

COMPARATIVE ANALYSIS OF LOGISTIC REGRESSION AND RANDOM FOREST FOR PREDICTING HEALTH RISKS AMONG THE ELDERLY WITH COMORBIDITIES IN BANGKA

Putri Mentari Endraswari^{1*}, Nurhaeka Tou², Syilva Qonita Yuriko³

Program Studi Teknologi Informasi, Fakultas Sains dan Teknik,
Universitas Bangka Belitung, Indonesia
putrimentari@ubb.ac.id^{1*}, nurhaeka@ubb.ac.id², syilvaqonita60@gmail.com³
*Corresponding author

Received 31 March 17, 2026; Revised April 25, 2026; Accepted April 29, 2026; Published April 30, 2026

ABSTRACT

The comorbidity of hypertension and diabetes in the elderly is a significant health problem because it increases the risk of complications and the burden on medical services. In resource-constrained healthcare facilities, machine learning approaches can facilitate more efficient risk-based screening. This study aimed to evaluate the ability of machine learning models to predict diabetes, hypertension, and their comorbidities in elderly patients in Bangka, and to identify key clinical factors contributing to these risks. This study used medical record data from 279 elderly patients in the emergency department. Two models were compared: logistic regression and Random Forest, with predictor variables including age, sex, BMI, waist circumference, and blood pressure. The results showed that the Random Forest model had moderate discrimination ability for diabetes (AUC 0.67) and hypertension (AUC 0.64), but poor discrimination ability for comorbidities (AUC 0.56–0.57). Adiposity (BMI and waist circumference) and blood pressure were the main determinants of risk. Overall, this model has the potential to serve as an initial screening tool, though it still requires recalibration, the addition of clinical features, and external validation before operational implementation.

Keywords: Elderly, Hypertension, Diabetes, Comorbidities, Machine learning, Risk screening

ABSTRAK

Komorbiditas hipertensi dan diabetes pada lansia merupakan permasalahan kesehatan yang signifikan karena meningkatkan risiko komplikasi dan beban layanan medis. Dalam konteks fasilitas kesehatan dengan sumber daya terbatas, pendekatan machine learning menawarkan potensi untuk membantu proses skrining berbasis risiko secara lebih efisien. Penelitian ini bertujuan untuk mengevaluasi kemampuan model machine learning dalam memprediksi diabetes, hipertensi, dan komorbiditas keduanya pada pasien lansia di Bangka, serta mengidentifikasi faktor klinis utama yang berkontribusi terhadap risiko tersebut. Studi ini menggunakan data rekam medis dari 279 pasien lansia di instalasi gawat darurat. Dua model dibandingkan, yaitu regresi logistik dan Random Forest, dengan variabel prediktor meliputi usia, jenis kelamin, IMT, lingkar perut, dan tekanan darah. Hasil menunjukkan bahwa model Random Forest memiliki kemampuan diskriminasi moderat terhadap diabetes (AUC 0,67) dan hipertensi (AUC 0,64), namun rendah pada komorbiditas (AUC 0,56–0,57). Variabel adipositas (IMT dan lingkar perut) dan tekanan darah menjadi determinan utama risiko. Secara keseluruhan, model ini berpotensi digunakan sebagai alat bantu skrining awal, meskipun masih memerlukan kalibrasi ulang, penambahan fitur klinis, serta validasi eksternal sebelum diimplementasikan secara operasional.

Kata Kunci: Lansia, Hipertensi, Diabetes, Komorbid, Machine learning, Skrining risiko

PENDAHULUAN

Kesehatan adalah salah satu penentu kesejahteraan yang mencakup aspek fisik, mental, dan sosial. Menurut Organisasi Kesehatan Dunia atau *World Health Organization* (WHO), kesehatan bukan berarti tidak memiliki penyakit atau kelemahan sama sekali, namun menandakan suatu keadaan yang sejahtera secara optimal. Dalam kesehatan masyarakat, kelompok lanjut usia (lansia) perlu memperoleh perhatian khusus, dibandingkan kelompok usia muda (Qonita et al., 2021). Lansia cenderung memiliki risiko lebih tinggi untuk mengalami beragam penyakit, khususnya penyakit degeneratif maupun penyakit komorbiditas (hipertensi dan diabetes) (Ariyanti et al., 2020).

Tekanan darah tinggi (hipertensi) didefinisikan sebagai keadaan medis dengan tekanan darah mencapai 140/90 mmHg atau lebih, yang dapat berujung pada peningkatan jumlah penyakit (morbiditas) atau bahkan berujung kematian (mortalitas) (Agustina et al., 2014). Meskipun hipertensi merupakan salah satu penyakit tidak menular (PTM), namun hipertensi sering dijumpai pada lansia (Purwono et al., 2020; Wahidin et al., 2023). Sesuai dengan data Kementerian Kesehatan Republik Indonesia pada tahun 2018, menunjukkan bahwa dari 34,1% prevalensi hipertensi di Indonesia, proporsi tertinggi terjadi pada kelompok lansia (Kementerian Kesehatan Republik Indonesia, 2018).

Selain hipertensi, penyakit diabetes juga merupakan salah satu masalah yang kerap dijumpai pada lansia (Arini et al., 2022). Diabetes mellitus adalah gangguan metabolik yang ditandai oleh kadar glukosa darah yang tinggi akibat gangguan produksi atau fungsi insulin (Febriani & Fitri, 2019). Kedua penyakit (hipertensi dan diabetes) seringkali muncul bersamaan pada lansia dan saling memperparah kondisi kesehatan tubuh. Kombinasi keduanya dikenal sebagai komorbiditas yang berisiko tinggi terhadap munculnya komplikasi serius, termasuk gangguan kardiovaskular, gagal ginjal, serta neuropati (Wahyuni et al., 2023).

