



## TEKNIK MACHINE LEARNING UNTUK ANALISA KLASIFIKASI KUALITAS UDARA: A REVIEW

Haris Alfian<sup>1)</sup>, Sri Wahyuni<sup>2)</sup>, Aqil Revalino<sup>3)</sup>, M. Fitter Mirano<sup>4)</sup>, Elsa Rahmayana<sup>5)</sup>, Harun Mukhtar<sup>6\*)</sup>

<sup>1,2,3,4,5,6</sup>Fakultas Ilmu Komputer, Universitas Muhammadiyah Riau

<sup>1</sup>[230401364@student.umri.ac.id](mailto:230401364@student.umri.ac.id) <sup>2</sup>[230401384@student.umri.ac.id](mailto:230401384@student.umri.ac.id) <sup>3</sup>[230401357@student.umri.ac.id](mailto:230401357@student.umri.ac.id)

<sup>4</sup>[230401384@student.umri.ac.id](mailto:230401384@student.umri.ac.id) <sup>5</sup>[230401381@student.umri.ac.id](mailto:230401381@student.umri.ac.id) <sup>6</sup>[harunmukhtar@umri.ac.id](mailto:harunmukhtar@umri.ac.id)\*

### Abstract

*Air quality has a significant impact on human health and the environment, making its monitoring and classification extremely important. This review explores the application of machine learning techniques in analyzing and classifying air quality. Various methods such as decision trees, support vector machines, neural networks, and ensemble learning are evaluated to assess their effectiveness in processing complex and multidimensional air sensor data. This study also discusses challenges in data collection and preprocessing, selection of relevant features, and interpretation of classification results. Furthermore, this review identifies recent trends and future research opportunities in the use of machine learning to improve the accuracy and efficiency of air quality monitoring systems. The analysis results show that machine learning techniques have great potential to enhance our understanding of air quality dynamics and support better decision-making in environmental management*

**Keywords :** Machine Learning, Air Classification, Air Polution, Air Quality, Data Classification

### Abstrak

*Kualitas udara memiliki dampak signifikan terhadap kesehatan manusia dan lingkungan, sehingga pemantauan dan klasifikasinya menjadi sangat penting. Tinjauan ini mengeksplorasi penerapan teknik machine learning dalam menganalisis dan mengklasifikasikan kualitas udara. Berbagai metode seperti decision trees, support vector machines, neural networks, dan ensemble learning dievaluasi untuk menilai efektivitasnya dalam mengolah data sensor udara yang kompleks dan multidimensi. Penelitian ini juga membahas tantangan dalam pengumpulan dan preprocessing data, pemilihan fitur yang relevan, serta interpretasi hasil klasifikasi. Lebih lanjut, tinjauan ini mengidentifikasi tren terbaru dan peluang penelitian di masa depan dalam penggunaan machine learning untuk meningkatkan akurasi dan efisiensi sistem pemantauan kualitas udara. Hasil analisis menunjukkan bahwa teknik machine learning memiliki potensi besar untuk meningkatkan pemahaman kita tentang dinamika kualitas udara dan mendukung pengambilan keputusan yang lebih baik dalam pengelolaan lingkungan.*

**Keywords:** Machine Learning, Klasifikasi Udara, Pulusi Udara, Kualitas Udara, Klasifikasi Data

### PENDAHULUAN

Udara merupakan sebuah elemen tak kasat mata yang menjadi peran vital dalam kehidupan di bumi. Salah satu komponen utama dan penting dalam kehidupan makhluk hidup adalah udara atau oksigen (Octaviano et al. 2022b). Tanpa udara, makhluk hidup tidak akan dapat bertahan. Udara adalah bentuk transparan yang dapat kita rasakan,

meskipun tidak dapat dilihat. Udara bersih menyehatkan saluran pernapasan, mengurangi risiko penyakit jangka panjang, membantu memperpanjang umur, meningkatkan daya tahan tubuh dan kefokuskan, dan mood akan semakin meningkat (Kirono, Asror, and Wibowo 2022). "Menghirup udara di bawah standar kualitas udara 0–100 g/m<sup>3</sup> dapat menyebabkan penyakit jangka panjang" (Ibnus Pasedja and Mohammad Yanuar

Hariyawan 2021) “Poor air quality is damage to air. It can harm people, animals, and plants, interfere with aesthetics, comfort, or properties, and threaten human health”(Ramadhani, Prasetyowati, and Sibaroni 2025). Namun, untuk sekarang ini sangat sulit menentukan kualitas udara di wilayah tertentu, contohnya ditempat industri dan kota kota besar yang Dimana kasus pencemaran udaranya cukup besar dikarenakan asap kendaraan dan lainnya (Susrama and Diyasa 2024).

Salah satu Pencemaran udara kota disebabkan oleh terciptanya perkembangan serta pertumbuhan kota yang padat. Pencemaran udara menyebabkan masalah Kesehatan bagi banyak Masyarakat dunia, karena adanya kandungan senyawa didalamnya (Wahyudiayanta and Supriyati 2024). Gas yang berbahaya, partikel padat, ozon, nitrogen dioksida, sulfur, karbon monoksida, dan zat berbahaya lainnya adalah beberapa polutan yang dapat memengaruhi kesehatan makhluk hidup dan lingkungan(Wiranata et al. 2023). Penuaan dini, eksim , munculnya jerawat karena bakteri dan rhinitis adalah akibat dari polusi udara dalam jangka pendek, dan kanker paru paru, penyumbatan pembuluh darah asma, stroke dalam jangka panjangnya (Khusna et al. 2023). Namun, polutan dapat masuk ke udara secara tidak langsung(Prihastuti and Utami 2023) . Polusi udara merupakan sekumpulan partikulat asing yang terkandung pada komposisi udara yang mampu mengubah komposisi udara menjadi tidak normal,dan dengan adanya partikulat tersebut akan mengakibatkan kerusakan ekosistem makhluk hidup(Arkadia, Hananto, and Prasvita 2022). Keputusan tentang kualitas udara nantinya harus mempertimbangkan fakta dan data kualitas udara sebelumnya(Rifqi and Aldisa 2023) “*Air pollution is evil pollution that surrounds us and is very dangerous and detrimental to health. Air pollution can come from various sources, such as burning rubbish, fossil fuels, cigarette smoke, transportation fumes, industrial activities, etc. According to Waluyo (xxxx), EC, The entry or entry of a substance, energy, or other component into the air due to human activities that exceed established air quality standards is known as air pollution*”(Evitania 2023).

Pengembangan dari machine learning berbasis Neural network merupakan metode pembelajaran mendalam yang digunakan dalam penelitian ini. Metode ini menggunakan banyak lapisan tersembunyi yang memungkinkan pembelajar untuk mempelajari fitur data. Metode pembelajaran mendalam sudah ada di dalamnya dan III banyak digunakan di berbagai bidang, seperti pemrosesan bahasa alami, penglihatan komputer, dan audio, antara lain(Karyadi 2022), Teknologi seperti sensor polusi udara dan sistem pemantauan real-time dapat membantu mengatasi masalah polusi udara.

Mereka dapat memantau kualitas udara secara akurat dan memberikan informasi yang dapat diandalkan untuk pengambilan keputusan(Rizal, Arifin, Rasyd, Asvin, et al. 2023).

