IMPLEMENTASI LIGHTGBM DENGAN KNN *IMPUTATION* UNTUK DETEKSI DINI RISIKO KEHAMILAN

SKRIPSI

Disusun Oleh

SYAHNUR ALAWIYAH

09040222072

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2026

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama	:	SYAHNUR ALAWIYAH
NIM	:	09040222072
Program Studi	:	Matematika
Angkatan	:	2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "IMPLEMENTASI LIGHTGBM DENGAN KNN *IMPUTATION* UNTUK DETEKSI DINI RISIKO KEHAMILAN". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 01 April 2026

Yang menyatakan,



SYAHNUR ALAWIYAH

NIM. 09040222072

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama	:	SYAHNUR ALAWIYAH
NIM	:	09040222072
Judul skripsi	:	IMPLEMENTASI LIGHTGBM DENGAN KNN *IMPUTATION* UNTUK DETEKSI DINI RISIKO KEHAMILAN

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I

Dian Yuliati, M.Si
NIP. 198707142020122015

Pembimbing II

Nurissaidah Ulinnuha, M.Kom
NIP. 199011022014032004

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya

Dr. Yuniar Farida, M.T
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama	:	SYAHNUR ALAWIYAH
NIM	:	09040222072
Judul Skripsi	:	IMPLEMENTASI LIGHTGBM DENGAN KNN *IMPUTATION* UNTUK DETEKSI DINI RISIKO KEHAMILAN

Telah dipertahankan di depan Tim Penguji
pada tanggal 01 April 2026

Mengesahkan,
Tim Penguji

Penguji I

Aris Fanani, M.Kom
NIP. 198701272014031002

Penguji II

Putroue Keumala Intan, M.Si
NIP. 198805282018012001

Penguji III

Dian Yuliati, M.Si
NIP. 198707142020122015

Penguji IV

Nurissaidah Ulinnuha, M.Kom
NIP. 199011022014032004

Mengetahui,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya

Dr. A. Saepul Hamdani, M.Pd
NIP. 196507312000031002

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

UIN SUNAN AMPEL
S U R A B A Y A

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : SYAHNUR ALAWIYAH
NIM : 09040222072
Fakultas/Jurusan : Fakultas Sains dan Teknologi / Matematika
E-mail address : 09040222072@student.uinsby.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
☑ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :
Implementasi LightGBM dengan KNN Imputation untuk Deteksi Dini Risiko Kehamilan

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 08 April 2026

Penulis

(SYAHNUR ALAWIYAH)
nama terang dan tanda tangan

DAFTAR ISI

HALAMAN JUDUL . . . . . **i**

LEMBAR PERSETUJUAN PEMBIMBING . . . . . **ii**

PENGESAHAN TIM PENGUJI SKRIPSI . . . . . **iii**

HALAMAN PERNYATAAN KEASLIAN . . . . . **iv**

MOTTO . . . . . **v**

PERSEMBAHAN . . . . . **vi**

KATA PENGANTAR . . . . . **vii**

DAFTAR ISI . . . . . **ix**

DAFTAR TABEL . . . . . **xii**

DAFTAR GAMBAR . . . . . **xv**

ABSTRAK . . . . . **xvi**

ABSTRACT . . . . . **xvii**

I PENDAHULUAN **1**

1.1. Latar Belakang Masalah 1
1.2. Rumusan Masalah 6
1.3. Tujuan Penelitian 7
1.4. Manfaat Penelitian 7
1.5. Batasan Masalah 8
1.6. Sistematika Penelitian 8

II TINJAUAN PUSTAKA **10**

2.1. Risiko Kehamilan 10
2.2. Faktor-Faktor yang Mempengaruhi Risiko Kehamilan 11
2.2.1. Usia Ibu 11

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

2.2.2. Tekanan Darah 12
2.2.3. Gula Darah 12
2.2.4. Suhu Tubuh 13
2.2.5. *Body Mass Index* (BMI) 13
2.2.6. Riwayat Komplikasi Kehamilan 13
2.2.7. Diabetes 14
2.2.8. Kesehatan Mental 15
2.2.9. Denyut Jantung 15
2.3. *Preprocessing* 15
2.3.1. *Missing Value* 16
2.3.2. Normalisasi 18
2.4. *Stratified K-Fold Cross Validation* (SKCV) 18
2.5. *Light Gradient Boosting Machine* (LightGBM) 19
2.6. *Confusion Matrix* 23
2.7. Integrasi Keilmuan 24

III METODE PENELITIAN 29
3.1. Jenis Penelitian 29
3.2. Sumber Data 29
3.3. Variabel Penelitan 29
3.4. Tahapan Penelitian 30
3.5. Tahapan Uji Coba 34

IV HASIL DAN PEMBAHASAN 36
4.1. Deskripsi Data 36
4.2. *Preprocessing* 40
4.2.1. Penanganan *Missing Value* 40
4.2.2. Normalisasi 43
4.3. *Stratified K-Fold Cross Validation* (SKCV) 43
4.4. Proses Klasifikasi 44
4.4.1. Tahap *Training* 45
4.4.2. Tahap *Testing* 60
4.5. Evaluasi Hasil 63
4.6. Skema Uji Coba Parameter 65
4.7. Analisis Hasil 86
4.8. Integrasi Nilai Islam dalam Klasifikasi 89

