

PREDIKSI DINI DIABETES

Berbasis
MACHINE LEARNING



Muhammad Rosdiana M.Kom

PREDIKSI DINI DIABETES BERBASIS MACHINE LEARNING

Studi Perbandingan dan Evaluasi Kinerja

Muhammad Rosdiana M.Kom



Penerbit KBM Indonesia

Adalah penerbit dengan misi memudahkan proses penerbitan buku-buku penulis di tanah air indonesia, serta menjadi media *sharing* proses penerbitan buku

PREDIKSI DINI DIABETES BERBASIS MACHINE LEARNING

Studi Perbandingan dan Evaluasi Kinerja

Copyright @2026 by Muhammad Rosdiana M.Kom

All rights reserved

PENERBIT KARYA BAKTI MAKMUR (KBM) INDONESIA

Anggota IKAPI (Ikatan Penerbit Indonesia) **NO. IKAPI 279/JTI/2021**

Depok, Sleman-Jogjakarta (Kantor) **081357517526 (Tlpn/WA)**

Penulis

Muhammad Rosdiana M.Kom

Desain Sampul

Eca Art

Tata Letak

Lailatul Nur Ula

Editor Naskah

Dr. Muhamad Husein Maruapey, Drs., M.Sc

ISBN

978-634-278-315-3

15,5 x 23 cm, vi + 127 halaman

Cetakan ke-1, April 2026

Isi buku diluar tanggungjawab penerbit

Hak cipta merek KBM Indonesia sudah terdaftar di

DJKI-Kemenkumham dan isi buku dilindungi undang-undang

Dilarang keras menerjemahkan, memfotokopi, atau

Memperbanyak sebagian atau seluruh isi buku ini

Tanpa seizin penerbit karena beresiko sengketa hukum

Website

<https://penerbitkbm.com>

www.penerbitbukumurah.com

Instagram

@penerbit.kbmindonesia

@penerbitbukujogja

Email

naskah@penerbitkbm.com

Distributor

<https://penerbitkbm.com/toko-buku/>

Youtube

Penerbit KBM Sastrabook

Sanksi Pelanggaran Pasal 113

Undang-Undang No. 28 Tahun 2014 Tentang Hak Cipta

- (i) Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp 100.000.000 (seratus juta rupiah).
- (ii) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).
- (iii) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp 1.000.000.000,00 (satu miliar rupiah).
- (iv) Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp 4.000.000.000,00 (empat miliar rupiah).

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh,

Segala puji syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat dan karunia-Nya sehingga buku berjudul “Prediksi Dini Diabetes Berbasis Machine Learning: Studi Perbandingan dan Evaluasi Kinerja” ini dapat diselesaikan dengan baik.

Diabetes merupakan salah satu penyakit tidak menular dengan prevalensi yang semakin meningkat, baik secara global maupun nasional. Dampak komplikasinya yang serius menjadikan deteksi dini sebagai langkah krusial dalam penanganan dan pencegahan kondisi lebih lanjut. Seiring berkembangnya teknologi, pemanfaatan *machine learning* membuka peluang baru dalam mempercepat proses identifikasi risiko diabetes secara lebih akurat dan efisien.

Berangkat dari kebutuhan tersebut, buku ini disusun guna memberikan wawasan mendalam mengenai integrasi ilmu kesehatan dan teknologi kecerdasan buatan. Materi disajikan secara sistematis, dimulai dari pemahaman dasar tentang diabetes, konsep dan metode machine learning, prapemrosesan data, seleksi fitur, perbandingan algoritma, hingga aplikasi praktis menggunakan dataset medis. Penulis berharap pembaca dapat memperoleh gambaran utuh mengenai bagaimana membangun sistem prediksi berbasis data yang dapat membantu tenaga medis maupun peneliti.

Buku ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada keluarga, rekan dosen, mahasiswa, serta semua pihak yang turut memberikan dukungan baik secara langsung maupun tidak langsung. Semoga Allah SWT membalas kebaikan mereka dengan pahala yang berlipat ganda.

