

SISTEM MONITORING KUALITAS UDARA BERBASIS IOT DAN MACHINE LEARNING UNTUK PREDIKSI POLUSI MENGGUNAKAN RANDOM FOREST

¹Raja Syahlul Idhamsyah Ruzfi, ²Yohanni Syahra
Program Studi Teknologi Informasi
Universitas Muhammadiyah Sumatera Utara
Raja56ruzfi@gmail.com

Submit : 07 Juni 2026 | **Diterima** : 19 Juni 2026 | **Terbit** : 19 Juli 2026

ABSTRAK

Kualitas udara merupakan faktor penting yang memengaruhi kesehatan manusia dan lingkungan. Peningkatan aktivitas industri dan transportasi menyebabkan meningkatnya konsentrasi polutan udara sehingga diperlukan sistem monitoring yang mampu memberikan informasi secara real-time dan akurat. Penelitian ini bertujuan untuk mengembangkan sistem monitoring kualitas udara berbasis Internet of Things (IoT) yang terintegrasi dengan algoritma machine learning Random Forest untuk melakukan prediksi kualitas udara. Sistem menggunakan mikrokontroler ESP32 dengan sensor DHT22, MQ-135, dan MQ-7 untuk memperoleh data suhu, kelembaban, serta konsentrasi gas pencemar. Data dikirim ke server dan ditampilkan dalam dashboard berbasis web. Dataset historis digunakan untuk membangun model prediksi menggunakan Random Forest. Hasil penelitian menunjukkan sistem mampu melakukan monitoring real-time dan menghasilkan prediksi dengan akurasi sebesar 95,91%. Sistem ini dapat menjadi solusi efektif dalam pemantauan dan prediksi kualitas udara.

Kata Kunci: IoT; kualitas udara; machine learning; monitoring; random forest.

PENDAHULUAN

Kualitas udara merupakan salah satu indikator penting dalam menentukan kesehatan lingkungan dan kualitas hidup manusia. Peningkatan aktivitas manusia, khususnya di wilayah perkotaan seperti transportasi, industri, dan pembakaran bahan bakar fosil, menyebabkan meningkatnya konsentrasi polutan udara. Paparan polusi udara dalam jangka panjang dapat menimbulkan berbagai gangguan kesehatan, seperti penyakit pernapasan, kardiovaskular, hingga kematian dini. Oleh karena itu, diperlukan sistem monitoring kualitas udara yang mampu memberikan informasi yang akurat, cepat, dan relevan terhadap kondisi lingkungan sekitar.

Sistem monitoring kualitas udara yang tersedia saat ini umumnya masih bersifat konvensional dan memiliki keterbatasan dalam penyajian data secara real-time serta cakupan wilayah yang terbatas. Hal ini menyebabkan data yang dihasilkan belum mampu merepresentasikan kondisi kualitas udara secara lokal secara optimal. Selain itu, keterbatasan jumlah stasiun pemantauan juga membuat informasi kualitas udara kurang merata dan tidak dapat diakses secara langsung oleh masyarakat (Putro et al., 2024). Kondisi tersebut menunjukkan bahwa sistem monitoring yang ada belum sepenuhnya efektif dalam mendukung pengambilan keputusan berbasis data lingkungan.

Perkembangan teknologi Internet of Things (IoT) memberikan solusi dalam mengatasi keterbatasan tersebut. IoT memungkinkan integrasi antara sensor lingkungan, mikrokontroler, dan jaringan internet untuk melakukan pengumpulan dan pengiriman data secara otomatis dan real-time. Dengan adanya IoT, sistem monitoring kualitas udara dapat dikembangkan menjadi lebih fleksibel, efisien, dan mudah diakses oleh pengguna (Saputra et al., 2023). Namun demikian, sebagian besar sistem monitoring berbasis IoT masih berfokus pada penyajian data secara deskriptif dan belum memiliki kemampuan untuk memprediksi kondisi kualitas udara di masa mendatang.

