

**PERBANDINGAN *RANDOM FOREST* DAN *XGBOOST*
PADA PREDIKSI DAMPAK TRANSPORTASI PUBLIK
TERHADAP KONSENTRASI PM2.5 UNTUK
PENGENDALIAN KUALITAS UDARA**

SKRIPSI SARJANA SISTEM INFORMASI

Disusun Oleh:

Endien Stevani Susarno

227006516049



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI &
INFORMATIKA**

UNIVERSITAS NASIONAL

2026

**PERBANDINGAN *RANDOM FOREST* DAN *XGBOOST*
PADA PREDIKSI DAMPAK TRANSPORTASI PUBLIK
TERHADAP KONSENTRASI PM2.5 UNTUK
PENGENDALIAN KUALITAS UDARA**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh
Endien Stevani Susarno
227006516049



PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL
2026

HALAMAN PEGESAHAN

HALAMAN PENGESAHAN

TUGAS AKHIR

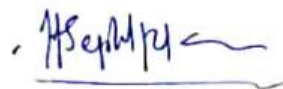
PERBANDINGAN RANDOM FOREST DAN XGBOOST PADA PREDIKSI
DAMPAK TRANSPORTASI PUBLIK TERHADAP KONSENTRASI PM2.5
UNTUK PENGENDALIAN KUALITAS UDARA



Endien Stevani Susarno

227006516049

Dosen Pembimbing



(Prof. Dr. Septi Andryana, S.Kom., MMSI)

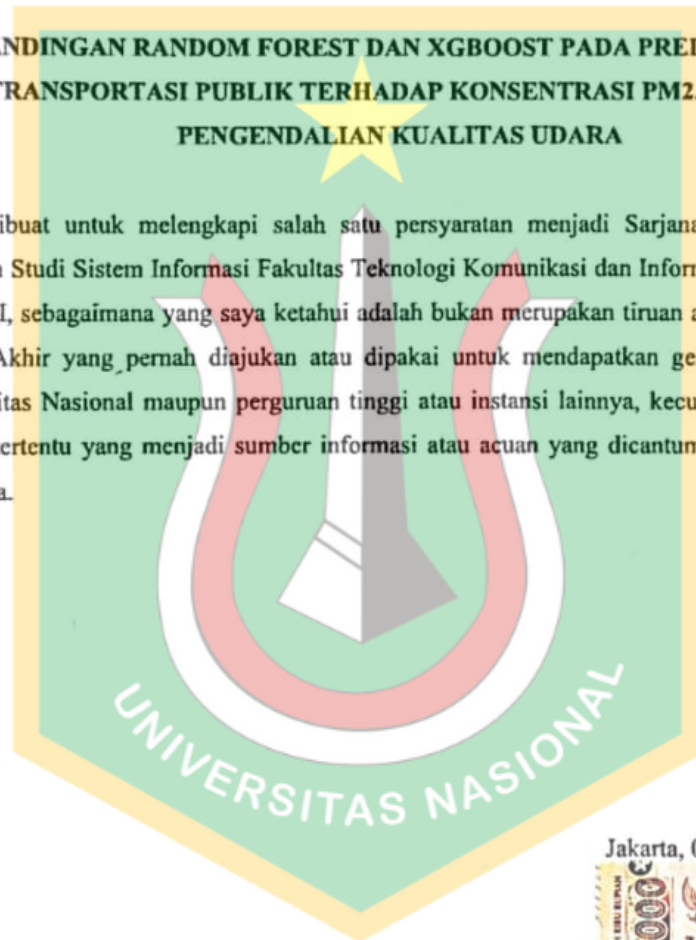
PERNYATAAN KEASLIAN TUGAS AKHIR

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

PERBANDINGAN RANDOM FOREST DAN XGBOOST PADA PREDIKSI DAMPAK TRANSPORTASI PUBLIK TERHADAP KONSENTRASI PM2.5 UNTUK PENGENDALIAN KUALITAS UDARA

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.



Jakarta, 02 Maret 2026



Endien Stevani Susarno

227006516049

LEMBAR PERSETUJUAN REVIEW AKHIR

LEMBAR PERSETUJUAN REVIEW AKHIR

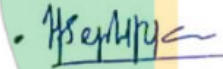
Tugas Akhir dengan judul :

PERBANDINGAN RANDOM FOREST DAN XGBOOST PADA PREDIKSI DAMPAK TRANSPORTASI PUBLIK TERHADAP KONSENTRASI PM2.5 UNTUK PENGENDALIAN KUALITAS

UDARA

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2025-2026 pada tanggal 24 Februari 2026

Dosen Pembimbing I



Prof. Dr. Septi Andryana, S.Kom., MMSI.

