



Pengembangan Algoritma *Random Forest* untuk Deteksi Dini Hipertensi Berbasis Web di Puskesmas Melintang

Alya Mistiara Salsabila¹, Yudistira Bagus Pratama², Arvi Pramudyantoro³

^{1,2,3} Universitas Muhammadiyah Bangka Belitung

Email: alyaamistiaraslsa@gmail.com

Article Info

Article history:

Received April 23, 2026

Revised Mei 04, 2026

Accepted Mei 18, 2026

Keywords:

Random Forest, Hypertension, Early Detection, Machine Learning

ABSTRACT

Hypertension is a chronic medical condition that requires early detection to prevent severe complications. This study aims to develop a web-based early detection system for hypertension using the Random Forest algorithm, specifically designed for Puskesmas Melintang. The study utilized a secondary dataset comprising 1,985 patient records with 11 clinical and lifestyle features. Hyperparameter optimization was performed using GridSearchCV. The model achieved an accuracy of 95.21% and an Area Under the Curve (AUC) of 0.9874. The system was deployed using Flask for the backend, Streamlit for the frontend, and hosted on Google Cloud Platform (GCP). User Acceptance Testing (UAT) indicated high satisfaction in usability, functionality, output, and utility. The implementation of this system provides an effective and efficient digital solution to accelerate the cardiovascular screening process and reduce administrative burdens at primary healthcare facilities.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Article Info

Article history:

Received April 23, 2026

Revised Mei 04, 2026

Accepted Mei 18, 2026

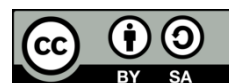
Keywords:

Random Forest, Hipertensi, Deteksi Dini, Machine Learning

ABSTRACT

Hipertensi merupakan kondisi medis kronis yang memerlukan deteksi dini untuk mencegah komplikasi berat. Penelitian ini bertujuan untuk mengembangkan sistem deteksi dini hipertensi berbasis web menggunakan algoritma Random Forest yang dirancang khusus untuk Puskesmas Melintang. Penelitian menggunakan dataset sekunder berjumlah 1.985 rekam medis dengan 11 fitur klinis dan gaya hidup. Optimasi hyperparameter dilakukan menggunakan metode GridSearchCV. Hasil pengujian menunjukkan bahwa model mencapai tingkat akurasi sebesar 95,21% dan nilai Area Under the Curve (AUC) sebesar 0,9874. Sistem ini diimplementasikan menggunakan Flask sebagai backend, Streamlit sebagai frontend, dan disebarluaskan melalui infrastruktur Google Cloud Platform (GCP). Pengujian penerimaan pengguna (UAT) menunjukkan tingkat kepuasan yang sangat positif pada aspek kemudahan penggunaan, fungsionalitas, keluaran, dan utilitas. Implementasi sistem ini memberikan solusi digital yang efektif dan efisien untuk mengakselerasi proses skrining kardiovaskular dan mengurangi beban administrasi di fasilitas layanan kesehatan primer.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Alya Mistiara Salsabila

Universitas Muhammadiyah Bangka Belitung

Email: alyaamistiaraslsa@gmail.com



PENDAHULUAN

Hipertensi merupakan kondisi medis kronis yang ditandai dengan peningkatan tekanan darah secara menetap dan sering dijuluki sebagai *silent killer* (World Health Organization, 2023). Penyakit ini menjadi kontributor utama komplikasi berat seperti stroke, gagal jantung, hingga penyakit ginjal kronis bagi penderitanya (Kementerian Kesehatan RI, 2023). Selain dampak kesehatan, hipertensi juga menimbulkan beban ekonomi yang signifikan akibat penurunan produktivitas dan tingginya biaya pengobatan jangka panjang (Harahap et al., 2021).

Prevalensi hipertensi di Indonesia menunjukkan tren peningkatan yang mencapai angka lebih dari 34% berdasarkan Survei Kesehatan Indonesia (Kementerian Kesehatan RI, 2023). Di Puskesmas Melintang sendiri, tercatat sebanyak 3.694 kasus pada tahun 2019 yang merupakan angka tertinggi di wilayah Kota Pangkalpinang (Nurjanah et al., 2022). Proses skrining di puskesmas tersebut saat ini masih dilakukan secara manual menggunakan tensimeter sehingga kurang efisien untuk menangani populasi pasien yang besar (Septian et al., 2025).

