



**ANALISIS PERFORMA EXTREME GRADIENT BOOSTING(XGBOOST) DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA DI JAKARTA**

LAPORAN SKRIPSI

YERICHO IMMANUEL RUMAHORBO

41520010162

Pembimbing :

Harni Kusniyati, ST., M.Kom

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024



**ANALISIS PERFORMA EXTREME GRADIENT BOOSTING(XGBOOST) DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA DI JAKARTA**

LAPORAN SKRIPSI

YERICHO IMMANUEL RUMAHORBO

41520010162

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Yericho Immanuel Rumahorbo
NIM : 41520010162
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Performa Extreme Gradient Boosting (XGBOOST)
Dalam Klasifikasi Tingkat Pencemaran Udara Di Jakarta

Menyatakan bahwa laporan Aplikatif/Tugas Akhir/Jurnal/Media Imiah ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Yericho Immanuel Rumahorbo

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Yericho Immanuel Rumahorbo
NIM : 41520010162
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Performa Extreme Gradient Boosting (XGBOOST) Dalam Klasifikasi Tingkat Pencemaran Udara Di Jakarta

Telah berhasil dipertahankan pada sidang di hadapan Dewan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana

Disahkan oleh:

Pembimbing : Harni Kusniyati, ST., M.Kom. ()
NIDN : 0324068101
Ketua Penguji : Harni Kusniyati, ST., M.Kom. ()
NIDN : 0324068101
Penguji 1 : Mohamad Yusuf, S.Kom., MCS. ()
NIDN : 0307097606
Penguji 2 : Siti Maesaroh, S.Kom., M.TI. ()
NIDN : 0413059003

Jakarta, 19 Juli 2024

Mengetahui,

Dekan



Nama : Dr. Bambang Jokonowo, S.Si., M.Ti
NIDN : 0320037002

Ketua Program Studi



Nama : Dr. Hadi Santoso, S.Kom., M.Kom.
NIDN : 0225067701

KATA PENGANTAR

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

1. Prof.Dr. Ir. Andi Adriansyah. M.Eng / selaku Rektor Universitas Mercu Buana
2. Dr. Bambang Jokonow, S.Si.MTI / selaku Dekan Fakultas Ilmu Komputer
3. Dr. Hadi Santoso, S.Kom., M.Kom / selaku Kepala Program Studi Teknik Informatika Di Universitas Mercu Buana
4. Harni Kusniyati, ST., M.Kom / selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
5. Mohamad Yusuf, S.Kom., MCS selaku Dosen Penguji 1 dan Siti Maesaroh, S.Kom., M.TI. sebagai Penguji 2 Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Orang tua dan keluarga yang telah memberikan dukungan serta doa yang tak henti-hentinya
7. Teman-teman serta seluruh pihak yang telah membantu penyelesaian Tugas Akhir ini dengan baik

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

Jakarta, 7 Juli 2024

UNIVERSITAS
MERCU BUANA



Yericho Immanuel Rumahorbo

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Yericho Immanuel Rumahorbo
NIM : 41520010162
Program Studi : Teknik Informatika
Judul Laporan Skripsi : ANALISIS PERFORMA EXTREME GRADIENT BOOSTING (XGBOOST) DALAM
KLASIFIKASI TINGKAT PENCEMARAN UDARA DI JAKARTA

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 Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 13 Juli 2024

Yang menyatakan,



Yericho Immanuel Rumahorbo

ABSTRAK

Nama : Yericho Immanuel Rumahorbo
Nim : 41520010162
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Analisis Performa Extreme Gradient Boosting (XGBOOST)
Dalam Klasifikasi Pencemaran Udara Di Jakarta
Pembimbing : Harni Kusniyati, ST., M.Kom