Berdasarkan penelitian dari Febriani & Fitri (2019), penyakit degeneratif yang sering terjadi pada lansia antara lain hipertensi, obesitas, dan diabetes melitus (Febriani & Fitri, 2019). Hasil tersebut sesuai dengan Profil Kesehatan di Dinas Kesehatan Provinsi Kepulauan Bangka Belitung pada tahun 2015, dimana hipertensi menjadi penyakit dengan kasus tertinggi di wilayah Bangka. Sementara itu, diabetes menempati posisi kelima dalam daftar sepuluh penyakit paling banyak dialami masyarakat di provinsi tersebut pada tahun yang sama (Pemerintah Provinsi Kepulauan Bangka Belitung, 2015).

Meskipun hipertensi dan diabetes merupakan penyakit kronis yang sering terjadi pada lansia, tidak semua lansia mengalami kedua penyakit tersebut secara bersamaan. Perbedaan ini menunjukkan adanya faktor-faktor tertentu yang memengaruhi terjadinya komorbiditas. Penting untuk membedakan antara penyakit tunggal dan komorbiditas, karena keberadaan dua penyakit secara bersamaan dapat meningkatkan risiko komplikasi, memperburuk kondisi kesehatan, serta menambah kompleksitas dalam penanganan klinis. Selain itu, komorbiditas juga berimplikasi pada meningkatnya beban layanan kesehatan dan biaya perawatan. Oleh karena itu, diperlukan analisis yang lebih mendalam untuk mengidentifikasi pola hubungan dan faktor risiko yang berkontribusi terhadap munculnya komorbid hipertensi dan diabetes pada lansia, sehingga dapat mendukung

upaya pencegahan dan pengelolaan yang lebih efektif (Herdiani, 2019).

Data pada penelitian ini selain menggunakan data faktor medis, digunakan juga informasi faktor fisik seperti usia, indeks massa tubuh (IMT), tekanan darah, kadar gula darah, dan kebiasaan hidup. Faktor fisik tersebut juga sama pentingnya dalam menentukan tingkat risiko kesehatan lansia (Rahmatillah et al., 2020). Oleh karena itu, selain dilakukan analisis pada aspek medis (hipertensi dan diabetes), digunakan juga aspek fisik sebagai pelengkap analisis agar memperoleh gambaran yang komprehensif mengenai kondisi kesehatan lansia, khususnya di wilayah dengan populasi lansia yang terus meningkat seperti Kabupaten Bangka.

Penelitian ini bertujuan untuk membandingkan kinerja metode Regresi Logistik dan Random Forest dalam memprediksi risiko kesehatan pada lansia di Kabupaten Bangka dengan memanfaatkan pendekatan *machine learning* (ML), khususnya terkait hipertensi, diabetes, dan komorbiditas keduanya. Selain itu, penelitian ini juga bertujuan untuk mengidentifikasi faktor-faktor prediktor utama yang berkontribusi terhadap kondisi risiko tinggi (*HighRisk*), yang didefinisikan berdasarkan tekanan darah tidak terkontrol, indeks massa tubuh (IMT) ≥ 27 kg/m² (kategori obesitas untuk populasi Asia), atau kombinasi komorbid hipertensi-diabetes. Melalui pendekatan ini, penelitian diharapkan dapat memberikan model prediksi yang akurat sekaligus interpretatif guna mendukung skrining dini dan pengambilan keputusan klinis di tingkat pelayanan kesehatan primer.

Penelitian ini menggabungkan metode Random Forest sebagai model utama dalam pemodelan prediktif dengan Regresi Logistik sebagai model interpretatif untuk menghasilkan *Odds Ratio* dan pengujian signifikansi terhadap kovariat kunci seperti usia, jenis kelamin, IMT, serta status hipertensi dan diabetes. Regresi Logistik memungkinkan estimasi besaran asosiasi antar variabel, sedangkan Random Forest mewakili keluarga model non-linier yang mampu menangkap interaksi dan hubungan kompleks antar variabel, serta menghasilkan ukuran kepentingan fitur (*feature importance*) yang bermanfaat dalam praktik triase (Aditya et al., (2024); Oktafian et al., 2024). Perbandingan dilakukan untuk mengevaluasi kinerja masing-masing metode dalam memprediksi risiko kesehatan secara empiris. Sementara itu, pendekatan kombinasi digunakan untuk memanfaatkan keunggulan masing-masing metode, yaitu Random Forest dalam meningkatkan akurasi prediksi dan Regresi Logistik dalam memberikan interpretasi terhadap faktor-faktor risiko. Dengan demikian, penelitian ini tidak hanya menghasilkan model yang optimal secara performa, tetapi juga informatif dan dapat digunakan dalam pengambilan keputusan klinis.

Meskipun pendekatan ini telah banyak digunakan dalam studi prediksi kesehatan, penelitian ini menawarkan kontribusi spesifik melalui penerapannya pada populasi lansia di Kabupaten Bangka yang masih terbatas dalam literatur, serta dengan memanfaatkan data kesehatan lokal. Selain itu, penelitian ini tidak hanya berfokus pada satu luaran, tetapi juga mengkaji prediksi risiko secara terintegrasi, termasuk hipertensi, diabetes, dan komorbiditas keduanya dalam satu kerangka analisis. Pendekatan ini diharapkan dapat memberikan pemahaman yang lebih kontekstual serta mendukung pengambilan

keputusan berbasis data di lingkungan layanan kesehatan dengan sumber daya terbatas (Yasmin, 2023).

