

INTEGRASI IOT DAN ALGORITMA CATBOOST UNTUK DETEKSI KUALITAS UDARA SECARA REAL-TIME DI WILAYAH KOTA PADANG

Muhammad Habib Yuhandri¹, Hasri Awal²

¹Fakultas Ilmu Komputer, Sistem Informasi, Universitas Putra Indonesia, Padang, Indonesia

²Fakultas Ilmu Komputer, Sistem Komputer, Universitas Putra Indonesia, Padang, Indonesia

Email: ¹mhabibyuhandri@upiyptk.ac.id, ²hasriawal@upiyptk.ac.id

Abstrak

Penelitian ini mengembangkan sistem monitoring kualitas udara berbasis *Internet of Things (IoT)* dan algoritma machine learning *CatBoost* untuk mendeteksi tingkat pencemaran udara di Kota Padang secara *real-time*. Sistem ini menggunakan sensor Nova PM (PM2.5 dan PM10) dan MQ-7 (karbon monoksida/CO) yang dihubungkan ke mikrokontroler Wemos D1, lalu mengirimkan data secara otomatis ke *Google Sheets* sebagai basis penyimpanan. Data yang terkumpul diproses melalui tahapan preprocessing menggunakan *StandardScaler*, kemudian dilatih menggunakan model *CatBoost*. Hasil evaluasi menunjukkan bahwa model *CatBoost* mampu mengklasifikasikan lima kategori kualitas udara (Baik, Sedang, Tidak Sehat, Sangat Tidak Sehat, dan Berbahaya) dengan akurasi mencapai 95%, serta nilai *precision* dan *f1-score* rata-rata di atas 0.90. Sistem ini juga diimplementasikan dalam bentuk antarmuka pengguna (*GUI*) berbasis *Streamlit* yang menampilkan data sensor terkini dan hasil prediksi secara visual. Hasil penelitian ini menunjukkan bahwa integrasi *IoT* dan *machine learning* dapat menjadi solusi yang efektif dalam mendeteksi dan memantau kualitas udara secara real-time di kawasan perkotaan.

Kata Kunci: Kualitas udara, *Internet of Things*, *CatBoost*, *Machine Learning*, *Streamlit*

Abstract

This study developed an Internet of Things (IoT)-based air quality monitoring system using the CatBoost machine learning algorithm to detect air pollution levels in Padang City in real time. The system utilizes Nova PM sensors (PM2.5 and PM10) and MQ-7 (carbon monoxide/CO) connected to a Wemos D1 microcontroller, which transmits data automatically to Google Sheets as a data storage platform. The collected data is preprocessed using StandardScaler and trained with the CatBoost model. Evaluation results show that the CatBoost algorithm can classify five air quality categories (Good, Moderate, Unhealthy, Very Unhealthy, and Hazardous) with an accuracy of 95%, and an average precision and f1-score above 0.90. The system is also implemented in a user-friendly Streamlit-based graphical interface that displays real-time sensor data and prediction results visually. The findings demonstrate that integrating IoT and machine learning offers an effective solution for real-time air quality detection and monitoring in urban environments.

Keywords: Air quality, Internet of Things, CatBoost, Machine Learning, Streamlit

1. PENDAHULUAN

Udara merupakan campuran dari berbagai jenis gas seperti oksigen (O₂), karbon dioksida (CO₂), dan ozon (O₃) yang berperan penting dalam menjaga keseimbangan kehidupan di bumi[1]. Oksigen dibutuhkan untuk pernapasan, CO₂ dimanfaatkan oleh tumbuhan dalam proses fotosintesis, dan ozon melindungi bumi dari radiasi ultraviolet[2]. Komposisi udara tidak bersifat tetap, tetapi sangat dipengaruhi oleh kondisi suhu, tekanan, dan keberadaan zat pencemar dari

lingkungan[3]. Di tengah pesatnya pembangunan kota dan pertumbuhan populasi, kualitas udara di daerah perkotaan kian menurun akibat peningkatan emisi dari kendaraan bermotor, kegiatan industri, pembakaran sampah, dan penurunan ruang terbuka hijau.

