

**KLASIFIKASI KUALITAS UDARA DKI JAKARTA
BERDASARKAN INDEKS STANDAR PENCEMAR UDARA
MENGUNAKAN METODE REGRESI LOGISTIK
MULTINOMIAL DAN *K-NEAREST NEIGHBOR***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
di Jurusan Matematika pada Fakultas MIPA**

Oleh :

RISKI WULANDARI FUTRI

08011382025109



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2024

HALAMAN PENGESAHAN

**KLASIFIKASI KUALITAS UDARA DKI JAKARTA
BERDASARKAN INDEKS STANDAR PENCEMAR UDARA
MENGUNAKAN METODE REGRESI LOGISTIK
MULTINOMIAL DAN *K-NEAREST NEIGHBOR***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
di Jurusan Matematika pada Fakultas MIPA**

Oleh :

RISKI WULANDARI FUTRI

08011382025109

Indralaya, 14 Mei 2024

Pembimbing Kedua



Des Alwine Zayanti, S.Si., M.Si

NIP. 197012041998022001

Pembimbing Utama



Prof. Yulia Resti, S.Si., M.Si., Ph.D

NIP. 197307191997022001

Mengetahui,

Ketua Jurusan Matematika



Dr. Dian Cahyawati Sukanda, S.Si., M.Si.

NIP. 197303212000122001

SURAT PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan dibawah ini :

Nama Mahasiswa : Riski Wulandari Fitri
NIM : 08011382025109
Fakultas/Jurusan : MIPA/Matematika

Menyatakan bahwa skripsi ini adalah karya sendiri dan karya ilmiah ini belum pernah diajukan sebagai penentuan persyaratan untuk memperoleh gelar kesarjana strata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat dalam skripsi ini berasal dari penulisan pihak lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikian surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 3 Juni 2024



Riski Wulandari Fitri
NIM. 08011382025109

HALAMAN PERSEMBAHAN

MOTTO

“Allah tidak akan membebani seseorang melainkan sesuai dengan kesanggupannya”

(Q.S Al-Baqarah, 2:286)

“Maka sesungguhnya bersama kesulitan itu ada kemudahan.
Sesungguhnya bersama kesulitan itu ada kemudahan”

(Q.S Al-Insyirah, 94:5-6)

“Orang lain tidak akan paham *struggle* dan masa sulitnya kita, yang mereka ingin tahu hanya bagian *success stories*nya saja. Jadi berjuanglah untuk diri sendiri meskipun tidak ada yang tepuk tangan, kelak diri kita di masa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini”

Skripsi ini kupersembahkan kepada :

- 1. Allah SWT**
- 2. Alm. Ayah dan Mama**
- 3. Saudara**
- 4. Keluarga Besar**
- 5. Dosen**
- 6. Sahabat dan Teman**
- 7. Almamater**

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Dengan mengucapkan Alhamdulillah, segala puji bagi Allah Subhanahu wa Ta'ala yang telah memberikan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "**Klasifikasi Kualitas Udara DKI Jakarta Berdasarkan Indeks Standar Pencemar Udara Menggunakan Metode Regresi Logistik Multinomial dan *K-Nearest Neighbor***" sebagai salah satu syarat memperoleh gelar sarjana sains studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Penulis menyadari bahwa proses pembuatan skripsi ini sebagai proses pembelajaran yang sangat berharga yang tak lepas dari kekurangan dan keterbatasan. Dengan segala hormat dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan yang tulus penulis sampaikan kepada kedua orang tuaku tercinta, **Alm. Ayah Heriyanto** dan **Mama Saimah** yang tak pernah lelah mendidik, membimbing, menasehati, dan mendukung serta tak henti untuk mendoakan anaknya. Terima kasih atas segala perjuangan dan pengorbanan hingga detik ini dan sampai kapanpun. Penulis juga menyampaikan terima kasih dan penghargaan kepada:

1. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph.D** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
2. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