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

V PENUTUP 92
5.1. Kesimpulan 92
5.2. Saran 93

DAFTAR PUSTAKA 93

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

DAFTAR TABEL

2.1 *Confusion Matrix* 23

3.1 Variabel Prediktor Penelitian 30
3.2 Kombinasi Parameter LightGBM 35

4.1 Deskripsi Data Kategorik 36
4.2 Statistik Deskriptif Data Numerik pada Variabel Penelitian 37
4.3 Sampel Data Penelitian 39
4.4 Hasil Perhitungan Jarak 41
4.5 Tetangga Terdekat 41
4.6 Data Setelah Pengolahan *Missing Value* 42
4.7 Contoh Data Hasil Normalisasi 43
4.8 Distribusi Data *Training* dan *Testing* pada Setiap *Fold* 44
4.9 Nilai Residual 45
4.10 Nilai *Hessian* 46
4.11 Nilai *Information Gain* pada *Node* 1 50
4.12 Nilai *Information Gain* pada *Node* 2 Kanan 54
4.13 Nilai *Information Gain* pada *Node* 3 Kanan 58
4.14 Sampel Data *Testing* 60
4.15 Nilai *leaf* Tahap *Testing* 61
4.16 Hasil *Confusion Matrix* Model 63
4.17 Uji Coba pada Parameter *n_estimators* 100 dan *max_depth* 5 dengan Menghapus *Missing Value* 65
4.18 Uji Coba pada Parameter *n_estimators* 100 dan *max_depth* 10 dengan Menghapus *Missing Value* 66
4.19 Uji Coba pada Parameter *n_estimators* 100 dan *max_depth* 15 dengan Menghapus *Missing Value* 67

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

4.20 Uji Coba pada Parameter *n_estimators* 250 dan *max_depth* 5 dengan Menghapus *Missing Value* 68
4.21 Uji Coba pada Parameter *n_estimators* 250 dan *max_depth* 10 dengan Menghapus *Missing Value* 69
4.22 Uji Coba pada Parameter *n_estimators* 250 dan *max_depth* 15 dengan Menghapus *Missing Value* 70
4.23 Uji Coba pada Parameter *n_estimators* 500 dan *max_depth* 5 dengan Menghapus *Missing Value* 71
4.24 Uji Coba pada Parameter *n_estimators* 500 dan *max_depth* 10 dengan Menghapus *Missing Value* 72
4.25 Uji Coba Parameter *n_estimators* 500 dan *max_depth* 15 dengan Menghapus *Missing Value* 73
4.26 Kombinasi Parameter Terbaik LightGBM dengan Menghapus *Missing Value* 74
4.27 *Confusion Matrix* Kombinasi Parameter Terbaik LightGBM dengan Menghapus *Missing Value* 74
4.28 Uji Coba Parameter *n_estimators* 100 dan *max_depth* 5 dengan KNN *Imputation* 76
4.29 Uji Coba Parameter *n_estimators* 100 dan *max_depth* 10 dengan KNN *Imputation* 77
4.30 Uji Coba Parameter *n_estimators* 100 dan *max_depth* 15 dengan KNN *Imputation* 78
4.31 Uji Coba Parameter *n_estimators* 250 dan *max_depth* 5 dengan KNN *Imputation* 79
4.32 Uji Coba Parameter *n_estimators* 250 dan *max_depth* 10 dengan KNN *Imputation* 80
4.33 Uji Coba Parameter *n_estimators* 250 dan *max_depth* 15 dengan KNN *Imputation* 81
4.34 Uji Coba Parameter *n_estimators* 500 dan *max_depth* 5 dengan KNN *Imputation* 82
4.35 Uji Coba Parameter *n_estimators* 500 dan *max_depth* 10 dengan KNN *Imputation* 83

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

4.36 Uji Coba Parameter *n_estimators* 500 dan *max_depth* 15 dengan KNN *Imputation* 84
4.37 Kombinasi Parameter Terbaik LightGBM dengan KNN *Imputation* . 85
4.38 *Confusion Matrix* Kombinasi Parameter Terbaik LightGBM dengan KNN *Imputation* 85
4.39 Perbandingan Hasil Penelitian Ini dengan Penelitian Sebelumnya . . 88

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

DAFTAR GAMBAR

2.1 *Stratified K-Fold Cross Validation* 19
2.2 *Leaf-Wise Growth* 20

3.1 Diagram Alir Penelitian 31
3.2 Diagram Alir *Preprocessing* 31
3.3 Diagram Alir LightGBM 33

4.1 Visualisasi *Missing Value* 40
4.2 *Node* 1 pada $X_1 \leq 0.191$ 46
4.3 *Node* 1 pada $X_4 \leq 0.305$ 47
4.4 *Node* 1 pada $X_8 \leq 0$ 48
4.5 *Node* 2 pada $X_3 \leq 0.425$ 50
4.6 *Node* 2 pada $X_7 \leq 0$ 51
4.7 *Node* 2 pada $X_{10} \leq 0$ 52
4.8 *Node* 3 pada $X_5 \leq 0.591$ 54
4.9 *Node* 3 pada $X_6 \leq 0.186$ 55
4.10 *Node* 3 pada $X_{11} \leq 0.545$ 56
4.11 Visualisasi Pohon Pertama 59
4.12 Contoh Jalur Pohon Keputusan 62
4.13 Perbandingan Hasil Penanganan *Missing Value* 87