Kondisi ini dapat berdampak buruk terhadap kesehatan masyarakat. Polutan seperti partikulat PM_{2.5} dan PM₁₀ diketahui dapat menembus saluran pernapasan dan menyebabkan gangguan sistem respirasi dan kardiovaskular. Karbon monoksida (CO), yang dihasilkan dari proses pembakaran tidak sempurna, bersifat toksik dan berbahaya jika

terhirup dalam konsentrasi tinggi. Di Kota Padang, menunjukkan bahwa rata-rata konsentrasi PM₁₀ di beberapa kawasan masih berada dalam ambang batas baku mutu udara ambien menurut PP No. 41 Tahun 1999, namun tetap menunjukkan tren peningkatan yang mengkhawatirkan[4]. Hal ini mengindikasikan perlunya sistem pemantauan kualitas udara yang bersifat real-time dan dapat menjangkau masyarakat secara luas.

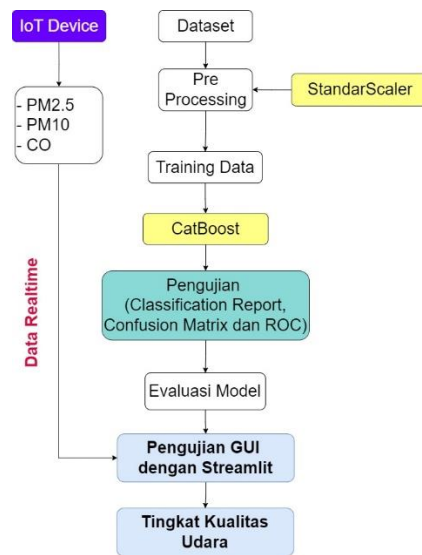
Sayangnya, sistem pemantauan udara resmi seperti ISPU hanya terpasang di titik-titik tertentu dan tidak mampu memberikan informasi langsung kepada masyarakat yang tinggal jauh dari lokasi alat. Untuk itu, diperlukan sistem monitoring kualitas udara berbasis teknologi Internet of Things (IoT) yang mampu membaca, mengirim, dan menampilkan data secara otomatis tanpa interaksi manual. Sensor seperti Nova PM digunakan untuk mengukur konsentrasi PM_{2.5} dan PM₁₀, sementara sensor MQ-7 mampu mendeteksi kadar karbon monoksida. Data yang dikirim melalui jaringan nirkabel ke server dapat diolah dan divisualisasikan melalui antarmuka web agar mudah diakses oleh masyarakat umum.

Seiring dengan perkembangan teknologi, pemanfaatan algoritma machine learning semakin banyak digunakan untuk meningkatkan akurasi prediksi kualitas udara. Model CatBoost, salah satu algoritma berbasis gradient boosting yang dikembangkan oleh Yandex, telah terbukti efektif dalam menangani data tabular yang kompleks dan berisi variabel kategorik[5]–[7]. Beberapa penelitian menunjukkan bahwa CatBoost memiliki kinerja yang kompetitif dalam memprediksi tingkat polusi udara, bahkan melebihi akurasi model seperti Random Forest dan XGBoost[8]–[10]. Penelitian lain menunjukkan bahwa kombinasi sensor CO dan suhu dengan model CatBoost mampu menghasilkan prediksi indeks kualitas udara dengan akurasi lebih dari 93%[7].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan sistem monitoring kualitas udara berbasis IoT yang dilengkapi dengan model machine learning CatBoost untuk klasifikasi tingkat kualitas udara di Kota Padang. Sistem ini dirancang untuk bekerja secara real-time, menggunakan sensor Nova PM dan MQ-7, serta menyimpan data ke Google Sheets sebagai basis penyimpanan cloud. Prediksi hasil klasifikasi CatBoost kemudian divisualisasikan melalui dashboard Streamlit yang dapat diakses oleh pengguna umum. Dengan demikian, diharapkan sistem ini mampu memberikan informasi yang cepat, akurat, dan dapat membantu masyarakat serta pemangku kebijakan dalam mengantisipasi risiko kesehatan dan kebakaran akibat polusi udara yang meningkat.