Perkembangan teknologi Internet of Things (IoT) memungkinkan sistem monitoring dilakukan

secara real-time dan lebih efisien dalam pengumpulan data lingkungan (Alam & Rahman, 2021). Selain itu, penerapan IoT dalam monitoring kualitas udara juga telah terbukti mampu meningkatkan akurasi dan kemudahan akses data bagi pengguna (Huang et al., 2021).

Di sisi lain, metode machine learning banyak digunakan dalam prediksi kualitas udara karena kemampuannya dalam mengolah data yang kompleks dan bersifat non-linear (Gao et al., 2022). Beberapa penelitian juga menunjukkan bahwa model berbasis machine learning dapat meningkatkan akurasi prediksi secara signifikan (Wang et al., 2023).

Kemampuan prediksi kualitas udara menjadi sangat penting sebagai bentuk peringatan dini terhadap potensi pencemaran lingkungan. Dalam hal ini, penerapan machine learning dapat digunakan untuk menganalisis pola data historis dan menghasilkan prediksi yang lebih akurat. Salah satu algoritma yang banyak digunakan dalam prediksi data lingkungan adalah Random Forest, yang dikenal mampu menangani data non-linear dan multivariat dengan tingkat akurasi yang tinggi (Li et al., 2024). Selain itu, Random Forest juga memiliki keunggulan dalam mengurangi overfitting melalui pendekatan ensemble learning, sehingga menghasilkan model yang lebih stabil dan memiliki kemampuan generalisasi yang baik (Nugroho & Wijaya, 2024).

Berdasarkan kajian penelitian terdahulu, sebagian besar penelitian masih memisahkan antara sistem monitoring berbasis IoT dan sistem prediksi berbasis machine learning. Penelitian monitoring umumnya hanya berfokus pada pengumpulan dan visualisasi data secara real-time, sedangkan penelitian prediksi kualitas udara sering menggunakan dataset sekunder yang tidak terintegrasi langsung dengan perangkat sensor. Hal ini menunjukkan adanya kesenjangan penelitian (*research gap*), yaitu belum adanya sistem yang mengintegrasikan monitoring real-time berbasis IoT dengan prediksi kualitas udara berbasis machine learning menggunakan data primer dari sensor secara langsung.

Keunikan penelitian ini terletak pada integrasi sistem monitoring kualitas udara berbasis IoT dengan algoritma Random Forest dalam satu sistem yang terpadu. Selain itu, penelitian ini menggunakan data primer yang diperoleh langsung dari sensor lingkungan secara real-time serta dilengkapi dengan mekanisme penyimpanan data lokal sebagai cadangan. Pendekatan ini diharapkan mampu meningkatkan keandalan sistem serta menghasilkan prediksi kualitas udara yang lebih akurat dan relevan terhadap kondisi lingkungan aktual.

Berdasarkan latar belakang tersebut, tujuan penelitian ini adalah untuk mengembangkan sistem monitoring kualitas udara berbasis Internet of Things (IoT) yang terintegrasi dengan algoritma Random Forest untuk melakukan prediksi kualitas udara. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem cerdas untuk pemantauan lingkungan, serta menjadi solusi alternatif dalam menyediakan informasi kualitas udara yang real-time, akurat, dan prediktif.

TINJAUAN PUSTAKA

Sistem Monitoring

Sistem monitoring merupakan sistem yang digunakan untuk mengamati dan mengumpulkan data dari suatu objek atau lingkungan secara berkelanjutan guna mendukung proses analisis dan pengambilan keputusan. Sistem ini mencakup proses pengumpulan, penyimpanan, dan penyajian data dalam bentuk yang mudah dipahami, seperti grafik atau dashboard. Dengan perkembangan teknologi, sistem monitoring telah berkembang menjadi sistem otomatis berbasis digital yang mampu bekerja secara real-time dan terintegrasi dengan jaringan internet, sehingga meningkatkan efisiensi dan akurasi dalam pemantauan kondisi lingkungan (Pebralia et al., 2024).