## METODOLOGI PENELITIAN

Metode penelitian ini adalah systematic literature review(SLR).Pendekatan ini dipilih untuk meninjau studi tentang tehnik machine learning yang digunakan untuk menganalisa kualitas udara, Systematic literature review adalah sebuah metode tinjauan yang terdiri dari mengidentifikasi, mengevaluasi dan memahami bukti penelitian yang tersedia dengan tujuan untuk menjawab researce question yang telah ditentukan. Teknik yang digunakan merupakan tinjauan pustaka sistematis untuk mengelompokkan beberapa hasil penelitian(Lubis 2023)

## PERNYATAAN PENELITIAN

Untuk memandu kita untuk mengidentifikasi,mengevaluasi dan memahami bukti penelitian yang tersedia dari penelitian sebelumnya,kami telah menentukan 4 pertanyaan yang akan dijawab daalam systematic literature riview ini. 4 researce question tersebut kami rangkum sebagai berikut:

1. RQ1 –Data set apa saja yang digunakan dalam menganalisa kualitas udara?
2. RQ2 – metode apa yang sering dan paling banyak digunakan para penulis sebelumnya dalam penelitian menganalisa kualitas udara?
3. RQ3 - Apaa saja faktor pengaruh machine learning digunakan dalam menganalisa kualitas udara?
4. RQ4 - Apa saja ukuran kinerja yang didapat terhadap analisa kualitas udara dengan menggunakan machine learning?

RQ1 Bertujuan untuk mengetahui data set mana saja dan apa bentuk data set yang sering digunakan dalam menggunakan machine learning untuk menganalisa kualitas udara, RQ2 bertujuan untuk mengidentifikasi dan menelaah apa saja tehnik machine learning yang bisa digunakan dalam menganalisa kualitas udara mengetahui data set mana saja dan apa bentuk data set yang sering digunakan dalam menggunakan machine learning untuk menganalisa kualitas udara, RQ3 bertujuan untuk mengetahui apa saja sih peran yang bisa diberikan machine learning dalam menganalisa kualitas udara, RQ4 bertujuan untuk melihat dan menelaah apa saja hasil kinerja yang dihasilkan para penulis sebelumnya setelah melakukan penelitian menganalisa kualitas udara.

## PROTOKOL PENINJAUAN

Protokol peninjauan ini bertujuan untuk mengevaluasi secara sistematis literatur yang ada mengenai penggunaan machine learning untuk

menganalisa kualitas udara. Kami akan menganalisa , menilai dan menemukan metode metode yang digunakan dalam penelitian menganalisa kualitas udara dari studi studi primer yang diterbitkan 5 tahun terakhir. Pencarian akan dilakukan didatabase elektronik utama yang mencakup google scholar, garuda, IEEE Explore dan sainece dirrect, dengan menggunakan keyword yang sudah ditentukan sebelumnya. Hasil peninjauan ini diharapkan dapat memberi wawasan berharga dan bermanfaat.

Untuk memilih studi utama dari daftar awal, kriteria inklusi dan eksklusi telah kami rancang,kriteria tersebut kami cantum dibawah ini:

- A. Kriteria Inklusi
- a) Studi yang membahas tentang penggunaan algoritma machine learning dalam menganalisa kualitas udara.

b) Studi yang membahas tentang perbandingan dua atau lebih algoritma dalam menganalisa kualitas udara.

c) Studi yang membahas tentang peran penting machine learning dalam menganalisa kualitas udara.

d) Studi yang membahas dan menjelaskan dari RQ1 , RQ2 , RQ3, RQ4 yang sudah kami jabarkan.

- B. Kriteria Pengecualian
- a) Penelitian yang tidak membahas tentang menganalisa / meneliti kualitas udara.

b) Kajian yang tidak membahas penggunaan machine learning.

c) Studi yang tidak menjawab RQ1, RQ2, RQ3, RQ4 yang telah kami jabarkan.

EKSTRAKSI DATA

Bentuk ekstraksi data dirancang untuk mengunpulkan jurnal junral utama yang kami temukan untuk dapat menjawab empat researc question yang telah kami jabarkan sebelumnya. Kumpulan jumlah jurnal akhir yang relevan dengan SLR ini kami jabarkan pada gambar 1 Dibawah ini, dan tabel 1 untuk menunjukkan hubungan antara penelitian utama dan reserace question, apakah penelitian tersebut bisa menjawab pertanyaan / researce question kita.



Gambar 1. Sumber Jurnal Yang Digunakan Dan Jumlah Ekstraksi Jurnal Yang Relevan

Tabel 1. Daftar Jurnal Yang Relevan

| Reference                                                        | RQ1 | RQ2 | RQ3 | RQ4 |
|------------------------------------------------------------------|-----|-----|-----|-----|
| Hidayatullah et al. 2022)                                        | ✓   | ✓   |     | ✓   |
| (Wiranata et al. 2023)                                           | ✓   |     | ✓   | ✓   |
| (Rifqi and Aldisa 2023)                                          | ✓   | ✓   |     |     |
| (Bagye 2020)                                                     | ✓   |     | ✓   |     |
| (Wahyudiyanta and Supriyati 2024)                                | ✓   | ✓   | ✓   |     |
| (Octaviano et al. 2022a)                                         | ✓   | ✓   | ✓   | ✓   |
| (Evitania 2023)                                                  | ✓   | ✓   | ✓   | ✓   |
| (Khumaidi, Raafi'udin, and Solihin 2020)                         | ✓   |     |     | ✓   |
| (Karyadi 2022)                                                   | ✓   | ✓   | ✓   | ✓   |
| (Valentino Jayadi, Handhayani, and Dolok Lauro 2023)             | ✓   | ✓   | ✓   | ✓   |
| (Sentinel- et al. 2023)                                          | ✓   | ✓   | ✓   | ✓   |
| (Rizal, Arifin, Rasyd, Suradi, et al. 2023)                      | ✓   |     | ✓   |     |
| (Ramadhani, Prasetyowati, and Sibaroni 2025)                     | ✓   | ✓   | ✓   | ✓   |
| (Perdana and Muklason 2023)                                      | ✓   | ✓   | ✓   | ✓   |
| (Kirono, Asror, and Wibowo 2022)                                 | ✓   | ✓   | ✓   | ✓   |
| (Adityo Nugroho, Ibnu Asror, and Yanuar Firdaus Arie Wibowo2023) | ✓   | ✓   | ✓   | ✓   |
| (Arkadi, Hananto, and Prasvita 2022)                             | ✓   | ✓   |     |     |
| (Ridho and Mahalisa 2023)                                        | ✓   | ✓   |     | ✓   |
| (Susrama and Diyasa 2024)                                        | ✓   | ✓   | ✓   | ✓   |
| (Ibnus Pasedja and Mohammad Yanuar Hariyawan 2021)               | ✓   | ✓   | ✓   | ✓   |
| (Roni et al. 2022)                                               | ✓   | ✓   |     | ✓   |
| (A. E. Putra et al. 2023)                                        | ✓   | ✓   | ✓   | ✓   |
| (R. E. Putra, Kalista, and ... 2023)                             | ✓   | ✓   | ✓   |     |
| (Rahmawati 2023)                                                 | ✓   | ✓   | ✓   | ✓   |
| (Rizky and Hakim 2020)                                           | ✓   | ✓   | ✓   | ✓   |
| (Khusna et al. 2023)                                             | ✓   | ✓   |     | ✓   |
| (Prihastuti and Utami 2023)                                      | ✓   | ✓   | ✓   | ✓   |
| (Bagye 2020)                                                     | ✓   |     | ✓   | ✓   |
| (Z. Zhang et al. 2024)                                           | ✓   | ✓   | ✓   | ✓   |
| (Imam et al. 2024)                                               | ✓   | ✓   |     | ✓   |
| (Wood 2022)                                                      | ✓   | ✓   | ✓   | ✓   |

|                                        |   |   |   |   |
|----------------------------------------|---|---|---|---|
| (Livingston et al. 2023)               | ✓ | ✓ |   | ✓ |
| (Al-Eidi et al. 2023)                  | ✓ | ✓ | ✓ | ✓ |
| (D. Zhang and Woo 2020)                | ✓ | ✓ |   | ✓ |
| (Farhadi, Palacin, and Blythe 2023)    | ✓ |   |   |   |
| (Liu et al. 2023)                      |   | ✓ | ✓ | ✓ |
| (Iskandaryan, Ramos, and Trilles 2023) | ✓ | ✓ | ✓ | ✓ |
| (Yang et al. 2023)                     | ✓ | ✓ | ✓ |   |
| (Nematirad, Behrang, and Pahwa 2024)   | ✓ | ✓ |   | ✓ |