Akhir kata, penulis menyadari bahwa buku ini masih memiliki keterbatasan. Kritik dan saran yang konstruktif sangat diharapkan demi penyempurnaan karya di masa mendatang. Semoga buku ini bermanfaat bagi pembaca dan dapat menjadi salah satu referensi dalam upaya mengembangkan solusi berbasis teknologi untuk meningkatkan kualitas pelayanan kesehatan.

Wassalamu'alaikum warahmatullahi wabarakatuh,

Muhammad Rosdiana M.Kom

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	iii
PENDAHULUAN	1

BAGIAN 1

FONDASI DASAR	3
---------------------	---

BAB 1

EPIDEMIOLOGI DAN LATAR BELAKANG KLINIS DIABETES ...	4
---	---

- 1.1. Distribusi dan Prevalensi Diabetes secara Global dan Nasional 4
- 1.2. Karakteristik Klinis Utama serta Klasifikasi Diabetes 5
- 1.3. Faktor Risiko, Mekanisme Biologis, dan Strategi Pencegahan 6

BAB 2

DASAR-DASAR <i>MACHINE LEARNING</i> DALAM BIDANG KESEHATAN.....	8
---	---

- 2.1. Konsep Fundamental *Machine Learning* untuk Analisis Kesehatan 8
- 2.2. Kategori Algoritma Relevan untuk Sistem Diagnosis dan Prediksi 9
- 2.3. Aspek Etika, Regulasi, dan Tanggung Jawab Penggunaan AI dalam Kesehatan 10

BAB 3

PRAPEMROSESAN DATA UNTUK ANALITIK MEDIS	12
---	----

- 3.1. Teknik Pembersihan dan Transformasi Data dalam Rekam Medis..... 12
- 3.2. Penanganan Missing Values, Outliers, dan Ketidakseimbangan Data 13

3.3. Normalisasi, Standardisasi, dan Pengodean Variabel Klinis	14
---	----

BAGIAN 2

METODOLOGI DAN PENDEKATAN PEMODELAN.....	17
---	-----------

BAB 4

ANALISIS KOMPARATIF ALGORITMA KLASIFIKASI	18
4.1. Studi Perbandingan Kinerja Berbagai Algoritma Klasifikasi.....	18
4.2. Kelebihan, Keterbatasan, dan Konteks Penggunaan Metode	19
4.3. Strategi Pemilihan Algoritma Optimal untuk Prediksi Diabetes	20

BAB 5

TEKNIK REKAYASA FITUR DAN SELEKSI ATRIBUT	21
5.1. Identifikasi Fitur Klinis yang Berpengaruh pada Prediksi.....	21
5.2. Pendekatan Seleksi Fitur untuk Data Medis	22
5.3. Optimasi Kombinasi Fitur untuk Peningkatan Performa Model	22

BAB 6

PERBANDINGAN METODE-METODE MACHINE LEARNING..	24
6.1. <i>Logistic Regression</i>	25
6.2. <i>Random Forest</i>	44
6.3. <i>Support Vector Machine (SVM)</i>	51
6.4. <i>K-Nearest Neighbors (KNN)</i>	60
6.5. <i>XGBoost (Extreme Gradient Boosting)</i>	70

BAGIAN 3

IMPLEMENTASI DAN APLIKASI PRAKTIS	81
--	-----------

BAB 7

KERANGKA EVALUASI DAN VALIDASI MODEL	82
7.1. Metrik Evaluasi Utama untuk Model Klasifikasi	82
7.2. Validasi Silang, Pengujian Model, dan Strategi Penghindaran Overfitting	85
7.3. Interpretasi Hasil Evaluasi untuk Konteks Klinis.....	95

BAB 8

STUDI KASUS: PIMA INDIANS DIABETES DATASET 97

 8.1. Analisis Karakteristik Dataset Pima Indians 97

 8.2. Implementasi Model Prediksi dan Evaluasi Performa 98

 8.3. Kesimpulan 118

DAFTAR PUSTAKA.....121

PROFIL PENULIS..... 127

PENDAHULUAN

Diabetes melitus merupakan salah satu penyakit tidak menular dengan angka pertumbuhan tercepat di dunia. Jumlah penyandang diabetes meningkat secara signifikan setiap tahun dan diperkirakan dapat berlipat ganda dalam dua dekade mendatang. Kondisi ini menjadikan diabetes sebagai ancaman serius bagi kesehatan masyarakat global maupun nasional, bukan hanya karena prevalensinya yang tinggi, tetapi juga karena komplikasi jangka panjang yang dapat menurunkan kualitas hidup hingga meningkatkan angka kematian. Komplikasi seperti penyakit jantung, gagal ginjal, retinopati, serta kerusakan saraf menjadikan diabetes sebagai penyakit kronis yang membutuhkan penanganan komprehensif secara berkelanjutan.

