



## Implementasi Algoritma XGBoost, CatBoost, dan LGBM untuk Klasifikasi Pencemaran Udara

Gilang Enggar Saputra<sup>1</sup>, Made Hanindia Prami Swari<sup>2</sup>, Afina Lina Nurlaili<sup>3</sup>

<sup>1,2,3</sup>Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia

E-mail: 21081010237@student.upnjatim.ac.id, madehanindia.fik@upnjatim.ac.id, afina.lina.if@upnjatim.ac.id

Article Info	Abstract
<b>Article History</b> Received: 2025-10-07 Revised: 2025-11-13 Published: 2025-12-06  <b>Keywords:</b> <i>Air Pollution;</i> <i>ISPU;</i> <i>XGBoost;</i> <i>CatBoost;</i> <i>LightGBM;</i> <i>Machine Learning</i> <i>Classification.</i>	This study aims to classify air pollution levels in Surabaya City using supervised learning methods based on ensemble boosting algorithms, namely XGBoost, CatBoost, and LightGBM. The dataset consists of five air quality parameters (PM10, SO <sub>2</sub> , CO, O <sub>3</sub> , and NO <sub>2</sub> ) processed through data cleaning, encoding, and normalization. Model performance was evaluated using accuracy, precision, recall, and F1-score. The results show that the three models successfully learned the underlying patterns, although their performance was affected by class imbalance. The XGBoost model achieved optimal performance after applying ClassWeight adjustment and minority class merging techniques, with an accuracy of 0.9594, precision of 0.8632, recall of 0.7787, and an F1-score of 0.8098. CatBoost and LightGBM also demonstrated strong performance, achieving their highest F1-scores of 0.7549 and 0.7263, respectively, after data balancing. These findings highlight the effectiveness of imbalance-handling techniques and provide a foundation for developing a more accurate and adaptive air quality prediction system.
Artikel Info	Abstrak
<b>Sejarah Artikel</b> Diterima: 2025-10-07 Direvisi: 2025-11-13 Dipublikasi: 2025-12-06  <b>Kata kunci:</b> <i>Pencemaran Udara;</i> <i>ISPU;</i> <i>XGBoost;</i> <i>CatBoost;</i> <i>LightGBM;</i> <i>Machine Learning;</i> <i>Klasifikasi.</i>	Penelitian ini bertujuan untuk mengklasifikasikan tingkat pencemaran udara di Kota Surabaya menggunakan metode supervised learning berbasis algoritma ensemble boosting, yaitu XGBoost, CatBoost, dan LightGBM. Dataset terdiri atas lima parameter kualitas udara (PM10, SO <sub>2</sub> , CO, O <sub>3</sub> , dan NO <sub>2</sub> ) yang diproses melalui tahapan pembersihan, encoding, dan normalisasi. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa ketiga model mampu mempelajari pola data dengan baik, meskipun performanya dipengaruhi oleh ketidakseimbangan kelas. Model XGBoost memberikan performa optimal setelah penerapan teknik ClassWeight dan penggabungan kelas minoritas, dengan akurasi 0.9594, precision 0.8632, recall 0.7787, dan F1-score 0.8098. Model CatBoost dan LightGBM juga menunjukkan performa baik, masing-masing mencapai F1-score tertinggi 0.7549 dan 0.7263 setelah penyeimbangan data. Temuan ini menegaskan efektivitas teknik penanganan ketidakseimbangan data serta memberikan dasar bagi pengembangan sistem prediksi kualitas udara yang lebih akurat dan adaptif.

### I. PENDAHULUAN

Pencemaran udara merupakan salah satu permasalahan lingkungan paling serius di Indonesia dan menjadi ancaman bagi kesehatan masyarakat serta keseimbangan ekosistem. Faktor penyebab utamanya berasal dari aktivitas manusia, terutama peningkatan jumlah kendaraan bermotor, aktivitas industri, dan pembangunan kawasan perkotaan (Umri et al., 2021). Sebagai kota metropolitan dengan mobilitas tinggi, Surabaya menghadapi tantangan besar terkait kualitas udara, terutama dari polutan seperti partikulat (PM10), sulfur dioksida (SO<sub>2</sub>), karbon monoksida (CO), ozon (O<sub>3</sub>), dan nitrogen dioksida (NO<sub>2</sub>) yang berdampak negatif terhadap kesehatan masyarakat (Nababan et al., 2023). Oleh karena itu, dibutuhkan analisis yang akurat untuk

memahami pola pencemaran udara sebagai dasar dalam mendukung pengambilan kebijakan lingkungan yang tepat.

