

**IMPLEMENTASI LIGHT GRADIENT BOOSTING MACHINE
UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS
DENGAN DETEKSI OUTLIER ROBUST RANDOM CUT FOREST**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
SITI NURUS SHOFIYAH
09020222045

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2026

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : SITI NURUS SHOFTYAH
NIM : 09020222045
Program Studi : Matematika
Angkatan : 2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "IMPLEMENTASI LIGHT GRADIENT BOOSTING MACHINE UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS DENGAN DETEKSI OUTLIER ROBUST RANDOM CUT FOREST". 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, 19 Juni 2026

Yang menyatakan,



SITI NURUS SHOFTYAH
NIM. 09020222045

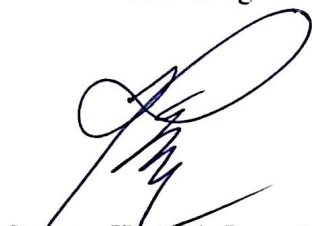
LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : SITI NURUS SHOFTYAH
NIM : 09020222045
Judul skripsi : IMPLEMENTASI LIGHT GRADIENT BOOSTING
MACHINE UNTUK KLASIFIKASI PENYAKIT
DIABETES MELITUS DENGAN DETEKSI
OUTLIER ROBUST RANDOM CUT FOREST

telah diperiksa dan disetujui untuk diujikan.

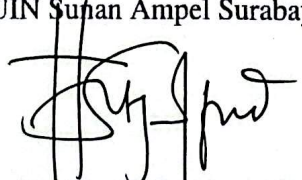
Pembimbing I


Putroue Keumala Intan, M.Si
NIP. 198805282018012001

Pembimbing II


Dr. Abdulloh Hamid, M.Pd
NIP. 198508282014031003

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya


Dr. Yuniar Farida, M.T
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

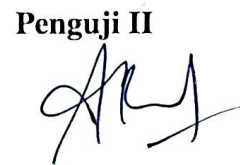
Nama : SITI NURUS SHOFIYAH
NIM : 09020222045
Judul skripsi : IMPLEMENTASI LIGHT GRADIENT BOOSTING
MACHINE UNTUK KLASIFIKASI PENYAKIT
DIABETES MELITUS DENGAN DETEKSI
OUTLIER ROBUST RANDOM CUT FOREST

Telah dipertahankan di depan Tim Penguji
pada tanggal 19 Juni 2026.

Mengesahkan,
Tim Penguji

Penguji I

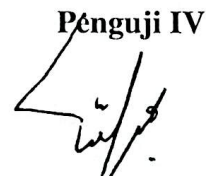

Dr. Yuniar Farida, M.T
NIP. 197905272014032002

Penguji II


Aris Fanani, M.Kom
NIP. 198701272014031002

Penguji III

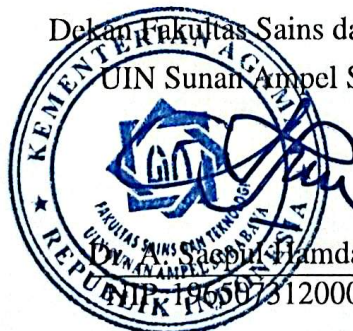

Putrouz Keumala Intan, M.Si
NIP. 198805282018012001

Penguji IV


Dr. Abdullah Hamid, M.Pd
NIP. 198508282014031003

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya


Sagduh Hamdani, M.Pd
NIP. 196007312000031002

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

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

Nama : SITI HURUS SHOFIYAH
NIM : 09020222045
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / MATEMATIKA
E-mail address : shofiaja763@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Skripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

IMPLEMENTASI LIGHT GRADIENT BOOSTING MACHINE
UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS DENGAN
DETEKSI OUTLIER ROBUST RANDOM CUT FOREST

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif 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, 26 JUNI 2021

Penulis


(Siti Hurus Shofiayah)
nama terang dan tanda tangan

ABSTRAK

IMPLEMENTASI LIGHT GRADIENT BOOSTING MACHINE UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS DENGAN DETEKSI OUTLIER ROBUST RANDOM CUT FOREST