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

ABSTRAK

IMPLEMENTASI LIGHTGBM DENGAN KNN *IMPUTATION* UNTUK DETEKSI DINI RISIKO KEHAMILAN

Risiko kehamilan merupakan salah satu permasalahan utama dalam kesehatan maternal yang berkontribusi terhadap tingginya angka kesakitan dan kematian ibu serta bayi. Risiko tersebut dipengaruhi oleh berbagai faktor sehingga diperlukan pendekatan analisis yang akurat dan efisien untuk mendukung deteksi dini. Penelitian ini bertujuan untuk membangun dan mengevaluasi model klasifikasi dalam mendeteksi tingkat risiko kehamilan dengan memanfaatkan teknik *K-Nearest Neighbor Imputation* (KNN *Imputation*) untuk menangani *missing value* serta algoritma LightGBM sebagai metode klasifikasi utama. Proses pengembangan model dilakukan melalui serangkaian uji coba parameter guna memperoleh kombinasi terbaik yang mampu meningkatkan performa model. Evaluasi model dilakukan menggunakan skema *Stratified K-Fold Cross Validation* (SKCV) dengan metrik akurasi, sensitivitas, dan spesifisitas. Hasil penelitian menunjukkan bahwa model yang diusulkan menghasilkan akurasi sebesar 97.64%, sensitivitas sebesar 96.81%, dan spesifisitas sebesar 98.18%. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengklasifikasikan tingkat risiko kehamilan. Dengan demikian, metode yang digunakan berpotensi dikembangkan sebagai sistem pendukung keputusan untuk membantu tenaga medis dalam melakukan deteksi dini risiko kehamilan.

Kata kunci: Kehamilan Berisiko Tinggi, KNN *Imputation*, LightGBM, SKCV.

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

ABSTRACT

IMPLEMENTATION OF LIGHTGBM WITH KNN IMPUTATION FOR EARLY DETECTION OF PREGNANCY RISK

Pregnancy risks are one of the primary issues in maternal health, contributing to high rates of maternal and infant morbidity and mortality. These risks are influenced by various factors, necessitating an accurate and efficient analytical approach to support early detection. This study aims to develop and evaluate a classification model for detecting pregnancy risk levels by utilizing the K-Nearest Neighbor Imputation (KNN *Imputation*) technique to handle missing values and the LightGBM algorithm as the primary classification method. The model development process involved a series of parameter trials to obtain the optimal combination capable of enhancing model performance. Model evaluation was conducted using the Stratified K-Fold Cross Validation (SKCV) scheme with accuracy, sensitivity, and specificity metrics. The results of the study show that the proposed model achieved an accuracy of 97.64%, a sensitivity of 96.81%, and a specificity of 98.18%. These results indicate that the model has excellent capability in classifying pregnancy risk levels. Thus, the method used has the potential to be developed as a decision support system to assist medical personnel in the early detection of pregnancy risks.

Keywords: High-Risk Pregnancy, KNN *Imputation*, LightGBM, SKCV.

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

DAFTAR PUSTAKA

Adha, S. (2022). *Pedoman Program Pelayanan Kesehatan Ibu Maternal engan Pendekatan Keluarga Melalui Gerakan Masyarakat Peduli Gender (Pak Bamega) dalam Upaya Percepatan Penurunan Angka Kematian Ibu*. CV Widina Media Utama.

Adhitya, R. R., Witanti, W., and Yuniarti, R. (2023). Perbandingan Metode CART Dan Naïve Bayes Untuk Klasifikasi Customer Churn. *INFOTECH journal*, 9(2):307–318.

Akbar, I., Supriadi, F., and Junaedi, D. I. (2025). Pemanfaatan Machine Learning di Bidang Kesehatan. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1):1744–1749.

Akhmad, Y. A., Wiryono, S. K., and Sukarno, S. (2024). Mitigating Default Risks in Peer-To-Peer Lending Platforms: The Role of Information Asymmetry in Indonesia. *African Journal of Biomedical Research*, 27(3S):1201–1218.

Allorerung, P. P., Erna, A., Bagussahrir, M., and Alam, S. (2024). Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 9(3):178–191.

Amalia, F. and Sayuti, M. I. (2023). Peningkatan Upaya Penurunan Angka Kematian Ibu dan Anak Melalui Penggunaan Lembar Deteksi Risiko Kehamilan pada Pelayanan Antenatal Care (ANC) Terpadu di Kabupaten Sinjai. *Journal of Governance and Policy Innovation*, 3(1):58–72.

Ardiana, M. (2024). *Skrining Penyakit Kardiovaskular pada Kehamilan*. Airlangga University Press.

Arif, M. (2025). Explainable AI in Maternal Health: Utilizing XGBoost and SHAP Values for Enhanced Risk Prediction and Interpretation. *International Journal of Emerging Multidisciplinaries: Computer Science & Artificial Intelligence*, 4(1):16.

Arsenault, C., Mfeka-Nkabinde, N. G., and et al. (2024). Antenatal Care Quality and Detection of Risk Among Pregnant Women: An Observational Study in Ethiopia, India, Kenya, and South Africa. *PLOS Medicine*, 21(8):e1004446.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Astarani, I. G. A. R. and Rahayuda, I. G. S. (2025). Analisis Perbandingan XGBoost dan LightGBM dalam Prediksi Penjualan Ritel Walmart Store Sales. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(4):717–728.