3. Ibu **Des Alwine Zayanti, S.Si., M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya serta selaku Pembimbing Kedua yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan dan arahan sehingga skripsi ini dapat diselesaikan.
4. Ibu **Prof. Yulia Resti, S.Si., M.Si., Ph.D** selaku Dosen Pembimbing Utama yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan dan arahan dengan penuh perhatian, pengertian, dan kesabaran sehingga skripsi ini dapat diselesaikan.
5. Ibu **Dr. Yuli Andriani, S.Si., M.Si** dan Ibu **Dr. Evi Yuliza, S.Si., M.Si** selaku Dosen Pembahas yang telah bersedia memberikan kritik dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
6. Ibu **Dra. Ning Eliyati, M.Pd** dan Ibu **Eka Susanti, S.Si., M.Sc** selaku Ketua Seminar dan Sekretaris seminar yang telah bersedia meluangkan waktu dan tenaga dalam seminar penulis.
7. Ibu **Dr. Sisca Octarina, S.Si., M.Sc** selaku Dosen Pembimbing Akademik yang sangat baik telah memberikan saran, membimbing, membantu, dan mengarahkan urusan akademik penulis setiap semester.
8. **Seluruh Dosen di Jurusan Matematika FMIPA** yang telah memberikan ilmu, nasihat, motivasi, serta bimbingan selama proses perkuliahan.
9. Pak **Irwansyah** selaku admin dan Ibu **Hamidah** selaku pegawai tata usaha Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah banyak membantu penulis selama perkuliahan.

10. Adikku tersayang, **Aditya Prayoga** yang selalu memberikan semangat dan membantu penulis selama perkuliahan ini, serta **Keluarga Besar** yang telah memberikan semangat serta doa untuk keberhasilan penulis.
11. **Adrian Zulkah Putra, A.Md.T** yang menjadi support system, telah banyak berkontribusi dalam penulisan skripsi ini, selalu menemani, meluangkan waktu, tenaga, pikiran, serta memberikan semangat untuk pantang menyerah sehingga skripsi ini dapat terselesaikan dengan baik.
12. **Inayah, Devi, Feni, Emveb, dan Novia** sahabat dari masa sekolah hingga saat ini yang telah membantu, mengingatkan, dan mendukung dengan tulus selama ini.
13. **Sylvia, Risa, Indah, dan Seluruh Teman Angkatan 2020** yang telah berjuang bersama-sama serta memberi semangat untuk menyelesaikan perkuliahan dan mendapatkan gelar yang diimpikan.
14. **Kakak tingkat angkatan 2018 dan 2019 dan adik tingkat angkatan 2021 dan 2022** atas segala kebaikan dan bantuan.
15. **Seluruh pihak** yang tidak dapat penulis sebutkan satu per satu. Semoga segala kebaikan yang diberikan mendapatkan balasan terbaik dari Allah.
16. Terakhir, kepada diri saya sendiri. **Riski Wulandari Futri**, terima kasih telah bertahan sejauh ini walaupun sering merasa putus asa jika sesuatu yang dilakukan belum berhasil, namun terima kasih untuk selalu berusaha. Terima kasih karena memutuskan untuk tidak menyerah sesulit apapun proses penyusunan skripsi ini sampai akhirnya bisa menyelesaikannya dengan semaksimal dan sebaik mungkin.

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan seluruh pihak yang membutuhkan. Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, 20 Mei 2024



Penulis

**CLASSIFICATION OF AIR QUALITY IN JAKARTA
BASED ON AIR POLLUTANT STANDARD INDEX USING
MULTINOMIAL LOGISTIC REGRESSION METHOD
AND K-NEAREST NEIGHBOR**

By :