Salah satu tantangan utama dalam pengendalian diabetes adalah tingginya jumlah kasus yang tidak terdiagnosis. Banyak individu menjalani kehidupan sehari-hari tanpa mengetahui bahwa mereka memiliki kadar gula darah yang tinggi atau berada pada kondisi prediabetes. Fenomena ini menyebabkan keterlambatan penanganan, sehingga meningkatkan risiko komplikasi yang lebih parah. Karena itu, deteksi dini menjadi langkah strategis dalam upaya pencegahan dan penanganan diabetes. Proses deteksi yang tepat dan cepat dapat membantu masyarakat memperoleh intervensi medis yang lebih baik, sekaligus mengurangi beban biaya pengobatan jangka panjang.

Perkembangan teknologi data dan kecerdasan buatan, khususnya machine learning, membuka peluang besar dalam meningkatkan efektivitas upaya deteksi dini diabetes. Machine learning memungkinkan sistem untuk mempelajari pola dari data klinis dan memprediksi kondisi kesehatan secara otomatis tanpa instruksi eksplisit. Hal ini menjadikannya alat pendukung keputusan

yang penting, terutama ketika data medis yang tersedia melimpah dan kompleks. Dengan pemanfaatan model prediksi, tenaga kesehatan dapat terbantu dalam mengidentifikasi pasien berisiko dan menentukan langkah intervensi lebih awal.

Namun, keberhasilan sistem prediksi berbasis machine learning sangat bergantung pada kualitas data serta pemilihan algoritma yang tepat. Setiap algoritma memiliki karakteristik, kelebihan, dan keterbatasan yang berbeda dalam menangani data klinis. Oleh karena itu, perbandingan berbagai metode sangat diperlukan untuk mendapatkan model terbaik yang mampu memberikan prediksi akurat dan dapat diandalkan dalam konteks medis. Selain itu, proses prapemrosesan data seperti penanganan data hilang, normalisasi, serta seleksi fitur memegang peran penting dalam menghasilkan model yang optimal.

Buku ini hadir untuk memberikan pemahaman komprehensif mengenai pengembangan model prediksi dini diabetes berbasis machine learning. Bahan pembahasan disusun secara sistematis, dimulai dari pengenalan diabetes, landasan teoretis, hingga implementasi praktis dengan studi kasus menggunakan dataset Pima Indians Diabetes. Pembaca akan diajak memahami konsep fundamental, teknik pengolahan data, evaluasi algoritma, serta interpretasi hasil model. Fokus utama buku ini adalah membandingkan beberapa algoritma klasifikasi populer seperti Logistic Regression, KNN, Random Forest, SVM, dan XGBoost untuk menentukan model paling efektif dalam konteks prediksi diabetes.

Melalui pembahasan yang mendalam dan aplikatif, buku ini diharapkan dapat menjadi referensi bagi mahasiswa, peneliti, akademisi, serta praktisi kesehatan yang ingin mengeksplorasi potensi teknologi kecerdasan buatan dalam sektor medis. Lebih dari itu, buku ini diharapkan dapat memfasilitasi kolaborasi antara bidang teknologi dan kesehatan, guna menghasilkan sistem prediksi yang akurat, efisien, dan layak diterapkan demi meningkatkan kualitas layanan kesehatan masyarakat.