Kualitas Udara

Kualitas udara merupakan kondisi udara yang ditentukan oleh kandungan zat pencemar di dalamnya, seperti karbon monoksida (CO), nitrogen dioksida (NO₂), dan partikel debu. Di Indonesia, kualitas udara diukur menggunakan Indeks Standar Pencemaran Udara (ISPU) yang mengklasifikasikan kondisi udara ke dalam beberapa kategori, mulai dari baik hingga berbahaya. Kualitas udara yang buruk dapat berdampak negatif terhadap kesehatan manusia, seperti gangguan pernapasan dan penyakit kronis, sehingga pemantauan kualitas udara menjadi sangat penting (World Health Organization, 2022).

Internet of Things (IoT)

Internet of Things (IoT) merupakan konsep teknologi yang memungkinkan perangkat fisik dilengkapi dengan sensor dan konektivitas jaringan untuk saling berkomunikasi dan bertukar data melalui internet. Dalam sistem monitoring, IoT memungkinkan pengumpulan dan pengiriman data dilakukan secara otomatis dan real-time tanpa intervensi manusia. Penggunaan IoT dalam monitoring kualitas udara memberikan keunggulan berupa fleksibilitas, efisiensi, serta kemudahan akses data oleh pengguna melalui platform berbasis web atau aplikasi (Saputra et al., 2023).

Machine Learning

Machine learning adalah metode dalam kecerdasan buatan yang memungkinkan sistem untuk belajar dari data historis dan menghasilkan prediksi atau keputusan tanpa pemrograman eksplisit. Metode ini sangat efektif dalam menganalisis data yang kompleks dan bersifat non-linear. Dalam penelitian ini, digunakan pendekatan supervised learning, yaitu model dilatih menggunakan data yang telah diberi label untuk memprediksi kategori kualitas udara berdasarkan parameter lingkungan yang diperoleh dari sensor (Sarker, 2021).

Algoritma Random Forest

Random Forest merupakan algoritma machine learning berbasis ensemble learning yang bekerja dengan membangun banyak decision tree dan menggabungkan hasil prediksinya untuk meningkatkan akurasi. Algoritma ini memiliki keunggulan dalam menangani data multivariat dan non-linear serta mampu mengurangi overfitting melalui teknik bootstrap sampling dan pemilihan fitur secara acak. Oleh karena itu, Random Forest banyak digunakan dalam prediksi kualitas udara karena mampu menghasilkan model yang stabil dan memiliki tingkat akurasi tinggi (Nugroho et al., 2024)

METODE PENELITIAN

Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan metode eksperimen dengan pendekatan pengembangan sistem monitoring kualitas udara berbasis Internet of Things (IoT) yang terintegrasi dengan algoritma machine learning. Pendekatan ini dipilih karena IoT mampu melakukan pengumpulan data secara real-time, sedangkan algoritma Random Forest efektif dalam mengolah data yang bersifat non-linear dan multivariat dengan tingkat akurasi yang tinggi (Saputra et al., 2023).

Perancangan Sistem

Sistem yang dikembangkan terdiri dari perangkat keras dan perangkat lunak yang saling terintegrasi. Perangkat keras menggunakan mikrokontroler ESP32 yang terhubung dengan sensor DHT22 untuk mengukur suhu dan kelembaban, serta sensor MQ-135 dan MQ-7 untuk mendeteksi gas pencemar. Perangkat lunak meliputi sistem pengiriman data ke server dan dashboard berbasis web untuk menampilkan hasil monitoring. Data dari sensor dikirim secara real-time melalui jaringan internet dan disimpan dalam basis data, serta dilengkapi dengan modul MicroSD sebagai penyimpanan cadangan.

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh langsung dari sensor yang terpasang pada sistem. Proses pengumpulan data dilakukan secara kontinu dalam interval waktu tertentu sehingga menghasilkan data historis yang dapat digunakan untuk analisis. Parameter yang dikumpulkan meliputi suhu, kelembaban, serta konsentrasi gas pencemar yang menjadi variabel utama dalam menentukan kualitas udara.