BKKBN (2024). Laporan Kependudukan Indonesia 2024. Technical report, Badan Kependudukan dan Keluarga Berencana Nasional.

Chahid, I., Elmiad, A. K., and Badaoui, M. (2023). Data Preprocessing For Machine Learning Applications in Healthcare: A Review. *2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA)*, pages 1–6.

Chen, F., Yu, L., Mao, J., Yang, Q., Wang, D., and Yu, C. (2024). A Novel Data-Characteristic-Driven Modeling Approach for Imputing Missing Value in Industrial Statistics: A Case Study of China Electricity Statistics. *Applied Energy*, 373:123854.

Chen, X., Aljrees, T., Umer, M., Saidani, O., Almuqren, L., Mzoughi, O., Ishaq, A., and Ashraf, I. (2023). Cervical Cancer Detection Using K Nearest Neighbor Imputer and Stacked Ensemble Learning Model. *Digital Health*, 9:20552076231203802.

Crump, C., Sundquist, J., and Sundquist, K. (2024). Adverse Pregnancy Outcomes and Long-Term Mortality in Women. *JAMA Internal Medicine*, 184(6):631–640.

Cífková, R. (2023). Hypertension in Pregnancy: A Diagnostic and Therapeutic Overview. *High Blood Pressure and Cardiovascular Prevention*, 30(4):289–303.

Damayanti, N. A., Wulandari, R. D., and Ridlo, I. A. (2023). Maternal Health Care Utilization Behavior, Local Wisdom, and Associated Factors Among Women in Urban and Rural Areas, Indonesia. *International Journal of Women's Health*, 15:665–677.

Data, W. B. (2023). Stillbirth Rate (per 1,000 Total Births).

Deng, L., Lu, K., and Hu, H. (2025). An Interpretable LightGBM Model for Predicting Coronary Heart Disease: Enhancing Clinical Decision-Making with Machine learning. *PLOS One*, 20(9 September).

Dol, J., Hughes, B., Bonet, M., Dorey, R., Dorling, J., Grant, A., Langlois, E. V., Monaghan, J., Ollivier, R., Parker, R., Roos, N., Scott, H., Shin, H. D., and Curran, J. (2022). Timing of Maternal Mortality and Severe Morbidity During the

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Postpartum Period: A Systematic Review. *JBI Evidence Synthesis*, 20(9):2119–2194.

Dol, J., Hughes, B., Bonet, M., Dorey, R., Dorling, J., Grant, A., Langlois, E. V., Monaghan, J., Ollivier, R., Parker, R., Roos, N., Scott, H., Shin, H. D., and Curran, J. (2023). Timing of Maternal Mortality and Severe Morbidity During the Postpartum Period: A Systematic Review. *JBI Evidence Synthesis*, 21(1):98–199.

Dubey, P., Dwivedi, A. K., Sharma, K., Martin, S. L., Thompson, P. M., and Reddy, S. Y. (2025). Associations of Mental Disorders with Maternal Health Outcomes. *Communications Medicine*, 5(1):1–8.

Ekawati, F. M., Muchlis, M., Iturrieta-Guaita, N. G., and Putri, D. A. D. (2023). Recommendations for improving maternal health services in Indonesian primary care under the COVID-19 pandemic: Results of a systematic review and appraisal of international guidelines. *Sexual & Reproductive Healthcare*, 35:100811.

Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., and Tabona, O. (2021). A Survey on Missing Data in Machine Learning. *Journal of Big Data*, 8(1):1–37.

Evranata Pardede, E., Anggraeny, F. T., and Junaidi, A. (2025). Prophet–LightGBM Hybrid Model Implementation in Cafe Menu Sales Prediction: Implementasi Model Hybrid Prophet–LightGBM dalam Prediksi Penjualan Menu Kafe. *JATI EMAS (Jurnal Aplikasi Teknik dan Pengabdian Masyarakat)*, 9(4):413–418.

Fadlil, A., Herman, and Praseptian, M. D. (2022). K Nearest Neighbor Imputation Performance on Missing Value Data Graduate User Satisfaction. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(4):570–576.

Fatkhurrohman, Nugroho, B. I., and Fadillah, N. (2023). Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 4(3):3906–3917.

Febriantoro, E., Setyati, E., and Santoso, J. (2023). Pemodelan Prediksi Kuantitas Penjualan Mainan Menggunakan LightGBM. *SMARTICS Journal*, 9(1):7–13.

Gan, M., Pan, S., Chen, Y., Cheng, C., Pan, H., and Zhu, X. (2021). Application of The Machine Learning LightGBM Model to The Prediction of The Water Levels of The Lower Columbia River. *Journal of Marine Science and Engineering*, 9(5).

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Ganie, S. M., Pramanik, P. K. D., and Zhao, Z. (2024). Improved Liver Disease Prediction from Clinical Data Through An Evaluation of Ensemble Learning Approaches. *BMC Medical Informatics and Decision Making*, 24(1):160.

Goetzl, L. (2023). Maternal Fever in Labor: Etiologies, Consequences, and Clinical Management. *American Journal of Obstetrics and Gynecology*, 228(5):S1274–S1282.

Goyal, R., Singhal, M., and Jialal, I. (2025). *Type 2 Diabetes*. StatPearls Publishing, Treasure Island (FL).

Hamdani, I. M., Nurhidayat, Karman, A., H, N. F. A., and Julyaningsih, A. H. (2024). Edukasi dan Pelatihan Data Science dan Data Preprocessing. *Intisari: Jurnal Inovasi Pengabdian Masyarakat*, 2(1):19–26.