Bagian 1

FONDASI DASAR

BAB 1

EPIDEMIOLOGI DAN LATAR BELAKANG KLINIS DIABETES

1.1. Distribusi dan Prevalensi Diabetes secara Global dan Nasional

Diabetes melitus merupakan salah satu penyakit tidak menular dengan pertumbuhan tercepat di dunia. Prevalensi diabetes—yakni proporsi total kasus pada suatu populasi pada waktu tertentu—telah meningkat tajam dalam beberapa dekade terakhir, terutama pada diabetes tipe 2. Peningkatan prevalensi ini dipengaruhi oleh perubahan gaya hidup, urbanisasi, serta pola konsumsi yang semakin tinggi akan kalori dan lemak jenuh. Berbeda dari prevalensi, insidensi menggambarkan jumlah kasus baru dalam suatu periode waktu tertentu dan menjadi indikator yang penting untuk mengevaluasi laju penularan atau pengembangan penyakit. Data terbaru organisasi kesehatan dunia menunjukkan bahwa lebih dari 500 juta orang hidup dengan diabetes dan diperkirakan angka ini dapat meningkat hingga hampir dua kali lipat dalam dua dekade ke depan.

Distribusi diabetes tidak merata di seluruh dunia, dengan kawasan Pasifik Barat, Asia Tenggara, dan Timur Tengah mencatat angka prevalensi tertinggi. Negara berpendapatan rendah dan menengah menunjukkan peningkatan paling signifikan, dipengaruhi oleh perubahan pola hidup, industrialisasi, serta rendahnya akses terhadap edukasi kesehatan dan layanan medis. Kondisi ini menunjukkan bahwa diabetes kini bukan hanya penyakit negara maju, tetapi telah menjadi isu global yang memerlukan intervensi lintas sektor. Urbanisasi yang cepat mengubah pola diet dan aktivitas

fisik masyarakat, sehingga risiko diabetes semakin tinggi terutama pada masyarakat usia produktif.

Di Indonesia, prevalensi diabetes terus menunjukkan peningkatan dari tahun ke tahun. Data Riset Kesehatan Dasar (Riskesdas) mencatat perkembangan signifikan pada kasus diabetes, yang sejalan dengan meningkatnya obesitas dan perubahan gaya hidup masyarakat. Namun, fenomena *iceberg* menyebabkan banyak kasus tidak terdiagnosis, sehingga angka prevalensi sesungguhnya kemungkinan lebih tinggi dibandingkan data resmi. Distribusi diabetes di Indonesia juga bervariasi antarwilayah, dipengaruhi oleh faktor ekonomi, etnisitas, intensitas urbanisasi, pola makan, serta akses terhadap fasilitas kesehatan. Tren peningkatan prevalensi dan distribusi yang tidak merata menegaskan pentingnya strategi preventif, deteksi dini, serta pengendalian komprehensif di tingkat nasional.

1.2. Karakteristik Klinis Utama serta Klasifikasi Diabetes

Secara klinis, diabetes melitus ditandai oleh hiperglikemia kronis akibat gangguan sekresi insulin, kerja insulin, atau keduanya. Kondisi hiperglikemia berkepanjangan ini berdampak pada kerusakan organ target seperti jantung, ginjal, mata, dan jaringan saraf. Gejala klinis klasik meliputi poliuria (sering buang air kecil), polidipsia (sering haus), dan polifagia (sering lapar). Selain itu, pasien dapat mengalami penurunan berat badan tanpa sebab yang jelas, kelelahan, pandangan kabur, serta komplikasi kronis seperti nefropati dan neuropati. Pemeriksaan laboratorium menggunakan parameter gula darah puasa, gula darah postprandial, HbA_{1c}, dan gula darah sewaktu menjadi dasar diagnosis diabetes.

Secara umum, diabetes diklasifikasikan menjadi diabetes tipe 1, diabetes tipe 2, dan diabetes gestasional. Diabetes tipe 1 disebabkan oleh proses autoimun yang menghancurkan sel beta pankreas sehingga tubuh mengalami kekurangan insulin absolut. Kondisi ini umumnya muncul pada usia anak-anak atau remaja, meskipun dapat pula terjadi pada dewasa dengan onset lambat. Pasien tipe 1

mebutuhkan terapi insulin seumur hidup. Diabetes tipe 2 merupakan bentuk yang paling umum, ditandai oleh resistensi insulin disertai penurunan fungsi sekresi insulin. Kondisi ini kuat dipengaruhi oleh faktor gaya hidup dan obesitas. Tren saat ini menunjukkan peningkatan kasus tipe 2 pada usia muda akibat perubahan pola hidup masyarakat.