Permasalahan tersebut menuntut adanya pendekatan baru berbasis teknologi digital seperti *machine learning* dengan algoritma *Random Forest* untuk deteksi dini yang otomatis (Septian et al., 2025). Algoritma ini dinilai handal dalam memproses data medis pasien yang kompleks melalui teknik optimasi *hyperparameter* guna menemukan konfigurasi model terbaik (Septian et al., 2025). Analisis data cerdas ini memungkinkan diagnosis dilakukan lebih cepat dan efisien sehingga mendukung strategi pencegahan penyakit kronis di layanan kesehatan primer (Septian et al., 2025).

Kebaruan penelitian ini terletak pada penggunaan parameter klinis serta gaya hidup yang komprehensif dalam model untuk menghasilkan prediksi risiko hipertensi yang presisi. Sistem deteksi ini diintegrasikan ke dalam antarmuka berbasis web menggunakan *framework* Streamlit dan Flask serta disebarluaskan melalui layanan *cloud computing* Google Cloud Platform (GCP). Penggunaan teknologi awan memastikan sistem dapat diakses secara fleksibel tanpa harus membebani infrastruktur perangkat keras fisik yang terbatas di puskesmas.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasional berbasis *machine learning* untuk menghasilkan model prediksi yang objektif (Alpaydin, 2024). Sumber data yang dianalisis merupakan data medis sekunder dari repositori publik Kaggle yang terdiri atas 1.985 observasi dengan 11 atribut klinis dan gaya hidup. Setiap baris observasi tersebut telah dilengkapi dengan label kelas aktual berupa "Hipertensi" atau "Normal" yang menjadi target pembelajaran klasifikasi bagi algoritma.

Tahapan pra-pemrosesan data dilakukan untuk mengonversi data mentah menjadi format matematis melalui standarisasi fitur numerik menggunakan *StandardScaler* dan pengkodean fitur kategorikal menggunakan *One-Hot Encoding* (Liu et al., 2021). Seluruh himpunan data yang telah bersih kemudian dipartisi menjadi 80% data latih dan 20% data uji menggunakan teknik *stratified splitting*. Penerapan stratifikasi ini bertujuan secara spesifik untuk menjaga keseimbangan rasio kelas target pada kedua himpunan data agar model terhindar dari bias mayoritas saat pelatihan.



Pemodelan utama dilakukan menggunakan algoritma *Random Forest Classifier* yang dioptimasi melalui metode *GridSearchCV* dengan teknik validasi silang lima lipatan atau *5-Fold Cross-Validation* (James et al., 2023). Model terbaik yang dihasilkan kemudian diintegrasikan ke dalam arsitektur sistem deteksi dini berbasis web menggunakan *framework* Flask dan Streamlit yang beroperasi di infrastruktur *Google Cloud Platform* (Singh, 2021). Sebagai tahapan validasi akhir, kelayakan operasional sistem diukur melalui *User Acceptance Testing* (UAT) menggunakan kuesioner berskala *Likert* yang melibatkan 12 responden dari kalangan masyarakat (Radzi et al., 2022).

HASIL DAN PEMBAHASAN

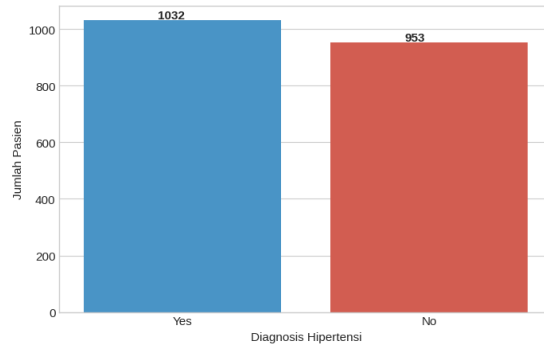
Hasil

Pelatihan model dalam penelitian ini menggunakan dataset rekam medis sekunder yang memuat 1.985 baris data pasien dengan sebelas parameter. Keseluruhan atribut tersebut merangkum variabel demografis, metrik klinis fisik, hingga rekam jejak gaya hidup yang relevan dengan kondisi kardiovaskular pasien. Penjelasan mengenai peran dan deskripsi dari setiap variabel penelitian disajikan secara rinci pada Tabel 1.