Diabetes Melitus merupakan salah satu penyakit tidak menular dengan kondisi kadar gula darah tinggi. Tingginya kasus Diabetes Melitus menjadikan Indonesia sebagai negara dengan jumlah penyandang diabetes melitus terbanyak kelima di dunia. Deteksi dini sangat dianjurkan dalam pencegahan komplikasi dan peningkatan jumlah penderita Diabetes Melitus. Pencegahan ini dapat diatasi dengan machine learning. Penelitian ini bertujuan untuk mengevaluasi efektivitas model LightGBM dalam mengklasifikasikan penyakit Diabetes Melitus. Metode yang digunakan pada penelitian ini yaitu metode klasifikasi LightGBM. Pembagian data menggunakan K-Fold Cross Validation dengan K=5. Adapun optimasi deteksi *outlier* menggunakan Robust Random Cut Forest dengan penanganan *outlier* 5%. Uji coba parameter LightGBM seperti *learning rate* 0.075, 0.05, dan 0.1 *num leaves* 2, 4, 6, 8, 10, 12, 14, 16 dan *min data in leaf* 10, 20, 40, 80 diterapkan pada penelitian ini untuk mendapatkan hasil yang optimal.

Hasil penelitian menunjukkan bahwa model LightGBM dengan parameter terbaik yang diperoleh melalui uji coba parameter mampu memberikan kinerja klasifikasi yang baik pada data diabetes melitus. Berdasarkan pengujian menggunakan K-Fold Cross Validation, model menghasilkan nilai akurasi sebesar 85.36%, sensitivitas sebesar 79.03% dan spesifitas sebesar 85.77%. Penerapan metode SMOTE berhasil menyeimbangkan distribusi kelas sehingga model dapat mengenali data pasien diabetes melitus dengan lebih optimal dan mengurangi bias terhadap kelas mayoritas. Dengan demikian, metode LightGBM terbukti efektif dalam mengklasifikasikan diabetes melitus dan berpotensi digunakan sebagai pendukung deteksi dini untuk membantu pengambilan keputusan di bidang kesehatan.

Kata kunci: Diabetes Melitus, Klasifikasi, LightGBM, RRCF

ABSTRACT

IMPLEMENTATION OF A LIGHT GRADIENT BOOSTING MACHINE FOR CLASSIFYING DIABETES MELLITUS USING A ROBUST RANDOM CUT FOREST FOR OUTLIER DETECTION

Diabetes mellitus is a noncommunicable disease characterized by high blood sugar levels. The high incidence of diabetes mellitus makes Indonesia the country with the fifth-highest number of people living with diabetes mellitus in the world. Early detection is highly recommended to prevent complications and reduce the number of people with diabetes mellitus. This prevention can be addressed using machine learning. This study aims to evaluate the effectiveness of the LightGBM model in classifying diabetes mellitus. The method used in this study is the LightGBM classification method. Data splitting was performed using K-Fold Cross Validation with K=5. Outlier detection was optimized using a Robust Random Cut Forest with a 5% outlier threshold. LightGBM parameter tests included learning rates of 0.075, 0.05, and 0.1, and the number of leaves 2, 4, 6, 8, 10, 12, 14, 16, and *min data in leaf* 10, 20, 40, 80 were applied in this study to obtain optimal results.

The results of the study show that the LightGBM model with the optimal parameters—determined through parameter tuning—delivers good classification performance on diabetes mellitus data. Based on testing using K-fold cross-validation, the model achieved an accuracy of 85.36%, a sensitivity of 79.03%, and a specificity of 85.77%. The application of the SMOTE method successfully balanced the class distribution, allowing the model to recognize diabetes mellitus patient data more effectively and reduce bias toward the majority class. Thus, the LightGBM method has proven effective in classifying diabetes mellitus and has the potential to be used as a tool for early detection to aid decision-making in the healthcare field.

Keywords: Diabetes Mellitus, Classification, LightGBM, RRCF

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	xi
DAFTAR GAMBAR	xiii
ABSTRAK	xiv
ABSTRACT	xv
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	8
1.3. Tujuan Penelitian	8
1.4. Manfaat Penelitian	8
1.5. Batasan Masalah	9
1.6. Sistematika Penulisan	9
II TINJAUAN PUSTAKA	11
2.1. Penyakit Diabetes Melitus	11
2.1.1. Faktor Penyebab Diabetes Melitus	11
2.2. Robust Random Cut Forest	15
2.3. <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	17
2.4. K-fold Cross Validation	18
2.5. LightGBM	19
2.6. Confussion Matrix	22
2.7. Integrasi Keilmuan	23
2.7.1. Diabetes dalam Islam	23
2.7.2. Klasifikasi dalam Islam	26
III METODE PENELITIAN	29

