

**IMPLEMENTASI RANDOM OVERSAMPLING PADA
ALGORITMA RANDOM FOREST UNTUK KLASIFIKASI
INDEKS STANDAR UDARA DI DKI JAKARTA 2024**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

HARSEN AARGRETYA

22.11.4958

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**IMPLEMENTASI RANDOM OVERSAMPLING PADA
ALGORITMA RANDOM FOREST UNTUK KLASIFIKASI INDEKS
STANDAR UDARA DI DKI JAKARTA 2024**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

HARSENA ARGRETYA

22.11.4958

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**IMPLEMENTASI RANDOM OVERSAMPLING PADA
ALGORTIMA RANDOM FOREST UNTUK KLASIFIKASI INDEKS
STANDAR UDARA DI DKI JAKARTA 2024**

yang disusun dan diajukan oleh

Harsena Argretya

22.11.4958

telah disetujui oleh Dosen Pembimbing
pada tanggal 30 Januari 2026

Dosen Pembimbing,



Kusnawi, S.Kom, M.Eng
NIK. 19030202112

HALAMAN PENGESAHAN

JALUR NON-REGULER

**IMPLEMENTASI RANDOM OVERSAMPLING PADA
ALGORTIMA RANDOM FOREST UNTUK KLASIFIKASI INDEKS
STANDAR UDARA DI DKI JAKARTA 2024**

yang disusun dan diajukan oleh

Harsena Argretya
22.11.4958

Telah dipertahankan di depan Dewan Penguji
pada tanggal 26 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dr. Hartatik, S.T., M.Cs.
NIK. 190302232



Windha Mega Pradnya Dhubita, S.Kom., M.Kom.
NIK. 190302185



Kusnawi, S.Kom., M.Eng.
NIK. 19030202112



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 26 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.kom
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Harsena Argretya

NIM : 22.11.4958

Menyatakan bahwa Laporan dengan judul berikut:

Implementasi Random Oversampling pada Algoritma Random Forest untuk Klasifikasi Indeks Standar Udara di DKI Jakarta 2024

Dosen Pembimbing : Kusnawi, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 26 Januari 2026

Yang Menyatakan,



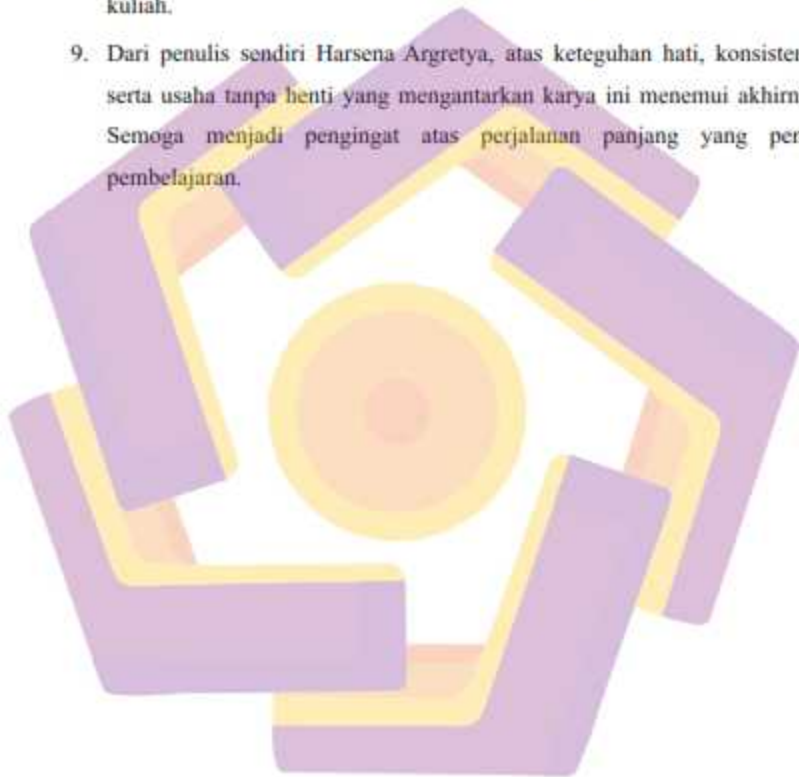
Harsena Argretya

HALAMAN PERSEMBAHAN

Puji syukur kehadirat Tuhan yang Maha Esa. Dengan izin-Nya, penulis dapat menyelesaikan Tugas Akhir ini sebagai wujud tanggung jawab akademik sekaligus langkah akhir dalam menempuh pendidikan sarjana. Kehadiran karya ini menjadi bukti perjalanan panjang dan penuh doa, usaha, dan pengorbanan. Dengan penuh rasa syukur, penulis mempersembahkan karya ini kepada:

1. Segala puji dan terima kasih penulis ucapkan kepada Tuhan Yang Maha Esa atas limpahan rahmat, kesehatan, serta kemudahan yang diberikan sehingga penulis dapat menyelesaikan penelitian ini.
2. Artoyo Joko Suharjo sebagai ayah dan Miming Nofianti sebagai ibu tercinta, yang menjadi pilar utama kekuatan penulis. Terima kasih atas kasih sayang tak terbatas, doa restu tanpa henti, dorongan materi dan semangat, serta segala pengorbanan untuk mendukung perjalanan hidup dan pendidikan penulis.
3. Hermintyas Febularina, adik perempuan tercinta, yang tak pernah lelah memberikan motivasi, doa, dan penyemangat di setiap tahapan perjuangan penulis.
4. Bapak Kusnawi, S.Kom, M.Eng selaku Dosen Pembimbing Kolaborasi Riset, atas bimbingan penuh kesabaran, arahan yang teliti, dan dorongan yang sangat berarti hingga tugas akhir ini dapat terselesaikan.
5. Bapak Majid Rahard, S.kom, M.Eng, selaku Dosen Wali, atas perhatian, arahan, dan dukungan yang konsisten selama masa perkuliahan.
6. Najwa Ar-Rumy Zabine Wahab, sosok yang Tuhan hadirkan sebagai penguat langkah penulis. Terima kasih telah menjadi tempat pulang di tengah lelah, menjadi tenang saat pikiran berisik, dan menjadi semangat saat penulis hampir menyerah. Doa, dukungan, kesabaran, serta kehadiranmu adalah cahaya kecil yang selalu menuntun penulis untuk tetap bertahan dan menyelesaikan tugas akhir ini.

7. Teman-teman dari kendang, yang telah menjadi ruang berbagi tawa, dukungan, dan kebersamaan yang memperkaya perjalanan hidup serta studi penulis.
8. Teman-teman 22 IF 07, atas solidaritas, kerja sama, serta atmosfer positif yang senantiasa menguatkan selama menempuh pendidikan di bangku kuliah.
9. Dari penulis sendiri Harsena Argretya, atas keteguhan hati, konsistensi, serta usaha tanpa henti yang mengantarkan karya ini menemui akhirnya. Semoga menjadi pengingat atas perjalanan panjang yang penuh pembelajaran.



KATA PENGANTAR

Dengan memanjatkan puji syukur ke hadirat Tuhan Yang Maha Esa, penulis menyampaikan rasa terima kasih atas limpahan rahmat dan karunia-Nya sehingga laporan tugas akhir ini dapat diselesaikan dengan baik. Penyusunan laporan ini tidak terlepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak yang telah memberikan kontribusi berarti. Oleh karena itu, dengan penuh hormat penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Kusnawi, S.Kom., M.Eng., selaku Dosen Pembimbing, atas bimbingan, arahan, serta motivasi yang senantiasa diberikan hingga terselesaikannya penelitian ini.
4. Bapak Artoyo Joko Suharjo dan Ibu Miming Nofianti, selaku orang tua penulis, atas cinta, kasih sayang, doa, dukungan, dan pengorbanan yang tidak ternilai harganya.

Akhir kata, penulis menyadari bahwa laporan ini masih memiliki kekurangan dan keterbatasan. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun demi penyempurnaan karya ini pada masa yang akan datang. Semoga laporan ini dapat memberikan manfaat bagi pembaca serta menjadi kontribusi kecil dalam mendukung pengembangan ilmu pengetahuan.