Tabel 1. Variabel Penelitian

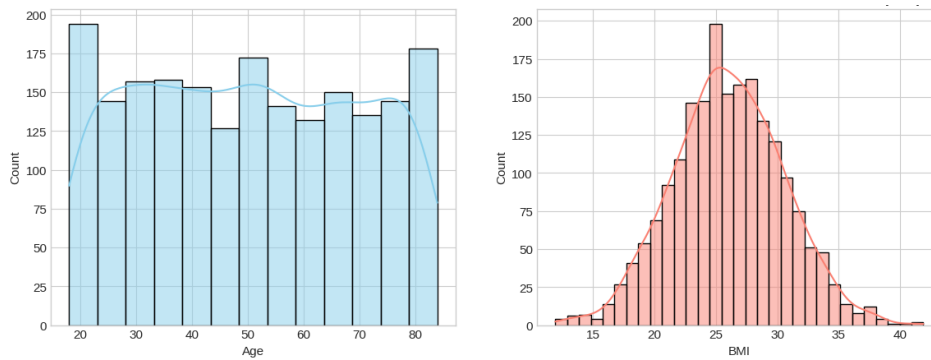
Nama Variabel	Tipe Data	Peran	Keterangan / Deskripsi
Age	Integer	Prediktor	Usia pasien dalam hitungan tahun.
Salt_Intake	Float	Prediktor	Estimasi asupan garam harian pasien.
Stress_Score	Integer	Prediktor	Tingkat stres subjektif (skala numerik).
BP_History	Object	Prediktor	Riwayat tekanan darah masa lalu.
Sleep_Duration	Float	Prediktor	Rata-rata jam tidur pasien per hari.
BMI	Float	Prediktor	Indeks Massa Tubuh (Body Mass Index).
Medication	Object	Prediktor	Obat kardiovaskular yang sedang dikonsumsi.
Family_History	Object	Prediktor	Riwayat keturunan hipertensi dalam keluarga.
Exercise_Level	Object	Prediktor	Intensitas aktivitas fisik atau olahraga.
Smoking_Status	Object	Prediktor	Status kebiasaan merokok pasien.
Has_Hypertension	Object	Target	Diagnosis akhir status risiko hipertensi pasien.
Nama Variabel	Tipe Data	Peran	Keterangan / Deskripsi

Sebelum memasuki tahapan pelatihan algoritma klasifikasi, sekumpulan data ini dibedah melalui *Exploratory Data Analysis* (EDA) untuk memetakan distribusi polanya. Evaluasi terhadap variabel dependen menunjukkan distribusi kelas target berada dalam kondisi yang sangat seimbang, yakni 1.032 terdiagnosis hipertensi dan 953 kategori normal. Keseimbangan proporsi ini menjamin model terhindar dari anomali bias mayoritas sehingga algoritma dapat mempelajari pola pembeda secara objektif.



Gambar 1. Distribusi Kelas Pasien Normal vs Hipertensi

Analisis distribusi fitur numerik mengungkap bahwa sebaran usia pasien sangat representatif dan mendekati pola distribusi normal. Rentang usia yang dicakup sangat luas, mulai dari 18 tahun hingga 84 tahun, dengan pemusatan data pada kelompok usia paruh baya sekitar 50 tahun. Sementara itu, distribusi indeks massa tubuh juga menunjukkan pemusatan pada angka 26,0 yang secara klinis dikategorikan sebagai faktor risiko hipertensi.