HASIL DAN PEMBAHASAN  
DATA SET

Ada banyak dataset yang digunakan dalam penelitian dalam menganalisa kualitas udara yang telah kami ambil dari banyak artikel jurnal terdahulu mulai dengan menggunakan keyword “teknik” “machine learning” “analisa” “kualitas” “udara”, dan kami menggunakan artiker jurnal dari tahun 2020-2024.kami menemukan jumlah total dengan keyword yang telah kami tentukan diatas tadi dan memperoleh 39 jurnal yang relevan dan kami tetapkan sebagai jurnal utama yang akan kami gunakan untuk menjawab researce question yang kami jabarkan seblummnya. sebuah data adalah unsur yang paling utama dalam penelitian karena data merupakan objek utama yang akan diolah.(Hidayatullah et al. 2022). Beberapa dataset yang digunakan adalah :

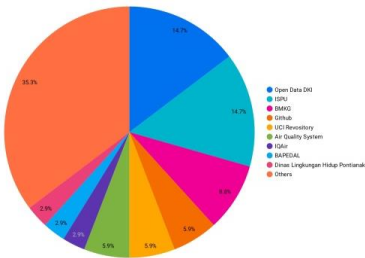
Dilakukan dengan metode regresi vektor support vector berdasarkan bentuk indeks standar pencemaran udara(ISPU) yang diperoleh dari website dinas lingkungan hidup DKI jakarta Jakarta open data(Hidayatullah et al. 2022), penelitian lainnya tentang prediksi kualitas udara menggunakan Metode LSTM dataset yang digunakan dataset yang diambil dari kualitas udara harian kota bandung musim 2019 yang diperoleh dari sensor dan perangkat IoT milik pemerintah kota bandung. Informasi juga dapat diperoleh dari website informasi terbuka pemerintahan DKI Jakarta. Data yang diambil merupakan data kkualitas udara DKI Jakarta pada bulan febuari 2021 hingga oktober 2021 (Adityo Nugroho, Ibnu Asror, and Yanuar Firdaus Arie Wibowo 2023).

Sekumpulan dataset juga diambil dari data cuaca yang dimana penulisnya membahas tentang kondisi cuaca yang dapat menyebabkan kebakaran hutan,maka pengambilan data cuaca juga bisa digunkan untuk mengidentifikasi dan menganalisa kualitas udara contohnya Data cuaca diperoleh dari IBM Cloud[28] dimana data ini berisikan sebuah data cuacapada rentang waktu januari 2015 sampai september 2019 diprovinsi riau dengan total

keseluruhan data 1733 melalui riset yang dilakukan oleh (Sentinel- et al. 2023).

Kajian klasifikasi kualitas udara menggunakan algoritma naive bayes menggunakan data hasil aplikasi terbuka sekunder.data sekunder itu sendiri merupakan sebuah data yang diperoleh bukan dari subjek utama penelitiannya, melainkan di ambil dari sumber sumber yang sudah ada sebelumnya. Open government data(OGP) telah menjadi platfrom bagi negara negara peserta untuk mengembangkan data kelola yang mengedepankan keterbukaan. Partisipasi masyarakat,akuntablilitas,dan penggunaan teknologi untuk memperkuat tata kelola(Kirono, Asror, and Wibowo 2022).

Dan masih banyak dataset yang digunakan oleh penulis terlebih dahulu dalam menganalisa kualitas udara seperti uci repository machine learning, CSV (Comma Separated Values) dari jakarta open data yangdigunakan untuk menganalisa kualitas udara, dataset dari data primer dinas lingkungan hidup provinsi yang digunakan untuk peramalan kualitas udara. Dan juga contohnya dalam artikel international yang dilakukan oleh (Farhadi, Palacin, and Blythe 2023) The dataset from the Newcastle Urban Observatory is used. Dan ada juga peneliti yang menggunakan datset kaggle seperti penelitian yang dilakukan oleh(Ridho and Mahalisa 2023) Pengumpulan dataset dengan cara pengunduhan pada API Command kaggle.



Gambar 2. Presentase Jumlah Dataset Yang Umum Dan Banyak Digunakan.

METODE / TEKNIK

Dalam menganalisa kualitas udara banyak terdapat metode yang digunakan oleh peneliti dengan memanfaatkan machine learning.dari 39 jurnal

yang relevan yang kami gunakan sebagai jurnal utama untuk acuan kami, kami menemukan sebanyak 59 metode yang digunakan dalam menganalisa kualitas udara.

Metode yang paling umum dan paling banyak digunakan adalah support vector machine (SVM) kami menemukan 8 jurnal yang menggunakan metode SVM dalam menganalisa kualitas udara. - Support vector machine merupakan metode klasifikasi yang menggunakan metode pembelajaran mesin untuk memprediksi kelas berdasarkan model hasil proses pelatihan, yang menggunakan prinsip minimalisasi struktural. Klasifikasi biasanya diterapkan pada bidang batas, yang kemudian diklasifikasikan menjadi kelas positif dan negatif (Valentino Jayadi, Handhayani, and Dolok Lauro 2023). Support vector machine merupakan system pembelajaran yang menggunakan ruang hipotesis berupa fungsi linear pada fitur berdimensi tinggi dan dilatih oleh algoritma pembelajaran berdasarkan teori optimasi (Maulina Putri, Arhami, and Hendrawaty 2023). Proses klasifikasi menggunakan algoritma SVM memberikan hasil evaluasi kinerja yang sangat baik, dengan akurasi tinggi, presisi, recall, dan f1-score yang memuaskan. Model ini dapat diandalkan untuk membuat prediksi dan mengolah data terkini terkait kualitas udara (Wahyudiyanta and Supriyati 2024). Dalam sebuah penelitian lain juga dilakukan oleh (Susrama and Diyasa 2024) dan menulis dalam artikel mereka bahwa SVM juga terlibat dalam beberapa hyperparameter seperti nilai  $c$  atau  $\gamma$ . Secara keseluruhan, nilai yang dipilih untuk hyperparameter memiliki dampak terhadap kinerja prediksi model.

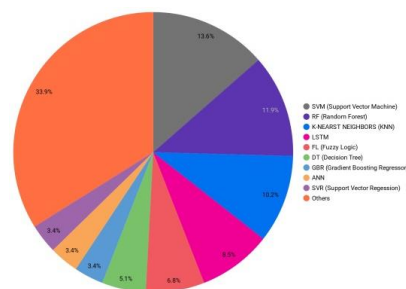
Metode banyak kedua yang sering digunakan dalam menganalisa kualitas udara adalah Random Forest (RF), RF adalah sebuah algoritma machine learning yang digunakan untuk mengklasifikasikan kumpulan suatu data besar. karena fungsi ini dapat digunakan dengan banyak dimensi berbeda dan performa tinggi. (Adityo Nugroho, Ibnu Asror, and Yanuar Firdaus Arie Wibowo 2023), metode Support vector machine (SVM) juga digunakan oleh (Ridho and Mahalisa 2023) dalam penelitiannya.

Metode paling banyak ketiga adalah K-NEAREST NEIGHBOARD kami menemukan ada 6 jurnal yang menggunakan metode itu. K-NEAREST NEIGHBOR (K-NN) merupakan sebuah algoritma yang populer dalam machine learning yang dapat digunakan sebagai regresi dan klasifikasi. Algoritma ini bekerja berdasarkan "nearest neighbor" dimana prediksi untuk instance baru berdasarkan pada pengidentifikasi atau nilai yang dimilikinya. Dalam nearest neighbor, nilai  $K$  adalah jumlah tetangga terdekat yang akan

dipertimbangkan untuk melakukan prediksi. (Wiranata et al. 2023).