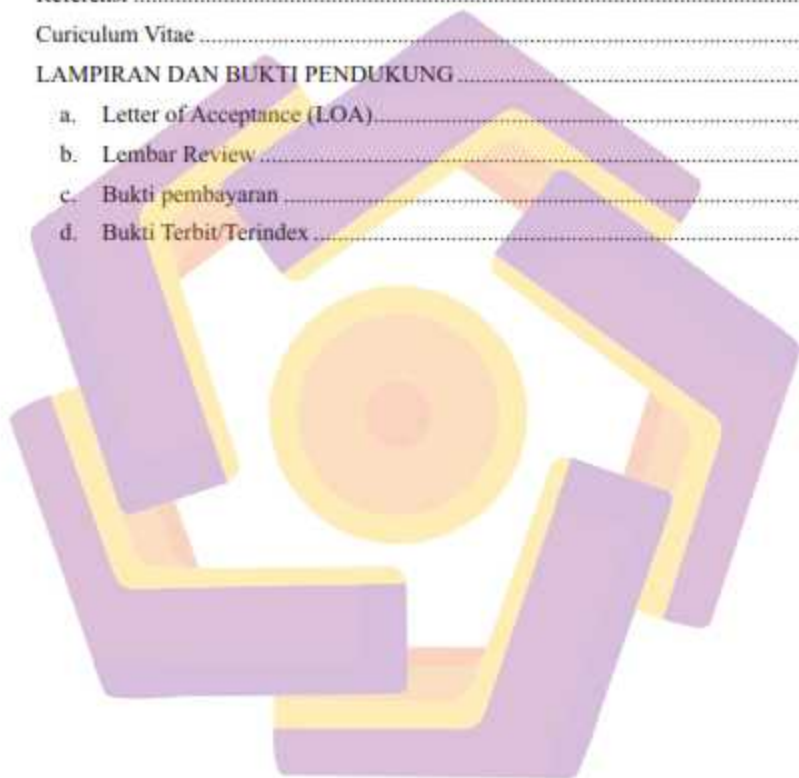
Yogyakarta, 6 Desember 2025

Penulis

DAFTAR ISI

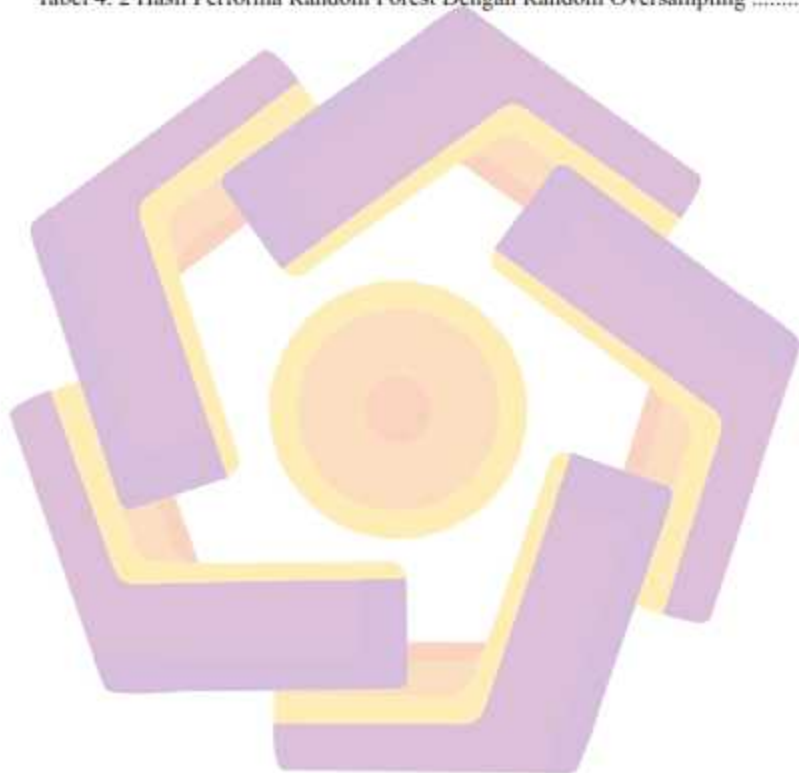
Halaman Judul.....	i
Halaman Persetujuan.....	ii
Halaman Pengesahan	iii
Halaman Pernyataan Keaslian Karya.....	iv
Halaman Persembahan	v
Kata Pengantar	vii
Daftar Isi.....	viii
Daftar Tabel.....	x
Daftar Gambar.....	xi
Daftar Lambang dan Singkatan.....	xii
Daftar Istilah.....	xiii
Intisari	xiv
<i>Abstract</i>	xv
Bab I Pendahuluan	1
1.1. Gambaran Umum	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	3
Bab II TINJAUAN PUSTAKA	4
2.1. Studi Literatur	4
2.2. Landasan Teori	5
BAB III METODE PENELITIAN	9
3.1. Metode.....	9
3.1.1. Identifikasi Masalah.....	10
3.1.2 Data Collection	10
3.1.3 Data Pre-processing	11
3.1.4 Data Preparation	11
3.1.5 Machine Learning Modelling	13
3.1.6 Machine Learning Model Evaluation	14
3.1.7 Deployment.....	16