Gambar 2. Distribusi Usia dan Indeks Massa Tubuh Pasien

Eksplorasi lebih lanjut dilakukan melalui uji matriks korelasi terhadap seluruh sepuluh fitur prediktor beserta label target untuk memastikan tidak adanya multikolinearitas. Mengingat matriks korelasi hanya memproses data numerik, seluruh fitur kategorikal bertipe objek ditransformasi sementara ke dalam bentuk angka menggunakan metode *Label Encoding*. Hasil visualisasi peta panas (*heatmap*) menunjukkan nilai korelasi silang antar fitur independen sangat rendah, sementara variabel riwayat tekanan darah (*BP_History*) dan tingkat stres (*Stress_Score*) mencatatkan korelasi positif paling menonjol terhadap target diagnosis.

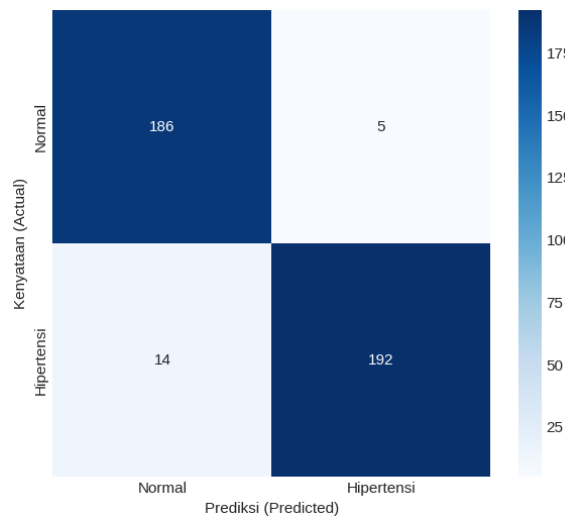
Tahap pra-pemrosesan data berhasil mentransformasi seluruh fitur heterogen menjadi format biner dan numerik yang terstandarisasi. Melalui teknik *One-Hot Encoding* dan *StandardScaler*, seluruh variabel prediktor kini memiliki skala yang seragam untuk diolah oleh fungsi objektif algoritma. *Dataset* akhir kemudian dibagi secara bertingkat menjadi 1.588 sampel untuk pelatihan dan 397 sampel independen untuk pengujian performa.

Optimasi konfigurasi model menggunakan metode *Grid Search Cross-Validation* berhasil mengidentifikasi parameter arsitektur yang paling akurat. Konfigurasi terbaik yang ditetapkan mencakup penggunaan 100 pohon keputusan dengan batasan kedalaman maksimal sepuluh tingkatan. Penentuan parameter ini terbukti secara empiris mampu memberikan skor akurasi rata-rata tertinggi sebesar 95,21% pada himpunan data latih.



Evaluasi final terhadap 397 sampel data uji independen menunjukkan bahwa model memiliki kemampuan generalisasi yang sangat impresif. Algoritma *Random Forest* yang telah dioptimasi berhasil mencetak nilai akurasi klasifikasi kumulatif sebesar 94,96%. Capaian angka ini memberikan bukti ilmiah bahwa model memiliki tingkat presisi yang sangat tinggi dalam mendeteksi risiko hipertensi pada data baru.

Pembedahan hasil melalui instrumen *Confusion Matrix* mengungkap efektivitas model dalam meminimalkan kesalahan tipe I dan tipe II. Model berhasil memberikan prediksi tepat pada 186 pasien normal dan 192 pasien yang benar-benar menderita hipertensi. Rendahnya jumlah kesalahan *False Negative* membuktikan bahwa sistem memiliki sensitivitas yang sangat baik dalam mengenali kondisi medis yang krusial.



Gambar 3. Confusion Matrix Model

Laporan klasifikasi menunjukkan nilai presisi yang sangat superior pada kelas hipertensi dengan pencapaian angka sebesar 97%. Meskipun nilai *recall* pada kelas tersebut sedikit lebih rendah di angka 93%, performa keseluruhan tetap konsisten pada skor F1 sebesar 95%. Stabilitas metrik evaluasi silang ini menegaskan bahwa algoritma tidak mengalami anomali bias ekstrem terhadap salah satu kelas diagnosis.

