



Analisis Prediksi Kadar CO₂ (PPM) Berdasarkan Suhu Dan Kelembapan Menggunakan Sensor DHT22 Dan MQ135 Dengan Perbandingan Model Machine Learning Berbasis IoT Pada ESP32

Bacilius Agung Suburdjati^{*1}, Lisda Sukoco², Xavier Cedric³

^{*1,2}Universitas Santo Borromeus

³SMA Santa Angela Bandung

Email: ^{*1}basnagung@gmail.com, ²lidasukoco2014@gmail.com, ³xaviercededa@gmail.com

Abstract

This study aims to analyze and predict carbon dioxide (CO₂) concentration in parts per million (PPM) based on temperature and humidity using DHT22 and MQ135 sensors integrated with an ESP32 microcontroller as an Internet of Things (IoT) system. Data were collected in two measurement sessions. The first session was conducted from 16:00 to 20:00 with a one-second sampling interval, during which significant spikes occurred due to exposure to insecticide fumes. The second session was carried out from 00:00 to 08:00 with a two-second sampling interval, producing approximately 13,000 more stable data points suitable for machine learning analysis. The collected data were processed using several machine learning algorithms, including Linear Regression, Random Forest, CatBoost, LightGBM, XGBoost, and Neural Network (Keras). The results indicate that the CatBoost model achieved the best performance, with an accuracy of 98.9%, a Mean Absolute Error (MAE) of 2.75, and a Mean Relative Error (MRE) of 1.08%. This study demonstrates that an ESP32-based IoT implementation using integrated gas and temperature sensors can be utilized as a simple air quality prediction system and serve as a foundation for developing future intelligent environmental monitoring devices.

Keywords : Internet of Things, CO₂ Prediction, CatBoost, ESP32, Machine Learning, MQ135, DHT22, Air Quality Monitoring.

Abstrak

Penelitian ini bertujuan untuk menganalisis dan memprediksi kadar karbon dioksida (CO₂) dalam satuan part per million (PPM) berdasarkan pengaruh suhu dan kelembapan menggunakan sensor DHT22 dan MQ135 yang diintegrasikan pada mikrokontroler ESP32 sebagai sistem Internet of Things (IoT). Data dikumpulkan dalam dua sesi pengujian. Sesi pertama dilakukan pada pukul 16.00–20.00 dengan interval 1 detik, di mana terdapat lonjakan signifikan akibat paparan asap insektisida. Sesi kedua dilakukan pada pukul 00.00–08.00 dengan interval 2 detik, menghasilkan sekitar 13.000 data yang lebih stabil untuk analisis pembelajaran mesin. Data yang diperoleh diolah menggunakan beberapa algoritma machine learning, yaitu Linear Regression, Random Forest, CatBoost, LightGBM, XGBoost, dan Neural Network (Keras). Hasil menunjukkan bahwa model CatBoost memberikan performa terbaik dengan tingkat akurasi mencapai 98,9%, nilai Mean Absolute Error (MAE) sebesar 2.75, dan Mean Relative Error (MRE) sebesar 1.08%. Penelitian ini menunjukkan bahwa implementasi IoT berbasis ESP32 dengan integrasi sensor gas dan suhu dapat digunakan sebagai sistem prediksi kualitas udara sederhana, serta menjadi dasar pengembangan perangkat pemantauan lingkungan cerdas di masa depan.

Kata Kunci : Internet of Things, CO₂ Prediction, CatBoost, ESP32, Machine Learning, MQ135, DHT22, Air Quality Monitoring.