Ada juga penulis yang membandingkan antara metode KNN dan SVM untuk klasifikasi kualitas udara. Sang peneliti atau penulis menjelaskan tentang KNN dan juga SVM sebagai metode klasifikasi kualitas udara. Sang penulis menuliskan bahwa KNN classifier adalah metode untuk mengklasifikasikan objek berdasarkan data latih yang paling dekat dengan objek tersebut. Data latih tersebut diproyeksikan kedalam ruang multidimensi, dimana setiap dimensi mewakili suatu properti dari data tersebut. Ruang dimensi tersebut dibagi berdasarkan klasifikasi dan pelatihan, dan bedanya dengan SVM adalah SVM (support vector machine) adalah sebuah metode klasifikasi yang menerapkan metode machine learning untuk memprediksi kelas berdasarkan pola hasil proses pelatihan dengan menggunakan prinsip minimalisasi resiko struktural. Algoritma support vector machine (SVM) bertujuan untuk menemukan cutoff yang paling baik sebagai garis pemisah antara kedua kelas, dan memperoleh hasil berdasarkan hasil SVM presisi, akurasi, recall dan score F1. Algoritma ini lebih baik dibandingkan algoritma KNN dalam mengklasifikasikan tingkat pencemaran kualitas udara. (Valentino Jayadi, Handhayani, and Dolok Lauro 2023)

Selain 3 metode diatas ada juga yang menggunakan metode environmental health risk assessment (EHRA), salah satunya (Rahmawati 2023) menganalisis risiko kesehatan yang diterima pekerja akibat paparan polutan udara dengan menggunakan metode Environmental Health Risk Assessment (EHRA). Dan dalam salah satu artikel internasional ada juga yang menggunakan metode yang namanya SHapley Additive exPlanations (SHAP) method (Nematirad, Behrang, and Pahwa 2024) Dan masih banyak metode lainnya yang digunakan dalam klasifikasi dan analisa kualitas udara ini, selengkapnya kami rangkum dalam sebuah grafik berikut ini:



Gambar 3. Persentase Penggunaan Metode Machine Learning Yang Digunakan

## FAKTOR YANG MEMPENGARUHI

Kualitas udara dipengaruhi oleh berbagai faktor kompleks yang saling terkait. Faktor faktor utama tersebut meliputi emisi dari sumber antropogenik seperti industri, transportasi, dan aktivitas rumah tangga, serta sumber alami seperti debu dan zat zat lainnya.

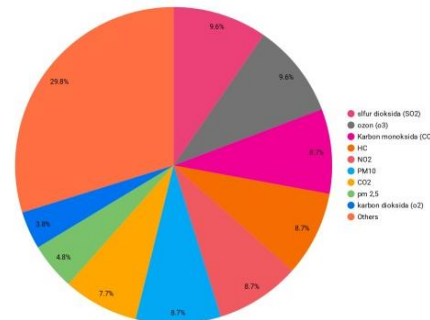
Dalam artikel yang dibuat oleh (Wiranata et al. 2023) kami menemukan faktor pengaruh yang digunakan sebagai faktor yang mempengaruhi PM10(partikel udara),SO2(sulfur dioksida), CO(karbon monoksida, O3(ozon) dan NO2(nitrogen dioksida) 1,dalam penelitian lainnya yang dilakukan oleh (Bagye 2020) dia menggunakan CO, CO<sub>2</sub>, HC, PM10, suhu dan kelembaban dalam penelitiannya.peneliti lainnya juga sama yaitu (Rizky and Hakim 2020) yang menetapkan faktor yang mempengaruhi dalam artikelnya adalah karbon monoksida (CO), sulfur dioksida (SO<sub>2</sub>), nitrogen dioksida(NO<sub>2</sub>),ozon permukaan (O<sub>3</sub>), dan partikel debu (PM10).

Dalam artikel lain yang ditulis oleh (R. E. Putra, Kalista, and ... 2023) yang menetapkan faktor pengaruh dalam penelitiannya adalah partikulat (PM10), sulfur dioksida (SO<sub>2</sub>), karbon monoksida (CO), ozon (O<sub>3</sub>), dan nitrogen dioksida (NO<sub>2</sub>), faktor pengaruh yang sama juga digunakan oleh (Roni et al. 2022) yaitu Hidrokarbon (HC), Karbon monoksida (CO), Sulfur dioksida (SO<sub>2</sub>), Nitrogen dioksida (NO<sub>2</sub>), Ozon (O<sub>3</sub>), dan Partikulat (PM10 dan PM<sub>2.5</sub>).

Pada salah satu penelitian international yang dilakukan oleh (Imam et al. 2024) *these material can be classified into three main group. The intial category includes pollutants that directly affect the AQI value and most notably include particular matter such as PM<sub>2.5</sub>, NO<sub>2</sub>, and other. The second category includes factor such as temperature and relative hummadity(RH)), which indirectly affect air quality. The third category contains information about the time and date of each sampling. Table 1 shows the statistical analisys of factor affecting each AQL Value” and (Wood 2022) used a dataset of o<sub>3</sub> (Ozone), CO(carbon monoxide), NO<sub>2</sub>(Nitrogen dioxide), SO<sub>2</sub>(Sulfur dioxide), C<sub>6</sub>H<sub>6</sub>(Benzene) and particullate matter 2.5(PM<sub>2.5</sub>). (Livingston et al. 2023) “The proposed model used eight parameters like NO<sub>2</sub>CO, O<sub>3</sub>, PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, TEMP, PRES, DEWP,*

*RAIN,WD, WSPM of Beijing dataset in order to predict AQI and it is tested with regional dataset”.*

Untuk data selengkapnya kami kumpulkan pada grafik berikut:



Gambar 4. Persentase Faktor Yang Mempengaruhi Kualitas Udara

Dari grafik diatas ozon dan sulfur dioksida(so<sub>2</sub>) menjadi faktor pengaruh yang paling banyak ditetapkan penulis sebagai faktor pengaruh dalam penelitiannya, lalu ada juga carbon monoksida, HC, NO<sub>2</sub> dengan jumlah faktor pengaruh yang kami temukan dalam jurnal yang relevan.

## UKURAN KINERJA

Dalam menganalisa kinerja machine learning untuk klasifikasi kualitas udara, berbagai ukuran kinerja digunakan untuk mengevaluasi efektivitas dan kendala model oleh para peneliti. Terdapat 36 ukuran kinerja yang kami temukan.

Salah satu ukuran kinerja terbanyak yang didapatkan adalah RMSE(Root mean square error). Kami menemukan terdapat 7 artikel yang memiliki ukuran kinerja RMSE. Root mean square error merupakan ukuran kinerja yang sangat umum digunakan untuk menilai keakuratan prediksi yang diperoleh dari suatu model.keuntungan utama menggunakan RMSE adalah memberikan pinalti terhadap kesalahan besar dan meskalakan skor dalam satuan yang sama dengan nilai prediksi (Perdana and Muklason 2023). Root mean square atau yang disingkat dengan RMSE merupakan akar kuadrat dari rata rata akar kuadrat semua nilai kesalahan. Penggunaan RMSE sangat umum dan dianggap sebagai ukuran kesalahan tujuan umum yang sangat baik untuk prediksi numerik(Hidayatullah et al. 2022).dalam sebuah penelitian yang dilakukan oleh (Khumaidi, Raafi'udin, and Solihin 2020). Hasil perhitungan RMSE data dan uji adalah 2,37 RMSE untuk train score dan 3,15 RMSE untuk test score. Kinerja yang cukup baik diperoleh karena nilai prediksi

RMSE lebih rendah dari nilai standar deviasi data eksperimen.

