

**ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS
UDARA DI JAKARTA MENGGUNAKAN ALGORITMA
ORDINAL LOGISTIC REGRESSION DAN *NEURAL
NETWORK***

SKRIPSI SARJANA SISTEM INFORMASI

Disusun Oleh:

Jepplin Sinaga

21700651078



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

2025

**ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS
UDARA DI JAKARTA MENGGUNAKAN ALGORITMA
ORDINAL LOGISTIC REGRESSION DAN *NEURAL
NETWORK***

SKRIPSI SARJANA SISTEM INFORMASI

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

Disusun Oleh:

Jepplin Sinaga

217006516078



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

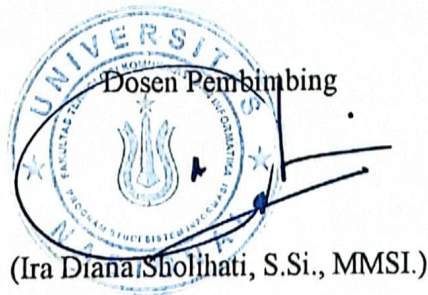
2025

HALAMAN PENGESAHAN

TUGAS AKHIR

**ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS
UDARA DI JAKARTA MENGGUNAKAN ALGORITMA
ORDINAL LOGISTIC REGRESSION DAN NEURAL
NETWORK**



Dosen Pembimbing

(Ira Diana Sholihati, S.Si., MMSI.)

PERNYTAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa tugas akhir dengan judul :

ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS UDARA DI JAKARTA MENGGUNAKAN ALGORITMA ORDINAL LOGISTIC REGRESSION DAN NEURAL NETWORK

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 dan publikasi dari tugas akhir yang sudah 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.



217006516078

LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas akhir dengan judul :

ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS UDARA DI JAKARTA MENGUNAKAN ALGORITMA ORDINAL LOGISTIC REGRESSION DAN NEURAL NETWORK

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjan Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas akhir ini diajukan pada Sidang Review Akhir Semester Genap 2025 – 2026 pada tanggal 27 Agustus 2025.

Dosen Pembimbing

Ira Diana Sholihati, S.Si., MMSI.

NIDN : 0328037304

Ketua Program Studi

Dr. Ir. Andrianingsih, S.Kom., MMSI

NIDN : 030309790

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU DIREVISI

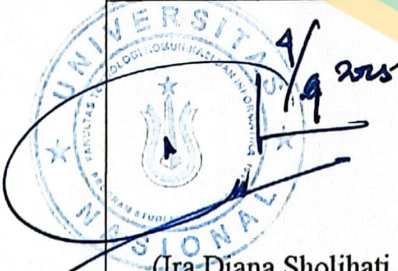
Nama : Jepplin Sinaga
NPM : 217006516078
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Infomatika
Program Studi : Sistem Informasi
Tanggal Sidang : 27 Agustus 2025

JUDUL DALAM BAHASA INDONESIA :

ANALISIS PERBANDINGAN KLASIFIKASI KUALITAS UDARA DI
JAKARTA MENGGUNAKAN ALGORITMA ORDINAL LOGISTIC
REGRESSION DAN NEURAL NETWORK

JUDUL DALAM BAHASA INGGRIS :

COMPARATIVE ANALYSIS OF AIR QUALITY CLASSIFICATION IN
JAKARTA USING ORDINAL LOGISTIC REGRESSION AND NEURAL
NETWORK ALGORITHMS

TANDA TANGAN DAN TANGGAL		
Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 1 September 2025	TGL : 2/9 2025	TGL : 1 September 2025
 (Ira Diana Sholihati, S.Si., MMSI.)	 (Dr. Ir. Andrianingsih, S.Kom., MMSI)	 (Jepplin Sinaga)

KATA PENGANTAR

Segala hormat, puji syukur dan kemuliaan bagi Tuhan Yesus Yang Maha Esa dengan kasih dan penyertaanNya telah menuntun setiap langkah dalam penyusunan skripsi ini. Tanpa kasih dan kekuatanNya, penulis tidak akan mampu menyelesaikan Skripsi dengan judul **“Analisis Perbandingan Klasifikasi Kualitas Udara Di Jakarta Menggunakan Algoritma Ordinal Logistic Regression Dan Neural Network”** sebagai salah satu syarat kelulusan Program Studi Sarjana Komputer Fakultas Teknologi Komunikasi dan Informatika. Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir, Ibu Ira Diana Sholihati, S.Si., MMSI yang telah meluangkan banyak waktu, tenaga, pikiran, bimbingan, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi. Penulis juga mengucapkan banyak terima kasih kepada:

1. Ayah, Ibu dan seluruh keluarga penulis yang telah banyak memberi dukungan dalam segala bentuk yang tak terhitung.
2. Seluruh dosen pengajar di Program Studi Sistem Informasi, FTKI maupun dosen di Program Studi lain yang memberikan banyak ilmu.
3. Teman-teman seangkatan dan sehimpuan berbagai angkatan yang telah membantu dan mendukung.
4. Tabulator IT UnasFest 2023 selaku teman yang telah mendukung dan memberikan semangat dalam proses penyusunan skripsi.
5. Lysda Deceline Rumapea selaku pacar penulis yang telah ikut serta membantu, mendukung, dan motivasi sepanjang masa pacaran dan perkuliahan.

Penulis menyadari bahwasanya masih banyak kekurangan dalam penyajian, pemilihan kata-kata maupun pembahasan hasil skripsi ini. Oleh karena itu, dengan

kerendahan hati penulis mengharapkan saran, kritik yang membangun dari semua pihak untuk perbaikan skripsi ini.

Akhir kata, kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan motivasi dalam penyusunan skripsi ini, penulis mengucapkan terima kasih banyak. Semoga skripsi ini dapat memberikan manfaat baik bagi pengembangan ilmu pengetahuan maupun bagi pihak-pihak yang berkepentingan serta dapat menjadi referensi yang berguna dan dapat diterima dengan baik bagi pembaca yang memberikan kontribusi yang berarti dalam bidang analisis data.