I. PENDAHULUAN

Perkembangan teknologi Internet of Things (IoT) telah mendorong penerapan sistem pemantauan

lingkungan yang mampu bekerja secara otomatis dan real-time. Salah satu bidang yang memperoleh manfaat signifikan dari perkembangan ini adalah



pemantauan kualitas udara, khususnya kadar karbon dioksida (CO_2) yang menjadi salah satu indikator utama kesehatan lingkungan. Konsentrasi CO_2 yang melebihi ambang batas normal diketahui dapat berdampak negatif terhadap kesehatan manusia, seperti menurunnya kemampuan kognitif, sakit kepala, sesak napas, dan penurunan performa belajar maupun bekerja (World Health Organization, 2023). Kondisi ini lebih rentan terjadi pada wilayah perkotaan dengan tingkat aktivitas manusia yang tinggi, sehingga diperlukan sistem monitoring yang akurat, terukur, dan mampu memberikan informasi secara langsung. Dalam konteks tersebut, IoT menawarkan solusi melalui integrasi sensor, mikrokontroler, dan platform komputasi yang memungkinkan pengukuran dan analitik berbasis data secara simultan. Pada penelitian ini dikembangkan sebuah sistem pemantauan kualitas udara berbasis IoT menggunakan mikrokontroler ESP32 yang dilengkapi sensor DHT22 untuk mengukur suhu dan kelembapan, serta sensor MQ-135 untuk mengukur konsentrasi CO_2 dalam satuan PPM (*parts per million*). Data yang diperoleh dikirim secara otomatis ke server dan dashboard berbasis web, serta disimpan dalam format CSV untuk keperluan analisis lanjutan.

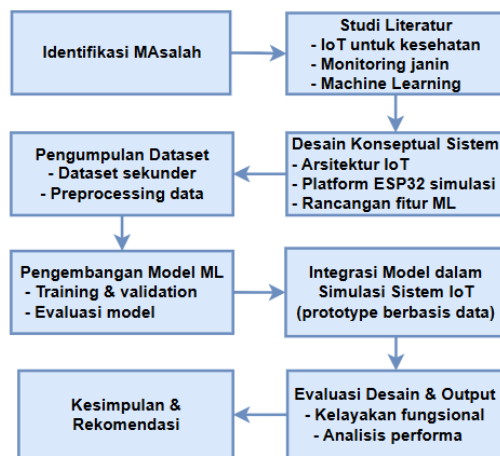
Selain melakukan pengukuran, penelitian ini juga berfokus pada pemanfaatan model machine learning untuk memprediksi kadar CO_2 berdasarkan parameter suhu dan kelembapan. Beberapa model pembelajaran mesin digunakan untuk mengevaluasi performa prediksi, antara lain Linear Regression, Random Forest, CatBoost, LightGBM, XGBoost, dan Neural Network (Keras). Data dikumpulkan melalui dua sesi pengujian luar ruangan: pengambilan data pertama dilakukan pada pukul 16.00–20.00 dengan interval 1 detik dan mengalami anomali peningkatan akibat paparan insektisida aerosol, sedangkan sesi kedua dilakukan pukul 00.00–08.00 dengan interval 2 detik dan menghasilkan sekitar 13.000 data yang lebih stabil.

Penelitian ini diharapkan dapat memberikan gambaran mengenai hubungan antara suhu, kelembapan, dan kadar CO_2 serta mengidentifikasi model prediksi dengan performa terbaik yang dapat diterapkan pada sistem pemantauan udara berbasis IoT. Dengan demikian, sistem ini tidak hanya berfungsi sebagai alat monitoring pasif, tetapi juga sebagai sistem prediktif yang berpotensi mendukung pengembangan teknologi lingkungan cerdas (*smart environmental monitoring system*) di masa depan.

II. METODE PENELITIAN

Tahapan Penelitian dan Alur Kerja

Tahapan penelitian pada studi ini disusun secara sistematis untuk memastikan proses penelitian berjalan terstruktur, terukur, dan sesuai dengan tujuan penelitian. Secara umum, alur penelitian terdiri dari beberapa tahap utama mulai dari identifikasi masalah hingga evaluasi akhir sistem.



Gambar 1. Tahapan dan Alur Penelitian

Tahap awal dimulai dengan **identifikasi masalah** dan **studi literatur** untuk memperoleh pemahaman teoretis mengenai teknologi Internet of Things (IoT), sensor gas lingkungan, serta algoritma pembelajaran mesin yang relevan. Setelah parameter dan kebutuhan sistem ditentukan, penelitian berlanjut pada perancangan sistem dan pengumpulan data menggunakan sensor MQ135 dan DHT22 yang terhubung dengan mikrokontroler ESP32.