Pencemaran udara menjadi tantangan serius di banyak kota metropolitan, termasuk Jakarta, Indonesia. Dalam konteks ini, penelitian ini mengevaluasi performa Extreme Gradient Boosting (XGBoost) dalam klasifikasi tingkat pencemaran udara dengan menggunakan dataset yang diambil dari Indeks Standar Pencemaran Udara (ISPU) data tersebut berasal dari Satu Data Jakarta. Dataset ini mencakup data kualitas udara yang dikumpulkan dari berbagai stasiun pengukuran di Jakarta. Langkah awal penelitian melibatkan pemrosesan data dari Satu Data Jakarta, termasuk pembersihan data dan pemilihan fitur yang relevan untuk analisis klasifikasi. Metode yang digunakan meliputi pendekatan kuantitatif dengan desain penelitian yang sistematis. Data yang digunakan diambil dari sumber open data Jakarta, yang mencakup informasi tentang kualitas udara. Pada proses ini melibatkan pemrosesan data untuk memastikan kualitas data yang baik, pemilihan fitur yang relevan, dan penerapan model XGBoost. Setelah itu, model XGBoost diterapkan dan disesuaikan dengan dataset untuk mengidentifikasi pola kompleks yang mungkin memengaruhi tingkat pencemaran udara. Hasil eksperimen menunjukkan bahwa XGBoost memberikan kinerja klasifikasi yang memuaskan berdasarkan Satu Data Jakarta. Analisis lebih lanjut mencakup evaluasi parameter model dan identifikasi fitur yang paling berpengaruh dalam menentukan tingkat pencemaran udara di Jakarta. Penelitian ini menggabungkan keunggulan analisis data dari Satu Data Jakarta dengan kecanggihan model XGBoost, memberikan kontribusi dalam pemahaman lebih lanjut tentang faktor-faktor yang memengaruhi kualitas udara di wilayah dan pemantauan yang lebih baik serta memberikan wawasan kepada pihak berkepentingan untuk merancang kebijakan mitigasi yang lebih efektif. Hasil penelitian ini menunjukkan bahwa model XGBoost memiliki performa yang baik dalam klasifikasi tingkat pencemaran udara di Jakarta, dengan akurasi mencapai 90.22% lalu presisi 90.55% dan nilai recall 90.22%. Dengan memanfaatkan dataset Satu Data Jakarta, penelitian ini memberikan masalah pencemaran udara di berbagai kota metropolitan.

Kata Kunci: Pencemaran udara, Klasifikasi Pencemaran udara, XGBoost, Jakarta, ISPU

ABSTRACT

Name : Yericho Immanuel Rumahorbo
NIM : 41520010162
Study Program : Teknik Informatika
Title Research Proposal : Analisis Performa Extreme Gradient Boosting (XGBOOST)
Dalam Klasifikasi Tingkat Pencemaran Udara Di Jakarta

Air pollution is a serious challenge in many metropolitan cities, including Jakarta, Indonesia. In this context, this research evaluates the performance of Extreme Gradient Boosting (XGBoost) in classifying air pollutant levels using a dataset taken from the Air Pollution Standard Index (ISPU). The data comes from Satu Data Jakarta. This dataset includes air quality data collected from various measurement stations in Jakarta. The initial steps of the research involved processing data from Satu Data Jakarta, including data cleaning and selecting relevant features for classification analysis. The method used includes a quantitative approach with a systematic research design. The data used is taken from Jakarta open data source, which include information about air quality. This process involves data processing to ensure good data quality, selecting relevant features, and applying and adjusted to the dataset to identify complex patterns that may influence air pollution levels. Experimental results show that XGBoost provides satisfactory classification performance based on Satu Data Jakarta. Further analysis includes evaluating model parameters and identifying the most influential features in determining the level of air pollution in Jakarta. This research combines the data analysis advantages of Satu Data Jakarta with the sophistication of the XGBoost model, contributing to further understanding of the factors that influence regional air quality and better monitoring and providing insight to interested parties to design more effective mitigation policies. The results of this research show that the XGBoost model has good performance in classifying air pollution levels in Jakarta, with an accuracy of 90.22% and a precision of 90.55% and a recall value of 90.22%. by utilizing the Satu Data Jakarta dataset, this research provides air pollution problems in various metropolitan cities.