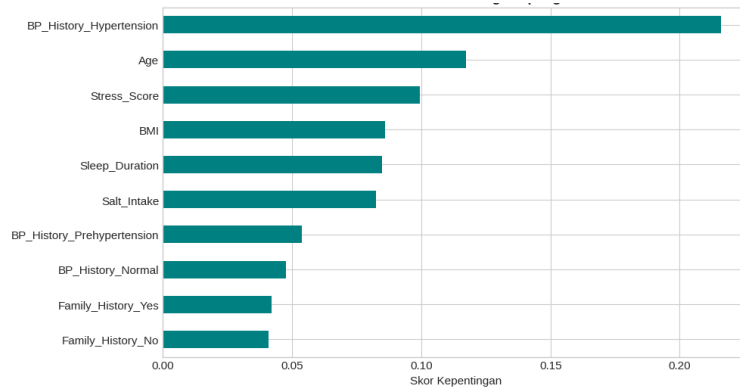
Tabel 2. Classification Report Performa Model

Kelas (Diagnosis)	Presisi (Precision)	Sensitivitas (Recall)	F1-Score	Jumlah Data (Support)	Kelas (Diagnosis)
0 (Normal)	0.93	0.97	0.95	191	0 (Normal)
1 (Hipertensi)	0.97	0.93	0.95	206	1 (Hipertensi)
Akurasi Global			0.95	397	Akurasi Global
Macro Avg	0.95	0.95	0.95	397	Macro Avg

Analisis tingkat kepentingan fitur menunjukkan bahwa riwayat tekanan darah masa lalu merupakan prediktor yang paling determinan dalam proses klasifikasi. Variabel demografis seperti usia dan indeks massa tubuh menempati urutan kontributor utama berikutnya sesuai dengan kaidah patologi kardiovaskular. Pemetaan ini memberikan wawasan objektif bagi

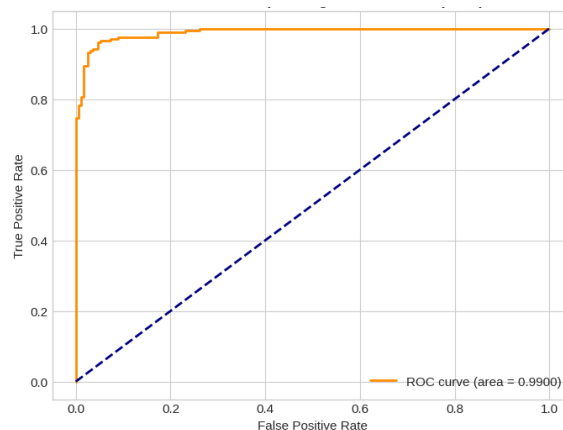


tenaga medis mengenai faktor-faktor risiko yang paling berpengaruh terhadap kenaikan tekanan darah.



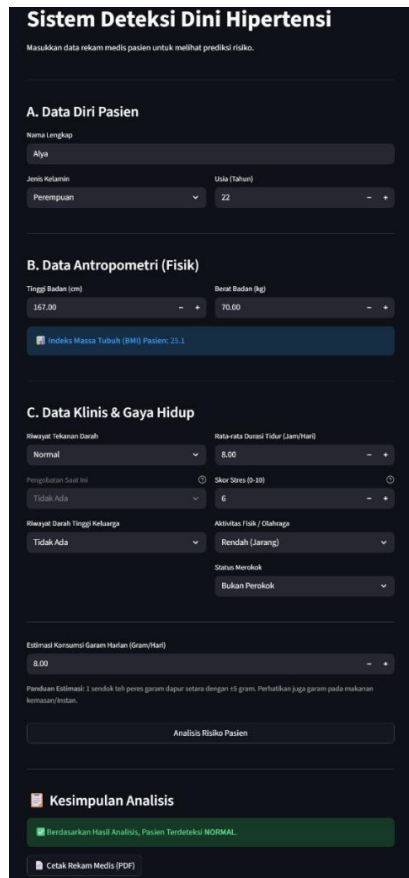
Gambar 4. Hierarki 10 Fitur Paling Berpengaruh

Validasi performa sistem secara grafis diperkuat oleh kurva *Receiver Operating Characteristic* yang melengkung tajam mendekati sudut kiri atas. Nilai *Area Under the Curve* yang dicapai menembus angka 0,99 yang merepresentasikan tingkat probabilitas diskriminasi kelas yang nyaris sempurna. Skor AUC yang sangat tinggi ini memberikan jaminan klinis bahwa model sangat layak diimplementasikan di fasilitas kesehatan primer.