Data yang dikumpulkan kemudian melalui proses **preprocessing**, meliputi pembersihan data, normalisasi, dan pembentukan fitur untuk memastikan dataset siap digunakan pada tahap pemodelan.

Tahap berikutnya adalah pelatihan model machine learning, evaluasi performa model, serta pemilihan model terbaik berdasarkan metrik akurasi, MAE, MRE, dan nilai koefisien determinasi (R^2). Model dengan performa terbaik kemudian diintegrasikan dalam prototipe sistem IoT sehingga mampu melakukan prediksi kadar CO_2 secara real-time.

Tahapan terakhir adalah evaluasi sistem, analisis performa model ketika berjalan pada prototipe, serta penyusunan kesimpulan dan rekomendasi pengembangan lanjutan.

Metode Pengambilan Data

Pengumpulan data dilakukan menggunakan perangkat IoT berbasis ESP32 yang terhubung dengan sensor MQ135 untuk mendeteksi kadar CO_2 dalam satuan PPM, serta sensor DHT22 untuk mengukur suhu ($^{\circ}C$) dan kelembapan (%).

Data direkam dalam dua sesi pengujian untuk memastikan variasi kondisi lingkungan:



Se si	Wakt u	Interval Samplin g	Kondisi Lingkungan
1	16:00 – 20:00	1 detik	Kondisi bervariasi, terdapat gangguan aktivitas luar
2	00:00 – 08:00	2 detik	Lingkungan tenang, data relatif stabil

Data disimpan dalam format **CSV** melalui web server ESP32 untuk dianalisis lebih lanjut menggunakan Python.

Pemrosesan Data (Preprocessing)

Sebelum digunakan dalam proses pelatihan model, data terlebih dahulu melalui tahap preprocessing untuk memastikan kualitas dan konsistensi dataset. Tahapan yang dilakukan meliputi: Menghapus nilai kosong (*missing values*), Menghilangkan duplikasi data, Menangani outlier dan noise, Normalisasi nilai, Pembentukan fitur (*feature engineering*), Membagi dataset menjadi **80% training dan 20% testing**. Tahapan ini memastikan data siap untuk proses pemodelan machine learning.

Pengembangan dan Evaluasi Model Machine Learning

Beberapa model machine learning diterapkan untuk memprediksi kadar CO₂ berdasarkan variabel suhu dan kelembapan. Model yang digunakan meliputi: Linear Regression, Random Forest, CatBoost, LightGBM, XGBoost,

dan Neural Network (Keras)

Evaluasi performa dilakukan menggunakan metrik berikut: Mean Absolute Error (MAE), Mean Relative Error (MRE), Nilai koefisien determinasi (R²), Akurasi prediksi (%). Model dengan performa tertinggi kemudian dipilih sebagai model final untuk diintegrasikan ke sistem IoT.

Integrasi Model ke Sistem IoT

Model terbaik kemudian diterapkan ke dalam sistem IoT berbasis ESP32 untuk memungkinkan prediksi kadar CO₂ secara real-time. Prototipe dilengkapi dengan: Output prediksi numerik, Klasifikasi kondisi kualitas udara, Rekomendasi kesehatan berbasis usia pengguna. Integrasi ini menghasilkan prototipe sistem monitoring kualitas udara yang tidak hanya mencatat data sensor, tetapi juga memberikan informasi prediktif dan rekomendasi berbasis machine learning.