Pengolahan dan Pemodelan Data

Data yang telah diperoleh kemudian melalui tahap preprocessing yang meliputi pembersihan data dari nilai yang tidak valid serta pelabelan berdasarkan kategori kualitas udara mengacu pada Indeks Standar Pencemaran Udara (ISPU). Dataset selanjutnya dibagi menjadi data pelatihan dan data pengujian untuk membangun model machine learning. Model dikembangkan menggunakan algoritma Random Forest dengan pendekatan supervised learning, di mana model dilatih menggunakan data

berlabel untuk menghasilkan prediksi kualitas udara. Parameter yang digunakan meliputi jumlah pohon ($n_estimators$), kedalaman maksimum pohon (max_depth), jumlah fitur ($max_features$), serta $random_state$ untuk menjaga konsistensi hasil (Nugroho et al., 2024).

Evaluasi Model

Evaluasi model dilakukan dengan menguji hasil prediksi menggunakan data pengujian dan membandingkannya dengan data aktual menggunakan confusion matrix. Confusion matrix digunakan untuk melihat jumlah prediksi yang benar dan salah yang dihasilkan oleh model dalam setiap kategori kualitas udara.

Berdasarkan confusion matrix, kinerja model diukur menggunakan nilai akurasi sebagai indikator utama. Akurasi menunjukkan tingkat ketepatan model dalam mengklasifikasikan data secara keseluruhan. Rumus akurasi dinyatakan sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan: TP (True Positive) adalah jumlah data yang diprediksi benar sebagai kelas positif, TN (True Negative) adalah jumlah data yang diprediksi benar sebagai kelas negatif, FP (False Positive) adalah jumlah data yang salah diprediksi sebagai positif, dan FN (False Negative) adalah jumlah data yang salah diprediksi sebagai negatif.

Nilai akurasi yang diperoleh digunakan untuk menilai tingkat keberhasilan model Random Forest dalam mengklasifikasikan kualitas udara. Semakin tinggi nilai akurasi, maka semakin baik kemampuan model dalam melakukan prediksi. Hasil evaluasi ini kemudian diintegrasikan ke dalam sistem monitoring sehingga pengguna dapat melihat hasil klasifikasi kualitas udara secara real-time.

HASIL DAN PEMBAHASAN

Hasil Implementasi Sistem

Hasil penelitian diawali dengan implementasi sistem monitoring kualitas udara berbasis Internet of Things (IoT). Sistem yang dikembangkan menggunakan mikrokontroler ESP32 yang terintegrasi dengan sensor DHT22, MQ-135, dan MQ-7 untuk mengukur parameter suhu, kelembaban, serta konsentrasi gas pencemar. Data yang diperoleh dari sensor berhasil dikirim secara real-time ke server dan disimpan dalam basis data. Selain itu, sistem juga mampu menyimpan data cadangan pada modul MicroSD sebagai antisipasi kegagalan jaringan. Secara keseluruhan, sistem yang dibangun mampu berjalan dengan stabil dan melakukan proses akuisisi data secara kontinu.

Hasil Monitoring Kualitas Udara

Data hasil monitoring ditampilkan melalui dashboard berbasis web dalam bentuk tabel dan grafik untuk memudahkan interpretasi pengguna. Visualisasi data menunjukkan bahwa sistem mampu menampilkan perubahan parameter lingkungan secara dinamis sesuai kondisi aktual.