RISKI WULANDARI FUTRI

08011382025109

ABSTRACT

DKI Jakarta's air quality has changed as a result of increased physical development of the city and industrial centers. Currently the air is drier and dirtier than before, if this is allowed to continue it will have a negative impact on living things. Therefore, research is needed on the classification of DKI Jakarta's air quality which is expected to help the DKI Jakarta Environment Agency to provide information to the public, so that people can prevent the adverse effects of air pollution when outdoors. Classification is a process to determine the level of air quality, classification testing in this study uses two different methods, namely Multinomial Logistic Regression and K-Nearest Neighbor. Researchers compared the two methods based on the level of classification accuracy, namely the accuracy, precision, recall, and f1-score values. The level of classification accuracy produced by the Multinomial Logistic Regression method is 82.45%, 60.56%, 60.47%, and 60.51%, respectively. The level of classification accuracy produced by the K-Nearest Neighbor method is 93.87%, 92.33%, 92.94%, and 92.63%, respectively. The results of this study indicate that the K-Nearest Neighbor method has better classification accuracy results compared to the Multinomial Logistic Regression method.

Keywords: Air Quality, Multinomial Logistic Regression, K-Nearest Neighbor

**KLASIFIKASI KUALITAS UDARA DKI JAKARTA
BERDASARKAN INDEKS STANDAR PENCEMAR UDARA
MENGUNAKAN METODE REGRESI LOGISTIK
MULTINOMIAL DAN *K-NEAREST NEIGHBOR***

Oleh :

RISKI WULANDARI FUTRI

08011382025109

ABSTRAK

Kualitas udara DKI Jakarta telah mengalami perubahan akibat dari meningkatnya pembangunan fisik kota dan pusat-pusat industri. Saat ini udara lebih kering dan kotor daripada sebelumnya, jika hal ini dibiarkan terus menerus maka akan berdampak buruk bagi makhluk hidup. Oleh karena itu, diperlukan penelitian mengenai klasifikasi kualitas udara DKI Jakarta yang diharapkan dapat membantu Dinas Lingkungan Hidup DKI Jakarta untuk memberikan informasi kepada masyarakat, sehingga masyarakat dapat melakukan pencegahan terhadap dampak buruk pencemaran udara ketika diluar ruangan. Klasifikasi merupakan proses untuk mengetahui tingkatan kualitas udara, pengujian klasifikasi dalam penelitian ini menggunakan dua metode yang berbeda yaitu Regresi Logistik Multinomial dan *K-Nearest Neighbor*. Peneliti membandingkan kedua metode tersebut berdasarkan tingkat ketepatan klasifikasinya yaitu dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Tingkat ketepatan klasifikasi yang dihasilkan oleh metode Regresi Logistik Multinomial secara berturut-turut yaitu 82,45%, 60,56%, 60,47%, dan 60,51%. Tingkat ketepatan klasifikasi yang dihasilkan oleh metode *K-Nearest Neighbor* secara berturut-turut yaitu 93,87%, 92,33%, 92,94%, dan 92,63%. Hasil dari penelitian ini menunjukkan bahwa metode *K-Nearest Neighbor* memiliki hasil tingkat ketepatan klasifikasi yang lebih baik dibandingkan dengan metode Regresi Logistik Multinomial.

Kata Kunci: Kualitas Udara, Regresi Logistik Multinomial, *K-Nearest Neighbor*

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERSEMBAHAN	iii
KATA PENGANTAR	iv
ABSTRACT	viii
ABSTRAK	ix
DAFTAR ISI	x
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan.....	4
1.5 Manfaat.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Udara	6
2.2 Kualitas Udara	6
2.3 Indeks Standar Pencemar Udara.....	6
2.4 Parameter Pencemaran Udara.....	7
2.4.1 Partikulat Matter (PM ₁₀)	7
2.4.2 Sulfur Dioksida (SO ₂)	7
2.4.3 Karbon Monoksida (CO).....	8
2.4.4 Ozon (O ₃)	8
2.4.5 Nitrogen Dioksida (NO ₂)	8
2.5 Klasifikasi.....	8
2.6 Diskritisasi.....	9
2.7 Analisis Regresi.....	10
2.8 Regresi Logistik	10
2.9 Regresi Logistik Multinomial	11
2.10 Penduga Parameter	12