III. HASIL DAN PEMBAHASAN

Data penelitian diperoleh melalui sistem berbasis *Internet of Things* (IoT) menggunakan ESP32, sensor DHT22, dan sensor MQ135. Data yang dikumpulkan mencakup 3 variable utama, yaitu:



Tabel 1 :Tabel variabel utama

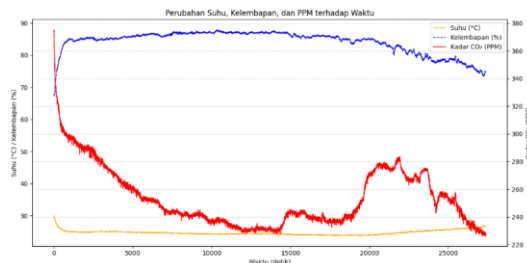
Variabel	Satuan	Sumber sensor	Keterangan
Suhu	°C	DHT22	Diukur untuk melihat pengaruh terhadap kadar CO ₂ .
Kelembapan	%RH	DHT22	Digunakan sebagai variable pembanding terhadap suhu.
Kadar CO ₂ (PPM)	ppm	MQ135	Nilai target (variable dependen) untuk prediksi model machine learning.

Dua sesi pengambilan data menghasilkan sekitar 13.000 data pengamatan.

Sesi pertama (16:00-20:00) menunjukkan banyak fluktuasi akibat penyemprotan insektisida dan aktivitas luar ruangan.

Sesi kedua (00:00-08:00) menghasilkan data yang jauh lebih stabil, dengan perubahan suhu bertahap dari 24°C hingga 29°C dan kelembapan dari 78% ke 62%.

Hasil pengamatan menunjukkan hubungan umum antara suhu dan kelembapan terhadap kadar CO₂. Ketika suhu meningkat, kadar CO₂ cenderung menurun secara perlahan, sedangkan pada kelembapan tinggi (di atas 75%) kadar CO₂ sering meningkat.



Gambar 2. Grafik garis 3 sumbu (waktu vs suhu, kelembapan dan ppm) dari data raw ESP32

Interpretasi:

Nilai suhu cenderung naik saat siang dan menurun menjelang malam.

Nilai kelembapan menunjukan pola berlawanan dengan suhu, meningkat ketika udara lebih dingin.

Nilai PPM CO₂ menunjukan fluktuasi yang signifikan pada periode tertentu dikarenakan adanya aktivitas di luar ruangan seperti mobil lewat, dan mobil masuk ke dalam halaman rumah.

Terlihat hubungan tidak *linear* antara suhu-kelembapan-ppm, sehingga pendekatan *machine learning* digunakan untuk membangun model prediktif yang lebih kompleks.

Data yang telah dikumpulkan kemudian digunakan untuk melatih beberapa model *machine learning*, yaitu:

Linear regression, Random Forest Regressor, CatBoost Regressor, LightGBM Regressor, XGBoost Regressor, Neural Network (keras)

Tabel 2. Tabel Hasil pelatihan model

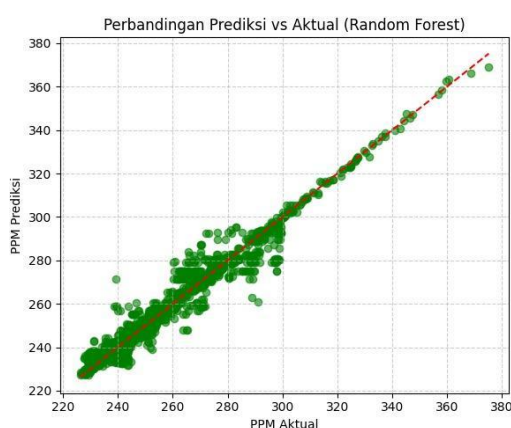
Model	MAE	R2	MRE(%)	Akurasi (%)
Linear regression	8.446	0.763	3.26	96.74
Random Forest	2.746	0.962	1.08	98.92
CatBoost	2.751	0.962	1.08	98.92
LightGBM	2.759	0.962	1.09	98.91
XGBOOST	2.762	0.960	1.09	98.91
Keras	7.913	0.641	2.96	97.04

Analisis hasil:

Model dengan performa terbaik adalah **Random Forest Regressor** dengan akurasi 98,92%, disusul CatBoost dan LightGBM.