Jakarta, ... Oktober 2025

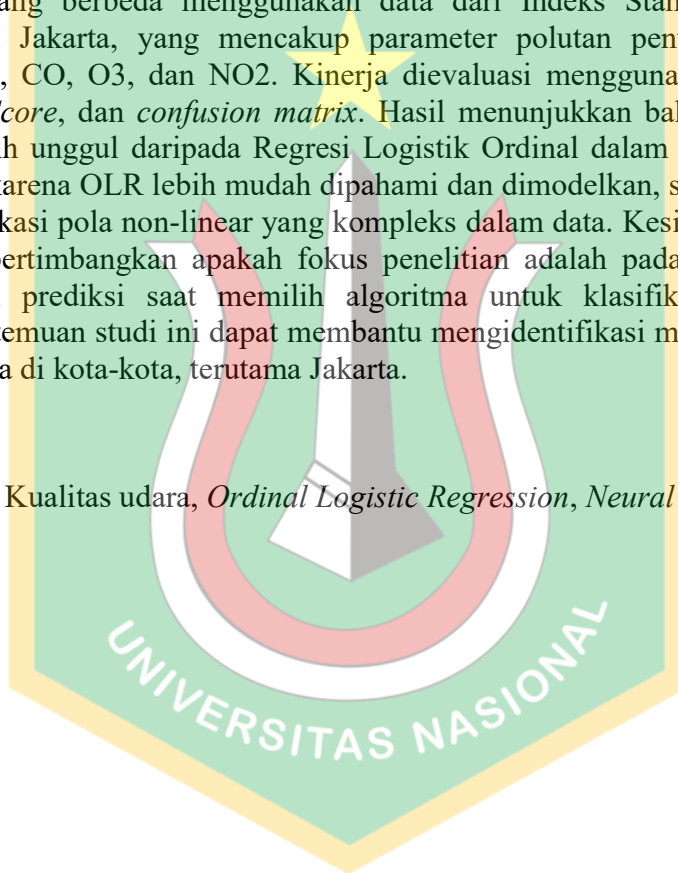
Jepplin Sinaga



ABSTRAK

Kualitas udara merupakan indikator utama untuk menilai kondisi lingkungan perkotaan. Pertumbuhan ekonomi Jakarta yang pesat dan populasi yang besar berkontribusi pada tingginya tingkat polusi udara di kota ini. Oleh karena itu, untuk mengklasifikasikan kualitas udara secara tepat, diperlukan metode analisis yang sesuai. Penelitian ini membandingkan cara algoritma *Neural Network* (NN) dan *Ordinal Logistic Regression* (OLR) mengklasifikasikan kualitas udara ke dalam kelompok yang berbeda menggunakan data dari Indeks Standar Polutan Udara (ISPU) DKI Jakarta, yang mencakup parameter polutan penting seperti PM10, PM2.5, SO2, CO, O3, dan NO2. Kinerja dievaluasi menggunakan *recall*, akurasi, presisi, *F1-Score*, dan *confusion matrix*. Hasil menunjukkan bahwa metode *Neural Network* lebih unggul daripada Regresi Logistik Ordinal dalam hal akurasi. Hal ini disebabkan karena OLR lebih mudah dipahami dan dimodelkan, sedangkan NN dapat mengidentifikasi pola non-linear yang kompleks dalam data. Kesimpulannya, penting untuk mempertimbangkan apakah fokus penelitian adalah pada interpretasi model atau akurasi prediksi saat memilih algoritma untuk klasifikasi kualitas udara. Diharapkan temuan studi ini dapat membantu mengidentifikasi metode untuk menilai kualitas udara di kota-kota, terutama Jakarta.

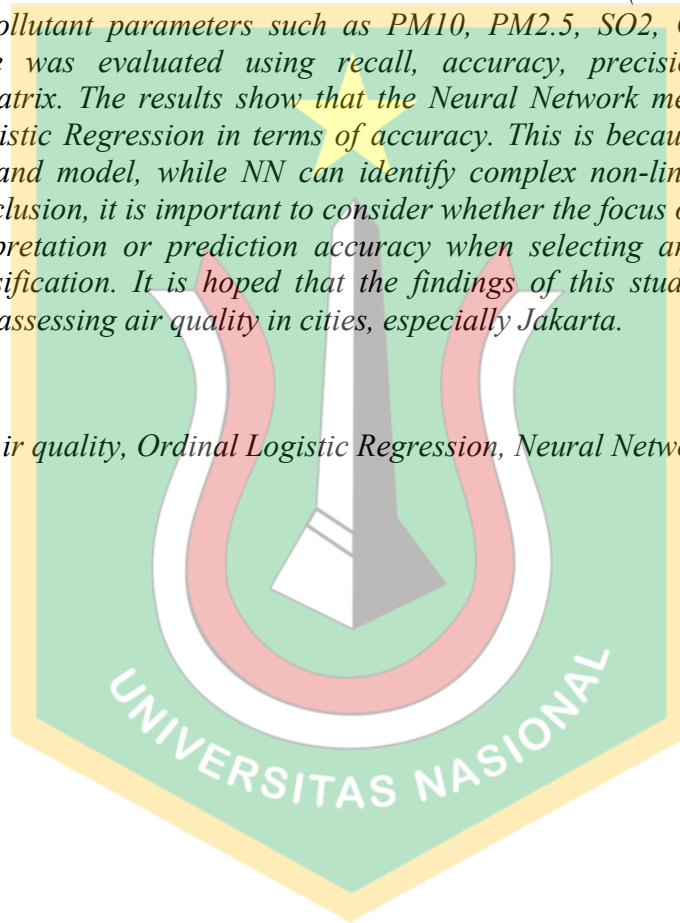
Kata kunci: Kualitas udara, *Ordinal Logistic Regression*, *Neural Network*, ISPU



ABSTRACT

Air quality is a key indicator for assessing urban environmental conditions. Jakarta's rapid economic growth and large population contribute to high levels of air pollution in the city. Therefore, to classify air quality accurately, an appropriate analysis method is needed. This study compares how Neural Network (NN) and Ordinal Logistic Regression (OLR) algorithms classify air quality into different groups using data from the DKI Jakarta Air Pollutant Standard Index (ISPU), which includes important pollutant parameters such as PM10, PM2.5, SO2, CO, O3, and NO2. Performance was evaluated using recall, accuracy, precision, F1-Score, and confusion matrix. The results show that the Neural Network method is superior to Ordinal Logistic Regression in terms of accuracy. This is because OLR is easier to understand and model, while NN can identify complex non-linear patterns in the data. In conclusion, it is important to consider whether the focus of the research is on model interpretation or prediction accuracy when selecting an algorithm for air quality classification. It is hoped that the findings of this study can help identify methods for assessing air quality in cities, especially Jakarta.