Diabetes gestasional merupakan hiperglikemia yang pertama kali terdeteksi selama kehamilan. Meskipun umumnya bersifat sementara, kondisi ini meningkatkan risiko komplikasi kehamilan dan persalinan, serta meningkatkan risiko ibu dan anak untuk mengembangkan diabetes tipe 2 di kemudian hari. Pengenalan klasifikasi diabetes menjadi penting untuk menentukan strategi penatalaksanaan yang tepat. Pemahaman klinis mengenai gejala, patofisiologi, dan perbedaan karakteristik tiap tipe membantu dalam menentukan intervensi medis maupun strategi pencegahan untuk menekan progresivitas penyakit.

1.3. Faktor Risiko, Mekanisme Biologis, dan Strategi Pencegahan

Faktor risiko diabetes dapat dibagi menjadi faktor yang tidak dapat dimodifikasi dan yang dapat dimodifikasi. Faktor risiko yang tidak dapat dimodifikasi meliputi riwayat keluarga, usia, dan predisposisi etnis tertentu. Individu dengan riwayat keluarga diabetes memiliki kemungkinan lebih tinggi mengembangkan penyakit ini, terutama tipe 2. Risiko juga meningkat seiring bertambahnya usia, khususnya di atas 40 tahun. Selain itu, kelompok etnis tertentu seperti Asia Selatan, Asia Tenggara, dan Afrika berisiko lebih tinggi dibandingkan kelompok lain. Faktor yang dapat dimodifikasi, termasuk obesitas, gaya hidup sedentari, pola makan tinggi gula dan lemak, merokok, serta stres berkelanjutan, memegang peran utama dalam peningkatan kasus diabetes tipe 2.

Secara biologis, mekanisme patogenesis diabetes berbeda antar tipe. Pada diabetes tipe 1, sistem imun menyerang sel beta pankreas sehingga produksi insulin menurun drastis. Kondisi ini menyebabkan hiperglikemia berat dan membutuhkan terapi insulin eksternal. Pada

diabetes tipe 2, terjadi kombinasi resistensi insulin dan disfungsi sel beta pankreas. Awalnya, tubuh dapat mengimbangi resistensi insulin dengan meningkatkan produksi insulin, tetapi seiring waktu, sel beta menjadi kelelahan dan gagal mempertahankan kadar glukosa normal. Proses ini tertanam dalam interaksi kompleks antara faktor genetika, lingkungan, dan gaya hidup.

Strategi pencegahan diabetes dapat dibagi menjadi pencegahan primer dan sekunder. Pencegahan primer bertujuan mencegah timbulnya penyakit melalui perubahan pola hidup, termasuk pengaturan berat badan ideal, penerapan aktivitas fisik teratur minimal 150 menit per minggu, pola makan seimbang tinggi serat dan rendah gula, serta menghindari kebiasaan merokok. Sementara pencegahan sekunder berfokus pada pengendalian glikemik bagi individu yang telah terdiagnosis diabetes, mencakup pemantauan kadar gula darah, pengobatan yang tepat, edukasi pasien, serta skrining rutin untuk deteksi dini komplikasi. Intervensi ini penting untuk mengurangi komplikasi jangka panjang dan meningkatkan kualitas hidup pasien. Pemahaman mengenai faktor risiko, mekanisme biologis, dan strategi pencegahan ini menjadi dasar yang kuat dalam mengembangkan model prediktif berbasis data untuk mendukung deteksi dini diabetes.

DAFTAR PUSTAKA

(Laporan ini menjadi sumber data epidemiologi dan prevalensi diabetes di Indonesia yang sangat penting).

Aditya, P. R., & Wibowo, A. (2021). *Prediksi Risiko Diabetes Melitus Menggunakan Extreme Gradient Boosting (XGBoost)*. Jurnal Ilmiah Teknologi Informasi Terapan, 7(2), 88-95. DOI: 10.33197/jitter.vol7.iss2.2021.693

Ali Linkon A, Rahman Noman I, Rashedul Islam M, Chakra Bortty J, Kumar Bishnu K, Islam A, et al. Evaluation of Feature Transformation and Machine Learning Models on Early Detection of Diabetes Mellitus. IEEE Access. 2024;12(October):165425-40.