Hartanto, A. D., Kholik, Y. N., and Pristyanto, Y. (2023). Stock Price Time Series Data Forecasting Using The Light Gradient Boosting Machine (LightGBM) Model. *JOIV: International Journal on Informatics Visualization*, 7(4):2270–2279.

Hippy, Y., Mapisamang, N. R., and Supu, R. D. (2023). Comparison of Blood Pressure in Sitting, Standing, and Lying Positions in Keeping Homeostatis. *Jurnal Ilmiah dr. Aloei Saboe*, 3(2):100–110.

Ibarra-Vazquez, G., Ramírez-Montoya, M. S., and Buenestado-Fernández, M. (2023). Forecasting Gender in Open Education Competencies: A Machine Learning Approach. *IEEE Transactions on Learning Technologies*, 17:1236–1247.

Ibrahim, T. and Ridwan, D. A. (2022). Determinan Penyebab Kematian Ibu dan Neonatal di Indonesia. *Jurnal Kedokteran Nanggroe Medika*, 5(2):43–48.

Imaroh, I. and Mesra, R. (2024). Peranan Dinas kesehatan dalam Peningkatan Kesehatan Mental Ibu Hamil untuk Mencegah Stunting pada Anak di Indonesia. *ETIC (Education and Social Science Journal)*, 1(5):411–423.

Immanuel, J. P. and Handhayani, T. (2025). Prediksi Harga Pangan Jayapura Menggunakan ELM, LSTM, LightGBM, dan GB. *Jurnal Ilmu Komputer dan Sistem Informasi*, 13(2).

Jarmakovica, A. (2025). Machine Learning-Based Strategies for Improving Healthcare Data Quality: An Evaluation of Accuracy, Completeness, and Reusability. *Frontiers in Artificial Intelligence*, 8:1621514.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Ji, Q., Zhang, S., Duan, Q., Gong, Y., Li, Y., Xie, X., Bai, J., Huang, C., and Zhao, X. (2022). Short-Term and Medium-Term Power Demand Forecasting with Multiple Factors Based on Multi-Model Fusion. *Mathematics*, 10(12):2148.

Joel, L. O., Doorsamy, W., and Paul, B. S. (2025). A Comparative Study of Imputation Techniques for Missing Values in Healthcare Diagnostic Datasets. *International Journal of Data Science and Analytics*, 20(7):6357–6373.

Juniarti, U., Nurhalizah, R. S., Islahati, A. K., Anandita, C. A.-M. R., and Setiawan, R. A. (2024). Perancangan Sistem Informasi Eksekutif untuk Meningkatkan Performa Kinerja Pemeriksaan Risiko Kesehatan Ibu Hamil pada Klinik Kebidanan dan Penyakit Kandungan. *Jurnal Ilmu Komputer dan Teknologi*, 5(1):26–30.

Kanber, B. M., Smadi, A. A., Noaman, N. F., Liu, B., Gou, S., and Alsmadi, M. K. (2024). LightGBM: A Leading Force in Breast Cancer Diagnosis Through Machine Learning and Image Processing. *IEEE Access*, 12:39811–39832.

Kasanova, E., Wijayanti, E., Simanjuntak, R., Ainiyah, N. H., Adriani, Patty, F. I. T., Apryanti, Y. P., and Amini, A. (2025). *Asuhan Kehamilan: Normal dan Komplikasi*. PT Bukuloka Literasi Bangsa.

Keerin, P. and Boongoen, T. (2021). Improved KNN Imputation for Missing Values in Gene Expression Data. *Computers, Materials and Continua*, 70(2):4009–4025.

Kemenkes (2024). *Laporan Kinerja Kementrian Kesehatan 2024*. Kementerian Kesehatan Republik Indonesia.

Kvalvik, L. G., Skjærven, R., Sulo, G., Singh, A., Harmon, Q. E., and Wilcox, A. J. (2024). Pregnancy History at 40 Years of Age as A Marker of Cardiovascular Risk. *Journal of the American Heart Association*, 13(5):30560.

Laia, M. (2023). Analisis Kinerja Algoritma K-Nearest Neighbor Imputation (KNNI) Untuk Missing Value Pada Klasifikasi Data Mining. *Journal of Informatics, Electrical and Electronics Engineering*, 2(3):92–98.

Lestari, F. D. and Hanum, S. M. F. (2025). Continuity of Midwifery Care for Pregnancy, Childbirth, Pstpartum, and Newborns: Kelanjutan pelayanan kebidanan untuk kehamilan, persalinan, pasca persalinan, dan bayi baru lahir. *Indonesian Journal on Health Science and Medicine*, 2(2):10–21070.

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Li, X., Lu, Y., Fu, X., and Qi, Y. (2021). Building the Internet of Things platform for smart maternal healthcare services with wearable devices and cloud computing. *Future Generation Computer Systems*, 118:282–296.

Liang, D., Jin, X., Yuan, Y., and Zou, R. (2023). Performance Analysis of Machine Learning Methods. *Journal of Physics: Conference Series*, 2428(1):012039.

Listiarini, U. D., Maryanti, E., and Sofiah, N. S. (2022). Status Gizi Ibu Hamil Berhubungan dengan Bayi Berat Badan Lahir Rendah (BBLR). *Jurnal Kesehatan Mahardika*, 9(2):10–15.