METODE PENELITIAN

Jenis dan Desain Penelitian

Penelitian ini merupakan studi observasional retrospektif berbasis data rekam medis IGD yang digunakan untuk pengembangan model prediksi risiko pada lansia. Pendekatan ini dipilih karena memungkinkan pemanfaatan data klinis rutin untuk mengidentifikasi pola dan memprediksi tiga luaran kesehatan secara simultan, yaitu diabetes, hipertensi, dan komorbiditas keduanya. Dalam penelitian ini, Random Forest digunakan sebagai model utama untuk klasifikasi dan prediksi risiko, sedangkan Regresi Logistik digunakan untuk mengidentifikasi dan menginterpretasikan faktor-faktor yang berasosiasi melalui estimasi *Odds Ratio* dan uji signifikansi. Dengan demikian, penelitian ini mengintegrasikan pendekatan prediktif dan interpretatif secara komplementer dalam satu kerangka analisis.

Jenis dan Sumber Data

Data yang digunakan merupakan jenis data Sekunder, yang berasal dari sistem rekam medis IGD Puskesmas Sungailiat, Kabupaten Bangka pada periode Januari hingga Maret 2025. Unit analisis adalah Poli Lansia.

Pengambilan data dalam penelitian ini telah memperoleh izin resmi dari Dinas Kesehatan Kabupaten Bangka berdasarkan surat izin Nomor 476/UN50/G.1.1/TU/2025. Seluruh data yang digunakan telah dianonimkan untuk menjaga kerahasiaan identitas pasien, sehingga tidak terdapat informasi yang dapat mengidentifikasi individu secara langsung.

Populasi, Kriteria inklusi/eksklusi, dan Ukuran Sampel

Target populasi dalam penelitian ini adalah pasien lanjut usia (≥ 50 tahun) yang datang ke Instalasi Gawat Darurat (IGD) Puskesmas Sungailiat. Unit analisis yang digunakan adalah kunjungan pasien, sehingga satu individu dimungkinkan memiliki lebih dari satu entri data apabila melakukan kunjungan berulang. Data penelitian diperoleh dari rekam medis dan diseleksi berdasarkan kriteria inklusi dan eksklusi untuk menjamin kualitas serta relevansi informasi. Kriteria inklusi meliputi rekam medis dengan informasi lengkap mengenai usia, jenis kelamin, tanda vital dasar (*vital signs*), serta diagnosis klinis. Sementara itu, kriteria eksklusi mencakup data duplikasi yang identik serta entri dengan ketidaklengkapan pada variabel prediktor utama.

Definisi diagnosis hipertensi dan diabetes dalam penelitian ini mengacu pada pencatatan diagnosis oleh tenaga medis pada rekam medis IGD, yang didasarkan pada standar klinis yang berlaku. Setelah proses penyaringan, diperoleh sebanyak 279 data kunjungan yang memenuhi kriteria dan digunakan sebagai sampel akhir penelitian. Distribusi data pada masing-masing luaran (hipertensi, diabetes, dan komorbiditas)

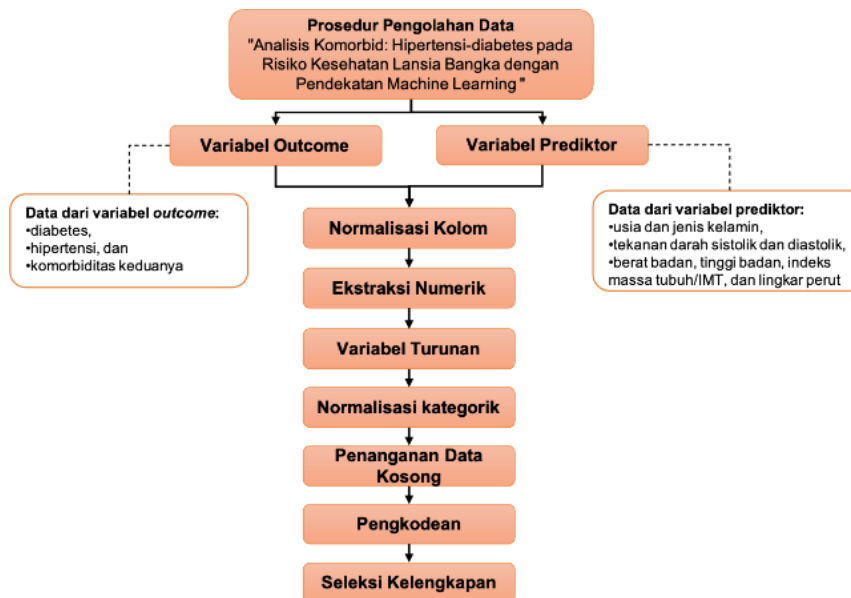
selanjutnya dianalisis untuk memastikan kecukupan representasi kelas dalam pengembangan model prediktif.