NIDN 0317097209

Ketua Program Studi



Dr. Ir. Andriahingsih, S.Kom., MMSI

NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Endien Stevani Susarno
NPM : 227006516049
Fakultas/Akademi : Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 24 Februari 2026

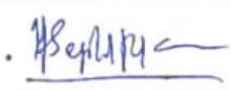

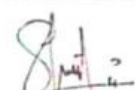
JUDUL DALAM BAHASA INDONESIA:

**PERBANDINGAN *RANDOM FOREST* DAN *XGBOOST* PADA
PREDIKSI DAMPAK TRANSPORTASI PUBLIK TERHADAP
KONSENTRASI PM2.5 UNTUK PENGENDALIAN KUALITAS
UDARA**

JUDUL DALAM BAHASA INGGRIS:

**COMPARISON OF *RANDOM FOREST* AND *XGBOOST* IN
PREDICTING THE IMPACT OF PUBLIC TRANSPORTATION
ON PM2.5 CONCENTRATION FOR AIR QUALITY CONTROL**

TANDA TANGAN DAN TANGGAL

Pembimbing I	Ka. Prodi	Mahasiswa
TGL: 02 Maret 2020	TGL: 02.03.2026	TGL:
 Prof. Dr. Septi Andryana	 Dr. Ir. Andrianingsih	 Endien Stevani Susarno

HALAMAN PERNYATAAN ORISINALITAS


HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Yang menyatakan,

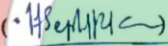
Nama : Endien Stevani Susarno

NIM : 227006516049

Tanda Tangan : 

Tanggal : 02 Maret 2026

Megetahui

Pembimbing I : Prof. Dr. Septi Andryana, S.Kom., MMSI. ()



KATA PENGANTAR

Segala hormat, puji dan Syukur saya panjatkan kepada Tuhan Yang Maha Esa karena atas izin, dan karunia-Nyalah sehingga saya sebagai penulis dapat melakukan dan menyelesaikan skripsi ini dengan waktu yang telah ditentukan dan direncanakan. Adapun topik Skripsi yang saya muat dengan judul **“PERBANDINGAN RANDOM FOREST DAN XGBOOST PADA PREDIKSI DAMPAK TRANSPORTASI PUBLIK TERHADAP KONSENTRASI PM2.5 UNTUK PENGENDALIAN KUALITAS UDARA”**. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi dan memperoleh gelar Sarjana Komputer dari Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional.

Penulis menyadari penelitian dan penyusunan skripsi ini tentu tidak terlepas dari dukungan serta bantuan berbagai pihak. Oleh karenanya, dengan penuh rasa hormat dan apresiasi, penulis ingin menyampaikan terima kasih yang sebesar-besarnya, terutama kepada dosen pembimbing Tugas Akhir, Prof. Dr. Septi Andryana, S.Kom., MMSI, yang telah dengan sabar membimbing, memberikan masukan dan ilmu pengetahuan yang sangat berarti. Dedikasi dan waktunya dalam membimbing penulis tidak hanya memberikan arah dalam penelitian namun juga menjadi sumber motivasi yang kuat bagi penulis untuk terus berusaha menyelesaikan tugas akhir ini sebaik mungkin. Selain itu, pada kesempatan ini penulis mengungkapkan rasa terima kasih yang setulus-tulusnya kepada:

1. Terima kasih kepada Dr. Agung Triayudi, S.Kom., M.Kom selaku dekan Fakultas FTKI dan Andrianingsih, S.Kom. MMSI selaku kaprodi Sistem Informasi atas segala arahan dan dukungan yang diberikan.
2. Terima kasih kepada kedua Orang tua saya Bapak Sarno dan Ibu Pani Winarsih, yang senantiasa memberikan dukungan, doa, serta perhatian dalam setiap Langkah yang penulis lalui.
3. Kepada seluruh keluarga besar yang tidak dapat penulis sebutkan namanya satu persatu. Terima kasih atas segala doa restu dan dukungan tiada henti yang selalu menyertai setiap langkah penulisan skripsi ini.

