

**SISTEM PREDIKSI TINGKAT KUALITAS UDARA DI JAKARTA
TIMUR BERBASIS *WEBSITE* MENGGUNAKAN ALGORITMA
*RANDOM FOREST REGRESSION***



**ANGELIA
NIM. 2110512110**

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
JAKARTA
2025**

**SISTEM PREDIKSI TINGKAT KUALITAS UDARA DI JAKARTA
TIMUR BERBASIS *WEBSITE* MENGGUNAKAN ALGORITMA
*RANDOM FOREST REGRESSION***

**ANGELIA
NIM. 2110512110**

Proposal Skripsi
sebagai salah satu syarat untuk melaksanakan
penelitian oleh mahasiswa pada
Program Studi S1 Sistem Informasi

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
JAKARTA
2025**

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Angelia

NIM : 2110512110

Tanggal : 4 Juli 2025

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan persyaratan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 4 Juli 2025

Yang Menyatakan,



Angelia

PERNYATAAN PERSETUJUAN PUBLIKASI

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Angelia

NIM : 2110512110

Fakultas : Ilmu Komputer

Program Studi : SI Sistem Informasi

Demi Pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**Sistem Prediksi Tingkat Kualitas Udara Di Jakarta Timur Berbasis Website
Menggunakan Algoritma Random Forest Regression**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti di Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formalitas, mengelola dalam bentuk pangkalan data (Basis Data), merawat dan mempublikasi Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 4 Juli 2025

Yang Menyatakan,



Angelia

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

Yang bertanda tangan di bawah ini:

Nama : Angelia
NIM. : 2110512110
Program Studi : ~~Informatika~~/Sistem Informasi Program Sarjana/~~Diploma—3~~
(*Coret yang tidak perlu)
Judul Tugas Akhir : Sistem Prediksi Tingkat Kualitas Udara di Jakarta Timur
Berbasis *Website* Menggunakan Algoritma *Random Forest Regression*

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti ujian sidang Tugas Akhir.

Jakarta, 16 April 2025

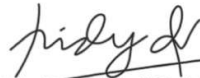
Menyetujui,

Dosen Pembimbing I,

Dosen Pembimbing II,



Andhika Octa Indarso, M.MSI



Nindy Irzavika, S.SI., M.T.

Mengetahui,

Koordinator Program Studi,



Anita Muliawati, S.Kom., MTI.

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Sistem Prediksi Tingkat Kualitas Udara di Jakarta Timur Berbasis
Website Menggunakan Algoritma *Random Forest Regression*
Nama : Angelia
NIM : 2110512110
Program Studi : SI Sistem Informasi

Disetujui oleh :

Penguji 1:

I Wayan Widi Pradnyana, M.TI

Penguji 2:

Mohamad Bayu Wibisono, S.Kom., M.M.

Pembimbing 1:

Andhika Octa Indarso, M.MSI

Pembimbing 2:

Nindy Irzavika, S.SI., M.T.

Diketahui oleh:

Koordinator Program Studi:

Anita Muliawati, S.Kom., MTL.

NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer:

Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM

NIP. 197605082003121002

Tanggal Ujian Tugas Akhir:

12 Juni 2025



ABSTRAK

Kualitas udara yang semakin memburuk, khususnya di wilayah Jakarta Timur, menjadi isu lingkungan yang signifikan karena berpotensi membahayakan kesehatan manusia dan ekosistem. Kurangnya informasi yang dapat diakses masyarakat mengenai kondisi kualitas udara serta belum tersedianya sistem prediksi yang akurat menjadi hambatan dalam mitigasi polusi. Oleh karena itu, penelitian ini bertujuan untuk membangun sistem prediksi kualitas udara berbasis *website* menggunakan algoritma *Random Forest Regression*. Sistem ini dikembangkan berdasarkan data Indeks Standar Pencemar Udara (ISPU) dan dirancang agar dapat memberikan hasil prediksi yang akurat, mudah dipahami, dan bermanfaat dalam pengambilan keputusan. Proses penelitian meliputi tahapan pengumpulan data ISPU, praproses data, pelatihan model, evaluasi menggunakan metrik RMSE, MAE, dan R^2 , serta pembangunan antarmuka sistem berbasis Streamlit. Hasil prediksi kemudian diklasifikasikan ke dalam kategori ISPU sesuai dengan regulasi PERMEN LHK No.14/MENLHK/SETJEN/KUM.1/7/2020. Sistem ini telah diuji menggunakan *scenario testing* dan menunjukkan bahwa seluruh fitur berjalan dengan baik dan sesuai kebutuhan pengguna, khususnya bagi Dinas Lingkungan Hidup Jakarta. Sistem ini ditujukan untuk mendukung pengambilan keputusan berbasis data dalam pemantauan dan evaluasi kualitas udara oleh pihak yang berwenang.

