



**PERBANDINGAN RANDOM FOREST, SVM DAN XGBOOST DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA AKIBAT
KEBAKARAN HUTAN BERBASIS DATA SENTINEL-5P DI
KALIMANTAN**

**TUGAS AKHIR
SKRIPSI**

Diaz Marlow Buana
41522010054

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**PERBANDINGAN RANDOM FOREST, SVM DAN XGBOOST DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA AKIBAT
KEBAKARAN HUTAN BERBASIS DATA SENTINEL-5P DI
KALIMANTAN**

**TUGAS AKHIR
SKRIPSI**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

**UNIVERSITAS
Diaz Marlow Buana
41522010054
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Diaz Marlow Buana
NIM : 41522010054
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

“Perbandingan Random Forest, SVM dan XGBoost Dalam Klasifikasi Tingkat Pencemaran Udara Akibat Kebakaran Hutan Berbasis Sentinel-5P di Kalimantan” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 10 Agustus 2026

UNIVERSITAS
MERCU BUANA



Diaz Marlow Buana

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name	: Diaz Marlow Buana
NIM /Student ID Number	: 41522010054
Program Studi /Study Program	: Teknik Informatika

Dengan Judul Tugas Akhir

/The title:

“PERBANDINGAN RANDOM FOREST, SVM DAN XGBOOST DALAM KLASIFIKASI TINGKAT PENCEMARAN UDARA AKIBAT KEBAKARAN HUTAN BERBASIS DATA SENTINEL-5P DI KALIMANTAN”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

23 Juli 2026

dengan nilai persentase sebesar :

/with a percentage value of:

24%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

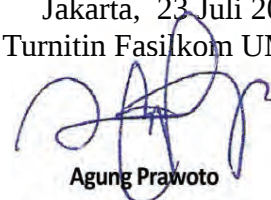
File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/168Um3JM4jMefQicL6kPouOI13wWBff/_view?usp=drive_link



Jakarta, 23 Juli 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

HALAMAN PENGESAHAN

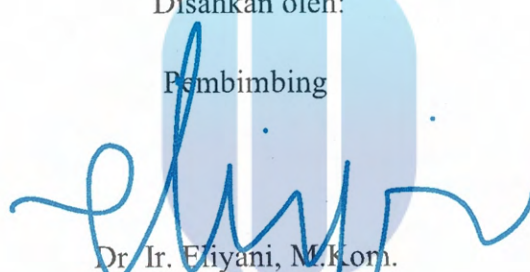
Tugas Akhir ini diajukan oleh:

Nama : Diaz Marlow Buana
NIM : 41522010054
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir : Perbandingan Random Forest, SVM dan XGBoost
Dalam Klasifikasi Tingkat Pencemaran Udara Akibat
Kebakaran Hutan Berbasis Sentinel-5P di Kalimantan

Telah berhasil dipertahankan pada sidang tanggal 10 Agustus 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing



Dr. Ir. Eliyani, M.Kom.

NIDN/NUPTK: 0321026901

UNIVERSITAS

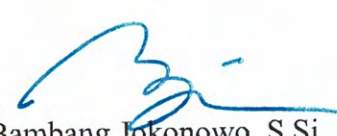
Jakarta, 10 Agustus 2026

Mengetahui,

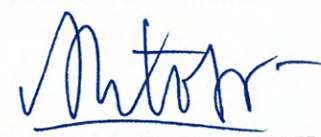
MERCU BUANA

Dekan Fakultas Ilmu Komputer

Ketua Program Studi
Teknik Informatika



Dr. Bambang Jokonowo, S.Si., MTI
NIDN/NUPTK: 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Dr. Ir. Eliyani, M.Kom. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Bersyukur kepada Allah SWT, yang telah memberikan saya rezeki untuk dapat berkuliah memperoleh gelar sarjana. Semoga memperolehnya ilmu selama di perguruan tinggi ini dapat memberikan manfaat untuk orang banyak, dan semoga kerja keras menuntut ilmu selama saya hidup ini dapat dijadikan pahala yang sebesar-besarnya untuk memperoleh surga-Nya.
6. Untuk Ayah, Mama dan kedua Abang saya. Terima kasih sudah memberikan saya kesempatan untuk berkuliah, terima kasih kerja kerasnya untuk membiayai anak dan adik paling kecil ini agar memperoleh gelar sarjana, terima kasih selalu mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana. Saya selalu doakan untuk kalian agar usaha dan upaya diganti berkali-kali lipat oleh Allah SWT.
7. Teruntuk teman teman Teknik Informatika 22. Terimakasih sudah memberikan warna selama masa kuliah yang kita jalani bersama-sama. Kalian selalu memberikan canda tawa, susah senang bersama-sama,