2. METODOLOGI PENELITIAN

Sistem yang dikembangkan dalam penelitian ini menggabungkan perangkat IoT untuk pengambilan data kualitas udara secara real-time dan model machine learning CatBoost untuk klasifikasi tingkat kualitas udara. Seluruh alur mulai dari akuisisi data hingga prediksi dan visualisasi disusun secara sistematis dalam tahapan berikut. Adapun alur lengkap metode penelitian ini dapat dilihat pada gambar berikut:



Gambar 1. Alur Metode

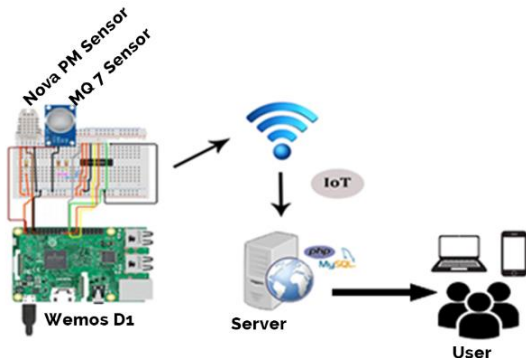
Alur metode dimulai dari perangkat IoT berbasis sensor Nova PM (untuk PM_{2.5} dan PM₁₀) dan MQ-7 (untuk CO) yang terpasang pada mikrokontroler Wemos D1. Sensor ini mengirimkan data udara secara *real-time* ke server melalui koneksi nirkabel. Data kemudian disimpan dalam bentuk dataset dan dilakukan preprocessing. Pada tahap preprocessing, dilakukan standarisasi fitur menggunakan StandardScaler untuk memastikan semua variabel numerik berada dalam skala yang seragam.

Setelah preprocessing, data dibagi menjadi data latih dan data uji. Model CatBoost digunakan untuk proses pelatihan klasifikasi tingkat kualitas udara. Model ini dipilih karena kemampuannya dalam menangani fitur numerik dan kategorikal, serta performanya yang unggul dalam menangani data tabular. Setelah pelatihan, model diuji menggunakan metrik evaluasi seperti *classification report*, *confusion matrix*, dan ROC curve untuk mengukur performa klasifikasi[11]–[15].

Tahap selanjutnya adalah implementasi model ke dalam aplikasi antarmuka menggunakan Streamlit. Aplikasi ini menampilkan data sensor terkini, hasil klasifikasi tingkat kualitas udara, serta memungkinkan pengguna melakukan prediksi

manual. Hasil klasifikasi akhir yang dihasilkan oleh sistem ditampilkan dalam kategori tingkat kualitas udara seperti baik, sedang, tidak sehat, sangat tidak sehat, atau berbahaya.

Sistem monitoring kualitas udara ini dirancang menggunakan pendekatan *Internet of Things (IoT)* yang memungkinkan pengambilan dan pengiriman data secara real-time ke server pusat. Adapun blok diagram dari rancangan sistem IoT yang digunakan dapat dilihat pada gambar berikut.



Gambar 2. Blok Diagram Perancangan IoT

Pada gambar ditunjukkan bahwa sistem terdiri dari beberapa komponen utama. Sensor Nova PM digunakan untuk mendeteksi partikel udara PM2.5 dan PM10, sedangkan sensor MQ-7 digunakan untuk mendeteksi gas karbon monoksida (CO). Kedua sensor ini terhubung dengan mikrokontroler Wemos D1 yang berfungsi sebagai unit pemrosesan data awal dan pengirim data.

Selanjutnya, data hasil pembacaan sensor dikirim secara real-time melalui jaringan Wi-Fi ke layanan Google Sheets yang berfungsi sebagai media penyimpanan berbasis cloud. Data yang tersimpan di Google Sheets kemudian dapat diakses dan divisualisasikan oleh pengguna melalui perangkat komputer maupun smartphone. Dengan pendekatan ini, pengguna dapat memantau kondisi kualitas udara kapan saja dan di mana saja secara praktis dan efisien.

HASIL DAN PEMBAHASAN

3.1. Dataset dan Preprocessing

Dataset yang digunakan dalam penelitian ini terdiri dari 1.100 data kualitas udara yang dikumpulkan dari berbagai wilayah di Indonesia melalui situs resmi pemantauan Indeks Standar Pencemaran Udara (ISPU). Data mencakup konsentrasi tiga parameter utama, yaitu PM2.5, PM10, dan CO, serta label kualitas udara berdasarkan kategori ISPU. Adapun cuplikan data yang digunakan dapat dilihat pada Tabel berikut.