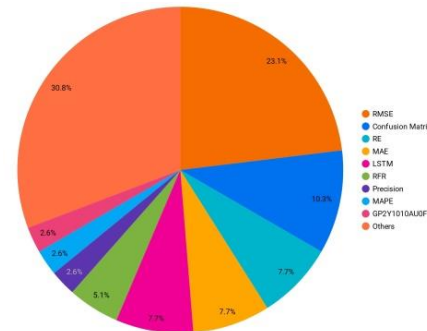
In international journal articles, several common performance measure are also generally used, one of which is (Al-Eidi et al. 2023). Root Mean Square Error (RMSE): Banyak digunakan dalam memperkirakan model regresi. Ini digunakan untuk menghitung deviasi rata-rata antara nilai model prediksi dan aktual. Nilai RMSE yang lebih rendah menunjukkan bahwa model mencapai kinerja yang lebih baik and several other studies also use RMSE as a performance measure such as in the research conducted by (D. Zhang and Woo 2020) and (Liu et al. 2023) and (Yang et al. 2023)

And confusion matrix as the second most commonly used performance measure in air quality classification analysis that we found in several previous journal we used in our article. Salah satu artikel yang ditulis oleh (Sentinel- et al. 2023). Pada pengujian model klasifikasi dengan algoritma random forest digunakan metode confusion matrix dengan perbandingan 90% data latih dan 10% data uji yang memberikan akurasi 71%, nilai presisi 0,30, recall 0,45, dan F1. Skor 0,36 inci. Kelas kosong maka nilai akurasi 0,88 ingat 0,69 dan skor f1 0,77 pada kelas kecil nilai akurasi 0,12. Matriks konfusi adalah metode untuk mengukur efektivitas metode klasifikasi. Intinya matriks konfusi berisi informasi yang membandingkan hasil klasifikasi yang dihasilkan system dengan hasil klasifikasi yang harusnya dihasilkan (Roni et al. 2022).

Matriks konfusi adalah sebuah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi pada kumpulan data yang kelasnya diketahui. Matriks konfusi memiliki 4 kombinasi nilai prediksi dan nilai aktual yang berbeda (A. E. Putra et al. 2023). nilai akurasi pada matriks konfusi dihitung dengan cara membagi jumlah perkiraan benar (true positive dan true negative) dengan jumlah total yang ada pada dataset. Juga dijelaskan dalam penelitian yang dilakukan oleh (Kirono, Asror, and Wibowo 2022) dan (Adityo Nugroho, Ibnu Asror, and Yanuar Firdaus Arie Wibowo 2023) matriks konfusi merupakan sebuah table matriks yang digunakan sebagai perhitungan performa dari suatu model data atau algoritma.

Ukuran kinerja paling banyak ketiga adalah LSTM (long short term memory) terdapat 3 artikel yang menggunakan ukuran kinerja LSTM salah satunya artikel internasional dari (Z. Zhang et al. 2024) and (Farhadi, Palacin, and Blythe 2023) and (Iskandaryan, Ramos, and Trilles 2023) long short-term memory (LSTM) network are a type of RNN designed to learn sequence prediction problem.

LSTM have four components: a memory cell that stores for a certain period of time, and three gates (input, output, and forget gates) that control the flow of data in and out of the cell.



Gambar 5: Persentase Ukuran Kinerja Yang Umum Digunakan Dalam Menganalisa Kualitas Udara

## DISKUSI

Dalam penelitian ini, kami telah menelaah beberapa jurnal publikasi dari beberapa sumber tentang Teknik Machine Learning untuk Analisis Kualitas Udara yang diterbitkan dalam lima tahun terakhir (2019-2024). Maksud dan tujuan dilakukannya penelitian ini untuk memberikan gambaran analisis kualitas udara menggunakan machine learning dan membahas metode yang terdapat dalam berbagai referensi yang sudah kami analisis. Kami telah melakukan pencarian di berbagai repositori digital sehingga kami menemukan bermacam faktor pengaruh yang dapat mempengaruhi kualitas udara menurut referensi yang kami gunakan, termasuk juga perbedaan metode machine yang digunakan.

Berdasarkan pengamatan kami, kualitas udara dapat diukur dengan beberapa cara dan faktor pengaruhnya yang beragam meskipun terdapat beberapa metode yang belum bisa dipahami dengan baik.

Beberapa metode yang banyak digunakan antara lain:

1. Support vector machine (SVM) : SVM telah terbukti handal dalam klasifikasi dan regresi kualitas udara. Misalnya, menggunakan SVM untuk mengklasifikasikan tingkat polusi udara di kota Cilacap. (Fadil Danu Rahman, Mulki, and Taryana 2024)
2. Hybrid Models: Penggabungan beberapa teknik machine learning sering menghasilkan performa yang lebih baik. Contohnya, menggabungkan

LSTM dengan Gaussian process regression untuk meningkatkan akurasi prediksi

3. Algoritma Random Forest: metode ini menjadi alternatif menarik dalam meningkatkan akurasi prakiraan cuaca berdasarkan data BMKG dari Stasiun Meteorologi Tunggul Wulung Cilacap (Fadil Danu Rahman, Mulki, and Taryana 2024)

4. K-Nearest Neighbors (KNN): penerapan algoritma K-NN dapat menyelesaikan masalah berupa pembagian data (sampel) dan hasil uji data, kemudian data uji dikelompokkan menggunakan K-nearest neighbour. Langkah awalnya adalah mengumpulkan seluruh data parameter kualitas udara dan kemudian membagikan pelatihan data, dilakukan pengujian data. Algoritma K-NN sangat relevan dengan jenis informasi yang digunakan dalam penelitian ini. (Sanmorino et al. 2022) 5. Deep learning : merupakan metode untuk mempelajari representasi pada beberapa level representasi yang diperoleh dengan menggabungkan modul sederhana namun non-linier, yang masing-masing mengubah representasi pada level yang sedikit lebih tinggi dan canggih. (yoshua bengio 2019) mengembangkan demonstrate profound learning untuk prediksi PM2.

5. jangka panjang dengan hasil yang menjanjikan. (yoshua bengio 2019) mengembangkan model deep learning untuk prediksi PM2.5 jangka panjang dengan hasil yang menjanjikan.

Meskipun teknik-teknik ini menunjukkan hasil yang menjanjikan, masih ada tantangan yang perlu diatasi. Misalnya, ketersediaan data yang berkualitas dan representatif masih menjadi kendala di banyak daerah. Selain itu, interpretabilitas model machine learning yang kompleks juga perlu ditingkatkan agar dapat lebih mudah diadopsi oleh pembuat kebijakan.

Ke depannya, integrasi data dari berbagai sumber yang digunakan seperti IEEE, Google Scholar, Garuda, ScienceDirect dan sumber data lainnya dengan teknik machine learning yang canggih diharapkan dapat semakin meningkatkan akurasi dan kehandalan analisis kualitas udara. Data biner mengikuti distribusi Binomial.

Model konseptual yang dipakai untuk mengidentifikasi dan mengevaluasi permasalahan dalam penelitian sistem analisis udara menggunakan machine learning dengan metode yang beragam. Support Vector Machine (SVM) adalah metode yang paling umum digunakan dalam penelitian ini, SVM merupakan klasifikasi yang

menerapkan teknik machine learning untuk memprediksi kelas berdasarkan model dari hasil suatu proses pelatihan, dengan menggunakan prinsip minimalisasi risiko struktural. (Chowanda et al. 2021) Dataset yang digunakan dari Website Jakarta Open Data dalam format comma-separated value (CSV) dalam file tersendiri setiap bulannya.