Perkembangan teknologi machine learning telah membuka peluang baru dalam analisis data lingkungan. Berbagai algoritma ensemble boosting seperti Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), dan Light Gradient Boosting Machine (LightGBM) dikenal efektif dalam menangani data berukuran besar, kompleks, dan memiliki pola nonlinear. XGBoost unggul dalam efisiensi komputasi dan mekanisme regularisasi untuk mencegah overfitting (Yulianti et al., 2022), CatBoost efektif dalam menangani fitur kategorikal serta meminimalkan target leakage (Istianto et al., 2024), sementara LightGBM memiliki performa cepat dan efisien melalui strategi leaf-wise

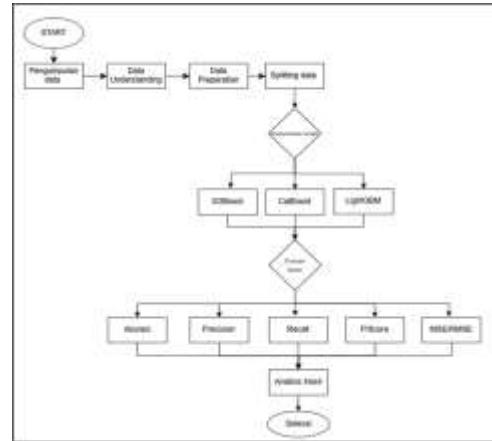
growth (Pillai et al., 2020). Algoritma-algoritma ini telah digunakan pada berbagai penelitian sebelumnya, termasuk prediksi kualitas udara (Syukur, 2024) dan perbandingan metode boosting pada kasus klasifikasi (Arif, 2023).

Meskipun demikian, penelitian terkait klasifikasi kualitas udara di Indonesia masih menghadapi masalah ketidakseimbangan kelas pada kategori ISPU, yang berpotensi menurunkan performa model dan menyebabkan bias terhadap kelas mayoritas. Sebagian besar penelitian sebelumnya hanya mengandalkan teknik oversampling seperti SMOTE (Nababan et al., 2023), sehingga metode alternatif seperti ClassWeight adjustment dan minority class merging belum banyak dievaluasi dalam konteks pencemaran udara. Kondisi tersebut menjadi celah penelitian yang penting untuk dikembangkan agar model mampu memberikan hasil klasifikasi yang lebih stabil dan representatif.

Berdasarkan permasalahan tersebut, penelitian ini menyajikan pendekatan komparatif antara tiga algoritma ensemble boosting—XGBoost, CatBoost, dan LightGBM—dalam klasifikasi tingkat pencemaran udara di Kota Surabaya berdasarkan parameter PM10, SO<sub>2</sub>, CO, O<sub>3</sub>, dan NO<sub>2</sub>. Fokus utama penelitian ini adalah menganalisis pengaruh teknik penyeimbangan data terhadap performa model, khususnya melalui penerapan ClassWeight dan penggabungan kelas minoritas sebagai kontribusi baru yang jarang digunakan pada studi pencemaran udara di Indonesia.

Rumusan masalah dalam penelitian ini difokuskan pada dua aspek utama. Pertama, bagaimana performa algoritma XGBoost, CatBoost, dan LightGBM dalam mengklasifikasikan tingkat pencemaran udara berdasarkan kategori ISPU pada kondisi data yang tidak seimbang. Kedua, sejauh mana teknik ClassWeight dan minority class merging mampu meningkatkan kinerja model dalam menangani ketidakseimbangan kelas, sehingga hasil klasifikasi menjadi lebih akurat dan representatif. Tujuan dari penelitian ini adalah untuk melakukan perbandingan kinerja ketiga algoritma ensemble boosting tersebut dan menentukan metode yang paling efektif dalam klasifikasi kualitas udara. Selain itu, penelitian ini juga bertujuan memberikan kontribusi terhadap pengembangan sistem pemantauan pencemaran udara yang lebih akurat dan adaptif di Kota Surabaya melalui penerapan teknik penyeimbangan data yang lebih optimal.