Variabel Penelitian

Penelitian ini menggunakan dua kelompok variabel utama, yaitu variabel *outcome* (dependen) dan variabel prediktor (independen). Variabel *outcome* diturunkan dari diagnosis utama pasien yang tercatat dalam kode ICD-10 (*International Classification of Diseases, 10th Revision*) maupun teks diagnosis rekam medis. Seorang pasien dikategorikan memiliki diabetes apabila salah satu kolom kode diagnosis berada dalam rentang E10–E14 atau teks diagnosis memuat kata kunci seperti “diabetes”, “mellitus”, atau “DM”. Kategori hipertensi ditetapkan apabila terdapat kode I10–I15 atau ditemukan kata kunci “hipertensi” maupun “*hypertension*” pada diagnosis. Adapun komorbiditas didefinisikan sebagai kondisi ketika kedua diagnosis tersebut terpenuhi secara bersamaan. Pendekatan ini digunakan untuk mengatasi variasi kualitas pencatatan klinis, terutama ketika kode ICD tidak selalu dicantumkan namun teks diagnosis tersedia, serta untuk meningkatkan sensitivitas pelabelan data tanpa menimbulkan risiko kebocoran data dari variabel pascadiagnosis (Widayanti et al., 2025).

Variabel prediktor dalam penelitian ini mencakup beberapa kelompok data, yaitu karakteristik demografis (usia dalam tahun dan jenis kelamin), tanda vital (tekanan darah sistolik dan diastolik), serta antropometri (berat badan, tinggi badan, indeks massa tubuh/IMT, dan lingkar perut). Selain itu, data perilaku seperti kebiasaan merokok, konsumsi alkohol, serta konsumsi sayur dan buah juga disertakan sebagai variabel prediktor, karena faktor-faktor tersebut diketahui berkontribusi terhadap peningkatan risiko penyakit tidak menular, termasuk hipertensi dan diabetes. Variabel-variabel ini digunakan untuk memperkaya model dalam menangkap aspek gaya hidup yang dapat memengaruhi kondisi kesehatan lansia. Informasi poli atau ruangan asal pasien juga direkam sebagai konteks layanan yang dapat mencerminkan kondisi klinis awal atau tingkat keparahan pasien saat datang ke fasilitas kesehatan.

Prosedur Analisis Data



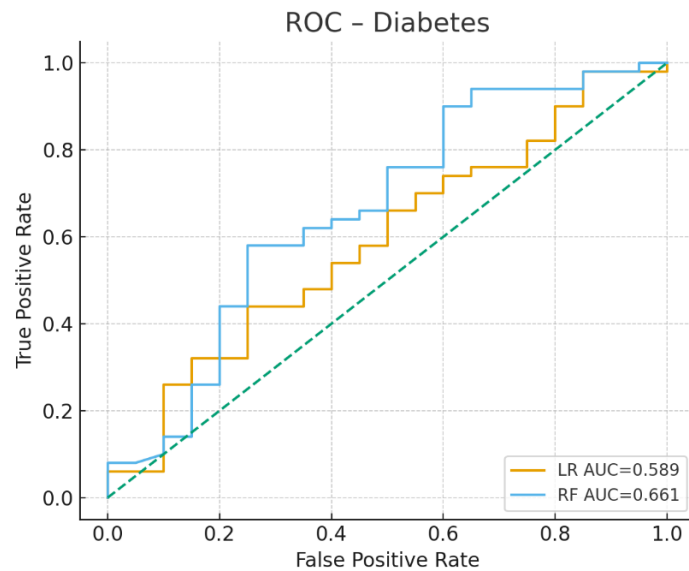
Gambar 1. Alur Prosedur Analisis Data

Pengolahan data dilakukan secara berurutan melalui beberapa tahapan seperti pada Gambar 1. Langkah pertama, seluruh nama kolom dinormalisasi (menggunakan huruf kecil tanpa spasi atau simbol) untuk memastikan konsistensi skema data. Kedua, nilai berformat teks yang memuat angka dan satuan dikonversi menjadi bentuk numerik murni. Ketiga, variabel turunan, khususnya indeks massa tubuh (IMT), dihitung ulang apabila belum tercantum dalam data asli. Keempat, kategori pada variabel nominal diseragamkan untuk menjaga konsistensi label. Kelima, nilai hilang pada prediktor numerik ditangani menggunakan imputasi median, sedangkan pada prediktor kategorik menggunakan modus. Keenam, variabel kategorik dikonversi menjadi representasi numerik melalui pengkodean one-hot, dengan penyesuaian terhadap kategori yang tidak dikenal selama proses pemodelan. Untuk menjaga kualitas data, prediktor dengan proporsi nilai hilang melebihi 60% dikeluarkan dari model dasar. Seluruh analisis dilakukan menggunakan bahasa pemrograman *Python*, dengan modul utama yaitu *pandas*, *scikit-learn*, *statsmodels*, dan *matplotlib*.

HASIL DAN PEMBAHASAN

Analisis dilakukan terhadap 279 pasien lansia di Puskesmas Sungailiat. ditemukan prevalensi Diabetes (DM) sebesar 70,97%, Hipertensi (HT) sebesar 45,52%, dan komorbid: diabetes-hipertensi sebesar 32,97%. Angka ini menunjukkan bahwa sepertiga populasi lansia IGD memiliki lebih dari satu penyakit kronis metabolik, menggambarkan beban ganda penyakit tidak menular pada kelompok usia lanjut di wilayah ini. Model *machine learning* dievaluasi menggunakan 25% data uji, dengan perbandingan antara Random Forest (RF) dan Regresi Logistik (LR).