Model Neural Network cenderung lebih sensitif terhadap variasi data, sehingga menghasilkan fluktuasi loss yang lebih tinggi meskipun secara konsep lebih kompleks.

Model *Linear Regression* sederhana memiliki performa paling rendah karena tidak mampu menangkap hubungan *non-linear* antar variabel (suhu, kelembapan, dan PPM).



Gambar 3. Gambar scatter plot “PPM aktual dengan PPM Prediksi (Random Forest)”

Grafik berikut memperlihatkan perbandingan antara nilai aktual dan hasil prediksi dari model terbaik *Random Forest*. Model ini menunjukkan sebaran data yang mendekati garis ideal, menandakan kemampuan prediksi yang baik terhadap nilai CO₂ (PPM).

Nilai PPM yang dihasilkan oleh sensor MQ135 ini terkadang **Tidak Stabil**. Hal ini disebabkan oleh:

- Harga dan sensitivitas sensor. MQ135 termasuk sensor gas dengan harga ekonomis yang memiliki tingkat *noise* dan *drift* lebih tinggi dibandingkan sensor laboratorium profesional. Seperti dinyatakan oleh Han et al. (2021), sensor gas berbasis semikonduktor “memiliki variabilitas respons akibat suhu lingkungan dan kelembapan udara.”
- Kondisi lingkungan. Pengukuran dilakukan di luar ruangan, sehingga variasi angin, cahaya, dan debu turut mempengaruhi hasil.
- Kalibrasi sederhana. Nilai R₀ (Resistansi sensor pada udara bersih) hanya dikalibrasi sekali, sehingga akurasinya menurun jika kondisi lingkungan berubah.

Meskipun demikian, hasil model menunjukkan bahwa dengan *machine learning*, data dari sensor murah dapat tetap digunakan untuk memprediksi kadar CO₂ dengan tingkat akurasi yang tinggi setelah dilatih dengan benar.

Setelah model terbaik (*Random Forest Regressor*) diperoleh, sistem dilengkapi dengan fungsi *post-processing* berupa klasifikasi tingkat kualitas udara dan rekomendasi kesehatan otomatis.

Kedua fungsi yang digunakan adalah:



```
def klasifikasi_ppm(ppm):  
    if ppm < 400:  
        return "Aman (Udara bersih)"  
    elif ppm < 1000:  
        return "Sedang (Masih normal)"  
    elif ppm < 2000:  
        return "Tidak Sehat (Batasi paparan lama)"  
    else:  
        return "Berbahaya (Ventilasi segera dibuka)"
```

Gambar 4. Fungsi klasifikasi ppm

Fungsi ini mengelompokkan nilai CO₂ (PPM) menjadi empat kategori utama berdasarkan ambang atas batas yang direferensikan dari standar kualitas udara dalam ruangan (EPA, 2022).

Klasifikasi ini penting karena nilai PPM saja tidak selalu intuitif untuk pembaca umum.

Selanjutnya fungsi berikut digunakan untuk memberikan rekomendasi spesifik berdasarkan umur pengguna, suhu, dan kelembapan.

```
def prediksi_kesehatan(temp, hum, umur):  
    ...
```

Gambar 5. Fungsi yang digunakan untuk rekomendasi spesifik umur, suhu, kelembapan.

```
Suhu: 29°C | Kelembapan: 70% | Umur: 9 tahun  
Prediksi PPM: 340.79  
Kategori Udara: Aman (Udara bersih)  
Rekomendasi: Anak-anak sangat sensitif terhadap udara kotor!  
  
Suhu: 31°C | Kelembapan: 63% | Umur: 25 tahun  
Prediksi PPM: 368.68  
Kategori Udara: Aman (Udara bersih)  
Rekomendasi: Dewasa masih toleran, tapi ventilasi harus cukup.  
  
Suhu: 27°C | Kelembapan: 85% | Umur: 65 tahun  
Prediksi PPM: 304.79  
Kategori Udara: Aman (Udara bersih)  
Rekomendasi: Lansia sebaiknya hindari paparan lama pada udara tidak sehat.
```

Gambar 6. Output dari 2 fungsi yang sudah didefinisikan.