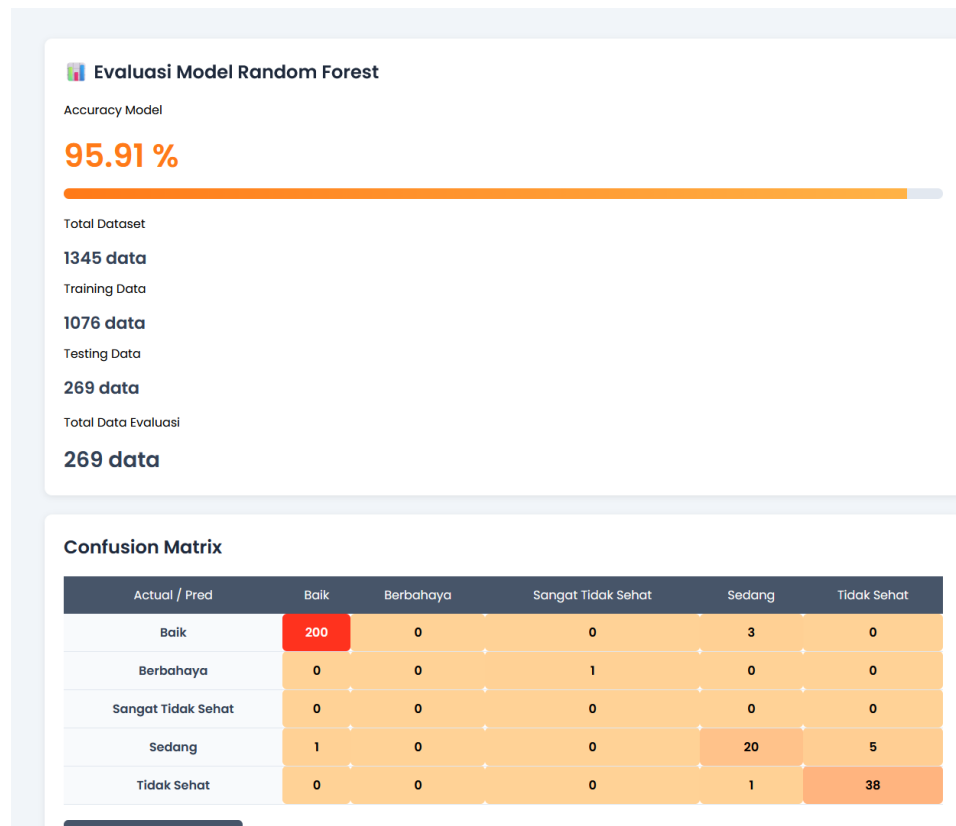
2026-03-16 18:43:19	26.7	87.1	185	184	Baik
2026-03-16 18:38:19	26.8	85.5	196	180	Baik
2026-03-16 18:33:19	31	75.6	507	450	Sedang
2026-03-16 18:28:19	31.6	79.9	498	456	Sedang

Gambar 1. Hasil Monitoring Kualitas Udara

Selain itu, grafik monitoring menunjukkan tren perubahan kualitas udara dalam periode waktu tertentu sehingga pengguna dapat memahami pola perubahan lingkungan dengan lebih mudah. Hal ini menunjukkan bahwa sistem tidak hanya mampu mengumpulkan data, tetapi juga menyajikan informasi secara informatif dan mudah dipahami.

Hasil Prediksi Model

Model machine learning yang dibangun menggunakan algoritma Random Forest diuji menggunakan data pengujian yang telah dipisahkan dari data pelatihan. Hasil evaluasi model menggunakan confusion matrix menunjukkan bahwa model mampu mengklasifikasikan kualitas udara dengan baik.



Gambar 2. Hasil Evaluasi Model Random Forest

Berdasarkan hasil pengujian, diperoleh nilai akurasi sebesar 95,91%, yang menunjukkan bahwa model memiliki tingkat ketepatan yang tinggi dalam mengklasifikasikan kualitas udara. Nilai ini menunjukkan bahwa sebagian besar prediksi yang dihasilkan oleh model sesuai dengan kondisi aktual.

Pembahasan

Hasil penelitian menunjukkan bahwa sistem yang dikembangkan mampu mengintegrasikan proses monitoring dan prediksi kualitas udara dalam satu platform. Implementasi IoT memungkinkan pengumpulan data secara real-time, sedangkan algoritma Random Forest mampu mengolah data tersebut menjadi informasi prediksi yang akurat. Tingginya nilai akurasi menunjukkan bahwa model mampu menangkap pola hubungan antara parameter lingkungan dengan kategori kualitas udara.