Atwany MZ, Sahyoun AH, Yaqub M. Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey. IEEE Access. 2022;10:28642-55.

Cahyono Y, Rosyani P, Syah FS, Putri FS, Ashari I, Sofian K. Liver Disease Classification Using Decision Tree and Random Forest Algorithms. *Int J Integr Sci*. 2025;4(1):135-40.

Dewi, C., & Utami, E. (2019). Sistem Pendukung Keputusan untuk Diagnosa Awal Diabetes Menggunakan Logistic Regression. *Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)*, 1(1), 1-8. DOI: 10.35134/komputa.viii.12

Fadli, A., et al. (2020). Deteksi Dini Diabetes Berbasis Machine Learning: Studi Kasus Dataset Pima Indian. *Jurnal Teknologi dan Sistem Komputer*, 8(2), 112-119. DOI: 10.14710/jtsiskom.8.2.2020.112-119

Hakim, L., & Sari, I. P. (2019). Penerapan Algoritma C4.5 untuk Prediksi Penyakit Diabetes. *Jurnal Sains dan Informatika*, 5(2), 148-156. DOI: 10.22216/jsi.v5i2.4293

- Jia L, Wang Z, Lv S, Xu Z. PE_DIM: An Efficient Probabilistic Ensemble Classification Algorithm for Diabetes Handling Class Imbalance Missing Values. IEEE Access. 2022;10(September):107459–76.
- Kementerian Kesehatan RI. (2018). *Laporan Nasional Riskesdas 2018*. Lembaga Penerbit Badan Penelitian dan Pengembangan Kesehatan (LPB). DOI: 10.13140/RG.2.2.36444.95361
- Khan FA, Zeb K, Al-Rakhami M, Derhab A, Bukhari SAC. Detection and Prediction of Diabetes Using Data Mining: A Comprehensive Review. IEEE Access. 2021;9:43711–35.
- Kurniastuti I, Andini A, Dwisapta MR. *Implementation of Neural Network for Classification of Diabetes Mellitus through Finger Nail Image. Procedia Comput Sci [Internet]*. 2024;234(2023):1625–32. Available from: <https://doi.org/10.1016/j.procs.2024.03.166>
- Kurniawan, H., & Arof, H. (2019). Pengaruh Pra-Pemrosesan Data terhadap Kinerja Algoritma Klasifikasi pada Dataset Medis. *Jurnal Teknologi Informasi dan Komunikasi*, 10(1), 1–10. DOI: 10.30646/tik-sin.v10i1.386
- Lara-Abelenda FJ, Chushig-Muzo D, Acosta CB, Wägner AM, Granja C, Soguero-Ruiz C. Evaluating Time Series Classification Models for Nocturnal Hypoglycemia: From Predictive Performance to Environmental Impact. IEEE Access. 2025;13(September):150756–71.
- Liu T, Ren Z, Xiong C, Peng W, Wu J, Huang S, et al. Optoacoustic classification of diabetes mellitus with the synthetic impacts via optimized neural networks. *Heliyon [Internet]*. 2023;9(10):e20796. Available from: <https://doi.org/10.1016/j.heliyon.2023.e20796>
- Maulana, R., & Hikmah, N. (2021). Pemanfaatan Internet of Things (IoT) dan Machine Learning untuk Monitoring Faktor Risiko Diabetes Secara Real-Time. *Jurnal Teknologi Kesehatan*, 17(2), 112–122. DOI: 10.1331/jukesh.v17i2.21035