2.11	Pengujian Parameter	14
2.11.1	Uji Simultan	14
2.11.2	Uji Parsial.....	15
2.12	Uji Kesesuaian Model	15
2.13	<i>Odds Ratio</i>	16
2.14	Standarisasi Data	17
2.15	<i>K-Nearest Neighbor</i>	17
2.16	<i>Confusion Matrix</i>	18
BAB III METODOLOGI PENELITIAN		21
3.1	Tempat.....	21
3.2	Waktu	21
3.3	Alat	21
3.4	Data Penelitian	21
3.5	Metode Penelitian.....	22
BAB IV HASIL DAN PEMBAHASAN		24
4.1	Deskripsi Data	24
4.2	Diskritisasi Data	25
4.3	Partisi Data	27
4.4	Regresi Logistik Multinomial	28
4.4.1	Penduga Parameter.....	28
4.4.2	Pembentukan Model Awal Regresi Logistik Multinomial	29
4.4.3	Uji Simultan	30
4.4.4	Uji Parsial.....	31
4.4.5	Uji Kesesuaian Model.....	32
4.4.6	Interpretasi Koefisien Parameter.....	34
4.4.7	Pemilihan Model Terbaik.....	36
4.4.8	Tingkat Ketepatan Klasifikasi Metode Regresi Logistik Multinomial	37
4.5	<i>K-Nearest Neighbor</i>	39
4.5.1	Data Contoh <i>K-Nearest Neighbor</i>	39
4.5.2	Standarisasi Data.....	40

4.5.3	Perhitungan Jarak <i>Euclidean</i>	41
4.5.4	Mengurutkan Jarak <i>Euclidean</i>	42
4.5.5	Tingkat Ketepatan Klasifikasi Metode <i>K-Nearest Neighbor</i>	43
4.6	Perbandingan Hasil.....	45
BAB V KESIMPULAN DAN SARAN		46
5.1	Kesimpulan.....	46
5.2	Saran	46
DAFTAR PUSTAKA		47
LAMPIRAN		50

DAFTAR TABEL

Tabel 2.1 Kategori kualitas udara berdasarkan nilai ISPU.....	7
Tabel 2.2 <i>Confusion matrix multiclass</i>	18
Tabel 2.3 Kategori ketepatan klasifikasi	20
Tabel 3.1 Variabel prediktor	21
Tabel 3.2 Variabel respon	22
Tabel 4.1 Keterangan variabel dan parameter.....	24
Tabel 4.2 Data penelitian	25
Tabel 4.3 Nilai minimum, Q_1 , Q_2 , Q_3 , dan nilai maksimum.....	26
Tabel 4.4 Hasil diskritisasi data	27
Tabel 4.5 Data latihan.....	27
Tabel 4. 6 Data uji.....	28
Tabel 4.7 Penduga parameter	28
Tabel 4.8 Tabel uji simultan.....	30
Tabel 4.9 Uji parsial.....	32
Tabel 4.10 Uji kesesuaian model	33
Tabel 4.11 <i>Odds ratio</i>	34
Tabel 4.12 Hasil klasifikasi regresi logistik multinomial	37
Tabel 4.13 <i>Confusion matrix</i> regresi logistik multinomial	37
Tabel 4.14 Data contoh	39
Tabel 4.15 Standarisasi data.....	40
Tabel 4.16 Jarak <i>euclidean</i>	42
Tabel 4.17 Penentuan kategori kualitas udara data ke-15.....	42
Tabel 4.18 Hasil klasifikasi <i>k-nearest neighbor</i>	43
Tabel 4.19 <i>Confusion matrix k-nearest neighbor</i>	43
Tabel 4.20 Perbandingan tingkat ketepatan klasifikasi.....	45

BAB I

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi merupakan proses menganalisis data yang akan ditempatkan pada suatu kelas khusus dari semua kelas yang ada. Dalam klasifikasi melibatkan pembuatan model dari data latih sebelumnya lalu digunakan untuk mengklasifikasikan data baru (Utomo & Mesran, 2020). Klasifikasi dapat dilakukan dengan metode Regresi Logistik Multinomial dan *K-Nearest Neighbor*.