Kesimpulan analisis:

- d. Model *Random Forest* memberikan hasil paling akurat untuk prediksi kadar CO₂ dengan akurasi 98.91%.

- e. Data yang dikumpulkan pada malam hari yang dilakukan di sesi kedua lebih stabil dibandingkan dengan sore hari yang dilakukan pada sesi pertama.
- f. Faktor lingkungan terbuka dan kualitas sensor ekonomis menjadi penyebab utama fluktuasi data.
- g. Sistem IoT ini berpotensi dikembangkan menjadi dashboard pemantauan kualitas udara real-time yang hemat biaya dan informatif.

IV. KESIMPULAN

Kesimpulan:

- a. Penelitian menunjukkan bahwa suhu dan kelembapan berpengaruh terhadap kadar CO₂ dengan pola hubungan non-linear suhu meningkat cenderung menurunkan kadar CO₂ sedangkan kelembapan tinggi meningkatnya.
- b. Dari enam model *machine learning* yang diuji **Random Forest** memberikan hasil yang terbaik dengan akurasi 98,91%, R2 sebesar 0,962, dan MAE sebesar 2,74.
- c. Data yang diambil antara 00:00-08:00 lebih stabil dibandingkan dengan sore hari yang dilakukan pada sesi pertama karena minimnya gangguan lingkungan pada jam 00:00-08:00.



Saran:

- a. Kalibrasi sensor secara rutin perlu dilakukan terutama pada sensor MQ135, untuk mengurangi kesalahan pengukuran akibat perubahan suhu, kelembapan dan pengaruh gas lain seperti amonia atau etanol.
- b. Untuk penelitian selanjutnya, disarankan menggunakan sensor gas dengan tingkat presisi lebih tinggi, seperti MH-Z19B atau Senseair S8, yang memiliki kompensasi suhu internal dan kalibrasi otomatis, sehingga hasil pengukuran kadar CO₂ akan lebih akurat dan stabil.
- c. Disarankan untuk meningkatkan sistem pemantauan menjadi dashboard web real-time, yang menampilkan data historis, status kualitas udara, serta memberikan peringatan otomatis jika kadar CO₂ melewati ambang batas berbahaya.
- d. Penelitian lanjutan juga dapat melibatkan pengujian di berbagai kondisi lingkungan (ruangan tertutup, ruang publik, area industri) untuk memperluas generalisasi model prediksi yang di hasilkan.

DAFTAR PUSTAKA

- [1]. J. Han, H. Kim, and S. Park, "Performance variation of semiconductor gas sensors under humidity and temperature changes,"

Sensors and Actuators B: Chemical, vol. 329, pp. 129–141, 2021.

- [2]. World Health Organization, *Ambient (Outdoor) Air Pollution*. Geneva, Switzerland: WHO Press, 2023.
- [3]. Environmental Protection Agency, *Indoor Air Quality Standards for Carbon Dioxide (CO₂)*. Washington, DC, USA: U.S. EPA, 2022.
- [4]. Hanwei Electronics Co., Ltd., *MQ-135 Gas Sensor Datasheet: Technical Data and CO₂ Sensitivity Curve*. China: Hanwei Electronics, 2016.
- [5]. J. Brownlee, *Machine Learning Algorithms From Scratch: With Python Examples*. Machine Learning Mastery, 2019.
- [6]. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
- [7]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [8]. S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2*. Birmingham, UK: Packt Publishing, 2020.
- [9]. Arduino Documentation, *ESP32 WiFi and AsyncWebServer Library Reference*. Arduino.cc, 2024.
- [10]. Espressif Systems, *ESP32 Technical Reference Manual*. Espressif Systems Ltd., 2023.
- [11]. Aosong Electronics Co., Ltd., *DHT22 Sensor Datasheet: AM2302 Temperature and Humidity Sensor Specifications*. 2020.