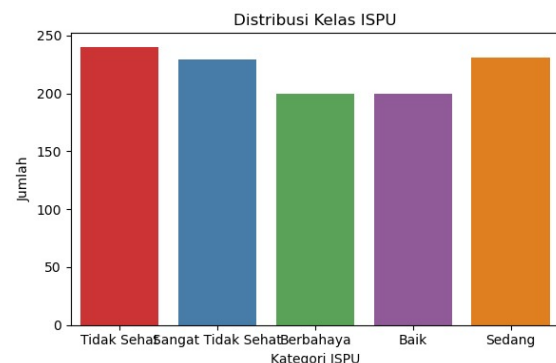
Tabel 1. Dataset Kualitas Udara

PM2.5	PM10	CO	ISPU_Label
43	128	12904	Tidak Sehat
144	243	18524	Sangat Tidak Sehat
63	294	16755	Sangat Tidak Sehat
113	254	17220	Sangat Tidak Sehat
41	129	12987	Tidak Sehat
100	253	17377	Sangat Tidak Sehat
108	205	18094	Sangat Tidak Sehat
39	188	12561	Tidak Sehat
238	397	33578	Berbahaya

Tabel 1 menunjukkan beragam variasi tingkat konsentrasi polutan yang diklasifikasikan ke dalam beberapa label ISPU seperti "Baik", "Tidak Sehat", "Sangat Tidak Sehat", dan "Berbahaya". Nilai PM2.5 dan PM10 mencerminkan partikel udara yang berukuran kecil dan besar, sedangkan nilai CO menunjukkan kadar gas karbon monoksida di udara. Kombinasi ketiga parameter ini digunakan sebagai input untuk menentukan klasifikasi tingkat kualitas udara.

Sebelum dilakukan pelatihan model, tahap preprocessing dilakukan untuk memastikan kualitas data dalam kondisi optimal. Beberapa langkah yang diterapkan antara lain adalah penanganan nilai kosong (missing value), konversi label ISPU menjadi representasi numerik menggunakan LabelEncoder, serta normalisasi data numerik menggunakan StandardScaler. Proses ini penting untuk menyamakan skala antar fitur dan mempercepat konvergensi model CatBoost selama proses pelatihan.

Distribusi data untuk masing-masing kategori ISPU perlu dianalisis guna mengetahui sebaran kelas target sebelum dilakukan pelatihan model. Adapun visualisasi dari distribusi kelas tersebut dapat dilihat pada gambar berikut.



Gambar 3. Distribusi Kelas ISPU pada Dataset

Gambar 3 menunjukkan distribusi jumlah data berdasarkan kategori ISPU, yaitu Baik,

Sedang, Tidak Sehat, Sangat Tidak Sehat, dan Berbahaya. Kategori "Tidak Sehat" memiliki jumlah data tertinggi, disusul oleh "Sedang" dan "Sangat Tidak Sehat". Sementara itu, kategori "Baik" dan "Berbahaya" memiliki jumlah data paling sedikit, namun masih cukup proporsional. Distribusi yang relatif seimbang ini menunjukkan bahwa data tidak terlalu bias terhadap satu kelas tertentu, sehingga model klasifikasi yang dibangun memiliki peluang yang baik untuk belajar secara optimal dari seluruh kelas yang ada.

3.2. Training Model

Untuk mengevaluasi performa model CatBoost dalam mengklasifikasikan tingkat kualitas udara berdasarkan parameter PM2.5, PM10, dan CO, dilakukan pengujian menggunakan metrik klasifikasi. Adapun hasil evaluasi klasifikasi dapat dilihat pada gambar berikut.

```

=== Classification Report ===