Analisis yang dilakukan dengan mengolah data adalah deskriptif dan analisis grafik yang dipakai dalam dataset. Analisis deskriptif membantu untuk mengatur data sedemikian rupa sehingga lebih mudah dipahami, menjelaskan hasil uji secara detail, dan menarik kesimpulan lebih lanjut. Analisis deskriptif yang dilakukan antara lain adalah dengan menampilkan data, ukuran kinerja, metode yang digunakan, dan sumber data set (Amiroh, Permata, and Rahmanti 2019)

Kualitas udara berdasarkan kadar O<sub>3</sub> diukur oleh stasiun pemantauan kualitas udara ambien (SUFA) dan dapat ditampilkan menggunakan Indeks Standar Kualitas Udara (ISPU) yang mencakup parameter O<sub>3</sub>. Apabila konsentrasi O<sub>3</sub> di udara mencapai Tingkat yang berbahaya bagi makhluk hidup atau nilai ISPU melebihi 50, maka nilai ISPU parameter O<sub>3</sub> bernilai 1, sebaliknya bernilai 0. Data kualitas udara berdasarkan parameter ISPU O<sub>3</sub> data biner karena berbasis nilai. (Kartikasari 2020)

Melihat dari pertanyaan penelitian yang sudah dijelaskan sebelumnya, pertanyaan pertama yaitu mengenai dataset yang sering digunakan dalam menganalisa kualitas udara. Banyak penelitian yang dilakukan dengan menggunakan metode regresi vektor pendukung atau support vector regression (SVR) berdasarkan pengumpulan data yang diperoleh dari website Jakarta Open Data dinas lingkungan hidup DKI Jakarta yang berupa indeks standar pencemaran udara (ISPU). (Hidayatullah et al. 2022)

Pertanyaan kedua yang dicantumkan adalah metode apa yang sering dan paling banyak digunakan para penulis sebelumnya dalam penelitian menganalisa kualitas udara? SVM dipilih sebagai teknik (CITASI). yang paling banyak digunakan untuk menganalisa udara berdasarkan data set yang dilampirkan dan SVR menjadi teknik yang paling sedikit digunakan dalam penelitian menurut referensi yang sudah kami analisis. Selanjutnya, kami mampu meningkatkan pemilihan teknik yang digunakan untuk analisis kualitas udara karena hanya menggunakan teknik yang paling banyak digunakan ini tidak menjamin peningkatan performa dalam menganalisa. Namun masih banyak referensi lain yang berfokus pada

teknik metode yang lain yang menghasilkan hasil yang baik.

Pertanyaan penelitian yang ketiga adalah tentang faktor pengaruh machine learning digunakan dalam menganalisa kualitas udara. Kualitas udara dipengaruhi oleh berbagai faktor kompleks yang saling terkait. Dalam penelitian ini (Bagye 2020) proses kalifikasi kualitas udara dilakukan dengan memanfaatkan metode support vector machine(SVM) dengan tiga node sensor. Data pengujian diperoleh dari pembacaan sensor menggunakan sistem pemantau kualitas udara yang mencakup parameter kualitas udara seperti CO, CO<sub>2</sub>, HC, PM<sub>10</sub>, suhu, dan kelembapan udara.

Dalam artikel yang dibuat oleh(Wiranata et al. 2023) kami menemukan faktor pengaruh yang digunakan sebagai untuk faktor yang mempengaruhi Partikel Udara (Pm<sub>10</sub>), Sulfur Dioxide (so<sub>2</sub>), carbon Monoxide(co), Ozon (o<sub>3</sub>), Nitrogen Dioxide (no<sub>2</sub>) dalam penelitian lainnya yang dilakukan oleh(Bagye 2020) dia menggunakan CO, CO<sub>2</sub>, HC, PM<sub>10</sub>, suhu dan kelembaban dalam penelitiannya.peneliti lainnya juga sama yaitu(Rizky and Hakim 2020) yang menetapkan faktor yang mempengaruhi dalam artikelnya adalah karbon monoxide(CO), sulfur dioxide (SO<sub>2</sub>), nitrogen dioxide (NO<sub>2</sub>), surface ozone (O<sub>3</sub>), dan partikel debu (PM<sub>10</sub>). Faktor diatas merupakan faktor utama dalam pembahasan kualitas udara yang dilakukan. Dalam penelusuran yang dilakukan pada penelitian ini kami mendapatkan gambaran kualitas udara berdasarkan faktorfaktor yang mempengaruhi lalu hasil dari pengujian menggunakan metode yang telah dipakai pada penelitian ini.

Pertanyaan terakhir yang kami uraikan dalam penelitian ini adalah ukuran kinerja yang didapat terhadap analisa kualitas udara dengan menggunakan machine learning. Dalam penelitian ini kami menemukan 36 ukuran kinerja yang digunakan. Dimana ukuran kinerja yang paling banyak digunakan yaitu RMSE (Root mean square error). Kami menemukan terdapat 7 artikel yang memiliki ukuran kinerja RMSE.RMSE merupakan sebuah matrik yang umum digunakan untuk mengevaluasi keakuratan prediksi yang diperoleh dari model. RSME populer digunakan karena pendekatan pengukuran kinerja ini dianggap sebagai metrik kesalahan umum yang baik untuk peramalan numerik.

Dalam studi ini, kami juga menyajikan beberapa table dan grafik yang menggambarkan beberapa pertanyaan yang kami bahas sebelumnya yang mana table dan grafik tersebut dapat menggambarkan dengan jelas persentase jawaban

dari pertanyaan-pertanyaan kami pada pembahasan diatas.

Penelitian ini bertujuan untuk menganalisis penelitian sebelumnya terkait analisis kualitas udara menggunakan mechine learning. Sebagian besar penelitian memiliki kumpulan data yang sangat banyak, namun kami tidak dapat memastikan apakah kumpulan data tersebut mewakili skenario nalisis kualitas udara menggunakan mechine learning. Untuk penelitian ini, kami tidak melakukan pembacaan manual atas semua judul makalah yang diterbitkan selama pembuatan makalah ini. Faktanya, kami melakukan string pencarian yang telah kami buat sebelumnya untuk menemukan studi atau data yang relevan dengan judul yang akan kami bahas dalam penelitian ini. Kami telah mencari penelitian sebanyak yang kami bisa sesuai inklusi dan kriteria. Namun ada kemungkinan kami telah melampaui penelitian lain yang mungkin lebih tepat sasaran dalam judul penelitian kami.

## SIMPULAN DAN SARAN

### KESIMPULAN

Dari tinjauan sistematis yang sudah dilakukan, dapat disimpulkan bahwa penggunaan machine learning dalam analisis kualitas udara telah menunjukkan perkembangan yang signifikan dalam beberapa tahun terakhir. Berbagai dataset digunakan dalam penelitian, mulai dari data sensor IoT hingga data open government, menunjukkan keragaman sumber informasi yang dapat dimanfaatkan. Metode yang paling sering digunakan adalah Support Vector Machine (SVM), Random Forest (RF), dan K-Nearest Neighbor (KNN), dengan masing-masing metode menunjukkan kelebihan dalam aspek klasifikasi dan prediksi yang berbeda.

Faktor-faktor yang mempengaruhi kualitas udara yang sering dianalisis meliputi partikel udara (PM<sub>10</sub>, PM<sub>2.5</sub>), gas-gas seperti sulfur dioxide (SO<sub>2</sub>), carbon monoxide (CO), ozone (O<sub>3</sub>), dan nitrogen dioxide (NO<sub>2</sub>). Ini menunjukkan kompleksitas di pemodelan kualitas udara yang melibatkan berbagai parameter. Dalam hal ukuran kinerja, Root Mean Square Error (RMSE) dan Confusion Matrix menjadi alat evaluasi yang dominan, memberikan gambaran tentang akurasi dan presisi model yang dikembangkan

Penelitian ini juga mengungkapkan bahwa penggunaan machine learning dalam analisis kualitas udara tidak hanya terbatas pada prediksi dan klasifikasi, tetapi juga mencakup aspek interpretasi model dan penilaian risiko kesehatan. Hal ini terlihat dari penggunaan metode seperti SHAP (SHapley Additive exPlanations) dan EHRA

(Environmental Health Risk Assessment) dalam beberapa studi.

Kesimpulannya, penggunaan machine learning dalam analisis kualitas udara menunjukkan potensi besar untuk meningkatkan pemahaman kita tentang dinamika polusi udara dan membantu dalam pengambilan keputusan yang lebih baik terkait manajemen kualitas udara. Namun, perlu diperhatikan bahwa keakuratan dan keandalan model sangat bergantung pada kualitas data input dan pemilihan metode yang tepat. Penelitian lebih lanjut diperlukan untuk mengoptimalkan penggunaan teknologi ini dalam konteks yang lebih luas dan kompleks dari manajemen kualitas udara perkotaan dan global.