Keywords: Air pollution, Air Pollution Classification, XGBoost, Jakarta, ISPU

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	v
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Batasan Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 Teori Utama.....	4
2.1.1 Kualitas Udara Di DKI Jakarta.....	4
2.1.2 Pencemaran Udara.....	5
2.1.3 Dampak pencemaran udara	5
2.1.4 Klasifikasi.....	6
2.2 Teori Pendukung	7
2.2.1 Extreme Gradient Boosting(XGBoost)	7
2.2.2 Google Colab.....	8
2.3 Penelitian Terdahulu	8

2.4	Critical Review.....	11
2.5	Gap Penelitian	13
2.5.1	Summarize	13
2.5.2	Synthesize	13
2.5.3	Comparation.....	14
2.5.4	Contribution	14
2.6	Confusion Matrix	15
BAB III METODE PENELITIAN.....		17
3.1	Jenis Penelitian.....	17
3.2	Desain Penelitian.....	17
3.2.1	Business Understanding.....	17
3.2.2	Data Understanding	18
3.2.3	Data Preparation.....	18
3.2.4	Modelling.....	18
3.2.5	Evaluation	18
3.3	Subjek Penelitian.....	19
3.4	Instrumen Penelitian.....	19
3.5	Teknik Pengumpulan Data	20
3.6	Pemrosesan awal data.....	20
3.7	Tahapan Penelitian.....	22
3.7.1	Business Understanding.....	23
3.7.2	Data Understanding	23
3.7.3	Data Preparation.....	23
3.7.4	Modelling.....	23
3.7.5	Evaluation	23
3.8	Evaluasi Hasil Penelitian.....	24
3.9	Timeline Penelitian	24
BAB IV HASIL DAN PEMBAHASAN		26
4.1	Pencapaian Tujuan Penelitian	26
4.2	Kontribusi Terhadap Bidang Teknologi Informasi	26
4.3	Implikasi dan Aplikasi	26

4.4	Dataset.....	27
4.5	Pengujian Akurasi	28
4.6	Implementasi dan Pengembangan.....	30
4.7	Pengujian Shap Value Kualitas Udara	31
4.8	Pengujian Dashboard Pencemaran Udara Di Jakarta.....	32
BAB V	KESIMPULAN DAN SARAN.....	34
5.1	Kesimpulan.....	34
5.2	Saran.....	35
DAFTAR PUSTAKA.....		36
LAMPIRAN.....		38



DAFTAR TABEL

Tabel 2 1. Penelitian Terdahulu	8
Tabel 2 2. Jumlah Penelitian Terdahulu (2020-2024)	13
Tabel 2 3. Comparison.....	14
Tabel 2 4. Nilai Matriks	15
Tabel 2 5. Perhitungan Modelling Matriks	16
Tabel 3 1. Perangkat Keras	19
Tabel 3 2. Perangkat Lunak	19
Tabel 3 3. Timeline Penelitian	24



DAFTAR GAMBAR

Gambar 2 1. Logo XGBoost.....	7
Gambar 2 2. Logo Google Colab.....	8
Gambar 2 3. Tahap Critical Review	12
Gambar 2 4. Pengumpulan Jurnal (Publis or Perish).....	12
Gambar 2 5. Rumus Accuracy.....	16
Gambar 2 6. Rumus Precision	16
Gambar 2 7. Rumus F1-Score.....	16
Gambar 2 8. Rumus Recall	16
Gambar 3 1. Alur Desain Penelitian	17
Gambar 3 2. Memeriksa Jumlah Nilai Yang Hilang Pada Fitur pm25.....	21
Gambar 3 3. Setelah Membersihkan Data Pada Fitur pm25.....	21
Gambar 3 4. Dataset Sebelum Pemilihan Fitur Pencemaran Udara	22
Gambar 3 5. Dataset Setelah Pemilihan Fitur Pencemaran Udara Di jakarta.....	22
Gambar 4 1. Kode Program Import Dataset	27
Gambar 4 2. Dataset Standar Pencemaran Udara DiProvinsi DKI Jakarta 2021	28
Gambar 4 3. Confusion Matriks.....	29
Gambar 4 4. Kode Program Visualisasi Confusion Matriks.....	29
Gambar 4 5. Visualisasi Grafik Train Data Pencemaran Udara	30
Gambar 4 6. Shap Value Mean Pencemaran Udara.....	31
Gambar 4 7. impact Pada Pencemaran Udara Di Jakarta	32
Gambar 4 8. Dashboard Pencemaran Udara Wilayah DKI Jakarta	33

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi Bimbingan	38
Lampiran 2. Curriculum Vitae	39
Lampiran 3. Surat Pernyataan HKI.....	39
Lampiran 4. Sertifikat BNSP	40
Lampiran 5. Formulir Revisi.....	41
Lampiran 6. Hasil Turnitin	44