Kata Kunci: dinas lingkungan hidup jakarta, indeks standar pencemar udara, prediksi, *random forest regression*, *website*

ABSTRACT

The deteriorating air quality, especially in East Jakarta, poses a serious environmental concern due to its negative impacts on human health and the environment. The lack of accessible information for the public and the absence of accurate forecasting systems remain major obstacles to air pollution mitigation. This study aims to develop a web-based air quality prediction system using the Random Forest Regression algorithm. The system is built on Air Pollution Standard Index (ISPU) data and designed to provide accurate and user-friendly predictions that support informed decision-making. The research includes stages such as data collection, preprocessing, model training, evaluation using RMSE, MAE, and R^2 metrics, and system development with the Streamlit framework. The predicted results are classified into ISPU categories according to the Indonesian Ministry of Environment and Forestry Regulation No.14/MENLHK/SETJEN/KUM.1/7/2020. The system has been tested through scenario-based evaluation, confirming that all features function correctly and meet user needs, particularly for the Jakarta Environmental Agency. This system is intended to support data-driven decision-making in monitoring and evaluating air quality by the relevant authorities.

Keywords: jakarta environmental agency, air pollution standard index, prediction, random forest regression, website

KATA PENGANTAR

Puji syukur peneliti panjatkan atas kehadiran Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga peneliti dapat menyelesaikan skripsi yang berjudul “Sistem Prediksi Tingkat Kualitas Udara di Jakarta Timur Berbasis *Website* Menggunakan Algoritma *Random Forest Regression*” dengan baik dan tepat waktu. Skripsi Tugas Akhir ini sebagai prasyarat dalam pengambilan Tugas Akhir atau Skripsi.

Dalam proses penyusunan skripsi ini, peneliti menyadari bahwa tidak akan dapat menyelesaikannya tanpa bantuan, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, peneliti ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Mama beserta keluarga yang senantiasa memberikan doa dan dukungan baik secara moral maupun material.
2. Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM selaku dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
3. Ibu Anita Muliawati, S.Kom., MTI selaku Koordinator Program Studi Sarjana Jurusan Sistem Informasi Fakultas Ilmu Komputer UPN Veteran Jakarta
4. Bapak Dr. Tjahjanto, S.Kom., M.M. selaku Dosen Pembimbing Akademik Program Studi Sistem Informasi Program Sarjana..
5. Bapak Andhika Octa Indarso, M.MSI selaku dosen pembimbing I yang telah memberikan bimbingan, arahan, serta masukan yang sangat berarti selama proses penyusunan tugas akhir ini.
6. Ibu Nindy Irzavika, S.SI., M.T. selaku dosen pembimbing II yang telah dengan sabar membimbing dan memberikan kontribusi pemikiran yang berharga dalam penyusunan tugas akhir ini.
7. Ibu Nofi Rahmawati Samputra selaku Staff Subkelompok Pemantauan Kualitas Lingkungan dan seluruh staff Dinas Lingkungan Hidup Jakarta atas izin, data, waktu pada kontribusinya di penelitian ini.
8. Seluruh Pengurus KSM Cyber Security 2024 yang telah menjadi wadah bagi penulis untuk bertumbuh, berbagi pengalaman, serta berkembang dalam memperluas relasi dan mengasah *soft skill* selama masa perkuliahan.
9. Seluruh sahabat peneliti yang senantiasa memberikan dukungan, semangat, dan kebersamaan selama perkuliahan hingga proses penyusunan skripsi, terutama teman-teman sepersirkelan yaitu Arsi, Frida, dan Risma, rekan satu dosen pembimbing Renatha dan Agung, serta teman seperjuangan dalam menjalani skripsi bersama Keisha, Brigita, Avif, dan Gathan, yang seluruh kehadirannya sangat berarti dalam setiap langkah perjalanan ini.