## II. METODE PENELITIAN



Gambar 1. Alur Penelitian

Metode penelitian ini diawali dengan pengumpulan data yang diperoleh dari sumber resmi terkait kualitas udara. Selanjutnya dilakukan tahap data understanding untuk memahami karakteristik, struktur, serta pola awal dari data yang digunakan. Setelah itu, pada tahap data preparation dilakukan proses pembersihan data, penanganan nilai hilang, normalisasi, dan encoding agar data siap digunakan dalam pemodelan. Data yang telah siap kemudian melalui proses splitting data menjadi data latih dan data uji. Tahap berikutnya adalah implementasi model dengan menggunakan tiga algoritma yaitu XGBoost, CatBoost, dan LightGBM untuk melakukan klasifikasi. Setelah model dibangun, dilakukan evaluasi model menggunakan metrik kinerja seperti akurasi, precision, recall, dan F1-score untuk menilai performa masing-masing model. Hasil dari evaluasi tersebut kemudian dianalisis pada tahap analisis hasil untuk membandingkan dan menentukan model terbaik, hingga akhirnya diperoleh kesimpulan pada tahap selesai.

XGBoost (Extreme Gradient Boosting) merupakan pengembangan dari metode Gradient Boosting yang diusulkan oleh Tianqi Chen pada tahun 2014[4]. Algoritma ini bekerja dengan membangun pohon keputusan secara berurutan untuk meminimalkan error secara bertahap. Kelebihan XGBoost dibandingkan metode boosting konvensional adalah efisiensi komputasi tinggi, regularisasi untuk mencegah overfitting, dan kemampuan menangani data besar[6]. Fungsi objektif XGBoost terdiri dari dua komponen utama, yaitu fungsi loss dan regularisasi, yang dinyatakan dengan persamaan:

$$obj(\theta) = L(\theta) + \Omega(\theta) \quad (1)$$

CatBoost (Categorical Boosting) adalah algoritma berbasis Gradient Boosted Decision Tree (GBDT) yang dikembangkan untuk menangani data kategorikal secara efisien tanpa memerlukan one-hot encoding[7]. Algoritma ini menggunakan pendekatan *ordered boosting* untuk mencegah *target leakage* dan struktur pohon simetris (*oblivious tree*) agar model lebih stabil dan terhindar dari overfitting[8]. Dalam prosesnya, CatBoost melakukan konversi fitur kategorikal menjadi nilai numerik melalui *mean target encoding* yang diregulasi, menggunakan persamaan:

$$Encoded(X_i) = \frac{\sum_{j=1}^{i-1} 1_{x_j=x_i} \cdot Y_j + \alpha \cdot \mu}{\sum_{j=1}^{i-1} 1_{x_j=x_i} + \alpha} \quad (2)$$

Selain itu, fungsi aktivasi *sigmoid* digunakan untuk klasifikasi biner, sedangkan *softmax* diterapkan pada klasifikasi multikelas. LightGBM (LGBM) adalah framework *gradient boosting* yang dikembangkan oleh Microsoft dan dirancang untuk efisiensi tinggi pada data berukuran besar[9]. Berbeda dari algoritma boosting lainnya, LGBM membangun pohon menggunakan strategi *leaf-wise growth*, yaitu memperluas daun dengan penurunan loss terbesar[5]. Pendekatan ini membuat model lebih cepat dan akurat, meskipun berisiko overfitting jika tidak diatur dengan baik. Untuk mengoptimalkan pembagian cabang pohon, LGBM menghitung *gain* dari setiap kemungkinan split dengan persamaan.

$$Gain = \frac{1}{2} \left[ \left( \frac{G_L^2}{(H_L + \lambda)} \right) + \left( \frac{G_R^2}{(H_R + \lambda)} \right) - \left( \frac{(G_L + G_R)^2}{(H_L + H_R + \lambda)} \right) \right] - \gamma \quad (2)$$

Rumus tersebut membantu menentukan pemisahan terbaik dalam membangun pohon keputusan. Secara keseluruhan, ketiga algoritma ini sama-sama berbasis boosting, namun XGBoost unggul dalam efisiensi dan regularisasi, CatBoost dalam penanganan data kategorikal, serta LGBM dalam kecepatan dan skalabilitas pada data besar.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil Penelitian