kenangan manis yang kalian berikan selama 4 tahun kita jalani. Terima kasih atas kerja keras kalian berkomitmen untuk lulus tepat waktu bersama-sama. Kedepannya semoga kita berkumpul dalam keadaan sukses dikehidupan yang akan mendatang dan meraih cita cita yang kalian inginkan.

8. Untuk saya Diaz Marlow Buana, terima kasih sudah menyelesaikan tanggung jawab sebagai mahasiswa yang diberikan oleh orang tua dengan hasil yang bagus, tanpa kendala apapun dan lulus tepat waktu. Saya akan berusaha kedepannya untuk membahagiakan kedua Orang Tua saya, Abang, Sagara, Biru dan ponakan saya yang lucu dan menggemaskan nantinya.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 10 Agustus 2026

UNIVERSITAS
MERCU BUANA
Diaz Marlow Buana

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Diaz Marlow Buana
NIM : 41522010054
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi : Perbandingan Random Forest, SVM dan XGBoost Dalam Klasifikasi Tingkat Pencemaran Udara Akibat Kebakaran Hutan Berbasis Sentinel-5P di Kalimantan

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 10 Agustus 2026

Yang menyatakan,



Diaz Marlow Buana

**PERBANDINGAN RANDOM FOREST, SVM DAN XGBOOST DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA AKIBAT
KEBAKARAN HUTAN BERBASIS DATA SENTINEL-5P DI
KALIMANTAN**

DIAZ MARLOW BUANA

ABSTRAK

Penelitian ini membahas klasifikasi tingkat kualitas udara di Kalimantan dengan pendekatan indeks kualitas udara analog ISPU berbasis data penginderaan jauh dan algoritma pembelajaran mesin. Data yang digunakan berupa grid spasial yang memuat konsentrasi CO, NO₂, indeks aerosol (AI) dari Sentinel-5P, serta informasi hotspot kebakaran hutan dan curah hujan rata-rata per kabupaten/kota untuk periode 2023–2024 yang diproses melalui pembersihan, penggabungan, normalisasi, dan pembentukan indeks pencemaran gabungan CO–NO₂ sebagai dasar empat kelas kualitas udara (Baik, Sedang, Tidak Sehat, dan Sangat Tidak Sehat). Model klasifikasi dikembangkan menggunakan algoritma Random Forest, Support Vector Machine (SVM), dan XGBoost dengan skema pembagian data latih dan uji yang seimbang, kemudian dievaluasi menggunakan confusion matrix, akurasi, precision, recall, dan macro F1-score, sekaligus dianalisis pentingnya masing-masing fitur untuk mengidentifikasi variabel yang paling berpengaruh terhadap kategori kualitas udara. Hasil yang diperoleh menunjukkan bahwa kombinasi fitur CO, NO₂, indeks aerosol, hotspot, dan curah hujan mampu menghasilkan performa klasifikasi yang tinggi. Dengan XGBoost sebagai model dengan kinerja terbaik, serta mengonfirmasi perbedaan karakteristik ketiga algoritma dalam menangani pola data spasial kualitas udara. Temuan ini diharapkan dapat menjadi dasar pengembangan sistem pemantauan kualitas udara berbasis data satelit dan pembelajaran mesin sebagai dukungan pengambilan keputusan dalam mitigasi dampak kebakaran hutan dan kabut asap di Kalimantan.

Kata kunci: kualitas udara, kebakaran hutan, Sentinel-5P, hotspot, indeks pencemaran, curah hujan, Random Forest, SVM, XGBoost.