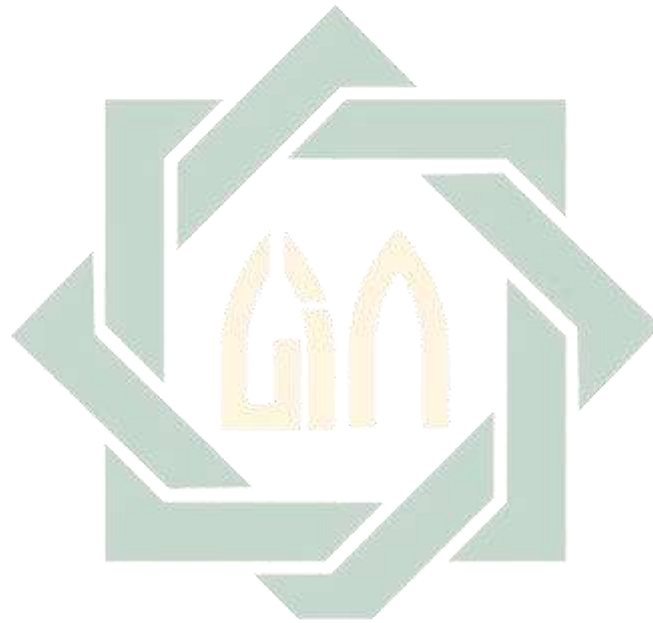
3.1. Jenis Penelitian	29
3.2. Sumber Data	29
3.3. Variabel Penelitian	30
3.4. Tahapan Penelitian	31
3.5. Skenario Uji Coba	34
IV HASIL DAN PEMBAHASAN	36
4.1. Pengolahan Data	36
4.2. Deteksi <i>Outlier</i> Menggunakan RRCF	36
4.3. Penanganan Data Tidak Seimbang	43
4.4. Pembagian Data Menggunakan K-Fold Cross Validation	45
4.5. LightGBM	45
4.5.1. Tahap Training	45
4.5.2. Tahap Testing	57
4.6. Uji Coba Parameter	60
4.7. Evaluasi Model	64
4.8. Diskusi Penelitian	66
4.9. Integrasi Keilmuan	67
V PENUTUP	70
5.1. Kesimpulan	70
5.2. Saran	71
DAFTAR PUSTAKA	71
A	80

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

2.1 Confusion Matrix	22
3.1 Sampel Data	29
3.2 Deskripsi Variabel Penelitian	30
3.3 Tuning Parameter	35
4.1 Contoh Data	36
4.2 Perhitungan Rentang Tiap Variabel	38
4.3 Variabel untuk Memotong Ruang Data	39
4.4 Nilai Displacement	41
4.5 Hasil Perhitungan CoDisp	42
4.6 Hasil Data Setelah Dilakukan Penanganan Outlier 5%	42
4.7 Pembagian Data menggunakan <i>k-Fold Cross Validation</i>	45
4.8 Information Gain Node 1	50
4.9 Information Gain Node 1	52
4.10 Nilai Leaf Seluruh Pohon	57
4.11 Perbandingan Data Aktual dan Data Prediksi	60
4.12 Hasil Pengujian Parameter LightGBM pada Learning Rate 0.075	61
4.13 Hasil Pengujian Parameter LightGBM pada Learning Rate 0.05	62
4.14 Hasil Pengujian Parameter LightGBM pada Learning Rate 0.1	62
4.15 Perbandingan Parameter Terbaik pada Setiap Learning Rate	63
4.16 Hasil Confusion Matrix Model	64
4.17 Nilai Evaluasi Tiap Fold	65
4.18 Perbandingan Model Tanpa dan Menggunakan RRCF serta SMOTE	66

1.1	Hasil Pengujian LightGBM dengan Num Leaves = 2		80
1.2	Hasil Pengujian LightGBM dengan Num Leaves = 4		82
1.3	Hasil Pengujian LightGBM dengan Num Leaves = 6		84
1.4	Hasil Pengujian LightGBM dengan Num Leaves = 8		87
1.5	Hasil Pengujian LightGBM dengan Num Leaves = 10		89
1.6	Hasil Pengujian LightGBM dengan Num Leaves = 12		91
1.7	Hasil Pengujian LightGBM dengan Num Leaves = 14		93
1.8	Hasil Pengujian LightGBM dengan Num Leaves = 16		96