Regresi Logistik Multinomial merupakan metode digunakan pada variabel respon yang mempunyai sifat *polychotomus* dengan skala data ordinal atau nominal yang lebih dari dua kategori (Hosmer & Lemeshow, 2000). *K-Nearest Neighbor* merupakan metode yang digunakan dalam mengklasifikasikan data berdasarkan dari tetangga (*K*) terdekat. *K-Nearest Neighbor* bekerja dengan menemukan jarak antara data latih dan data uji (Wang *et al.*, 2022).

Penelitian ini dilakukan dengan cara mengklasifikasikan kualitas udara DKI Jakarta pada metode Regresi Logistik Multinomial dan metode *K-Nearest Neighbor*. Penggunaan metode tersebut bertujuan untuk meminimumkan kesalahan yang mungkin terjadi pada saat pengklasifikasian serta untuk memperoleh nilai ketepatan yang lebih baik. Alasan memilih data kualitas udara karena udara sangat berpengaruh terhadap kehidupan manusia.

Udara merupakan salah satu elemen lingkungan terpenting yang berguna untuk menjaga kehidupan di seluruh muka bumi. Tubuh makhluk hidup membutuhkan oksigen dari atmosfer untuk menjalankan fungsi metabolisme

mereka. Selain oksigen, udara juga mengandung zat lain seperti formaldehida, karbon monoksida, karbon dioksida, jamur, dan virus (Fitria et al., 2008).

Masalah utama yang dihadapi oleh kota DKI Jakarta salah satunya yaitu pencemaran udara. Pencemaran udara terjadi ketika zat berbahaya tercampur ke dalam udara, hal ini dapat mengakibatkan kerusakan atau penurunan kualitas udara di DKI Jakarta. Oleh karena itu, akan terjadi gangguan kesehatan pada manusia (Simanjuntak, 2007).

Kualitas udara DKI Jakarta telah mengalami perubahan akibat dari perluasan fisik kota dan sentra industri. Kualitas udara yang bersih sangatlah dibutuhkan untuk aktivitas makhluk hidup dapat berjalan dengan lancar, diperlukan penelitian terkait dengan klasifikasi kualitas udara yang diharapkan dapat membantu Dinas Lingkungan Hidup DKI Jakarta untuk memberi himbauan terhadap masyarakat sekitar terhadap dampak buruk dari pencemaran udara.

Penelitian terdahulu mengenai klasifikasi kualitas udara yang dilakukan oleh Budiman (2023) menggunakan metode *Fuzzy Naive Bayes* dengan 1860 data masing-masing memperoleh hasil *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 92,32%, 54,05%, 63,02%, dan 57,56%. Penelitian oleh Safitri (2023) menggunakan metode *Fuzzy Random Forest* yang menggunakan 1825 data masing-masing memperoleh hasil *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 91,41%, 66,97%, 90%, dan 70,58%.

Penelitian terdahulu tentang klasifikasi menggunakan metode Regresi Logistik Multinomial oleh Rahmayani (2022) mengenai klasifikasi kualitas udara yang menggunakan 2.502 data masing-masing memperoleh hasil *accuracy*,

precision, *recall*, dan *f1-score* sebesar 99,31%, 78,45%, 78,63%, dan 78,54%. Penelitian oleh Saraswati (2022) mengenai klasifikasi hama dan penyakit tanaman jagung dengan 4.616 sampel masing-masing memperoleh hasil *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 94,75%, 69,79%, 59,58%, dan 64,16%.