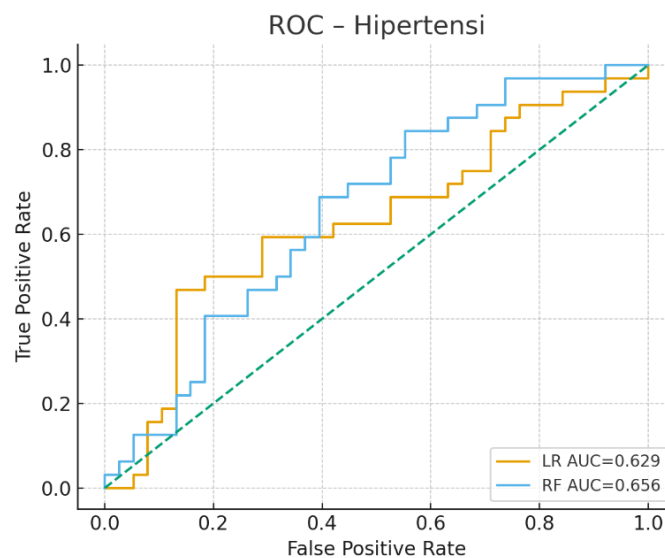
1. Prediksi Diabetes



Gambar 1. Kurva ROC Diabetes (perbandingan LR vs RF)

Berdasarkan Gambar 1, model Random Forest menunjukkan kemampuan diskriminasi yang moderat ($AUC \approx 0,67$), lebih baik dibanding Regresi Logistik ($AUC \approx 0,59$). Artinya, RF lebih mampu membedakan antara pasien lansia dengan dan tanpa diabetes berdasarkan variabel fisik dan vital yang tersedia. Perbedaan ini menunjukkan bahwa hubungan antarvariabel pada diabetes bersifat nonlinier, sehingga pendekatan berbasis pohon seperti RF lebih adaptif.

2. Prediksi Hipertensi

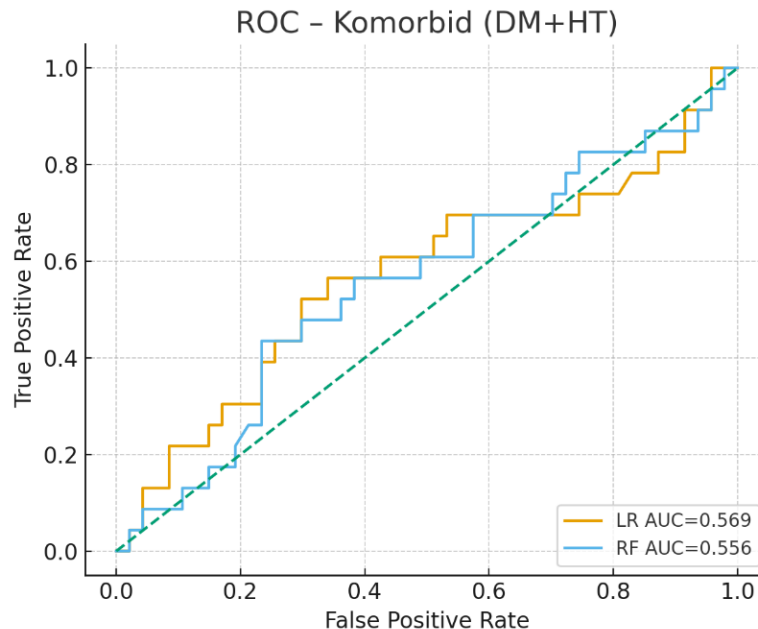


Gambar 2. Kurva ROC Hipertensi (perbandingan LR vs RF)

Untuk prediksi hipertensi (Gambar 2), kedua model menunjukkan kinerja yang relatif serupa (RF $AUC \approx 0,64$ dan LR $\approx 0,63$), dengan selisih yang sangat kecil.

Perbedaan ini tidak cukup signifikan untuk mendukung klaim mengenai keunggulan model tertentu maupun untuk menyimpulkan sifat hubungan antar variabel. Oleh karena itu, diperlukan evaluasi tambahan menggunakan metrik lain seperti sensitivitas, spesifisitas, presisi, serta kalibrasi model guna memperoleh pemahaman yang lebih komprehensif terhadap performa masing-masing metode.

3. Prediksi Komorbiditas

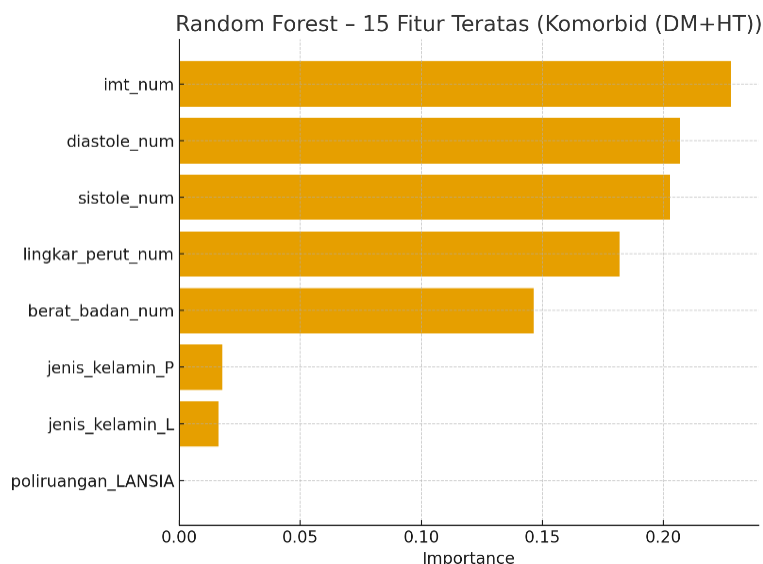


Gambar 3. Kurva ROC Komorbid (DM+HT) (perbandingan LR vs RF)

Pada kasus komorbid diabetes–hipertensi seperti pada gambar 3, kedua model menunjukkan kinerja yang relatif rendah ($AUC \approx 0,56-0,57$), yang mendekati kemampuan klasifikasi acak. Temuan ini mengindikasikan bahwa model belum mampu menangkap pola yang cukup kuat untuk membedakan kelompok komorbid secara akurat. Salah satu kemungkinan penyebabnya adalah keterbatasan variabel prediktor yang digunakan, di mana faktor fisik dasar seperti IMT, lingkaran perut, dan tekanan darah mungkin belum cukup merepresentasikan kompleksitas interaksi penyakit ganda pada lansia.