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

2.1	Pembagian Data Menggunakan K-Fold	18
2.2	Pohon Life-Wise LightGBM	19
3.1	Diagram Alir Penelitian	31
3.2	Diagram Alir RRCF	32
3.3	Diagram Alir LightGBM	34
4.1	Persebaran <i>Outlier</i>	37
4.2	Distribusi Data Setelah Data Outlier Dihapus	43
4.3	Distribusi Data Setelah Data Outlier Dihapus	44
4.4	Node 1	48
4.5	Node 1	50
4.6	Node 2 (Sisi Kiri Kedalaman 1)	53
4.7	Visualisasi Pohon Pertama	54

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Ahmed, W. (2024). Additive interaction of family medical history of diabetes with hypertension on the diagnosis of diabetes among older adults in india: longitudinal ageing study in india. *BMC Public Health*, 24(1):999.
- Al Tawil, A., Almazaydeh, L., Alqudah, B., Abualkishik, A. Z., and Alwan, A. A. (2024). Predictive modeling for breast cancer based on machine learning algorithms and features selection methods. *International Journal of Electrical and Computer Engineering*, 14(2):1937–1947.
- Alpaydin, E. (2014). *Introduction to Machine Learning*. MIT Press, Cambridge, MA, 3rd edition.
- Ayu, I., Wibudi, A., Panusunan, R., Ratna, M., Pande, I. M., Amir, H., L, D., Aron, M., Yensuari, and Tedjo, T. (2021). Pedoman pemantauan glukosa darah mandiri. Diakses tanggal 11 Februari 2026.
- Bahtiar, Y., Kartilah, T., and Cahyati, P. (2020). Edukasi masyarakat sehat sejahtera (emass): Jurnal pengabdian kepada masyarakat. *Jurnal Pengabdian Masyarakat*.
- Bakhtiari, S., Timbrell, N. E., and D’Almeida, S. M. (2025). Abnormally low hba1c caused by hemolytic anemia, a case report and literature review. *British Journal of Biomedical Science*, 81:13898.
- Barasa, N. R. T., Silitonga, H. T. H., Wijaya, E. D., Husna, H. N., Prayogo, B. R., Salvita, I. X., Rahmawati, L., Wartiningsih, M., Juliasih, N. D., and Isnaini, I. (2023). Pemeriksaan status gizi dan edukasi pencegahan diabetes melitus terhadap lansia di wilayah putukrejo, desa peniwen, kabupaten malang. *Jurnal Abdimas Kesehatan (JAK)*, 5(3):558–564.
- Bartos, M. D., Mullapudi, A., and Troutman, S. C. (2019). rrcf: Implementation of the robust random cut forest algorithm for anomaly detection on streams. *Journal of Open Source Software*, 4(35):1336.

- Cassim, N., Coetzee, L. M., and Glencross, D. K. (2020). Categorising specimen referral delays for cd4 testing: How inter-laboratory distances and travel times impact turn-around time across a national laboratory service in south africa. *African journal of laboratory medicine*, 9(1):1–7.
- Chen, H.-H. (2024). Understanding gradient boosting classifier: Training, prediction, and the role of gamma j. *arXiv preprint arXiv:2410.05623*.
- Choudhury, A. A. and Rajeswari, V. D. (2021). Gestational diabetes mellitus-a metabolic and reproductive disorder. *Biomedicine & Pharmacotherapy*, 143:112183.
- Cleveland Clinic (2025). Tes gula darah puasa. Diakses tanggal 11 Februari 2026.
- Collins, S. P., Storrow, A., Liu, D., Jenkins, C. A., Miller, K. F., Kampe, C., and Butler, J. (2021). *Buku Saku Diabetes Melitus untuk Awam*. Nama Penerbit, Kota Penerbit.
- Committee, A. D. A. P. P. and Committee:, A. D. A. P. P. (2022). 7. diabetes technology: standards of medical care in diabetes—2022. *Diabetes Care*, 45(Supplement.1):S97–S112.
- De la Cruz-Concepción, B., Flores-Cortez, Y. A., Barragán-Bonilla, M. I., Mendoza-Bello, J. M., and Espinoza-Rojó, M. (2023). Insulin: A connection between pancreatic β cells and the hypothalamus. *World journal of diabetes*, 14(2):76.
- 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):e0330377.
- Elisabeth, E., Suhartina, S., and Wienaldi, W. (2024). Hubungan gaya hidup dengan kejadian diabetes mellitus tipe 2 di rumah sakit royal prima medan. *Journal of Health Research Science*, 4(02):222–231.
- Fadillah, I., Nurahmi, N., and Kurniawan, L. B. (2025). Correlation between serum visfatin level with glycemic control in type 2 diabetes mellitus subjects. *Indonesian Journal of Clinical Pathology and Medical Laboratory*, 31(3):299–303.