```

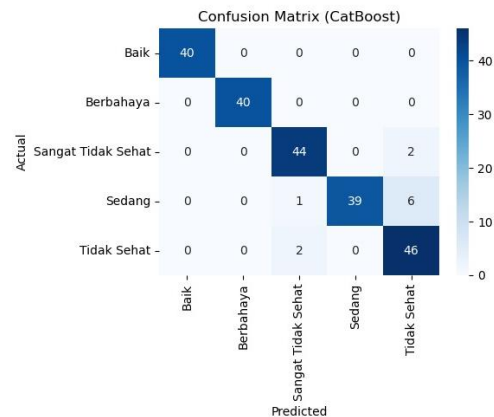
	precision	recall	f1-score	support
Baik	1.00	1.00	1.00	40
Berbahaya	1.00	1.00	1.00	40
Sangat Tidak Sehat	0.94	0.96	0.95	46
Sedang	1.00	0.85	0.92	46
Tidak Sehat	0.85	0.96	0.90	48
accuracy			0.95	220
macro avg	0.96	0.95	0.95	220
weighted avg	0.95	0.95	0.95	220

Gambar 4. Classification Report Model CatBoost

Gambar 4 menunjukkan performa model CatBoost dalam mengklasifikasikan lima kelas kategori ISPU. Model ini mampu menghasilkan akurasi keseluruhan sebesar 95%, dengan nilai precision, recall, dan f1-score yang cukup tinggi dan stabil di seluruh kelas. Kategori "Baik" dan "Berbahaya" menunjukkan performa sempurna dengan precision, recall, dan f1-score masing-masing sebesar 1.00, menandakan tidak ada kesalahan klasifikasi untuk kelas tersebut. Kategori "Sangat Tidak Sehat" memiliki f1-score sebesar 0.95, menunjukkan kemampuan deteksi yang kuat. Meskipun kategori "Sedang" memiliki recall lebih rendah (0.85), precision-nya tetap tinggi (1.00), yang berarti model cenderung berhati-hati dalam mengklasifikasikan ke kategori ini. Sementara itu, "Tidak Sehat" menunjukkan recall tinggi (0.96) namun precision lebih rendah (0.85), mengindikasikan adanya prediksi salah terhadap kelas ini. Secara keseluruhan, performa model sangat baik dan layak digunakan dalam sistem prediksi kualitas udara berbasis IoT secara real-time.

Untuk mendukung analisis evaluasi performa model CatBoost dalam klasifikasi tingkat kualitas udara, digunakan confusion matrix untuk

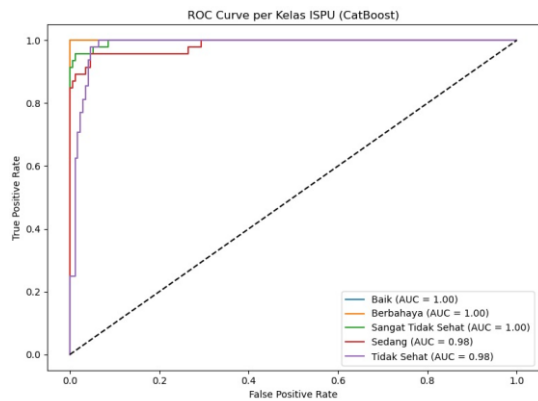
melihat akurasi prediksi per kelas secara lebih rinci. Adapun hasil confusion matrix dapat dilihat pada gambar berikut.



Gambar 5. Confusion Matrix Model CatBoost

Gambar 5 memperlihatkan bahwa model CatBoost memiliki tingkat akurasi klasifikasi yang sangat baik. Semua data untuk kategori "Baik" dan "Berbahaya" berhasil diklasifikasikan dengan benar sebanyak 40 data masing-masing, tanpa kesalahan prediksi. Untuk kategori "Sangat Tidak Sehat", terdapat 2 data yang salah diklasifikasikan sebagai "Tidak Sehat", sedangkan 44 lainnya berhasil diklasifikasikan dengan benar. Kelas "Sedang" menunjukkan kelemahan model, dengan 6 data diklasifikasikan sebagai "Tidak Sehat" dan 1 sebagai "Sangat Tidak Sehat", hanya 39 data yang diprediksi benar. Sementara itu, kelas "Tidak Sehat" memiliki 2 kesalahan prediksi, namun tetap mendominasi dengan 46 data yang diklasifikasikan dengan benar. Confusion matrix ini menegaskan bahwa meskipun terdapat sedikit misclassifikasi antar kelas yang berdekatan secara kategori, model secara umum telah menunjukkan performa klasifikasi yang sangat tinggi dan konsisten.