Penelitian terdahulu tentang klasifikasi menggunakan metode *K-Nearest Neighbor* oleh Amalia et al. (2022) masing-masing memperoleh hasil *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 96%, 92%, 95%, dan 93%. Penelitian oleh Kurniawan & Barokah (2020) mengenai klasifikasi penentuan pengajuan kartu kredit yang menggunakan 250 data masing-masing memperoleh hasil *accuracy*, *precision*, dan *recall* sebesar 93%, 92%, dan 83%.

Berdasarkan uraian permasalahan tersebut serta tingginya ketepatan klasifikasi dalam metode Regresi Logistik Multinomial dan *K-Nearest Neighbor* pada penelitian terdahulu, peneliti tertarik melakukan penelitian klasifikasi kualitas udara DKI Jakarta dengan membandingkan kedua metode tersebut.

1.2 Rumusan Masalah

Adapun rumusan masalah pada penelitian ini yaitu sebagai berikut.

1. Bagaimana tingkat ketepatan klasifikasi kualitas udara menggunakan metode Regresi Logistik Multinomial?
2. Bagaimana tingkat ketepatan klasifikasi kualitas udara menggunakan metode *K-Nearest Neighbor*?
3. Bagaimana perbandingan metode Regresi Logistik Multinomial dan *K-Nearest Neighbor* berdasarkan tingkat ketepatan klasifikasi kualitas udara?

1.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu sebagai berikut.

1. Penelitian ini menggunakan data kualitas udara DKI Jakarta tahun 2015-2021 dengan rasio data latih tahun 2015-2019 sebesar 75% dan data uji tahun 2020-2021 sebesar 25%.
2. Variabel yang digunakan yaitu Partikulat Matter (PM₁₀), Karbon Monoksida (CO), Nitrogen Dioksida (NO₂), Ozon (O₃), Sulfur Dioksida (SO₂), dan Kategori AQI.
3. Tingkat ketepatan klasifikasi dibatasi dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score*.

1.4 Tujuan

Adapun tujuan dari penelitian ini yaitu sebagai berikut.

1. Memperoleh tingkat ketepatan klasifikasi kualitas udara pada metode Regresi Logistik Multinomial.
2. Memperoleh tingkat ketepatan klasifikasi kualitas udara pada metode *K-Nearest Neighbor*.
3. Membandingkan metode Regresi Logistik Multinomial dan *K-Nearest Neighbor* berdasarkan tingkat ketepatan klasifikasi kualitas udara.

1.5 Manfaat

Adapun manfaat dari penelitian ini yaitu sebagai berikut.

1. Penelitian ini dapat menjadi sumber bacaan dan media pembelajaran dalam meningkatkan ilmu tentang klasifikasi kualitas udara berdasarkan Indeks Standar Pencemar Udara.

2. Penelitian dapat digunakan sebagai referensi untuk penelitian selanjutnya yang membahas klasifikasi objek dengan menggunakan metode Regresi Logistik Multinomial dan K-Nearest Neighbor.
3. Penelitian ini menjadi bahan masukan atau pertimbangan dalam menentukan kebijakan terkait informasi pengukuran kualitas udara untuk Dinas Lingkungan Hidup DKI Jakarta.