- Fauziyyah, M. H. and Utama, F. (2024). Literature review: Faktor-faktor yang berhubungan dengan kejadian diabetes mellitus di indonesia. *PREPOTIF: Jurnal Kesehatan Masyarakat*, 8(1):266–278.
- Han, Y., Wei, Z., and Huang, G. (2024). An imbalance data quality monitoring based on smote-xgboost supported by edge computing. *Scientific Reports*, 14(1):10151.
- He, C., Lu, S., Yu, H., Sun, Y., and Zhang, X. (2025). Global, regional, and national disease burden attributable to high systolic blood pressure in youth and young adults: 2021 global burden of disease study analysis. *BMC medicine*, 23(1):74.
- Hyun, U. and Sohn, J.-W. (2022). Autonomic control of energy balance and glucose homeostasis. *Experimental & molecular medicine*, 54(4):370–376.
- Indrahadi, D., Wardana, A., and Pierewan, A. C. (2021). The prevalence of diabetes mellitus and relationship with socioeconomic status in the indonesian population. *Jurnal Gizi Klinik Indonesia*, 17(3):103–112.
- International Diabetes Federation (2024). Diabetes in indonesia. Diakses tanggal 11 Februari 2026.
- Jiang, L., Xia, Z., Zhu, R., Gong, H., Wang, J., Li, J., and Wang, L. (2023). Diabetes risk prediction model based on community follow-up data using machine learning. *Preventive Medicine Reports*, 35:102358.
- Karolus, H., Butar Butar, S., Pangaribuan, S. M., Siregar, S. W., and Batubara, K. (2023). Hubungan aktivitas fisik dengan kadar glukosa darah pada pasien diabetes mellitus di ruang penyakit dalam rsud koja jakarta. *Jurnal Keperawatan Cikini*, 4(1):32–39.
- Kementerian Kesehatan Republik Indonesia (2025). Apa yang dimaksud dengan penyakit diabetes melitus. Diakses 2 Februari 2026.
- Kong, C.-C., Cheng, J.-D., and Wang, W. (2023). Neurotransmitters regulate β cells insulin secretion: a neglected factor. *World Journal of Clinical Cases*, 11(28):6670.

- Mader, J. K., Baumstark, A., Tüting, J., Sokol, G., Schuebel, R., Tong, Y., Roetschke, J., and Slingerland, R. J. (2024). Monitoring of the analytical performance of four different blood glucose monitoring systems: A post-market performance follow-up study. *Diabetes Therapy*, 15(12):2525–2535.
- Ma’u Luthfi, A. and Fauzi, F. (2024). Perbandingan klasifikasi random forest, support vector machines, dan lgbm pada klasifikasi kualitas udara di jakarta. *Jurnal Sistem Dan Teknologi Informasi Indonesia*, 9(2):99–108.
- 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 & Applied Science Research*, 14(5):17005–17010.
- Nugraha, F. S., Yulitasari, B. I., and Putri, T. I. Y. L. (2022). Stres dan kualitas tidur dengan kadar gula darah puasa pada lansia dm tipe ii. *Surya Medika: Jurnal Ilmiah Ilmu Keperawatan dan Ilmu Kesehatan Masyarakat*, 17(2):115–121.
- Olisah, C. C., Smith, L., and Smith, M. (2022). Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. *Computer Methods and Programs in Biomedicine*, 220:106773.
- Omotehinwa, T. O., Oyewola, D. O., and Dada, E. G. (2023). A light gradient-boosting machine algorithm with tree-structured parzen estimator for breast cancer diagnosis. *Healthcare Analytics*, 4:100218.
- Prama, T. T., Rahman, M. J., Zaman, M., Sarker, F., and Mamun, K. A. (2025). Diabd: A diabetes dataset for enhanced risk analysis and research in bangladesh. *Data in Brief*, page 111746.
- Puig-García, M., Caicedo-Montaña, C., Márquez-Figueroa, M., Chilet-Rosell, E., Montalvo-Villacis, G., Benazizi-Dahbi, I., Peralta, A., Torres-Castillo, A. L., and Parker, L. A. (2023). Prevalence and gender disparities of type 2 diabetes mellitus and obesity in esmeraldas, ecuador:

- a population-based survey in a hard-to-reach setting. *International journal for equity in health*, 22(1):124.
- Putri, M. G., Nugroho, H., and Adi, M. S. (2022). Hubungan indeks massa tubuh dan tingkat aktivitas fisik dengan kontrol glikemik diabetes melitus tipe 2. *Jurnal Epidemiologi Kesehatan Komunitas*, 7(1):341–350.
- Reubun, R. J. S., Tamtollahiru, C., and Yunita, M. (2022). Prevalensi retinopati diabetik pada penderita diabetes melitus di klinik utama provinsi maluku. *Care J. Ilm. Ilmu Kesehat*, 10(3):366–376.
- Rohmatulloh, V. R., Riskiyah, R., Pardjianto, B., and Kinasih, L. S. (2024). Hubungan usia dan jenis kelamin terhadap angka kejadian diabetes melitus tipe 2 berdasarkan 4 kriteria diagnosis di poliklinik penyakit dalam rsud karsa husada kota batu. *Prepotif: Jurnal Kesehatan Masyarakat*, 8(1):2528–2543.
- Rosares, V. E. and Boy, E. (2022). Pemeriksaan kadar gula darah untuk screening hiperglikemia dan hipoglikemia. *Jurnal Implementa Husada*, 3(2):65–71.
- Ruchika, R., Oktavia, E., Kartadinata, R., Wijaya, J., and Ratu, L. L. (2023). Prevalence of amputation events in diabetic ulcer patients at tarakan district hospital. *Jurnal MedScientiae*, 2(1).
- Sakr, H. F., Sirasanagandla, S. R., Das, S., Bima, A. I., and Elsamanoudy, A. Z. (2023). Insulin resistance and hypertension: mechanisms involved and modifying factors for effective glucose control. *Biomedicines*, 11(8):2271.
- Sami, A., Javed, A., Ozsahin, D. U., Ozsahin, I., Muhammad, K., and Waheed, Y. (2025). Genetics of diabetes and its complications: a comprehensive review. *Diabetology & Metabolic Syndrome*, 17(1):185.
- Saputra, S. I., Berawi, K. N., Susianti, S., and Hadibrata, E. (2023). Hubungan diabetes melitus dengan kejadian gagal ginjal kronik. *Medical Profession Journal of Lampung*, 13(5):787–791.

- SARI, M. L., Putra, F. A., and Widiyono, W. (2025). *Hubungan Gaya Hidup dengan Kadar Gula Darah pada Pasien Diabetes Mellitus Tipe II di Rumah Sakit Umum Diponegoro Dua Satu Klaten*. PhD thesis, Universitas Sahid Surakarta.
- Setiawan, I. and Dawis, A. M. (2023). Data science: Pendekatan dan langkah praktis dengan excel. *Journal of Innovation And Future Technology*, 5(1):11–22.
- Sihombing, P. R., Suryadiningrat, S., Sunarjo, D. A., and Yuda, Y. P. A. C. (2022). Identifikasi data outlier (pencilan) dan kenormalan data pada data univariat serta alternatif penyelesaiannya. *Jurnal Ekonomi Dan Statistik Indonesia*, 2(3):307–316.
- Simanjuntak, D. R. and Nugroho, R. A. (2025). Faktor risiko diabetes melitus dan prevalensi penyakit jantung koroner di rsu uki tahun 2021–2023. *Pro-Life*, 12(1):63–72.
- Sinoputro, D., Putri, F. R., Jomeiputri, G. H., Simbolon, L. A., Priscilla, M., Limbong, N. T., Ismail, N. A. B., Yolanda, R., Asteria, S., Hudyono, J., et al. (2015). Penggunaan insulin untuk pasien diabetes melitus dari generasi ke generasi. *Jurnal Kedokteran Meditek*.
- Son, J. and Accili, D. (2023). Reversing pancreatic β -cell dedifferentiation in the treatment of type 2 diabetes. *Experimental & Molecular Medicine*, 55(8):1652–1658.
- Sun, H., Saeedi, P., Karuranga, S., Pinkepank, M., Ogurtsova, K., Duncan, B. B., Stein, C., Basit, A., Chan, J. C., Mbanya, J. C., et al. (2022). Idf diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes research and clinical practice*, 183:109119.
- Supriyo, S., Ibrahim, M. Y., Santoso, P. N. D., and Dwiningsih, S. U. (2024). Diabetes self management education (dsme) upaya preventif untuk remaja terhindar diabetes melitus. *Jurnal Lintas Keperawatan*, 5(1):355–361.