Jika dibandingkan dengan penelitian sebelumnya, hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Saputra et al. (2023) yang menyatakan bahwa IoT efektif digunakan dalam monitoring lingkungan secara real-time. Selain itu, penelitian oleh Li et al. (2024) juga menunjukkan bahwa Random Forest memiliki performa tinggi dalam klasifikasi kualitas udara. Namun, keunggulan penelitian ini terletak pada integrasi antara sistem monitoring dan prediksi dalam satu sistem yang terpadu, sementara sebagian besar penelitian sebelumnya masih memisahkan kedua pendekatan tersebut.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan yang dapat menjadi ancaman terhadap validitas hasil. Pertama, jumlah sensor yang digunakan masih terbatas sehingga

parameter yang diamati belum mencakup seluruh jenis polutan udara. Kedua, data yang digunakan masih berasal dari satu lokasi sehingga belum merepresentasikan kondisi lingkungan yang lebih luas. Ketiga, kemungkinan adanya noise pada data sensor juga dapat memengaruhi akurasi model. Oleh karena itu, penelitian selanjutnya disarankan untuk menambahkan jenis sensor, memperluas cakupan lokasi pengambilan data, serta menerapkan teknik preprocessing yang lebih kompleks untuk meningkatkan kualitas data dan performa model.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem monitoring kualitas udara berbasis Internet of Things (IoT) yang dikembangkan mampu melakukan pengambilan dan pengiriman data secara real-time dengan baik serta menampilkan informasi kondisi lingkungan secara akurat melalui dashboard berbasis web. Integrasi algoritma Random Forest dalam sistem ini berhasil menjawab tujuan penelitian, yaitu menghasilkan model prediksi kualitas udara yang akurat dengan nilai akurasi sebesar 95,91%. Hasil tersebut menunjukkan bahwa model mampu mengenali pola data lingkungan dan mengklasifikasikan kualitas udara sesuai kondisi aktual. Kontribusi utama penelitian ini terletak pada integrasi sistem monitoring dan prediksi dalam satu platform yang terpadu, sehingga tidak hanya menyediakan informasi kondisi kualitas udara saat ini, tetapi juga memberikan prediksi sebagai bentuk peringatan dini. Dengan demikian, penelitian ini memberikan solusi yang lebih efektif dalam pemantauan dan prediksi kualitas udara berbasis teknologi IoT dan machine learning.

REFERENSI

- Alam, M. S., & Rahman, M. M. (2021). A review of smart air quality monitoring systems using Internet of Things. *Environmental Monitoring and Assessment*, 193(4), 1–15.
- Gao, J., Sun, Y., & Zhao, H. (2022). Machine learning-based air quality prediction: A comparative study. *Atmospheric Environment*, 273, 118956.
- Huang, C., Kuo, P., & Lin, Y. (2021). IoT-based air pollution monitoring and forecasting system. *Sensors*, 21(2), 456.
- Li, X., Zhang, Y., & Chen, L. (2024). Random forest-based air quality prediction model for urban environments. *Environmental Science Journal*, 18(2), 200–210.
- Nugroho, A., Sari, D., & Pratama, Y. (2024). Implementasi Metode Random Forest pada Kategori Konten Media Sosial. *Jurnal Jendela Edukasi*, 5(1), 78–89.
<https://ejournal.jendelaedukasi.id/index.php/JJM/article/view/633>
- Nugroho, A., & Wijaya, D. (2024). Ensemble learning for environmental data analysis. *Journal of Data Science*, 9(1), 33–41.
- Pebralia, R., Saputra, D., & Firmansyah, M. (2024). Monitoring systems in environmental applications. *Journal of Engineering Research*, 15(1), 78–85.
- Putro, H., Santoso, B., & Rahman, A. (2024). IoT-based air quality monitoring system. *International Journal of Technology*, 15(1), 85–94.
- Saputra, D., Prasetyo, A., & Nugraha, F. (2023). Implementation of Internet of Things in environmental monitoring. *Journal of Smart Systems*, 7(1), 12–20.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*. <https://doi.org/10.1007/s42979-021-00592-y>
- Wang, Y., Li, Z., & Guo, X. (2023). Air quality prediction based on machine learning models: A review. *Atmosphere*, 14(5), 789.
- World Health Organization. (2022). *WHO Global Air Quality Guidelines: Particulate Matter (PM_{2.5} and PM₁₀), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. World Health Organization.