**COMPARISON OF RANDOM FOREST, SVM AND XGBOOST IN
CLASSIFYING AIR POLLUTION LEVELS CAUSED BY FOREST FIRES
USING SENTINEL-5P DATA IN KALIMANTAN**

DIAZ MARLOW BUANA

ABSTRACT

This study investigates the classification of air quality levels in Kalimantan using an ISPU-like air quality index derived from remote sensing data and machine learning algorithms. The dataset consists of spatial grids containing CO and NO₂ concentrations, Aerosol Index (AI) from Sentinel-5P, hotspot information from forest and land fires, and average rainfall at the regency/city level for the 2023–2024 period, which are preprocessed through data cleaning, merging, normalization, and the construction of a combined CO–NO₂ pollution index as the basis for four air quality classes (Good, Moderate, Unhealthy, and Very Unhealthy). Classification models are developed using Random Forest, Support Vector Machine (SVM), and XGBoost with a balanced train–test split, and are evaluated using confusion matrix, accuracy, precision, recall, and macro F1-score, while feature importance analysis is performed to identify the most influential variables for each air quality category. The results show that the combination of CO, NO₂, Aerosol Index, hotspot counts, and rainfall achieves high classification performance, with XGBoost achieving the best overall metrics and highlighting the different characteristics of the three algorithms in handling spatial air quality patterns. These findings are expected to serve as a basis for developing satellite-based, machine-learning-driven air quality monitoring systems to support decision-making in mitigating the impacts of forest fires and haze in Kalimantan.

Keywords: air quality, forest fires, Sentinel-5P, hotspots, pollution index, rainfall, Random Forest, SVM, XGBoost.

DAFTAR ISI	
HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
PERNYATAAN SIMILARITY CHECK	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vii
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terdahulu	7
2.2 Teori Pendukung	12
2.2.1 Kebakaran Hutan dan Lahan	13
2.2.2 Kualitas Udara	13
2.2.3 Particulate Matter 2.5 (PM2.5)	14
2.2.4 Indeks Standar Pencemar Udara (ISPU).....	15
2.2.5 Sentinel-5P.....	16
2.2.6 Google Earth Engine (GEE)	17
2.2.7 Hotspot.....	17
2.2.8 Faktor Meteorologi	18
2.2.9 Machine Learning.....	18

2.2.10 Algoritma Klasifikasi Machine Learning	19
2.2.11 Random Oversampling	20
2.2.12 Mutual Information.....	21
2.2.13 Mutual Information.....	22
2.2.14 Principal Component Analysis (PCA).....	22
2.2.15 Stratified K-Fold Cross Validation.....	23
2.2.16 Stratified K-Fold Cross Validation.....	24
BAB III METODE PENELITIAN	27
3.1 Jenis Penelitian.....	27
3.2 Lokasi dan Cakupan Wilayah	27
3.3 Teknik Pengumpulan Data	29
3.3.1 Pengambilan Data	29
3.3.2 Teknik Pengolahan Data.....	31
3.3.3 Teknik Analisis Data	33
3.4 Pengolahan Data.....	35
3.4.1 Integrasi Dataset	35
3.4.2 Praproses Data	36
3.4.3 Pembentukan Label	38
3.5 Pemodelan dan Evaluasi	40
3.5.1 Pembentukan Data Training dan Testing	40
3.5.2 Stratified K-Fold Cross Validation	41
3.5.3 Penanganan Ketidakseimbangan Kelas	41
3.5.4 Pelatihan Model Klasifikasi.....	41
3.5.5 Evaluasi Model	42
3.5.6 Perbandingan Model.....	42
3.6 Perangkat dan Lingkungan Penelitian.....	42
3.7 Alur Penelitian	43
BAB IV PEMBAHASAN	45
4.1 Hasil Pengumpulan dan Integrasi Data	45
4.1.1 Hasil Pengumpulan Data	45
4.1.2 Hasil Integrasi Dataset.....	46
4.2 Hasil Praproses Data	48