Untuk mengevaluasi performa model CatBoost dalam hal kemampuan klasifikasi multikelas pada prediksi kualitas udara, digunakan kurva ROC (Receiver Operating Characteristic) untuk setiap kelas. Adapun tampilan kurva ROC per kelas ISPU dapat dilihat pada gambar berikut.



Gambar 6. ROC Curve

3.3. Akuisisi Data IoT

Untuk mendapatkan data kualitas udara secara real-time, digunakan perangkat berbasis Internet of Things (IoT) yang mampu mengukur nilai PM2.5, PM10, dan CO dari udara sekitar. Adapun bentuk fisik dari perangkat IoT yang digunakan dapat dilihat pada gambar berikut.



Gambar 7. Perangkat IoT untuk Akuisisi Data Kualitas Udara

Gambar 7 menunjukkan perangkat IoT yang dirancang dengan modul Wemos D1 sebagai mikrokontroler utama. Perangkat ini dilengkapi dengan sensor Nova PM untuk membaca konsentrasi PM2.5 dan PM10, serta sensor MQ-7 untuk mendeteksi kadar gas karbon monoksida (CO) di udara. Hasil pembacaan sensor ditampilkan secara lokal melalui LCD. Selain itu, data yang ditangkap oleh perangkat ini dikirim secara otomatis ke cloud dalam bentuk Google Sheets melalui koneksi Wi-Fi, sehingga dapat diakses dan dipantau secara real-time oleh pengguna. Mekanisme ini memungkinkan integrasi langsung antara data lapangan dan sistem machine learning untuk analisis dan klasifikasi kualitas udara.

Untuk menyimpan dan memantau data sensor secara daring, sistem ini mengirimkan hasil akuisisi

dari perangkat IoT ke Google Sheets secara otomatis. Adapun tampilan data yang telah terekam dapat dilihat pada gambar berikut.

	A	B	C	D
1	Timestamp	PM2.5	PM10	CO
2	6/10/2025 6:54:34	43	128	12904
3	6/10/2025 6:54:56	144	243	18524
4	6/10/2025 6:58:47	63	294	16755
5	6/10/2025 7:09:42	113	254	17220
6	6/10/2025 7:14:27	41	129	12987
7	6/10/2025 7:29:30	100	253	17377
8	6/10/2025 7:34:27	300	421	20159
9	6/10/2025 7:35:13	11	9	4849
10	6/10/2025 8:10:51	110	231	16386
11	6/10/2025 8:11:34	207	335	31022

Gambar 8. Tampilan Data Sensor Kualitas Udara di Google Sheets

Gambar 8 memperlihatkan data hasil pembacaan sensor yang telah tersimpan di Google Sheets. Terdapat tiga parameter utama yang direkam, yaitu konsentrasi PM2.5, PM10, dan CO, yang masing-masing tertera dalam kolom B, C, dan D. Sementara itu, kolom A mencatat timestamp atau waktu pengambilan data secara real-time. Data ini diperoleh dari pengukuran langsung perangkat IoT dan dikirimkan secara otomatis menggunakan koneksi Wi-Fi. Penggunaan Google Sheets sebagai media penyimpanan berbasis cloud memungkinkan integrasi langsung dengan aplikasi monitoring serta model machine learning, memudahkan proses klasifikasi dan visualisasi kualitas udara secara dinamis.

3.4. Implementasi GUI

Untuk menampilkan hasil klasifikasi kualitas udara secara real-time, sistem ini dilengkapi dengan antarmuka pengguna (Graphical User Interface/GUI) berbasis web menggunakan Streamlit. Adapun tampilan halaman utama dari GUI sistem dapat dilihat pada gambar berikut.



Gambar 9. Tampilan Antarmuka GUI Sistem Monitoring Kualitas Udara

Gambar 9 memperlihatkan implementasi GUI yang menampilkan hasil monitoring kualitas udara secara real-time di Kota Padang. Pada bagian atas, terdapat judul aplikasi dan deskripsi singkat mengenai fungsionalitas sistem. Data sensor terbaru ditampilkan secara langsung, mencakup nilai PM2.5, PM10, dan CO, serta hasil klasifikasi tingkat kualitas udara oleh model CatBoost. Informasi kualitas udara juga diberikan dalam bentuk notifikasi prediktif yang mudah dipahami pengguna, seperti "Berbahaya" atau "Tidak Sehat", berdasarkan parameter yang terbaca. Selain itu, riwayat data sensor dan hasil klasifikasinya disajikan dalam bentuk tabel untuk memberikan transparansi dan jejak rekam terhadap kondisi udara yang telah terdeteksi. Pendekatan GUI ini mendukung transparansi data lingkungan dan mempermudah proses pengambilan keputusan oleh masyarakat dan instansi terkait.