Namun demikian, kinerja yang rendah ini tidak hanya dapat dijelaskan oleh keterbatasan variabel. Faktor lain seperti ukuran sampel yang relatif terbatas, potensi ketidakseimbangan kelas, tumpang tindih distribusi antar kelompok, serta kualitas label dari data rekam medis retrospektif juga dapat memengaruhi performa model. Oleh karena itu, diperlukan pendekatan yang lebih komprehensif, baik melalui penambahan variabel klinis (misalnya kadar glukosa darah, profil lipid, riwayat keluarga, gaya hidup, dan kepatuhan pengobatan) maupun peningkatan kualitas dan kuantitas data, untuk menghasilkan model prediktif yang lebih akurat.

4. Variabel Penting (*Feature Importance*)



Gambar 4. 15 Fitur Teratas RF untuk Komorbid (DM+HT)

Analisis *feature importance* dari Random Forest pada Gambar 4 menunjukkan bahwa IMT, lingkar perut, serta tekanan darah sistolik dan diastolik merupakan prediktor dominan. Temuan ini konsisten dengan berbagai studi empiris yang menunjukkan bahwa obesitas, khususnya adipositas sentral yang diukur melalui lingkar perut, serta peningkatan tekanan darah merupakan faktor risiko utama untuk hipertensi dan diabetes tipe 2 pada populasi lansia. Penelitian Bai et al (2022) menunjukkan bahwa IMT dan lingkar perut berasosiasi signifikan dengan kejadian diabetes tipe 2 pada lansia, serta berkontribusi terhadap peningkatan risiko komorbiditas hipertensi dan diabetes melalui interaksi antara obesitas umum dan sentral. Selain itu, IMT dan lingkar perut juga telah diidentifikasi sebagai indikator utama obesitas yang berkaitan erat dengan hipertensi. Secara lebih luas, peningkatan kedua indikator tersebut dikaitkan dengan berbagai penyakit kronis, termasuk hipertensi dan diabetes pada kelompok usia lanjut (Zhang et al., 2021). Temuan ini juga sejalan dengan literatur Supadmi et al (2024) yang menyatakan bahwa adipositas sentral dan tekanan darah tinggi merupakan faktor risiko utama penyakit metabolik dan kardiovaskular pada lansia (Supadmi et al., 2024). Dengan demikian, hasil penelitian ini tidak hanya sejalan dengan temuan sebelumnya, tetapi juga memperkuat bukti bahwa faktor antropometri dan hemodinamik memainkan peran penting dalam menentukan risiko komorbiditas pada lansia.

Sementara itu, hasil Regresi Logistik juga memperkuat temuan tersebut, di mana sebagian besar variabel adipositas dan tekanan darah memiliki *Odds Ratio* (OR) > 1, yang berarti peningkatan nilai variabel tersebut meningkatkan kemungkinan munculnya diabetes, hipertensi, atau keduanya. Detail nilai OR beserta interval kepercayaan 95% (95% CI) dapat dilihat pada Tabel 1, Tabel 2, dan Tabel 3.

Tabel 1. *Odds Ratio* (OR) dengan 95% CI dan p untuk Diabetes

Variabel	Diabetes OR (95% CI)	Diabetes p
IMT (per 1 kg/m ²)	0.95 [0.83–1.08]	0.408
Lingkar perut (per 1 cm)	1.00 [0.98–1.03]	0.777
Sistolik (per 1 mmHg)	0.99 [0.97–1.01]	0.243
Diastolik (per 1 mmHg)	0.97 [0.93–1.01]	0.162
Perempuan vs Laki-laki	1.91 [0.99–3.71]	0.054

Keterangan: Nilai OR ditafsirkan per kenaikan 1 satuan sesuai unit pada kolom Variabel.
CI = *Confidence Interval*

Tabel 2. *Odds Ratio* (OR) dengan 95% CI dan p untuk Hipertensi

Variabel	Hipertensi OR (95% CI)	Hipertensi p
IMT (per 1 kg/m ²)	1.17 [0.99–1.39]	0.064
Lingkar perut (per 1 cm)	0.99 [0.96–1.01]	0.236
Sistolik (per 1 mmHg)	1.06 [1.03–1.09]	0.000
Diastolik (per 1 mmHg)	0.93 [0.89–0.98]	0.007
Perempuan vs Laki-laki	1.33 [0.69–2.58]	0.400

Keterangan: Nilai OR ditafsirkan per kenaikan 1 satuan sesuai unit pada kolom Variabel.
CI = *Confidence Interval*

Tabel 3. *Odds Ratio* (OR) dengan 95% CI dan p untuk Komorbid

Variabel	Komorbid OR (95% CI)	Komorbid p
IMT (per 1 kg/m ²)	1.03 [0.91–1.17]	0.636
Lingkar perut (per 1 cm)	1.00 [0.98–1.03]	0.710
Sistolik (per 1 mmHg)	1.03 [1.00–1.05]	0.019
Diastolik (per 1 mmHg)	0.97 [0.93–1.01]	0.177
Perempuan vs Laki-laki	1.78 [0.90–3.49]	0.095