- Mulyono, S., & Fanani, L. (2022). Prediksi Diabetes Melitus Menggunakan Deep Learning dengan Arsitektur Jaringan Saraf Tiruan. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 9(2), 345-354. DOI: 10.25126/jtiik.2022924567
- Nurfadilah, S., & Suryana, I. (2022). Interpretasi Model Prediksi Diabetes Menggunakan SHAP (SHapley Additive exPlanations). *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 8(1), 89-96. DOI: 10.26418/jp.v8i1.50523
- Otero Sanchez L, Zhan CY, Gomes da Silveira Cauduro C, Crenier L, Njimi H, Englebert G, et al. A machine learning-based classification of adult-onset diabetes identifies patients at risk of liver-related complications. *JHEP Reports* [Internet]. 2023;5(8):100791. Available from: <https://doi.org/10.1016/j.jhepr.2023.100791>
- Pangestu, A. A., & Zulianto, M. (2022). Analisis Kinerja Metode K-Nearest Neighbor (K-NN) dengan Normalisasi Data untuk Prediksi Penyakit Diabetes. *Jurnal Teknologi dan Manajemen Informatika*, 8(1), 23-30. DOI: 10.26905/jtmi.v8i1.7456
- Perkumpulan Endokrinologi Indonesia (PERKENI). (2021). *Pedoman Pengelolaan dan Pencegahan Diabetes Melitus Tipe 2 di Indonesia 2021*. PB PERKENI. (Meski seringkali tanpa DOI, pedoman ini adalah rujukan klinis utama di Indonesia. Dapat diakses langsung dari website PERKENI).
- Pratama, I. P. A. E., et al. (2022). Optimasi Random Forest untuk Prediksi Diabetes dengan Seleksi Fitur dan Tuning Hyperparameter. *Jurnal Media Informatika Budidarma*, 6(1), 405-414. DOI: 10.30865/mib.v6i1.3615
- Rahayu, P., & Setiawan, E. B. (2021). Optimasi Prediksi Diabetes dengan Metode Ensemble Learning: Voting Classifier. *Jurnal Siger Matematika*, 2(2), 77-85. DOI: 10.23960/jsm.v2i2.78
- Raj S, Raj S, Namdeo V, Srivastava A. Decoding the gene-disease associations in type 2 diabetes: A curated dataset for text mining-based classification. *Data Br.* 2024;54.

- Reza MS, Amin R, Yasmin R, Kulsum W, Ruhi S. *Improving diabetes disease patients classification using stacking ensemble method with PIMA and local healthcare data*. Heliyon [Internet]. 2024;10(2):e24536. Available from: <https://doi.org/10.1016/j.heliyon.2024.e24536>
- Rizki, A., et al. (2021). *Implementasi Metode SMOTE dan Random Forest untuk Klasifikasi Data Tidak Seimbang pada Deteksi Dini Diabetes*. Building of Informatics, Technology and Science (BITS), 3(3), 240-247. DOI: 10.47065/bits.v3i3.1234
- Saputra, D. A., & Priyanta, S. (2021). Tantangan dan Strategi Implementasi Artificial Intelligence dalam Sistem Layanan Kesehatan di Indonesia. *Jurnal Sistem Informasi dan Telematika*, 12(2), 123-134. DOI: 10.35194/jsi.v12i2.1456
- Saputra, R., & Hidayat, R. (2021). Komparasi Algoritma Klasifikasi untuk Prediksi Penyakit Diabetes. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 8(4), 831-838. DOI: 10.25126/jtiik.2021843151
- Sari, Y. W., & Fauzi, M. A. (2020). Prediksi Diabetes Menggunakan Algoritma Naive Bayes dan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(3), 556-563. DOI: 10.29207/resti.v4i3.1931
- Sinsirimongkhon M, Arwatchananukul S, Temdee P. Multi-Class Classification Method with Feature Engineering for Predicting Hypertension with Diabetes. *J Mob Multimed*. 2023;19(3):799-822.
- Siregar, B. G. T., & Andreswari, R. (2020). Perbandingan Metode Seleksi Fitur Filter, Wrapper, dan Embedded pada Klasifikasi Data Medis. *Jurnal Media Informatika Budidarma*, 4(4), 1024-1033. DOI: 10.30865/mib.v4i4.2356
- Wanto, A., et al. (2020). Analysis of the PIMA Indians Diabetes Dataset using Data Mining Methods with RapidMiner Tools. *Jurnal Teknik Informatika (Jutif)*, 1(1), 13-22. DOI: 10.20884/1.jutif.2020.1.1.11