Untuk mengakomodasi kebutuhan pengguna dalam melakukan simulasi prediksi kualitas udara secara manual, sistem menyediakan fitur input nilai PM2.5, PM10, dan CO secara langsung. Adapun tampilan antarmuka dari fitur uji prediksi manual dapat dilihat pada gambar berikut.



Gambar 10. Antarmuka Uji Prediksi Manual pada Sistem Monitoring Kualitas Udara

Gambar 10 menunjukkan fitur "Uji Prediksi Manual" yang memungkinkan pengguna memasukkan nilai parameter kualitas udara (PM2.5, PM10, dan CO) secara langsung untuk memperoleh hasil prediksi tingkat kualitas udara menggunakan model CatBoost. Setelah pengguna mengisi ketiga parameter tersebut dan menekan tombol Prediksi Manual, sistem akan menampilkan hasil klasifikasi secara instan. Contoh pada gambar memperlihatkan bahwa kombinasi nilai PM2.5 sebesar 100, PM10 sebesar 150, dan CO sebesar 12000 menghasilkan prediksi kualitas udara dengan kategori "Tidak Sehat". Fitur ini sangat bermanfaat untuk melakukan eksperimen atau validasi data dari sensor secara fleksibel tanpa menunggu input sensor real-time.

3. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem monitoring kualitas udara berbasis Internet of Things (IoT) dan model Machine Learning CatBoost yang dikembangkan dalam studi ini mampu memberikan solusi yang efektif dan akurat dalam memantau kondisi udara di Kota Padang secara real-time. Sistem ini mengandalkan sensor PM2.5, PM10, dan CO yang terhubung dengan mikrokontroler Wemos D1 untuk mendeteksi parameter pencemar udara, kemudian mengirimkan data secara otomatis ke Google Sheets. Data yang telah dikumpulkan diproses dan diprediksi menggunakan algoritma CatBoost, yang menunjukkan performa sangat baik dengan akurasi sebesar 95% dan nilai f1-score rata-rata sebesar 0.95. Hasil klasifikasi terhadap lima kategori kualitas udara Baik, Sedang, Tidak Sehat, Sangat Tidak Sehat, dan Berbahaya menunjukkan bahwa sistem ini dapat digunakan untuk deteksi dan klasifikasi secara akurat.

Selain itu, sistem ini juga telah diimplementasikan ke dalam antarmuka pengguna (GUI) berbasis Streamlit yang ramah pengguna. Antarmuka ini menampilkan data sensor terkini, hasil prediksi kualitas udara, riwayat data, serta menyediakan fitur prediksi manual berdasarkan input nilai PM2.5, PM10, dan CO. Implementasi GUI ini memungkinkan pengguna dari berbagai latar belakang untuk memantau dan memahami kualitas udara di sekitarnya secara mudah dan cepat. Sebagai arah penelitian lanjutan, sistem ini masih dapat dikembangkan lebih lanjut, seperti dengan menambahkan sensor untuk parameter lain (misalnya SO₂, NO₂, dan O₃).

4. REFERENCES

- [1] M. D. Safira, A. Syafiuddin, A. H. Z. Fasya, and B. Setianto, "Literature Review: Kualitas Udara Di Kawasan Industri Di Berbagai Lokasi Di Indonesia," *J. Public Heal.*, vol. 9, no. 2, pp. 38–47, 2022.
- [2] F. Tahir, W. Ridwan, and I. Z. Nasibu, "Monitor Kualitas Udara Berbasis Web Menggunakan Raspberry Pi dan Modul Wemos D1," *J. Tek.*, vol. 18, no. 1, pp. 35–44, 2020, doi: 10.37031/jt.v18i1.57.
- [3] M. Andini, Arif Budianto, Laili Mardiana, and Susi Rahayu, "Pengukuran Konsentrasi Emisi Partikulat di Ruang Tertutup Menggunakan Kit Pemantauan Kualitas Udara," *Lamda J. Ilm. Pendidik. MIPA dan Apl.*, vol. 5, no. 1 SE-Articles, pp. 133–139, Apr. 2025, doi: 10.58218/lambda.v5i1.1227.
- [4] R. S. Putra, H. Mariati, and Osonita, "Perbandingan Kualitas Udara Sebelum Dan Selama Psbb Covid-19 Di Kota Padang Berbasis