DAFTAR PUSTAKA

- Agarwal, S. (2013). Data mining: data mining concepts and techniques. *International Conference on Machine Intelligence and Research Advancement*, 203-207.
- Amalia, A., Zaidiah, A., & Isnainiyah, I. N. (2022). Analisis dan komparasi algoritma klasifikasi dalam indeks pencemaran udara di DKI Jakarta. *Jurnal Informatika Dan Komputer*, 7(2), 496–507.
- Asrudin, N., Assiddieq, M., & Rosdiana. (2023). Analisis nitrogen dioksida (NO₂) terhadap penurunan kualitas udara ambien. *Jurnal Teluk*, 3(1), 22–26.
- Budiman, A. (2023). *Klasifikasi kualitas udara DKI Jakarta berdasarkan indeks standar pencemar udara (ISPU) dengan metode fuzzy naive bayes*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
- Defiyanti, S. (2017). Integrasi metode clustering dan klasifikasi untuk data numerik. *Conference on Information Technology and Electrical Engineering*, 256–261.
- Fajaryanti, N. (2023). *Penerapan metode fuzzy decision tree untuk klasifikasi kualitas udara DKI Jakarta*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
- Fitria, L., Wulandari, R. A., Hermawati, E., & Susanna, D. (2008). Kualitas udara dalam ruang perpustakaan Universitas “X” ditinjau dari kualitas biologi, fisik, dan kimiawi. *Jurnal Kesehatan*, 12(2), 76–82.
- Han, J., & Kamber, M. (2013). *Data Mining : Concepts and Techniques*. United States of America: Morgan Kaufmann Publisher.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. New York: John Wiley & Sons.
- Kareem, I. A., & Duaimi, M. G. (2014). Improved accuracy for decision tree algorithm based on unsupervised discretization association rules mining. *International Journal of Computer Science and Mobile Computing*, 3(6), 176–183.
- Kementerian Lingkungan Hidup dan Kehutanan Republik Indonesia. (2020). *Indeks Standar Pencemar Udara*. Jakarta.
- Kurniawan, Y. I., & Barokah, T. I. (2020). Klasifikasi penentuan pengajuan kartu kredit menggunakan k-nearest neighbor. *Jurnal Ilmiah Matrik*, 22(1), 73–82.

- Normawati, D., & Prayogi, S. A. (2021). Implementasi naive bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada twitter. *Jurnal Sains Komputer & Informatika*, 5(2), 697–711.
- Oktaviani, Z., & Jawwad, M. A. S. (2023). Rencana pengelolaan dan pemantauan kualitas udara terhadap rencana kegiatan pembangunan pemukiman. *Jurnal Sains Dan Teknik Lingkungan*, 4(1), 154–159.
- Pratiwi, B. P., Handayani, A. S., & Sarjana. (2020). Pengukuran kinerja sistem kualitas udara dengan teknologi WSN menggunakan confusion matrix. *Jurnal Informatika Upgris*, 6(2), 66–75.
- Rahmayani, M. (2022). *Prediksi kualitas udara menggunakan metode ensemble pada model decision tree ID.3, random forest, dan regresi logistik multinomial*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
- Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied Regression Analysis*. New York: Springer Verlag.
- Rixson, L., & Stefanus, M. (2017). Konsentrasi radioaktivitas lingkungan dalam partikulat udara (PM₁₀) di kawasan nuklir pasar jumat (KNPJ). *Prosiding Seminar Nasional Teknologi Pengelolaan Limbah XV*, 187–194.
- Safitri, A. (2023). *Pengklasifikasian kualitas udara DKI Jakarta dengan metode fuzzy random forest*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
- Saraswati, D. H. (2022). *Implementasi model regresi logistik multinomial berdasarkan repeated k-fold cross validation untuk klasifikasi hama dan penyakit tanaman jagung*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
- Simanjuntak, A. G. (2007). Pencemaran Udara. *Buletin Limbah*, 11(1), 34–40.
- Sreemathy, J. (2012). An efficient text classification using KNN and naive bayesian. *Computer Science And Engineering Journal*, 4(3), 392–396.
- Sunu, P. (2001). *Melindungi Lingkungan Dengan Menerapkan ISO 14001*. Jakarta: Gramedia Widiasarana Indonesia.
- Utomo, D. P., & Mesran. (2020). Analisis komparasi metode klasifikasi data mining dan reduksi atribut pada data set penyakit jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437–444.
- Wang, H., Xu, P., & Zhao, J. (2022). Improved KNN algorithms of spherical regions based on clustering and region division. *Alexandria Engineering Journal*, 3571–3585.

Wibawa, A. P., Pernama, M. G. A., Akbar, M. F., & Dwiyanto, F. A. (2018). Metode-metode klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi*, 3(1), 134–138.