Keywords: *Air quality, Ordinal Logistic Regression, Neural Network, ISPU*



DAFTAR ISI

BAB 1 PENDAHULUAN.....	16
1.1 Latar Belakang.....	16
1.2 Batasan Masalah.....	18
1.3 Rumusan Masalah.....	19
1.4 Tujuan Penelitian.....	19
1.5 Manfaat Penelitian.....	19
BAB II TINJAUAN PUSTAKA	21
2.1 Landasan Teori	21
2.1.1 Data mining.....	21
2.1.2 Machine Learning	21
2.1.3 Classification.....	22
2.1.4 Google Colab	26
2.1.5 Kualitas Udara.....	26
2.1.6 Indeks Standar Pencemaran Udara	26
2.2 Studi literature.....	27
BAB III METODE PENELITIAN	36
3.1 Waktu dan lokasi penelitian.....	36
3.2 Penentuan Obyek dan subobyek penelitian	36
3.3 Fokus penelitian	36
3.4 Metode pengumpulan data	36
3.4.1 Informasi data	37
3.5 Desain penelitian.....	39

3.6 Metode klasifikasi	44
3.6.1 Ordinal Logistick Regression.....	44
3.6.2 Neural Network	48
3.7 Analisi evaluasi	49
3.7.1 Confussion matrix.....	49
BAB IV HASIL PEMBAHASAN	52
4.1 Pengumpulan dataset.....	52
4.2 Data preparation	53
4.3 Perhitungan Manual.....	55
4.3.1 Ordinal Logistick Regression.....	55
4.3.2 Neural Network	56
4.3 Penerapan dalam google colab.....	57
4.3.1 Upload dataset dan import library.....	58
4.3.2 Persiapan data.....	59
4.3.3 Penerapan OLR.....	60
4.3.4 Penerapan Neural Network.....	63
4.4 Analisis dan evaluasi hasil.....	68
4.5 Tampilan Dashboard Kualitas Udara	73
BAB V KESIMPULAN DAN SARAN.....	75
5.1 Kesimpulan.....	75
5.2 Saran	76
DAFTAR PUSTAKA	78

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Regresi Logistik Ordinal	24
Gambar 2.2 Ilustrasi Neural Network	25
Gambar 2.3 Kategori ISPU	27
Gambar 3.1 Sumber Data	37
Gambar 3.2 Dataset ISPU	38
Gambar 3.3 Desain Penelitian	40
Gambar 3.4 Missing values	42
Gambar 4.1 Tampilan awal Dataset	52
Gambar 4.2 Tampilan Nilai Kosong	53
Gambar 4.3 Tampilan Tabel Tidak Ada Data	54
Gambar 4.4 Tabel Fitur Dataset	54
Gambar 4.5 Import library	58
Gambar 4.6 Balencing data	60
Gambar 4.7 Membagi x dan y	61
Gambar 4.8 Penerapan OLR	61
Gambar 4.9 Performa OLR	62
Gambar 4.10 Performa OLR dengan data yang sedikit	63
Gambar 4.11 Library model neural network	64
Gambar 4.12 fitur dan label	64
Gambar 4.13 split dan scalling	65
Gambar 4.14 Penerapan Neural Network	65
Gambar 4.15 Implementasi grid search	66
Gambar 4.16 Table akurasi terbaik	67
Gambar 4.17 Performa NN	67
Gambar 4.18 Performa NN dengan data yang banyak	68
Gambar 4.19 Hasil Klasifikasi	69
Gambar 4.20 Parameter dominan	70

Gambar 4.21 Confusion matrix OLR.....	71
Gambar 4.22 Confusion matrix NN.....	72
Gambar 4.23 Dashboard ISPU	74



DAFTAR TABEL

Tabel 2.1 Studi Literatur	27
Tabel 2.2 Deskripsi Dataset	38
Tabel 3.1 Perhitungan jarak OLR	46
Tabel 4.1 Sampel OLR.....	55