Keterangan: Nilai OR ditafsirkan per kenaikan 1 satuan sesuai unit pada kolom Variabel.
CI = *Confidence Interval*

Berdasarkan hasil analisis yang telah diperoleh, Random Forest menunjukkan diskriminasi yang lebih baik daripada regresi logistik untuk luaran diabetes (AUC moderat) dan keunggulan tipis pada hipertensi, sedangkan performa kedua model terhadap komorbid masih rendah. Walau demikian, regresi logistik menyediakan interpretasi yang lebih transparan melalui *Odds Ratio* (95% CI), menegaskan peran IMT, lingkar perut, serta tekanan darah sebagai determinan utama. Dengan demikian, pendekatan yang diusulkan bersifat komplementer: Random Forest digunakan untuk tujuan prediktif (skrining/klasifikasi), sedangkan regresi logistik untuk tujuan inferensial dan komunikasi klinis.

Hasil penelitian ini memiliki implikasi langsung bagi praktik klinis, pihak manajemen, maupun pembuatan kebijakan, dengan menjadikan model yang dihasilkan dari Random Forest sebagai alat bantu skrining awal untuk mengidentifikasi risiko

diabetes dan hipertensi pada lansia, sehingga memungkinkan prioritas pemeriksaan lanjutan secara lebih efisien. Temuan bahwa faktor adipositas (IMT dan lingkar perut) serta tekanan darah merupakan prediktor terkuat menegaskan pentingnya intervensi terfokus pada pengendalian berat badan dan tekanan darah sebagai langkah preventif utama. Namun demikian, temuan ini perlu diinterpretasikan secara hati-hati mengingat beberapa keterbatasan penelitian. Data yang digunakan berasal dari satu puskesmas dan terbatas pada pasien IGD, sehingga berpotensi menimbulkan bias seleksi dan membatasi generalisasi hasil ke populasi lansia secara lebih luas. Selain itu, ukuran sampel yang relatif terbatas serta belum dilakukannya validasi eksternal juga dapat memengaruhi stabilitas dan keandalan model prediktif yang dihasilkan. Nilai akurasi yang masih moderat pada beberapa luaran juga menunjukkan bahwa model belum sepenuhnya optimal untuk penerapan langsung dalam praktik klinis.

Secara keseluruhan, penelitian ini tetap memberikan kontribusi dalam konteks literatur dengan menunjukkan potensi penerapan *machine learning* dalam pemodelan prediktif risiko kesehatan, khususnya melalui integrasi prediksi tiga luaran sekaligus yaitu diabetes, hipertensi, dan komorbiditas keduanya menggunakan data rekam medis lokal pada populasi lansia di tingkat pelayanan primer. Berbeda dengan penelitian sebelumnya yang umumnya berfokus pada satu luaran penyakit atau menggunakan dataset berskala besar dari negara maju. Selain itu, pendekatan yang menggabungkan Random Forest sebagai model prediktif dan Regresi Logistik sebagai model interpretatif memberikan keseimbangan antara akurasi prediksi dan keterjelasan klinis, yang masih menjadi tantangan dalam banyak studi sebelumnya. Oleh karena itu, penelitian lanjutan dengan cakupan data yang lebih luas, desain yang lebih representatif, serta validasi eksternal sangat diperlukan untuk meningkatkan generalisasi dan kesiapan implementasi model dalam praktik nyata.

Temuan ini juga menyoroti pentingnya sinergi antara tenaga klinis, manajemen fasilitas kesehatan, dan pembuat kebijakan dalam mengoptimalkan pemanfaatan data dan teknologi, khususnya di lingkungan dengan keterbatasan sumber daya. Dengan demikian, penelitian ini tidak hanya berkontribusi pada aspek metodologis, tetapi juga memberikan nilai praktis dalam mendukung pengambilan keputusan berbasis data untuk upaya promotif dan preventif di tingkat pelayanan kesehatan primer.

KESIMPULAN

Studi ini menunjukkan bahwa pada populasi lansia di Kabupaten Bangka, model *machine learning* mampu mengklasifikasikan kondisi diabetes dan hipertensi dengan tingkat diskriminasi moderat, dimana Random Forest menunjukkan performa lebih baik dibandingkan regresi logistik, khususnya dalam prediksi diabetes. Faktor adipositas (IMT dan lingkar perut) serta tekanan darah terbukti konsisten sebagai determinannya yang paling berpengaruh, sebagaimana ditunjukkan oleh *feature importance* pada Random Forest dan *odds ratio* pada regresi logistik.

UCAPAN TERIMAKASIH

Ucapan terima kasih disampaikan kepada Lembaga Peneliti dan Pengabdian kepada Masyarakat (LPPM) Universitas Bangka Belitung, atas dukungan dana melalui program Peneliti Muda tahun anggaran 2025 dengan nomor kontrak 1266/UN50/M/PP/2025, yang telah memfasilitasi terlaksananya penelitian dan penyusunan artikel ilmiah ini hingga tahap publikasi.