- Wei J, Xu Y, Wang H, Niu T, Jiang Y, Shen Y, et al. *Metadata information and fundus image fusion neural network for hyperuricemia classification in diabetes*. Comput Methods Programs Biomed. 2024;256(July):1–12.
- Wibowo, E. S., & Sari, Y. (2020). Aspek Etika dan Regulasi Kecerdasan Buatan (AI) dalam Diagnosis Medis: Tinjauan Pustaka Sistematis. *Jurnal Hukum Kesehatan Indonesia*, 2(1), 45–60. DOI: 10.15294/jhki.v2i1.45678
- Zhang Y, Chen L, Ouyang Y, Wang X, Fu T, Yan G, et al. A new classification method for gestational diabetes mellitus: a study on the relationship between abnormal blood glucose values at different time points in oral glucose tolerance test and adverse maternal and neonatal outcomes in pregnant women with gestational diabetes mellitus. *AJOG Glob Reports* [Internet]. 2024;4(4):100390. Available from: <https://doi.org/10.1016/j.xagr.2024.100390>
- Zhou H, Rahman S, Angelova M, Bruce CR, Karmakar C. *A robust and generalized framework in diabetes classification a heterogeneous environments*. Comput Biol Med [Internet]. 2025;186(January):109720. Available from: <https://doi.org/10.1016/j.compbimed.2025.109720>

PROFIL PENULIS



Muhammad Rosdiana M.Kom

Muhammad Rosdiana adalah seorang akademisi dan peneliti yang berdedikasi di bidang Teknik Informatika, berkarya di Universitas Pamulang (Unpam), Tangerang, Indonesia. Ia menempuh pendidikan Sarjana Komputer di Unpam (fakultas Teknik Informatika). Setelah lulus, ia melanjutkan kariernya di Unpam sebagai dosen dan peneliti, aktif dalam mengembangkan penelitian dan publikasi di bidang cloud computing, kecerdasan buatan dan aplikasi digital.

Buku ini membahas pendekatan modern dalam mendeteksi dini penyakit diabetes menggunakan teknologi machine learning. Melalui perpaduan kajian medis, teknik pengolahan data, dan analisis algoritmik, penulis menyajikan pemahaman komprehensif terkait bagaimana model kecerdasan buatan dapat membantu mempercepat proses diagnosis, meningkatkan efisiensi, serta mendukung pengambilan keputusan klinis.

Dimulai dengan pengantar epidemiologi dan karakteristik klinis diabetes, buku ini menjelaskan faktor risiko, mekanisme biologis, serta urgensi deteksi dini. Pembaca kemudian diarahkan pada pemahaman fundamental machine learning, mencakup konsep, klasifikasi algoritma, peluang, serta tantangan etis dalam implementasinya di bidang kesehatan.

Tahapan penting seperti prapemrosesan data klinis meliputi pembersihan data, penanganan missing values, normalisasi, hingga seleksi fitur dipaparkan secara sistematis agar pembaca memahami pentingnya kualitas data dalam pembentukan model prediktif yang andal. Pembahasan kemudian berlanjut pada perbandingan algoritma klasifikasi populer seperti Logistic Regression, KNN, Random Forest, SVM, dan XGBoost. Setiap metode dijelaskan dari sisi teori, kelebihan, kekurangan, serta konteks penggunaannya.

Buku ini juga memuat studi kasus nyata, yaitu prediksi diabetes menggunakan Pima Indians Diabetes Dataset. Implementasi model, evaluasi kinerja, serta interpretasi hasil disajikan secara terstruktur sehingga mudah diikuti. Ulasan mengenai berbagai pendekatan evaluasi, seperti akurasi, recall, F1-score, dan AUC memberikan gambaran menyeluruh mengenai keandalan model dalam konteks klinis.

Dengan bahasa yang jelas dan praktis, buku ini menjadi rujukan berharga bagi mahasiswa, akademisi, peneliti, maupun praktisi kesehatan yang tertarik menggali potensi machine learning dalam mendukung pencegahan dan deteksi dini diabetes. Kehadiran buku ini diharapkan mampu menjadi jembatan antara dunia kesehatan dan teknologi, serta mendorong terciptanya sistem diagnosis yang lebih cepat, akurat, dan efektif.



PENERBIT KBM INDONESIA
Anggota IKAPI
0813 5751 7526 / (0353) 3234874
Kantor : Maguwoharjo, Depok, Sleman-Yogyakarta
@penerbitbukujogja
@penerbit.kbm

ISBN 978-634-278-315-3