## DAFTAR PUSTAKA

- Adityo Nugroho, Ibnu Asror, and Yanuar Firdaus Arie Wibowo. 2023. "Klasifikasi Tingkat Kualitas Udara DKI Jakarta Berdasarkan Open Government Data Menggunakan Algoritma Random Forest." *e-Proceeding of Engineering* 10(2): 1824.
- Al-Eidi, Shorouq, Fathi Amsaad, Omar Darwish, Yahya Tashtoush, Ali Alqahtani, and Niveshitha Niveshitha. 2023. "Comparative Analysis Study for Air Quality Prediction in Smart Cities Using Regression Techniques." *IEEE Access* 11(September): 115140–49. doi:10.1109/ACCESS.2023.3323447.
- Amiroh, Khodijah, Oktavia Ayu Permata, and Farah Zakiyah Rahmanti. 2019. "Analisis Kualitas Udara Untuk Monitoring Kesehatan Lingkungan Rumah Sakit." *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)* 4(1): 29–36. doi:10.30743/infotekjar.v4i1.1549.
- Arkadia, Arvi, Bayu Hananto, and Desta Sandya Prasvita. 2022. "Optimasi Long Short Term Memory Dengan Adam Menggunakan Data Udara Kota DKI Jakarta." *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*: 92–101.
- Bagye, Wire. 2020. "Klasifikasi Kualitas Udara Dengan Metode Support Vector." *JIRE (Jurnal Informatika & Rekayasa Elektronika)* 3(2): 187–99.
- Chowanda, Andry, Rhio Sutuyo, Meiliana, and Sansiri Tanachutiwat. 2021. "Exploring TextBased Emotions Recognition Machine Learning Techniques on Social Media Conversation." *Procedia Computer Science* 179(2019): 821–28. doi:10.1016/j.procs.2021.01.099.
- Evitania, Claudyana Gabrillia. 2023. "Implementation of the K-Nearest Neighbor Algorithm to Predict Air Pollution." 1(1): 45–54.
- Fadil Danu Rahman, Mulki Indana Zulfa Mulki, and Acep Taryana. 2024. "Clustering Dan Klasifikasi Data Cuaca Cilacap Dengan Menggunakan Metode K-Means Dan Random Forest." *Jurnal SINTA: Sistem Informasi dan Teknologi Komputasi* 1(2): 90–97. doi:10.61124/sinta.v1i2.15.
- Farhadi, Farzaneh, Roberto Palacin, and Phil Blythe. 2023. "Machine Learning for Transport Policy Interventions on Air Quality." *IEEE Access* 11(May): 43759–77. doi:10.1109/ACCESS.2023.3272662.
- Hidayatullah, Benny Khalid, Meta Kallista, Cati Setianingsih, Prodi S1, and Teknik Komputer. 2022. "Prediksi Indeks Standar Pencemar Udara Menggunakan Metode Support Vector Regression Berbasis Web (Studi Kasus Pada Kota Jakarta) Prediction of Air Pollutant Standard Index Using the Web-Based Long Short-Term Memory Method (Case Study in the Jakarta Capita)." *e - Proceeding of Engineering* 9(3): 1247–55. <https://data.jakarta.go.id/>.
- Ibnus Pasedja, Diemas, and dan Mohammad Yanuar Hariyawan. 2021. "Sistem Monitoring Kualitas Udara Berbasis Nodemcu Esp8266 Dengan Metode Fuzzy Logic (Software)." *Abec* 9: 1050–62. [www.bmkg.go.id](http://www.bmkg.go.id).
- Imam, Mohsin, Sufiyan Adam, Soumyabrata Dev, and Nashreen Nesa. 2024. "Air Quality Monitoring Using Statistical Learning Models for Sustainable Environment." *Intelligent Systems with Applications* 22(January): 200333. doi:10.1016/j.iswa.2024.200333.
- Iskandaryan, Ditsuhi, Francisco Ramos, and Sergio Trilles. 2023. "Graph Neural Network for Air Quality Prediction: A Case Study in Madrid." *IEEE Access* 11(December 2022): 2729–42. doi:10.1109/ACCESS.2023.3234214.
- Kartikasari, Diah. 2020. "Analisis Faktor-Faktor Yang Mempengaruhi Level Polusi Udara Dengan Metode Regresi Logistik Biner." *MATHunesa: Jurnal Ilmiah Matematika* 8(1): 55–59. doi:10.26740/mathunesa.v8n1.p55-59.
- Karyadi, Yadi. 2022. "Prediksi Kualitas Udara Dengan Metoda LSTM, Bidirectional LSTM, Dan GRU." *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)* 9(1): 671–84. doi:10.35957/jatisi.v9i1.1588.
- Khumaidi, Ali, Ridwan Raafi'udin, and Indra Permana Solihin. 2020. "Pengujian Algoritma Long Short-Term Memory Untuk Prediksi Kualitas Udara Dan Suhu Kota Bandung." *Jurnal Telematika* 15(1): 13–18.
- Khusna, Nida Faoziatun, Syifa Aulia, Shinta Amaria, Alfidha Rahmah, Safril Ahmadi Sanmas, and Fatkhurokhman Fauzi. 2023. "Peramalan Kualitas Udara Di Semarang Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) Forecasting Air Quality in Semarang Using the Autoregressive Integrated Moving Average (ARIMA) Method." *Prosiding Seminar Nasional UNIMUS*: 426–35.
- Kirono, Abdul Aziiz Hendrie, Ibnu Asror, and Yanuar Firdaus Arie Wibowo. 2022. "Klasifikasi Tingkat Kualitas Udara Dki Jakarta Menggunakan Algoritma Naïve Bayes." *e-Proceeding of Engineering* 9(3): 1962.
- Liu, Chunhao, Guangyuan Pan, Dongming Song, and Hao Wei. 2023. "Air Quality Index Forecasting via Genetic Algorithm-Based Improved Extreme Learning Machine." *IEEE Access* 11(July): 67086–97. doi:10.1109/ACCESS.2023.3291146.
- Livingston, S. John, S. Deepa Kanmani, A. Shamila Ebenezer, Dahlia Sam, and A. Joshi. 2023. "An Ensembled Method for Air Quality Monitoring and Control Using Machine Learning." *Measurement: Sensors* 30(March): 100914. doi:10.1016/j.measen.2023.100914.
- Lubis, Arifin Shaleh. 2023. "Perilaku Sarapan Mempengaruhi Pada Pelajar Dalam Beraktivitas ; Literatur Rivew." *Journal Transformation of Mandalika* 4(6): 260–66. <http://ojs.cahayamandalika.com/index.php/jtm/issue/archive>.
- Maulina Putri, Salma Sheila, Muhammad Arhami, and Hendrawaty Hendrawaty. 2023. "Penerapan Metode SVM Pada Klasifikasi Kualitas Air." *Journal of Artificial Intelligence and Software Engineering (J-AISE)* 3(2): 93. doi:10.30811/jaise.v3i2.4630.
- Nematirad, Reza, Mehdi Behrang, and Anil Pahwa. 2024. "Acoustic-Based Online Monitoring of Cooling Fan Malfunction in Air-Forced Transformers Using Learning Techniques." *IEEE Access* 12(February): 26384–400. doi:10.1109/ACCESS.2024.3366807.
- Octaviano, Alvino, Sofa Sofiana, Dion Orlando Agustino, and Perani Rosyani. 2022a. "Pemantauan Kualitas Udara Berbasis Internet O Things." *Media Online* 3(2): 147–56.
- Octaviano, Alvino, Sofa Sofiana, Dion Orlando Agustino, and Perani Rosyani. 2022b. "Pemantauan Kualitas Udara Berbasis Internet Of Things." 3(2).
- Perdana, Diaz, and Ahmad Muklason. 2023. "Machine Learning Untuk Peramalan Kualitas Indeks Standar Pencemar Udara DKI Jakarta Dengan Metode Hibrid ARIMAX-LSTM." *ILKOMNIKA: Journal of Computer Science and Applied Informatics* 5(3): 209–22. doi:10.28926/ilkomnika.v5i3.588.
- Prihastuti, Santi, and Anisa Utami. 2023. "Pengaruh Kualitas Udara Perkotaan Terhadap Willingness to Pay Menggunakan Metode Hedonic Price The Effect of Urban