- Penginderaan Jauh,” *J. Azimut*, vol. 5, no. 2, pp. 90–103, 2023, [Online]. Available: <https://ojs.units-pdg.ac.id/index.php/azimut>
- [5] D. Daimari, S. Mondal, B. Brahma, and A. Nag, “Favorite Book Prediction System Using Machine Learning Algorithms,” *J. Appl. Eng. Technol. Sci.*, vol. 4, no. 2, pp. 983–991, 2023, doi: 10.37385/jaets.v4i2.1925.
- [6] I. D. Mienye and Y. Sun, “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” *IEEE Access*, vol. 10, no. September, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
- [7] P. Pothupogu, Y. Santhakumar, A. Yenmanagandla, and A. Sriram, “Health Monitoring System Based on LightGBM, CatBoost and Explainable Boost Classifier,” in *2024 2nd World Conference on Communication & Computing (WCONF)*, 2024, pp. 1–7. doi: 10.1109/WCONF61366.2024.10692251.
- [8] R. Irfannandhy, L. B. Handoko, and N. Ariyanto, “Analisis Performa Model Random Forest dan CatBoost dengan Teknik SMOTE dalam Prediksi Risiko Diabetes,” *Edumatic J. Pendidik. Inform.*, vol. 8, no. 2, pp. 714–723, 2024, doi: 10.29408/edumatic.v8i2.27990.
- [9] M. T. Syamkalla, S. Khomsah, and Y. S. R. Nur, “Implementasi Algoritma Catboost Dan Shapley Additive Explanations (SHAP) Dalam Memprediksi Popularitas Game Indie Pada Platform Steam,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 777–786, 2024, doi: 10.25126/jtiik.1148503.
- [10] O. Pahlevi, D. A. N. Wulandari, R. L. Kanti, H. Leidiyana, and Y. Handrianto, “Model Klasifikasi Risiko Stunting Pada Balita Menggunakan Algoritma CatBoost Classifier,” *Bull. Comput. Sci. Res.*, vol. 4, no. 6, pp. 414–421, 2024, doi: 10.47065/bulletincsr.v4i6.373.
- [11] A. Lubis, Y. Irawan, Junadhi, and S. Defit, “Leveraging K-Nearest Neighbors with SMOTE and Boosting Techniques for Data Imbalance and Accuracy Improvement,” *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 1625–1638, 2024, doi: 10.47738/jads.v5i4.343.
- [12] C. L. Fantasy, F. L. M. Simanjuntak, R. L. A. Purba, Andrean, and O. Sihombing, “Analisis Komparasi Algoritma C5.0 Dan Naive Bayes Penentuan Penerima Beasiswa Universitas Prima Indonesia,” *J. TEKINKOM*, vol. 6, no. 2, pp. 508–517, 2023, doi: 10.37600/tekinkom.v6i2.926.
- [13] Herianto, B. Kurniawan, Z. H. Hartomi, Y. Irawan, and M. K. Anam, “Machine Learning Algorithm Optimization using Stacking Technique for Graduation Prediction,” *J. Appl. Data Sci.*, vol. 5, no. 3, pp. 1272–1285, 2024.
- [14] M. S. Al Reshan *et al.*, “Enhancing Breast Cancer Detection and Classification Using Advanced Multi-Model Features and Ensemble Machine Learning Techniques,” *Life*, vol. 13, no. 10, 2023, doi: 10.3390/life13102093.
- [15] K. A. Rashedi, M. T. Ismail, S. Al Wadi, A. Serroukh, T. S. Alshammari, and J. J. Jaber, “Multi-Layer Perceptron-Based Classification with Application to Outlier Detection in Saudi Arabia Stock Returns,” *J. Risk Financ. Manag.*, vol. 17, no. 2, 2024, doi: 10.3390/jrfm17020069.