Pada tahap awal penelitian, model XGBoost digunakan untuk mengklasifikasikan tingkat kualitas udara. Pengujian awal menunjukkan bahwa model memiliki akurasi tinggi sebesar 96,24% pada rasio 80:20. Namun, nilai precision, recall, dan F1-score masih rendah (0,38–0,48), yang mengindikasikan bahwa model terlalu fokus pada kelas mayoritas dan mengalami kesulitan dalam mengenali kelas

minoritas. Kondisi ini disebabkan oleh distribusi data yang tidak seimbang sehingga model cenderung bias terhadap kelas dengan jumlah data lebih besar.

**Tabel 1.** Pengujian XGBOOST

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
8	80:20	XGBoost	0.10	500	0.9624	0.4836	0.4878	0.4857
9	80:20	XGBoost	0.10	1000	0.9577	0.4856	0.4853	0.4854
11	70:30	XGBoost	0.01	500	0.9594	0.3835	0.3906	0.3866

Untuk mengatasi hal tersebut, dilakukan penyeimbangan data menggunakan metode *ClassWeight*, penggabungan kelas minoritas, serta kombinasi kedua teknik tersebut.

**Tabel 2.** Pengujian XGBOOST *ClassWeight* + *Penggabungan Data*

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
2	70:30	XGBoost + <i>ClassWeight</i>	0.10	500	0.9594	0.8632	0.7787	0.8098
3	70:30	XGBoost + <i>ClassWeight</i>	0.10	1000	0.9563	0.8612	0.7764	0.8077
12	70:30	XGBoost + <i>ClassWeight</i>	0.01	1000	0.9531	0.7537	0.7751	0.7631

Hasil terbaik diperoleh pada kombinasi *ClassWeight* dan penggabungan kelas dengan rasio 70:30, learning rate 0,10, dan iterasi 500. Model XGBoost kemudian mencapai akurasi 0,9594, precision 0,8632, recall 0,7787, dan F1-score 0,8098. Pencapaian ini menunjukkan peningkatan signifikan dibandingkan model dasar dan membuktikan efektivitas teknik penyeimbangan data dalam membantu model mengenali kelas minoritas dengan lebih baik.

Model CatBoost juga menunjukkan pola hasil yang serupa. Pengujian awal menghasilkan akurasi 96,24% pada rasio 80:20, namun nilai precision, recall, dan F1-score hanya sekitar 0,48. Penerapan *ClassWeight* tunggal tidak memberikan perubahan signifikan karena ketimpangan antar kelas masih terlalu besar.

**Tabel 3.** Pengujian CATBOOST

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
6	80:20	CatBoost	0.10	500	0.9577	0.4785	0.4860	0.4821
9	80:20	CatBoost	0.10	1000	0.9624	0.4809	0.4882	0.4845
18	80:20	CatBoost	0.01	1000	0.9624	0.4808	0.4886	0.4845

Kemudian diterapkan kombinasi *ClassWeight* dan penggabungan data untuk memperbaiki hasil evaluasi. Dengan kombi-

nasi tersebut, performa model meningkat secara signifikan.

**Tabel 4.** Pengujian CATBOOST *ClassWeight* + *Penggabungan Data*.

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
1	80:20	CatBoost + <i>ClassWeight</i> + <i>Penggabungan</i>	0.10	100	0.9286	0.7084	0.8442	0.7447
9	80:20	CatBoost + <i>ClassWeight</i> + <i>Penggabungan</i>	0.10	1000	0.9531	0.7271	0.7527	0.7378
16	70:30	CatBoost + <i>ClassWeight</i> + <i>Penggabungan</i>	0.01	1000	0.9406	0.7275	0.8306	0.7549

Ketika dikombinasikan dengan penggabungan data minoritas, performa model meningkat lebih baik. Hasil terbaik dicapai pada rasio 70:30 dengan learning rate 0,01 dan iterasi 1000, menghasilkan akurasi 0,9406, precision 0,7275, recall 0,8306, dan F1-score 0,7549. Nilai recall yang tinggi pada pengujian akhir menunjukkan bahwa CatBoost lebih mampu mengenali kelas minoritas setelah dilakukan penyeimbangan data.