REFERENCES

- Aditya, M. F. R., Lutvi, N., & Indahyanti, U. (2024). Prediksi Penyakit Hipertensi Menggunakan Metode Decision Tree dan Random Forest. *Jurnal Ilmiah Komputasi*, 23(1), 9-16. <https://doi.org/10.32409/jikstik.23.1.3503>
- Agustina, S., Sari, S. M., & Savita, R. (2014). Faktor-Faktor Yang Berhubungan Dengan Hipertensi Pada Lansia Di Atas Umur 65 Tahun. *Jurnal Kesehatan Komunitas*, 2(4), 180–186. <https://doi.org/10.25311/keskom.Vol2.Iss4.70>
- Arini, H. N., Anggorowati, A., & Pujiastuti, R. S. E. (2022). Dukungan keluarga pada lansia dengan Diabetes Melitus Tipe II: Literature review. *NURSCOPE: Jurnal Penelitian Dan Pemikiran Ilmiah Keperawatan*, 7(2), 172. <https://dx.doi.org/10.30659/nurscope.7.2.172-180>
- Ariyanti, R., Preharsini, I. A., & Sipolio, B. W. (2020). Edukasi Kesehatan Dalam Upaya Pencegahan Dan Pengendalian Penyakit Hipertensi Pada Lansia. *To Maega: Jurnal Pengabdian Masyarakat*, 3(2), 74-82. <https://doi.org/10.35914/tomaega.v3i2.369>
- Bai, K., Chen, X., Song, R., Shi, W., & Shi, S. (2022). Association Of Body Mass Index And Waist Circumference With Type 2 Diabetes Mellitus In Older Adults: A Cross-Sectional Study. *BMC geriatrics*, 22(1), 489.
- Febriani, R., & Fitri, M. (2019). Analisis Kejadian Hipertensi Pada Lansia Dengan Diabetes Melitus. *Jurnal'Aisyiyah Medika*, 4. <https://doi.org/10.36729/jam.v4i3.200>
- Herdiani, N. (2019). Hubungan IMT Dengan Hipertensi Pada Lansia Di Kelurahan Gayungan Surabaya. *Medical Technology and Public Health Journal*, 3(2), 183-189. <https://doi.org/10.33086/mtphj.v3i2.1179>
- Kementerian Kesehatan Republik Indonesia. (2018). *Hasil Riset Kesehatan Dasar (Riskesdas) 2018*. <https://layanandata.kemkes.go.id/katalog-data/riskesdas/ketersediaan-data/riskesdas-2018>
- Oktafian, W., Astuti, L. W., & Ramadhan, M. (2024). *Klasifikasi Penyakit Diabetes pada Lansia Menggunakan Model Random Forest* (Doctoral dissertation, Universitas Indo Global Mandiri).
- Pemerintah Provinsi Kepulauan Bangka Belitung. (2015). *Waspada penyakit hipertensi dan diabetes*. https://babelprov.go.id/siaran_pers/waspada-penyakit-hipertensi-dan-diabetes

- Purwono, J., Sari, R., Ratnasari, A., & Budianto, A. (2020). Pola konsumsi garam dengan kejadian hipertensi pada lansia. *Jurnal Wacana Kesehatan*, 5(1), 531-542. <https://doi.org/10.52822/jwk.v5i1.120>
- Qonita, F. N., Salsabila, N. A., Anjani, N. F., & Rahman, S. (2021). Kesehatan Pada Orang Lanjut Usia (Kesehatan Mental Dan Kesehatan Fisik). *Jurnal Psikologi Wijaya Putra (Psikowipa)*, 2(1), 10-19. <https://doi.org/10.38156/psikowipa.v2i1.42>
- Rahmatillah, V. P., Susanto, T., & Nur, K. R. M. (2020). Hubungan Karakteristik, Indeks Massa Tubuh (IMT) dengan Tekanan Darah pada Lanjut Usia di Posbindu. *Media Penelitian Dan Pengembangan Kesehatan*, 30(3), 233-240. <https://doi.org/10.22435/mpk.v30i3.2547>
- Supadmi, S., Sit, S., Kuntari, T., Hidayat, T., Khairunnisa, M., Ant, S., ... & Kom, M. (2024). *Sindrom Metabolik Pada Lanjut Usia (Lansia)*. Purbalingga: CV Eureka Media Aksara.
- Wahidin, M., Agustiya, R. I., & Putro, G. (2023). Beban Penyakit Dan Program Pencegahan Dan Pengendalian Penyakit Tidak Menular Di Indonesia. *Jurnal Epidemiologi Kesehatan Indonesia*, 6(2), 105–112. <https://doi.org/10.7454/epidkes.v6i2.6253>
- Wahyuni, S. D., Amalia, N., & Murti, B. (2023). Penyuluhan Hipertensi Dan Diabetes Melitus Pada Lansia Di Rw 12 Jebres, Surakarta. *Jurnal Pengabdian Komunitas*, 2(02), 70-76.
- Widayanti, J. R., Barus, J. F. A., Djuartina, T., Vanessa, I., Enumbi, M. A. Y., Setiati, S., & MI, W. (2025). Status Nutrisi, Komorbiditas dan Fungsi Kognitif pada Lansia Perempuan di Panti Wreda. *Jurnal Sains dan Kesehatan*, 6(2), 20-28. <https://doi.org/10.30872/jsk.v6i2.819>
- Yasmin, E. A. (2023). *Analisis Faktor-Faktor Yang Mempengaruhi Diabetes Mellitus Dengan Menggunakan Regresi Logistik Ordinal Di RSI Jemursari Surabaya* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- Zhang, W., He, K., Zhao, H., Hu, X., Yin, C., Zhao, X., & Shi, S. (2021). Association Of Body Mass Index And Waist Circumference With High Blood Pressure In Older Adults. *BMC geriatrics*, 21(1), 260.