10. Zayne, Dawnbreaker, Hua Cheng, Sylus, Caleb, Zhongli, dan Xie Lian yang telah hadir dan menjadi sumber hiburan, inspirasi, penenang di kala peneliti merasa lelah atau kehilangan semangat dalam menyelesaikan skripsi ini, serta telah menjadi ruang pelarian yang menyenangkan ketika dunia nyata terasa terlalu riuh.
11. Serta semua pihak yang tidak dapat peneliti sebutkan satu per satu tanpa mengurangi rasa hormat.

Semua dukungan dan bantuan yang telah diberikan menjadi pendorong bagi peneliti untuk terus berkembang karena peneliti menyadari bahwa masih banyak kekurangan dari Skripsi Tugas Akhir ini, baik dari penulisan maupun materi, mengingat terbatasnya pengetahuan dan pengalaman peneliti. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan untuk perbaikan di masa yang mendatang.

DAFTAR ISI

PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI	iii
LEMBAR PERSETUJUAN.....	iv
LEMBAR PENGESAHAN	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xv
DAFTAR SIMBOL.....	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	3
1.6 Luaran yang Diharapkan	4
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kajian Teori.....	6
2.1.1 Polusi Udara	6
2.1.2 Indeks Standar Pencemar Udara (ISPU)	7
2.1.3 Prediksi	9
2.1.4 Data Mining.....	10
2.1.5 Random Forest.....	11
2.1.6 Python.....	12
2.1.7 Evaluasi Model.....	13
2.1.8 Website	14
2.1.9 Streamlit	15

2.1.10	<i>Scenario Testing</i>	15
2.2	Penelitian Terdahulu	16
BAB III METODOLOGI PENELITIAN		20
3.1	Alur Penelitian	20
3.2	Tahapan Penelitian	21
3.2.1	Identifikasi Masalah	21
3.2.2	Persiapan Data	21
3.2.3	Pemahaman Data	21
3.2.4	Pre-Processing Data	22
3.2.5	Pengolahan Data	22
3.2.6	Implementasi Algoritma Random Forest Regression	23
3.2.7	Evaluasi Model	23
3.2.8	Implementasi Sistem	23
3.2.9	Pengujian Sistem	23
3.3	Alat dan Bahan	23
3.4	Jadwal Penelitian	24
BAB IV HASIL DAN PEMBAHASAN		26
4.1	Dataset	26
4.2	<i>Data Pre-Processing</i>	28
4.2.1	<i>Data Cleansing</i>	29
4.2.2	Pembagian Data	32
4.3	Implementasi Algoritma <i>Random Forest Regression</i>	33
4.4	Evaluasi Model	34
4.4.1	Rasio 80 : 20	35
4.4.2	Rasio 70 : 30	36
4.5	Implementasi Sistem Prediksi	36
4.5.1	Halaman Awal	37
4.5.2	Menu Input Manual	38
4.5.3	Menu Unggah CSV	40
4.5.4	Kategori ISPU	44
4.6	Pengujian Sistem	45
BAB V PENUTUP		48