Pengujian awal pada algoritma LGBM menunjukkan hasil akurasi tertinggi sebesar 0.9775 pada rasio 75:25 dengan *learning rate* 0.10 dan iterasi 100, namun nilai *precision*, *recall*, dan *F1-score* masih rendah, sekitar 0.39–0.49. Meskipun model memiliki akurasi tinggi, hal ini tidak mencerminkan kemampuan sebenarnya dalam mengenali kelas minoritas.

**Tabel 5.** Pengujian LGBM

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
2	75:25	LGBM	0.10	100	0.9775	0.3906	0.3972	0.3938
3	80:20	LGBM	0.10	100	0.9711	0.4858	0.4926	0.4892
6	75:25	LGBM	0.10	1000	0.9775	0.3906	0.3972	0.3938

Selanjutnya dilakukan penyeimbangan data menggunakan kombinasi *ClassWeight* dan penggabungan kelas, performa model meningkat signifikan.

**Tabel 6.** Pengujian LGBM *ClassWeight* + *Penggabungan Data*.

No	Rasio Splitting data	Model Pengujian	Learning Rate	Iterasi	Accuracy	Precision	Recall	F1-Score
2	70:30	LGBM	0.10	500	0.9563	0.7520	0.7117	0.7283
3	70:30	LGBM	0.10	1000	0.9563	0.7520	0.7117	0.7283
7	80:20	LGBM	0.10	100	0.8390	0.6956	0.7429	0.7098

Setelah diterapkan kombinasi *ClassWeight* dan penggabungan kelas, performa model meningkat signifikan. Hasil terbaik diperoleh pada rasio 70:30, learning rate 0,10, dan iterasi 500, dengan akurasi 0,9563, precision 0,7520, recall 0,7117, dan F1-score 0,7263. Performa ini lebih seimbang dibandingkan model dasar, meskipun masih berada sedikit di bawah XGBoost dan CatBoost.

## B. Pembahasan

Hasil penelitian menunjukkan bahwa ketiga algoritma boosting, yaitu XGBoost, CatBoost, dan LightGBM, pada awalnya memiliki akurasi tinggi namun kurang mampu mengenali kelas minoritas. Hal ini dipengaruhi oleh ketidakseimbangan data yang menyebabkan model lebih condong pada kelas mayoritas. Setelah diterapkan teknik penyeimbangan data berupa *ClassWeight* dan penggabungan kelas minoritas, performa ketiga model meningkat secara signifikan terutama pada metrik precision, recall, dan F1-score.

XGBoost menjadi algoritma dengan performa terbaik dengan F1-score tertinggi, yaitu 0,8098. CatBoost berada di posisi kedua dengan F1-score 0,7549, disusul oleh LightGBM dengan F1-score 0,7263. Keunggulan XGBoost dapat dikaitkan dengan proses optimasi gradien yang lebih stabil serta regularisasi yang efektif dalam mengendalikan kompleksitas model pada data tidak seimbang. Penerapan kombinasi *ClassWeight* dan penggabungan kelas terbukti menjadi strategi yang efektif karena mampu memperbaiki kemampuan model dalam mengenali kelas minoritas tanpa mengorbankan akurasi keseluruhan.

Hasil ini menunjukkan bahwa model berbasis ensemble boosting dapat digunakan sebagai sistem pendukung keputusan untuk pemantauan kualitas udara, khususnya dalam mendeteksi kategori pencemaran yang jarang muncul. Penelitian ini sekaligus memperkuat bukti bahwa teknik penyeimbangan data memainkan peran penting dalam memaksimalkan performa algoritma machine learning pada kasus klasifikasi dengan distribusi kelas yang tidak merata.

## IV. SIMPULAN DAN SARAN

### A. Simpulan

Penelitian ini membandingkan performa tiga algoritma ensemble boosting, yaitu XGBoost, CatBoost, dan LGBM, dalam

mengklasifikasikan tingkat kualitas udara berdasarkan kategori ISPU. Hasil pengujian menunjukkan bahwa XGBoost merupakan model dengan performa terbaik, diikuti oleh CatBoost dan LGBM. Rasio pembagian data 70:30 dan learning rate 0,10 terbukti menjadi konfigurasi yang paling optimal karena memberikan keseimbangan antara akurasi dan kemampuan generalisasi. Jumlah iterasi 500 menjadi titik stabil sebelum model mengalami overfitting. Selain itu, penerapan teknik ClassWeight dan penggabungan kelas terbukti efektif dalam menangani ketidakseimbangan data, sehingga meningkatkan kemampuan ketiga algoritma dalam mengenali kelas minoritas.