- Suryadin, K. A. M. J. U. A. and Zahara, H. (2022). *Metodologi Penelitian Kuantitatif*. Yayasan Penerbit Muhammad Zaini, Aceh.
- Susanti, N., Syahpira, D., Aulia, S., and Syahmala, A. (2024). Hubungan usia pada kejadian diabetes mellitus tipe-2 dengan pendekatan stepwise. *Jurnal Kesehatan Tambusai*, 5(2):4283–4288.
- Tarigan, M., Tarigan, R., Imelda, F., Jongudomkarn, D., et al. (2024). Identifying diabetes risks among indonesians: A cross-sectional study in a community setting. *Belitung Nursing Journal*, 10(1):41.
- Vardhan, H. and Sztipanovits, J. (2022). Reduced robust random cut forest for out-of-distribution detection in machine learning models. *arXiv preprint arXiv:2206.09247*.
- Vashisth, S. and Goyal, A. (2024). Dynamic anomaly detection using robust random cut forests in resource-constrained iot environments. *Informatica*, 48(23):107–120.
- Vassou, C., Tsiampalis, T., Georgousopoulou, E. N., Chrysohoou, C., Yannakoulia, M., Pitsavos, C., Cropley, M., and Panagiotakos, D. B. (2024). Association between family history of diabetes, irrational beliefs, and health anxiety with 10-year risk of type 2 diabetes mellitus: the attica epidemiological study (2002–2012). *International Journal of Behavioral Medicine*, 31(4):516–526.
- Wang, Y. and Wang, T. (2020). Application of improved lightgbm model in blood glucose prediction. *Applied Sciences*, 10(9):3227.
- Widiastuti, W., Zulkarnaini, A., Mahatma, G., et al. (2024). Review artikel: Pengaruh pola asupan makanan terhadap resiko penyakit diabetes. *Journal of Public Health Science*, 1(2):108–125.
- Williams, S., Raheim, S. A., Khan, M. I., Rubab, U., Kanagala, P., Zhao, S. S., Marshall, A., Brown, E., and Alam, U. (2022). Cardiac autonomic neuropathy in type 1 and 2 diabetes: epidemiology, pathophysiology, and management. *Clinical therapeutics*, 44(10):1394–1416.
- World Health Organization (2024). Diabetes. Diakses tanggal 11 Februari 2026.

- Xiong, X., Wang, A., He, J., Wang, C., Liu, R., Sun, Z., Zhang, J., and Zhang, J. (2024). Application of lightgbm hybrid model based on tpe algorithm optimization in sleep apnea detection. *Frontiers in Neuroscience*, 18:1324933.
- Xu, W., Ning, L., and Luo, Y. (2020). Wind speed forecast based on post-processing of numerical weather predictions using a gradient boosting decision tree algorithm. *Atmosphere*, 11:738.
- Yang, N., Zhang, W., Ji, C., Ge, J., Zhang, X., Li, M., Wang, M., Zhang, T., He, J., and Zhu, H. (2023). Metabolic alteration of circulating steroid hormones in women with gestational diabetes mellitus and the related risk factors. *Frontiers in Endocrinology*, 14:1196935.
- Yen, Y.-F., Wang, C.-C., Chen, Y.-Y., Hsu, L.-F., Hung, K.-C., Chen, L.-J., Ku, P.-W., Chen, C.-C., and Lai, Y.-J. (2022). Leisure-time physical activity and mortality risk in type 2 diabetes: a nationwide cohort study. *Diabetes & Metabolism*, 48(6):101378.
- Yeom, S. and JUNG, J. H. (2022). Weighted random cut forest algorithm for anomaly detection. In *Weighted random cut forest algorithm for anomaly detection*.
- Yuliawuri, H., Handoko, S. M., Simarmata, A. K., and Wibowo, J. (2024). Correlation between body mass index (bmi) and blood glucose level among people in rmci jakarta: A cross-sectional study. *Science Midwifery*, 12(3):1381–1386.
- Yun, H., Park, J. w., and Kim, J. K. (2023). A comparative evaluation of hba1c measurement methods and their implications for diabetes management. *Diagnostics*, 13(22):3449.