BAB IV HASIL.....	18
4.1.Hasil	18
BAB V KESIMPULAN	23
5.1.Kesimpulan	23
5.2.Saran.....	24
Referensi	26
Curiculum Vitae	28
LAMPIRAN DAN BUKTI PENDUKUNG	29
a. Letter of Acceptance (LOA).....	29
b. Lembar Review	30
c. Bukti pembayaran	33
d. Bukti Terbit/Terindex	34



DAFTAR TABEL

Tabel 2. 1 Kategori dan Batas ISPU.....	6
Tabel 3. 1 Dataset.....	10
Tabel 3. 2 Confusion Matrix	15
Tabel 4. 1 Hasil Performa Random Forest Tanpa Random Oversampling	19
Tabel 4. 2 Hasil Performa Random Forest Dengan Random Oversampling	20



DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian	9
Gambar 3. 2 Visualisasi algoritma RF	14
Gambar 3. 3 Arsitektur Sistem Berbasis Next js dan Flask	17
Gambar 4. 1 Ilustrasi distribusi kelas sebelum dan sesudah penerapan ROS	19
Gambar 4. 2 Tampilan Website Sistem Prediksi ISPU	21
Gambar 4. 3 Visualisasi waktu respons API dengan Postman.	22



DAFTAR LAMBANG DAN SINGKATAN



ISPU	Indeks Standar Pencemar Udara
PM10	Particulate Matter ukuran $\leq 10 \mu\text{m}$
PM2.5	Particulate Matter ukuran $\leq 2.5 \mu\text{m}$
NO ₂	Nitrogen dioksida
SO ₂	Sulfur dioksida
CO	Karbon monoksida
O ₃	Ozon
NCR	<i>Neighborhood Cleaning Rule</i>
RF	<i>Random Forest</i>
KNN	<i>K-Nearest Neighbors</i>
XGBoost	<i>eXtreme Gradient Boosting</i>
AI	<i>Artificial Intelligence</i>
API	<i>Application Programming Interface</i>
TP	<i>True Positive</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
FN	<i>False Negative</i>

DAFTAR ISTILAH



<i>Machine Learning</i>	Pemodelan otomatis berdasarkan pola data
<i>Artificial Intelligence</i>	Teknologi yang meniru kecerdasan manusia
<i>Random Forest</i>	Algoritma ensemble berbasis banyak decision tree
<i>Decision Tree</i>	Model keputusan berbentuk pohon
<i>Ensemble Learning</i>	Penggabungan beberapa model untuk hasil lebih baik
<i>Random Oversampling</i>	Duplikasi data kelas minoritas untuk seimbang
<i>Oversampling</i>	Penambahan data pada kelas minoritas
<i>Undersampling</i>	Pengurangan data pada kelas mayoritas
<i>Dataset</i>	Kumpulan data penelitian
<i>Imputation</i>	Pengisian nilai yang hilang
<i>Training Data</i>	Data untuk melatih model
<i>Testing Data</i>	Data untuk menguji model
<i>Stratified Splitting</i>	Pembagian data dengan proporsi kelas tetap
<i>Modelling</i>	Proses membangun model
<i>Bootstrapping</i>	Pengambilan sampel acak dengan pengembalian
<i>Overfitting</i>	Model terlalu cocok pada data latih
<i>Underfitting</i>	Model terlalu sederhana
<i>Entropy</i>	Ukuran ketidakpastian data
<i>Noise</i>	Data tidak sesuai pola
<i>Time To First Byte</i>	Waktu server memberikan byte pertama