5.1	Kesimpulan.....	48
5.2	Saran	48
DAFTAR PUSTAKA		50
RIWAYAT HIDUP.....		54
LAMPIRAN.....		55

DAFTAR TABEL

Tabel 1 Daftar Simbol.....	xvi
Tabel 2.1 Waktu Pengukuran Parameter-Parameter Dasar Indeks Standar Pencemar Udara (ISPU).....	8
Tabel 2.2 Rentang Kelas Indeks Standar Pencemar Udara (ISPU)	8
Tabel 2.3 Batas Setiap Parameter Indeks Standar Pencemar Udara (ISPU).....	9
Tabel 2.4 Penelitian Terdahulu	16
Tabel 3.1 Jadwal Rencana Penelitian.....	25
Tabel 4.1 Daftar Atribut yang Ada pada Dataset.....	26
Tabel 4.2 <i>Source Code</i> Pengisian <i>Missing Value</i>	31
Tabel 4.3 Rasio Pembagian <i>Data Training</i> dan <i>Data Testing</i>	32
Tabel 4.4 <i>Source Code</i> Pembagian Data.....	33
Tabel 4.5 Parameter pada Algoritma Random Forest.....	33
Tabel 4.6 Source Code Implementasi Algoritma <i>Random Forest Regression</i>	34
Tabel 4.7 Hasil Evaluasi Model dengan Rasio 80:20	35
Tabel 4.8 Hasil Evaluasi Model dengan Rasio 70:30	36
Tabel 4.9 <i>Scenario Testing</i>	45

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	20
Gambar 4.1 Dataset awal	30
Gambar 4.2 Dataset Setelah Penggabungan.....	30
Gambar 4.3 <i>Heat Map Correlation Matrix</i>	32
Gambar 4.4 Perbandingan Data Aktual dan Prediksi dari Model 80:20 dengan Grafik Matplotlib	35
Gambar 4.5 Perbandingan Data Aktual dan Prediksi dari Model 70:30 dengan Grafik Matplotlib	36
Gambar 4.6 Tampilan Halaman Awal pada Sistem Prediksi berbasis <i>Website</i>	37
Gambar 4.7 Tampilan Menu Input Manual pada Sistem	38
Gambar 4.8 Contoh Data Penginputan.....	39
Gambar 4.9 Tampilan Sistem Setelah Melakukan Prediksi.....	39
Gambar 4.10 Tampilan Sistem Setelah Memasukan Nilai Aktual	40
Gambar 4.11 Tampilan Menu Unggah CSV pada Sistem	41
Gambar 4.12 Contoh Data CSV Tanpa Nilai Max	42
Gambar 4.13 Contoh Data CSV Dengan Nilai Max.....	42
Gambar 4.14 Output Menu Unggah CSV Tanpa Nilai Aktual	42
Gambar 4.15 Output Menu Unggah CSV Dengan Nilai Aktual Pertama.....	43
Gambar 4.16 Output Menu Unggah CSV Dengan Nilai Aktual Kedua	43
Gambar 4.17 Tampilan Tabel Kategori ISPU pada Sistem	44

DAFTAR LAMPIRAN

Lampiran 1 Surat Permohonan Riset Mahasiswa	55
Lampiran 2 Lembar Dokumentasi Hasil Wawancara	56
Lampiran 3 Dokumentasi Wawancara	60
Lampiran 4 Dataset awal ISPU Tahun 2022	60
Lampiran 5 Dataset Awal ISPU Tahun 2023	61
Lampiran 6 Source Code model prediksi Random Forest Regression untuk Prediksi Tingkat Kualitas Udara di Jakarta Timur	61
Lampiran 7 Source Code Website Prediksi	68
Lampiran 8 Dokumentasi <i>testing</i> kepada pihak DLH Jakarta	72
Lampiran 9 Foto Stasiun Pemaantau Kualitas Udara	72
Lampiran 10 <i>User Acceptance Test</i> Responden 1	73
Lampiran 11 <i>User Acceptance Test</i> Responden 2	75
Lampiran 12 Hasil <i>Scenario Testing</i> Responden 1	77
Lampiran 13 Hasil <i>Scenario Testing</i> Responden 2	79
Lampiran 14 Surat Pemberian Izin Ttg Permintaan Data dan Wawancara	81
Lampiran 15 Hasil Turnitin.....	83

DAFTAR SIMBOL

Tabel 1 Daftar Simbol

No.	Simbol	Nama	Fungsi
1.		Terminator (Start/End)	Menandakan awal atau akhir pada <i>flowchart</i> .
2.		Process	Mewakili proses atau langkah yang dilakukan.
3.		Flowline	Menunjukkan arah aliran proses.
4.		Decision	Digunakan untuk percabangan logika, contohnya seperti "Ya" atau "Tidak."
5.		Input/Output	Digunakan untuk menandakan data yang masuk atau keluar.
6.		Document	Mewakili dokumen atau laporan.
7.		Grouping	Menandakan kelompok proses dalam <i>flowchart</i> .