Lokker, C., Abdelkader, W., Bagheri, E., Parrish, R., Cotoi, C., Navarro, T., Germini, F., Linkins, L. A., Haynes, R. B., Chu, L., Afzal, M., and Iorio, A. (2024). Boosting Efficiency in a Clinical Literature Surveillance System with LightGBM. *PLOS Digital Health*, 3(9):e0000299.

Magee, L. A., Bone, J., and et al. (2021). Pregnancy Outcomes and Blood Pressure Visit-to-Visit Variability and Level in Three Less-Developed Countries. *Hypertension (Dallas, Tex. : 1979)*, 77(5):1714.

Mahesh, T., Geman, O., Margala, M., Guduri, M., et al. (2023). The Stratified K-Folds Cross-Validation and Class-Balancing Methods With High-Performance Ensemble Classifiers for Breast Cancer Classification. *Healthcare Analytics*, 4:100247.

Misu, F., Gasbarro, D., and Alam, K. (2025). Inequality in Utilization of Maternal Healthcare Services in Low- and Middle-Income Countries: A Scoping Review of the Literature. *Maternal and Child Health Journal*, 29(6):741–766.

Modak, A., Ronghe, V., Gomase, K. P., Mahakalkar, M. G., and Taksande, V. (2023). A Comprehensive Review of Motherhood and Mental Health: Postpartum Mood Disorders in Focus. *Cureus*, 15(9):e46209.

Mojumdar, M. U., Assaduzzaman, M., Sarker, D., Shifa, H. A., Sajeeb, M. A. H., Bari, S., Chakraborty, N. R., and Alam, M. J. (2024). Maternal Health Risk Assessment Dataset.

Mondal, S., Nag, A., Barman, A. K., and Karmakar, M. (2023). Machine Learning-Based Maternal Health Risk Prediction Model for IoMT Framework. *International Journal of Experimental Research and Review*, 32(1):145–159.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Muludi, K., Setianingsih, R., Sholehurrohman, R., and Junaidi, A. (2024). Exploiting Nearest Neighbor Data and Fuzzy Membership Function to Address Missing Values in Classification. *PeerJ Computer Science*, 10:e1968.

Nanda, K. R. (2025). Maternal Age and Risk of Pregnancy Complications: A Qualitative Study. *Advances in Healthcare Research*, 3(2):132–147.

Nasir, F. and Irmawati (2025). Optimalisasi Pelayanan Kesehatan Ibu dan Anak untuk Menurunkan Angka Kematian Ibu dan Bayi di Indonesia. *Jurnal Kolaboratif Sains*, 8(7):4899–4903.

Nguyen, K. T., Tran, T. N., and Nguyen, H. T. (2024). Research on the Influence of Hyperparameters on the LightGBM Model in Load Forecasting. *Engineering, Technology and Applied Science Research*, 14(5):17005–17010.

Noviandy, T. R., Nainggolan, S. I., Raihan, R., Firmansyah, I., and Idroes, R. (2023). Maternal Health Risk Detection Using Light Gradient Boosting Machine Approach. *Infolitika Journal of Data Science*, 1(2):48–55.

Nugraha, I. G. W. and Siallagan, R. M. (2025). Preeclampsia Risk in Bali, Indonesia: The Role of Maternal Body Mass Index (BMI) and Gestational Weight Gain (GWG). *Bioscientia Medicina : Journal of Biomedicine and Translational Research*, 9(1):5923–5935.

Nyongesa, P., Ekhaguere, O. A., and et al. (2023). Maternal Age Extremes and Adverse Pregnancy Outcomes in Low-Resourced Settings. *Frontiers in Global Women's Health*, 4:1201037.

Pajala, G. W. and Hidayat, F. (2024). Hubungan Jarak dan Usia Kehamilan dengan Komplikasi Persalinan di RSUD Jhon Piet Wanane Sorong. *Jurnal Kedokteran dan Kesehatan*, 20(2):165–175.

Palanivinayagam, A. and Damaševičius, R. (2023). Effective Handling of Missing Values in Datasets for Classification Using Machine Learning Methods. *Information*, 14(2):92.

Park, J. and Hwang, E. (2021). A Two-Stage Multistep-Ahead Electricity Load Forecasting Scheme Based on LightGBM and Attention-BilSTM. *Sensors*, 21(22):7697.

Perejón, D., Bardalet, A., Gascó, I., Siscart, J., Serna, M. C., and Orós, M. (2024). Hypertension Subtypes and Adverse Maternal and Perinatal Outcomes - A

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Retrospective Population-Based Cohort Study. *BMC Pregnancy and Childbirth*, 24(1):1–9.

Pi, X., Wang, J., Chu, L., Zhang, G., and Zhang, W. (2025). Prediction of High-Risk Pregnancy Based on Machine Learning Algorithms. *Scientific Reports*, 15(1):15561.

Pranandita, A. A. G. A. and Widiartha, I. M. (2025). Optimasi Metode Support Vector Machine (SVM) Mengunakan Particle Swarm Optimization pada Permasalahan Klasifikasi Diabetes. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(4):879–888.

Pringandana, C. G. L. and Kusnawi, K. (2025). A Comparative Analysis of Hyperparameter-Tuned XGBoost and LightGBM for Multiclass Rainfall Classification in Jakarta. *Jurnal Teknik Informatika (JUTIF)*, 6(4):2467–2483.