INTISARI

Polusi udara merupakan permasalahan lingkungan yang berdampak luas terhadap kesehatan masyarakat dan ekosistem, terutama di wilayah perkotaan dengan tingkat kepadatan penduduk yang tinggi seperti DKI Jakarta. Tantangan utama dalam analisis kualitas udara tidak hanya terletak pada tingginya konsentrasi polutan, tetapi juga pada ketidakseimbangan kelas dalam data ISPU yang menyebabkan hasil klasifikasi menjadi kurang akurat. Data ISPU umumnya didominasi oleh kategori "Sedang", sedangkan kategori "Baik" dan "Tidak Sehat" tercatat jauh lebih sedikit sehingga model cenderung bias terhadap kelas mayoritas. Penelitian ini menerapkan teknik Random Oversampling (ROS) yang dikombinasikan dengan algoritma Random Forest untuk meningkatkan akurasi klasifikasi ISPU menggunakan data tahun 2024. Tahapan penelitian meliputi pengumpulan data dari portal Satu Data Jakarta, preprocessing dengan metode imputasi KNN, penyeimbangan kelas menggunakan ROS, pelatihan model Random Forest, serta evaluasi performa melalui metrik akurasi, presisi, recall, dan F1-score. Hasil eksperimen menunjukkan bahwa penggunaan ROS mampu meningkatkan akurasi model dari 96,71% menjadi 97,07%. Peningkatan juga terlihat pada nilai recall kelas "Tidak Sehat", yang naik dari 0,93 menjadi 0,94, sementara recall untuk kelas "Baik" tetap berada pada nilai 0,88. Meskipun peningkatannya relatif kecil, ROS memberikan proses pembelajaran yang lebih seimbang dan meningkatkan stabilitas model. Penelitian ini juga menegaskan keterbatasan ROS yang hanya menggandakan data tanpa menghasilkan variasi baru, sehingga metode lain seperti SMOTE atau ADASYN direkomendasikan untuk penelitian selanjutnya. Secara keseluruhan, integrasi Random Forest dan ROS mendukung pengembangan sistem pemantauan kualitas udara berbasis data untuk pengambilan keputusan lingkungan di wilayah perkotaan.

Kata kunci: Kualitas udara, Klasifikasi ISPU, Random Forest, Random Oversampling, Ketidakseimbangan Data

ABSTRACT

Air pollution remains one of the most pressing global environmental issues, posing serious risks to public health and ecosystems, especially in densely populated metropolitan areas such as DKI Jakarta. The difficulty in analyzing air quality data arises not only from high pollutant concentrations but also from the inherent class imbalance within the Air Pollution Standard Index (ISPU) dataset. Most samples fall within the "Moderate" category, while the "Good" and "Unhealthy" classes are underrepresented, leading machine learning models to develop a bias toward the majority class. This study applies the Random Oversampling (ROS) technique in combination with the Random Forest algorithm to improve the classification performance of ISPU categories using 2024 air quality data. The workflow consists of data acquisition from the SatuData Jakarta platform, preprocessing using KNN imputation, class balancing through ROS, model training using Random Forest, and performance evaluation using accuracy, precision, recall, and F1-score. Experimental results indicate that ROS enhances the overall classification accuracy from 96.71% to 97.07%. An improvement is also observed in the recall for the "Unhealthy" class, increasing from 0.93 to 0.94, while the recall for the "Good" class remains at 0.88. Although the performance gain is modest, the implementation of ROS leads to a more balanced learning process and improved model stability. The study also highlights a limitation of ROS, namely its tendency to duplicate existing samples without generating additional variation, suggesting that more advanced oversampling techniques such as SMOTE or ADASYN may be considered in future research. Overall, the integration of Random Forest and ROS strengthens data-driven air quality monitoring and supports environmental decision-making in urban settings.

Keywords: Air Quality, ISPU Classification, Random Forest, Random Oversampling, Class Imbalance