4. Seluruh dosen pengajar di Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional. Maupun dosen di Program Studi lainnya yang telah memberikan banyak ilmu pengetahuan dan motivasi.
5. Untuk teman-teman penulis yang bersedia memberikan bantuan, masukan, dan dukungan yang sangat berarti untuk penulis sehingga penulis mampu menyisipkan penulisan ini dengan sebaik-baiknya, yaitu Annisa, Aura, Azel, dan Tya.
6. Seluruh rekan-rekan di tim CDI UNAS yang telah memberikan penulis motivasi untuk penulisan tugas akhir ini.
7. Penulis juga mengucapkan terima kasih kepada Leonardo Edwin, Raditya Dika, dan Duo Pengacara melalui channel Youtubenya telah menjadi teman yang telah menemani, menghibur dan memberikan semangat bagi penulis selama proses penulisan skripsi ini.
8. Terakhir, penulis ingin berterima kasih kepada diri sendiri, Endien Stevani Susarno. Terima kasih karena telah memperjuangkan, merelakan waktu istirahatnya untuk senantiasa melanjutkan dan menyelesaikan skripsi ini hingga akhirnya selesai.

Penulis menyadari bahwasanya masih banyak kekurangan dalam penyajian dan pemilihan kata maupun pembahasan hasil dalam skripsi ini. Oleh karena itu, dengan kerendahan hati penulis mengharapkan saran, kritik yang membangun dari semua pihak untuk perbaikan skripsi ini. Semoga skripsi ini dapat memberikan manfaat baik bagi pengembang ilmu pengetahuan maupun pihak-pihak yang berkemungkinan,

Jakarta, 02 Maret 2026



Endien Stevani Susarno
227006516049

ABSTRAK

Pencemaran udara berpartikulat halus ($PM_{2.5}$) merupakan permasalahan lingkungan dan kesehatan utama di wilayah perkotaan padat, termasuk DKI Jakarta. Konsentrasi $PM_{2.5}$ di Jakarta secara konsisten melampaui pedoman World Health Organization (WHO), dengan sektor transportasi sebagai salah satu kontributor utama emisi. Penelitian ini bertujuan untuk menganalisis pengaruh aktivitas transportasi umum dan faktor meteorologi terhadap konsentrasi $PM_{2.5}$ serta membandingkan kinerja model statistik dan machine learning dalam memprediksi polusi udara. Data harian periode Januari 2023 - September 2025 digunakan, mencakup konsentrasi $PM_{2.5}$, jumlah penumpang transportasi umum, curah hujan, kecepatan angin, serta variabel waktu (hari dan bulan). Dua pendekatan pemodelan diterapkan, yaitu Random Forest, dan Extreme Gradient Boosting (XGBoost). Dengan menggunakan subset data 80:20 menunjukkan bahwa XGBoost mengungguli Random Forest dalam hal akurasi prediksi, dengan XGBoost menunjukkan hasil R^2 0.90 dan Random Forest sebesar 0.75. Analisis interpretabilitas menggunakan SHAP mengindikasikan bahwa faktor meteorologi, khususnya curah hujan dan kecepatan angin, memiliki pengaruh dominan. Simulasi skenario kebijakan menunjukkan bahwa peningkatan jumlah penumpang transportasi umum dapat memberikan dampak positif terhadap pengendalian polusi, meskipun efektivitasnya sangat dipengaruhi oleh kondisi meteorologi. Temuan ini menegaskan pentingnya integrasi kebijakan transportasi dan faktor iklim dalam strategi mitigasi pencemaran udara di Jakarta.

Kata kunci: $PM_{2.5}$, transportasi umum, machine learning, Random Forest, XGBoost, Jakarta.

ABSTRACT

Atmospheric particulate matter (PM_{2.5}) pollution is a major environmental and health problem in densely populated urban areas, including DKI Jakarta. PM_{2.5} concentrations in Jakarta consistently exceed World Health Organization (WHO) guidelines, with the transportation sector being one of the main contributors to emissions. This study aims to analyse the influence of public transportation activities and meteorological factors on PM_{2.5} concentrations and compare the performance of statistical and machine learning models in predicting air pollution. Daily data from January 2023 to September 2025 were used, covering PM_{2.5} concentrations, public transportation ridership, rainfall, wind speed, and time variables (day and month). Two modelling approaches were applied Random Forest, and Extreme Gradient Boosting (XGBoost). Using an 80:20 data subset shows that XGBoost outperforms Random Forest in terms of prediction accuracy, with XGBoost showing an R² result of 0.90 and Random Forest showing 0.75. Interpretability analysis using SHAP indicates that meteorological factors, particularly rainfall and wind speed, have a dominant influence. Policy scenario simulations show that an increase in public transportation ridership can have a positive impact on pollution control, although its effectiveness is greatly influenced by meteorological conditions. These findings emphasize the importance of integrating transportation policy and climate factors into air pollution mitigation strategies in Jakarta.

Keywords: PM_{2.5}, public transportation, machine learning, Random Forest, XGBoost, Jakarta.