Secara keseluruhan, kombinasi parameter optimal dan teknik penyeimbangan data menjadikan XGBoost sebagai model paling unggul dalam klasifikasi kualitas udara. Hasil penelitian ini memperkuat penerapan metode ensemble boosting sebagai pendekatan yang akurat dan efisien dalam sistem klasifikasi lingkungan berbasis data di Indonesia.

## B. Saran

Berdasarkan hasil penelitian, disarankan agar penelitian selanjutnya melakukan perbandingan dengan algoritma machine learning lainnya untuk mendapatkan model yang lebih efisien dalam klasifikasi pencemaran udara. Evaluasi model sebaiknya tidak hanya berfokus pada akurasi, tetapi juga mempertimbangkan metrik seperti recall, precision, dan F1-score agar hasil lebih representatif terhadap data tidak seimbang. Selain itu, penggunaan data yang lebih panjang dan variabel yang lebih beragam dapat meningkatkan akurasi serta kemampuan generalisasi model. Hasil penelitian ini diharapkan dapat menjadi dasar pengembangan sistem informasi kualitas udara yang lebih aplikatif, sehingga mampu membantu masyarakat dan pemerintah dalam pengambilan keputusan serta mitigasi dampak pencemaran udara.

## DAFTAR RUJUKAN

- A. A. Nababan, M. Jannah, M. Aulina, and D. Andrian, "Prediksi Kualitas Udara Menggunakan Xgboost Dengan Synthetic Minority Oversampling Technique (Smote) Berdasarkan Indeks Standar Pencemaran Udara (Ispu)," *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 7, no. 1, pp. 214–219, 2023, doi: 10.59697/jtik.v7i1.66.
- A. Febriansyah Istianto, A. Id Hadiana, and F. Rakhmat Umbara, "Prediksi Curah Hujan Menggunakan Metode Categorical Boosting (Catboost)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 4, pp. 2930–2937, 2024, doi: 10.36040/jati.v7i4.7304.
- F. D. U. Arif, *Perbandingan Kinerja Algoritma Random Forest, Xgboost Dan Lightgbm Dalam Klasifikasi Emosi Komentar Reddit*. 2023. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/handle/123456789/76259>
- L. Gladyza, A. Putri, S. A. Wicaksono, and B. zzRahayudi, "Analisis Klasifikasi Spam Email Menggunakan Metode Extreme Gradient Boosting ( XGBoost )," vol. 9, no. 2, pp. 1–8, 2025.
- M. A. A. Syukur, "Prediksi Indeks Kualitas Udara Menggunakan Metode CatBoost," *Undergrad. thesis, Univ. Islam Negeri Maulana Malik Ibrahim*, pp. 1–120, 2024, [Online]. Available: <https://ejournal.uin-suka.ac.id/saintek/JISKA/article/view/4609>
- N. L. Sabili and F. R. Umbara, "KLASIFIKASI PENYAKIT DIABETES MENGGUNAKAN ALGORITMA CATEGORICAL BOOSTING DENGAN FAKTOR RISIKO DIABETES," *J. Mhs. Tek. Inform.*, vol. 8, no. 6, pp. 11391–11398, 2024.
- N. Sunil Pillai, M. S. Nair, and D. M. Khan, "Air Quality Prediction System using LightGBM," *Int. Res. J. Eng. Technol.*, no. July, 2020, [Online]. Available: [www.irjet.net](http://www.irjet.net)
- S. E. Herni Yulianti, Oni Soesanto, and Yuana Sukmawaty, "Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," *J. Math. Theory Appl.*, vol. 4, no. 1, pp. 21–26, 2022, doi: 10.31605/jomta.v4i1.1792.
- S. S. A. Umri *et al.*, "Analysis and Comparison of Classification Algorithm in Air," *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 2, pp. 98–104, 2021, doi: 10.33387/jiko.