Prusty, S., Patnaik, S., and Dash, S. K. (2022). SKCV: Stratified K-Fold Cross-Validation on ML Classifiers for Predicting Cervical Cancer. *Frontiers in Nanotechnology*, 4.

Purnajaya, A. R., Darmawan, J., Yamin, V., and Charles, C. (2024). Pendekatan dengan Oversampling dan Undersampling untuk Meningkatkan Akurasi Diagnostik Kanker Tiroid. *Jurnal Pustaka Data (Pusat Akses Kajian Database, Analisa Teknologi, dan Arsitektur Komputer)*, 4(1):33–39.

Putri, R. H. (2025). *Persalinan Berisiko: Penanganan Tepat Kehidupan Lebih Sehat*. Penerbit NEM.

Rahim, M. R., Sanyoto, D. D., Istiqamah, E., Adhani, R., and Husaini, H. (2024). Analisis Faktor yang Berhubungan dengan Angka Kematian Ibu. *Jurnal Surya Medika*, 10(3):365–373.

Rahmi, R. M., de Oliveira, P., Selistre, L., Rezende, P. C., Pezzella, G. N., dos Santos, P. A., de Oliveira Pereira Vergani, D., Madi, S. R. C., and Madi, J. M. (2021). Continuous Glucose Monitoring in Obese Pregnant Women With No Hyperglycemia on Glucose Tolerance Test. *PLOS One*, 16(6):e0253047.

Ramadan, A., Nuswantoro, A., Triana, L., and Ihsan, B. M. (2024). Gambaran Kadar Glukosa Darah Sewaktu Pada Ibu Hamil di Puskesmas Tambelan Sampit Kota Pontianak. *Jurnal Ners*, 8(1):575–578.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Ramadanti, E., Dinathi, D. A., Sri, C., Aditya, K., and Chandranegara, R. (2024). Diabetes Disease Detection Classification Using Light Gradient Boosting (LightGBM) With Hyperparameter Tuning. *Jurnal dan Penelitian Teknik Informatika*, 8(2):956–963.

Raza, A., Siddiqui, H. U. R., Munir, K., Almutairi, M., Rustam, F., and Ashraf, I. (2022). Ensemble Learning-Based Feature Engineering to Analyze Maternal Health During Pregnancy and Health Risk Prediction. *PLOS One*, 17(11).

Reitzle, L., Heidemann, C., Baumert, J., Kaltheuner, M., Adamczewski, H., Icks, A., and Scheidt-Nave, C. (2023). Pregnancy Complications in Women With Pregestational and Gestational Diabetes Mellitus. *Deutsches Arzteblatt international*, 120(6):81–86.

Rizky, P. S., Hirzi, R. H., and Hidayaturrohman, U. (2022). Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(2):228–236.

Royana, A. N., Via, Y. V., and Putra, C. A. (2025). Evaluasi Kinerja LightGBM dan CatBoost Untuk Prediksi Churn Berdasarkan Dataset Pelanggan Layanan Streaming Musik. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(4):6977–6982.

Rufo, D. D., Debelee, T. G., Ibenthal, A., and Negera, W. G. (2021). Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM). *Diagnostics*, 11(9):1714.

Sarhaddi, F., Azimi, I., Axelin, A., Niela-Vilen, H., Liljeberg, P., and Rahmani, A. M. (2022). Trends in Heart Rate and Heart Rate Variability During Pregnancy and the 3-Month Postpartum Period: Continuous Monitoring in a Free-living Context. *JMIR mHealth and uHealth*, 10(6):e33458.

Seid, A. K., Morken, N. H., Klungsøyr, K., Kvalvik, L. G., Sorbye, L. M., Vatten, L. J., and Skjærven, R. (2023). Pregnancy Complications in Last Pregnancy and Mothers' Long-Term Cardiovascular Mortality: Does The Relation Differ From That of Complications in First Pregnancy? A Population-Based Study. *BMC Women's Health*, 23(1):1–8.

Selna, N., Artati, and Aryandi, R. (2024). Deteksi Kadar Gula Darah pada Ibu Hamil Trimester 2. *Jurnal Riset Sains dan Kesehatan Indonesia*, 1(4):150–156.

Sholikah, S. M., Nurwulansari, F., and Aini, E. N. (2024). *Faktor yang Berhubungan dengan Kejadian Kehamilan Risiko Tinggi*. Penerbit NEM.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Sinha, B. B., Ahsan, M., and Dhanalakshmi, R. (2023). LightGBM Empowered by Whale Optimization for Thyroid Disease Detection. *International Journal of Information Technology*, 15(4):2053–2062.

Soewarno, M. R. P., Ardianto, R., Suryani, R., Al-Hakim, R. R., and Wahyudi, C. I. (2025). Analisis Perbandingan Algoritma KNN dan SVM untuk Prediksi Risiko Kesehatan Ibu Hamil. *Jurnal Kolaborasi Riset Sarjana*, 2(3).

Sudrajat, W. and Cholid, I. (2023). K-nearest neighbor (k-nn) untuk penanganan missing value pada data umkm. *Jurnal Rekayasa Sistem Informasi dan Teknologi*, 1(2):54–63.

Syafarina, G. A., Amin, M., and Zaenuddin, Z. (2026). Perbandingan Tree-Based Ensemble (Random Forest, XGBoost, LightGBM) dan Deep Learning untuk Prediksi Klik Iklan Online. *Technologia: Jurnal Ilmiah*, 17(1):257–265.