- Air Quality on Willingness to Pay Using the Hedonic Price Method." *Geoplanart* 5(1): 30–39.
- Putra, Andika Eka, Tedy Rismawan, Jurusan Rekayasa, and Sistem Komputer. 2023. "Klasifikasi Kualitas Udara Berdasarkan Indeks Standar Pencemaran Udara ( ISPU ) Menggunakan Metode Fuzzy Tsukamoto." *Jurnal Komputer dan Aplikasi* 11(02): 190–96. <https://jurnal.untan.ac.id/index.php/jcskommipa/article/view/58704/75676598673>.
- Putra, R E, M Kalista, and ... 2023. "Klasifikasi Prediksi Kualitas Udara Menggunakan Metode Support Vector Machine (SVM)." *eProceedings ...* 10(4): 3790–96. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/20808%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/download/20808/20307>.
- Rahmawati, Ariska. 2023. "Analisis Kualitas Udara Ambien Pt X Menggunakan Metode Environmental Health Risk Assessment (Ehra)." *Jurnal Kesehatan Tambusai* 4(September): 3992–4001. <https://journal.universitaspahlawan.ac.id/index.php/jkt/article/view/18240/14238>.
- Ramadhani, Rizky Fauzi, Sri Suryani Prasetyowati, and Yuliant Sibaroni. 2025. "Performance Analysis of Air Pollution Classification Prediction Map with Decision Tree and ANN." 3(4): 536–43. doi:10.47065/josyc.v3i4.2117.
- Ridho, Ihda Innar, and Galih Mahalisa. 2023. "Analisis Klasifikasi Dataset Indeks Standar Pencemaran Udara (Ispu) Di Masa Pandemi Menggunakan Algoritma Support Vector Machine (Svm)." *Technologia : Jurnal Ilmiah* 14(1): 38. doi:10.31602/tji.v14i1.8005.
- Rifqi, Ahmad, and Rima Tamara Aldisa. 2023. "Penerapan Data Mining Untuk Clustering Kualitas Udara." *Jurnal Sistem Komputer dan Informatika (JSON)* 5(2): 289–97. doi:10.30865/json.v5i2.7145.
- Rizal, Muhammad, Arham Arifin, Muhammad Furqan Rasyd, Andi Asvin, Mahersatillah Suradi, and Akbar Bahtiar. 2023. "Design and Implementation of a Real-Time Air Pollution Monitoring System Based on Android at SMKS Darul Ulum Layoa Bantaeng Desain Dan Implementasi Sistem Pemantauan Polusi Udara Berbasis Android Real-Time Di SMKS Darul Ulum Layoa Bantaeng." 3(October): 143–52.
- Rizal, Muhammad, Arham Arifin, Muhammad Furqan Rasyd, Andi Asvin Mahersatillah Suradi, and Akbar Bahtiar. 2023. "Desain Dan Implementasi Sistem Pemantauan Polusi Udara Berbasis Android Real-Time Di SMKS Darul Ulum Layoa Bantaeng." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 3(2): 143–52. doi:10.57152/malcom.v3i2.894.
- Rizky, Robby robby, and Zaenal Hakim. 2020. "Implementasi Metode Fuzzy Sugeno Untuk Sistem Pengukuran Kualitas Udara Di Kota Pandeglang Berbasis Internet Of Things (IOT)." *Syntax : Jurnal Informatika* 9(1): 15–25. doi:10.35706/syji.v9i1.3439.
- Roni, I Komang, I Gusti Agung, Gede Arya, and I Wayan Supriana. 2022. "Implementasi Metode K-Nearest Neighbors Pada Sistem Klasifikasi Kualitas Udara Berdasarkan Partikulat Berbahaya Yang Terkandung." 1(November): 417–22.
- Sanmorino, Ahmad, Juhaini Alie, Nining Ariati, and Sanza Vittria Wulanda. 2022. "K-NN Based Air Classification as Indicator of the Index of Air Quality in Palembang." *Sinkron* 7(3): 853–59. doi:10.33395/sinkron.v7i3.11469.
- Sentinel-, Visualisasi Citra, Ahmad Efendi, Iwan Iskandar, Rahmad Kurniawan, and Muhammad Affandes. 2023. "Klasifikasi Kebakaran Hutan Riau Menggunakan Random Forest Dan." 4(3): 1602–12. doi:10.30865/klik.v4i3.1521.
- Susrama, I Gede, and Mas Diyasa. 2024. "PENGUKURAN INDEKS STANDAR PENCEMARAN UDARA MENGGUNAKAN SUPPORT VECTOR MACHINE Dimas Dzaky Daniswara , Aviolla Terza Damaliana , I Gede Susrama Mas Diyasa Pencemaran Udara Tertinggi ( Agus et Al ). Kualitas Udara Di Jakarta Terpengaruh." *Jurnal Penelitian Politeknik Penerbangan Surabaya* 9(1).
- Valentino Jayadi, Bryan, Teny Handhayani, and Manatap Dolok Lauro. 2023. "Perbandingan Knn Dan Svm Untuk Klasifikasi Kualitas Udara Di Jakarta." *Jurnal Ilmu Komputer dan Sistem Informasi* 11(2). doi:10.24912/jiksi.v11i2.26006.
- Wahyudiyanta, Soni Agung, and Supriyati. 2024. "Analisis Kualitas Udara Jakarta Dan Prediksi Tingkat Polusi Dengan Metode Mesin Pembelajaran SVM." *Prosiding Sains dan Teknologi* 3(1): 278. <https://predict-air-polution.streamlit.app/>.
- Wiranata, Ade Davy, Soleman Soleman, Irwansyah Irwansyah, I Ketut Sudaryana, and Rizal Rizal. 2023. "Klasifikasi Data Mining Untuk Menentukan Kualitas Udara Di Provinsi DKI Jakarta Menggunakan Algoritma K-Nearest Neighbors (K-Nn)." *Infotech: Journal of Technology Information* 9(1): 95–100. doi:10.37365/jti.v9i1.164.
- Wood, David A. 2022. "Local Integrated Air Quality Predictions from Meteorology (2015 to 2020) with Machine and Deep Learning Assisted by Data Mining." *Sustainability Analytics and Modeling* 2(October 2021): 100002. doi:10.1016/j.samod.2021.100002.
- Yang, Qian, Xinlei Zhang, Qing Luan, Shuai Sun, Yongqiang Zhao, Chenggege Fang, Xiaonan Mi, and Mengwei Li. 2023. "Optimization of Hourly PM2.5 Inversion Model Integrating Upper Air Meteorological Elements." *IEEE Access* 11(April): 35421–28. doi:10.1109/ACCESS.2023.3263790. yoshua bengio. 2019. "Deep Learning 简介 一、什么是 Deep Learning ?." *Nature* 29(7553): 1– 73.
- Zhang, Dan, and Simon S. Woo. 2020. "Real Time Localized Air Quality Monitoring and Prediction through Mobile and Fixed IoT Sensing Network." *IEEE Access* 8: 89584–94. doi:10.1109/ACCESS.2020.2993547.
- Zhang, Zhen, Shiqing Zhang, Caimei Chen, and Jiwei Yuan. 2024. "A Systematic Survey of Air Quality Prediction Based on Deep Learning." *Alexandria Engineering Journal* 93(February): 128–41. doi:10.1016/j.aej.2024.03.