yang ada dalam skripsi ini, maka dibuatlah sistematika yang merupakan kerangka dan pedoman penulisan skripsi. Adapun sistematika Penulisan adalah sebagai berikut:

1. Bagian awal skripsi

Bagian awal akan memuat halaman sampul depan, halaman judul, halaman pernyataan orisinalitas, halaman pengesahan, kata pengantar, halaman pernyataan persetujuan publikasi karya ilmiah untuk kepentingan akademis, abstrak (dalam Bahasa Indonesia dan Inggris), daftar isi, daftar tabel dan daftar gambar.

2. Bagian utama skripsi

Bagian utama terbagi atas bab dan sub bab yaitu sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini membahas identifikasi masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Pada bab ini membahas deskripsi wilayah studi, pencemaran udara dan $PM_{2.5}$, faktor-faktor yang mempengaruhi $pm_{2.5}$, algoritma pemodelan dan penelitian terdahulu

BAB III METODE PENELITIAN

Pada bab ini membahas, tahapan penelitian, lokasi dan waktu penelitian, jenis dan sumber data, variabel penelitian, metode pemodelan, evaluasi model, simulasi skenario kebijakan.

BAB IV HASIL DAN DISKUSI

Pada bab ini membahas, analisis deskriptif dan pola temporal. evaluasi kinerja model, analisa hubungan antar variabel, simulasi skenario kebijakan, dan Perancangan dashboard.

BAB V KESIMPULAN DAN SARAN

Pada bab ini membahas, kesimpulan dan saran.

3. Bagian akhir skripsi

Bagian akhir dari skripsi ini berisi tentang daftar pustaka dan daftar lampiran.