Szeghalmy, S. and Fazekas, A. (2023). A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors*, 23.

Thai, H. T. (2022). Machine learning for structural engineering: A state-of-the-art review. *Structures*, 38:448–491.

Tveiten, L., Diep, L. M., Halvorsen, T., and Markestad, T. (2024). Normal Range and Risk Factors for Deviating Body Temperatures During the First 24 Hours in Term-Born Infants Under Standardised Care: An Observational Study. *BMJ Paediatrics Open*, 8(1):e002596.

Uddin, M. F., Mim, S. A., Haque, M. A., Tariquajjaman, M., Jabeen, I., Latif, M. B., Sumi, A. U. H., Aarshe, F., Hossain, M. H., Nabi, S., Hossain, M. A., Khan, M. N. M., Saha, A., Ahmed, T., Alam, N. H., and Chisti, M. J. (2024). Sociodemographic and Maternal Health-Related Factors Associated with Mortality Among Children Under Three in Bangladesh: An Analysis of Data from Bangladesh Demographic and Health Survey 2017-18. *BMC Public Health*, 24(1):3324.

UNICEF (2025a). Bangladesh Faces Critical Challenges in Maternal and Newborn Care, with over 100,000 Child Deaths and 63,000 Stillbirths Annually.

UNICEF (2025b). Neonatal mortality.

UNICEF (2025c). Stillbirths and Stillbirth Rates.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

UNICEF (2025d). Uncertainty in AID Commitments Threatens Bangladesh's Progress in Maternal Health.

Wahyu Utami, E., Lumbantobing, P., Permatasari, A. S., , Wati, D. S., and Untari, S. (2025). *Buku Ajar Pemeriksaan Fisik Ibu dan Bayi*. Optimal Untuk Negeri.

Wang, Y. and Zhu, R. (2024). Breast Cancer Classification Based on KNNI Imputation of Missing Data. In *Artificial Intelligence Technologies and Applications*, pages 684–692. IOS Press.

Wenny, W. (2024). Normalisasi Data Kependudukan Dengan Model Min Max Dan Algoritma K-Means Untuk Pengelompokkan Tingkat Ekonomi Masyarakat. *Bulletin of Information System Research*, 2(2):53–63.

WHO (2024). Newborn Mortality.

WHO (2025). Maternal Mortality.

Widodo, S., Brawijaya, H., and Samudi, S. (2022). Stratified K-Fold Cross Validation Optimization on Machine Learning for Prediction. *Sinkron*, 7:2407–2414.

Wijayanto, D. and Hartato, B. P. (2024). Analisis Perbandingan Performa Algoritma XGBoost dan LightGBM pada Klasifikasi Kanker Payudara. *The Indonesian Journal of Computer Science*, 13(2).

Wiratma, D. Y., Kurniawan, B., and Syahputra, M. B. (2024). Analisa Glukosa Darah Ibu Hamil Trimester II di Rumah Sakit Estomihi Medan. *Jurnal Analis Laboratorium Medik*, 9(1):26–30.

Wyckoff, J. A., Lapolla, A., and et al. (2025). Preexisting Diabetes and Pregnancy: An Endocrine Society and European Society of Endocrinology Joint Clinical Practice Guideline. *The Journal of Clinical Endocrinology & Metabolism*, 110(9):2405–2452.

Yang, H., Chen, Z., Yang, H., and Tian, M. (2023). Predicting Coronary Heart Disease Using An Improved LightGBM Uodel: Performance Analysis and Comparison. *IEEE access*, 11:23366–23380.

Ye, W., Luo, C., Huang, J., Li, C., Liu, Z., and Liu, F. (2022). Gestational Diabetes Mellitus and Adverse Pregnancy Outcomes: Systematic Review and Meta-Analysis. *BMJ*, 377.

UIN SUNAN AMPEL
S U R A B A Y A
http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/

Ye, X., Baker, P. N., and Tong, C. (2024). The Updated Understanding of Advanced Maternal Age. *Fundamental Research*, 4(6):1719–1728.

Yunitasari, E. and Yuliana, I. (2025). Perkembangan Terkini dalam Kesehatan Maternal di Nusantara: Tinjauan Multidisiplin. *Jurnal Kesehatan dan Kebidanan Nusantara*, 3(1):1–6.

Zhang, S., Yuan, Y., Yao, Z., Yang, J., Wang, X., and Tian, J. (2022). Coronary Artery Disease Detection Model Based on Class Balancing Methods and LightGBM Algorithm. *Electronics (Switzerland)*, 11(9):1495.

Zhang, Y., Lu, M., Yi, Y., Xia, L., Zhang, R., Li, C., and Liu, P. (2024). Influence of Maternal Body Mass Index on Pregnancy Complications and Outcomes: A Systematic Review and Meta-Analysis. *Frontiers in Endocrinology*, 15:1280692.

Zhao, D., Liu, D., Shi, W., Shan, L., Yue, W., Qu, P., Yin, C., and Mi, Y. (2023). Association between Maternal Blood Glucose Levels during Pregnancy and Birth Outcomes: A Birth Cohort Study. *International Journal of Environmental Research and Public Health*, 20(3):2102.

UIN SUNAN AMPEL
S U R A B A Y A

http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/ http://digilib.uinsa.ac.id/