DAFTAR ISI

HALAMAN PEGESAHAN.....	iii
PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN REVIEW AKHIR	v
LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI vi	
HALAMAN PERNYATAAN ORISINALITAS.....	vii
KATA PENGANTAR	viii
ABSTRAK.....	x
<i>ABSTRACT</i>	xi
DAFTAR ISI.....	xii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR	xvi
BAB I PENDAHULUAN.....	1
1. 1 Latar Belakang.....	1
1. 2 Identifikasi Masalah.....	4
1. 3 Tujuan	4
1. 4 Manfaat	5
1. 5 Batasan Masalah.....	5
1. 6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Dekskripsi Wilayah Studi	8
2.2 <i>Particulate matter</i> 2.5 (PM _{2.5})	9
2.3 Faktor-Faktor yang Mempengaruhi PM _{2.5}	11
2.3.1 Aktivitas Transportasi dan Mobilitas Manusia	11
2.3.2 Faktor Metereologi.....	11

2.3.3	Pola Temporal	12
2.4	Algoritma Pemodelan.....	12
2.4.1	Random Forest (RF).....	12
2.4.2	XGBoost.....	13
2.5	Penelitian Terdahulu	14
BAB III METODE PENELITIAN.....		21
3.1	Tahapan Penelitian	21
3.2	Lokasi dan Waktu Penelitian	24
3.3	Jenis dan Sumber Data	25
3.4	Variabel Penelitian	26
3.5	Metode Pemodelan.....	27
3.5.1	Random Forest	27
3.5.2	XGBoost.....	27
3.6	Evaluasi Model.....	27
3.6.1	Mean Absolute Error (MAE)	27
3.6.2	Root Mean Square Error (RMSE).....	28
3.6.3	Coefficient of Determination (R^2).....	28
3.6.4	SHAP Values	29
3.7	Simulasi Skenario Kebijakan	29
BAB IV HASIL DAN DISKUSI		30
4.1	Analisis Deskriptif dan Pola Temporal	30
4.2.1	Evaluasi Model.....	32
4.2	Analisis Hubungan Antar Variabel	35
4.3	Simulasi Skenario Kebijakan	38
4.4	Perancangan Dashboard	39

BAB V KESIMPULAN DAN SARAN.....	43
5.1 Kesimpulan	43
5.2 Saran.....	44
DAFTAR PUSTAKA	46
LAMPIRAN.....	48
Lampiran 1 Cek Turnitin	48
Lampiran 2 Cek AI.....	49



DAFTAR TABEL

Tabel 2. 1 Studi Literatur	15
Tabel 3. 1 Daftar Data dan Sumber Data	25
Tabel 4. 4 Hasil Prediksi Konsentrasi $PM_{2,5}$	32
Tabel 4. 5 Hasil Evaluasi 3 Model	34
Tabel 4. 6 Pengaruh Variabel Terhadap Konsentrasi $PM_{2,5}$ pada Model Random Forest dan XGBoost	37



DAFTAR GAMBAR

Gambar 3 1 Tahapan Penelitian	21
Gambar 3 2 Peta Lokasi Stasiun Pemantauan Kualitas Udara.....	24
Gambar 4. 1 Rata-Rata PM _{2.5} per Bulan di DKI Jakarta	30
Gambar 4. 2 Rata-Rata Curah Hujan per Bulan di DKI Jakarta	30
Gambar 4. 3 Rata-Rata PM _{2.5} Per Musim di DKI Jakarta	30
Gambar 4. 4 Rata-Rata PM _{2.5} per Hari (senin-minggu)	31
Gambar 4. 5 Rata-Rata PM _{2.5} antara Hari Kerja dan Akhir Pekan	31
Gambar 4. 6 Pola Hubungan Variabel dengan Heatmap Correlation.....	35
Gambar 4. 7 SHAP Summary Plot dengan Random Forest.....	36
Gambar 4. 8 SHAP Summary Plot dengan XGBoost	36
Gambar 4. 9 Pengaruh Variabel Terhadap Konsentrasi PM _{2.5} pada Model Random Forest dan XGBoost	37
Gambar 4. 10 Simulasi Kenaikan penumpang transportasi umum	38
Gambar 4. 11 Dashboard Utama Sistem Analisis dan Prediksi Kualitas Udara PM _{2.5}	40
Gambar 4. 12 Dashboard Menu Analisi Pola Temporal.....	40
Gambar 4. 13 Dashboard Menu Simulasi Skenario Kebijakan.....	41
Gambar 4. 14 Dashboard Menu Simulasi Skenario Kebijakan, Posisi Prediksi skenario	42