4.2.1 Hasil Pemeriksaan Data	48
4.2.2 Hasil Penanganan Missing Value	49
4.2.3 Hasil Standarisasi Data	50
4.3 Hasil Pembentukan Tabel	51
4.3.1 Hasil Perhitungan ISPU	51
4.3.2 Distribusi Label	53
4.4 Hasil Pembagian dan Penyeimbangan Data.....	54
4.4.1 Hasil Pembagian Training dan Testing.....	54
4.4.2 Hasil Random Oversampling.....	55
4.5 Hasil Feature Selection dan PCA.....	56
4.5.1 Hasil Mutual Information	56
4.5.2 Hasil PCA	57
4.6 Hasil Pemodelan.....	59
4.6.1 Hasil Pemodelan Random Forest.....	59
4.6.2 Hasil Pemodelan XGBoost.....	61
4.6.3 Hasil Pemodelan SVM	61
4.7 Hasil Evaluasi Model	62
4.7.1 Stratified 5-Fold Cross Validation.....	63
4.7.2 Evaluasi Data Testing	63
4.7.3 Confusion Matrix.....	64
4.7.4 Perbandingan Hasil Cross Validation dan Data Testing.....	67
4.7.5 ROC-AUC	69
4.8 Perbandingan dan Pembahasan Model.....	71
4.8.1 Pembahasan Model.....	72
4.9 Analisis Feature Importance	73
4.10 Visualisasi Hasil Klasifikasi	74
4.10.1 Visualisasi PCA	74
4.10.2 Visualisasi Hasil Klasifikasi Random Forest	76
4.11 Pembahasan	77
BAB V KESIMPULAN DAN SARAN.....	79
5.1 Kesimpulan	79
5.2 Saran.....	80

DAFTAR PUSTAKA.....	81
LAMPIRAN.....	84



DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	7
Tabel 3. 1 Kumpulan Dataset.....	30
Tabel 3. 2 Struktur Dataset Hasil Integrasi	36
Tabel 3. 3 Informasi Label Kelas.....	38
Tabel 3. 4 Kriteria Pembentukan Label	39
Tabel 3. 5 Variabel Input dan Target	40
Tabel 4. 1 Hasil Pengumpulan Data	46
Tabel 4.2 Struktur Dataset Hasil Integrasi	47
Tabel 4.3 Kategori Nilai ISPU.....	52
Tabel 4.4 Hasil Pembagian Dataset	54
Tabel 4.5 Komponen Utama PCA	59
Tabel 4.6 Perbandingan Hasil Stratified 5-Fold Cross Validation.....	63



DAFTAR GAMBAR

Gambar 3. 1 Peta Lokasi Penelitian Pulau Kalimantan	28
Gambar 3. 2 FlowChat Alur Penelitian.....	44
Gambar 4.1 Peta Kalimantan	45
Gambar 4.2 Informasi Dataset	47
Gambar 4.3 Validasi Jumlah Data	49
Gambar 4.4 Hasil Imputasi	50
Gambar 4.5 Hasil Standarisasi	51
Gambar 4.6 Data Train Sebelum Balancing	53
Gambar 4.7 Validasi Data Test.....	53
Gambar 4.8 Random Over Sampling.....	55
Gambar 4.9 Feature Selection.....	56
Gambar 4.10 Hasil PCA	57
Gambar 4.11 PCA Component Loadings	58
Gambar 4.12 Pemodelan Random Forest	60
Gambar 4.13 Pemodelan XGBoost.....	61
Gambar 4.14 Pemodelan SVM	62
Gambar 4.15 Evaluasi Data Testing	64
Gambar 4.16 Confusion Matrix Random Forest.....	65
Gambar 4.17 Confusion Matrix XGBoost	66
Gambar 4.18 Confusion Matrix SVM	67
Gambar 4.19 Perbandingan 3 Algoritma	71
Gambar 4.20 Grafik Perbandingan 3 Algoritma.....	72
Gambar 4.21 Feature Importance Random Forest.....	73
Gambar 4.22 Visualisasi PCA	75
Gambar 4.23 Visualisasi Hasil Klasifikasi Random Forest.....	76

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi	84
Lampiran 2. Curriculum Vitae	85
Lampiran 3. Surat Pernyataan & Pengalihan HAKI	86
Lampiran 4. Sertifikat BNSP	88