Gambar 5. Kurva ROC dan Nilai AUC

Model terbaik yang telah disimpan dalam format biner berhasil diintegrasikan ke dalam antarmuka web interaktif yang disebarakan melalui infrastruktur GCP. Tampilan aplikasi menyajikan formulir rekam medis digital yang intuitif dan dilengkapi dengan sistem notifikasi visual untuk kesimpulan diagnosis. Kehadiran modul pelaporan otomatis juga memungkinkan petugas untuk mencetak hasil skrining ke dalam dokumen PDF secara instan.



PUSKESMAS MELINTANG
 Kecamatan Rangkui, Kota Pangkalpinang, Kepulauan Bangka Belitung
 Layanan Skrining Kesehatan Penyakit Tidak Menular (PTM)

LAPORAN HASIL DETEKSI DINI HIPERTENSI

A. IDENTITAS PASIEN

Nama Lengkap	Alya
Jenis Kelamin	Perempuan
Usia	22 Tahun
Tanggal Periksa	13 April 2026

B. PARAMETER MEDIS & GAYA HIDUP

Tinggi Badan	167.0 cm
Berat Badan	70.0 kg
BMI	25.1
Riwayat Tekanan Darah	Normal
Pengobatan	Tidak Ada
Riwayat Keluarga	Tidak Ada
Durasi Tidur	8.0 Jam
Tingkat Stres	Skor 6/10
Olahraga	Rendah (Jarang)
Status Merokok	Bukan Perokok
Konsumsi Garam	8.0 Gram/Hari

C. KESIMPULAN DETEKSI DINI

STATUS: NORMAL (TIDAK BERESIKO)

Rekomendasi: Pertahankan pola hidup sehat dan lakukan pemeriksaan rutin.

Petugas Pemeriksa,

(.....)

Gambar 6. Tampilan Antarmuka dan Laporan PDF

Pengujian penerimaan pengguna menunjukkan tingkat kepuasan yang sangat tinggi terhadap seluruh fungsionalitas yang disediakan oleh sistem. Distribusi jawaban kuesioner didominasi oleh pilihan "Sangat Setuju" pada metrik kemudahan penggunaan, fungsionalitas, keluaran, dan utilitas praktis. Mayoritas responden memberikan respons afirmatif yang mengonfirmasi bahwa alur aplikasi sangat jelas dan mampu mempercepat proses skrining.

Tabel 3. Distribusi Jawaban UAT

Kategori	Pernyataan	Sangat Setuju	Setuju	Netral	Tidak Setuju	Sangat Tdk Setuju
Kemudahan Penggunaan	Tampilan sistem mudah dipahami	6	5	1	0	0
	Alur pengisian data sangat jelas	6	5	1	0	0
	Bahasa medis mudah dimengerti	7	4	1	0	0
Fungsionalitas	Perhitungan otomatis BMI akurat	6	6	0	0	0
	Proses analisis AI sangat cepat	7	5	0	0	0
	Hasil prediksi ditampilkan jelas	7	5	0	0	0
Output dan Pelaporan	Fitur unduh dokumen (PDF) berfungsi	7	4	1	0	0
	Format dokumen PDF terlihat rapi	7	3	2	0	0
Manfaat Sistem	Membantu mempercepat skrining	8	4	0	0	0



Analisis umpan balik terbuka mengungkap apresiasi pengguna terhadap kesederhanaan desain antarmuka serta kecepatan respons prediksi kecerdasan buatan. Responden menilai bahwa otomatisasi perhitungan BMI dan visualisasi diagnosis sangat membantu dalam mentransformasi prosedur skrining manual. Di sisi lain, terdapat masukan kritis terkait perlunya penguatan lapisan keamanan privasi saat proses pengunduhan dokumen rekam medis digital.

Pembahasan

Penerapan algoritma *Random Forest* dalam penelitian ini terbukti sangat andal dengan pencapaian akurasi sebesar 95,21% dalam memprediksi risiko hipertensi. Tingginya nilai akurasi yang disertai skor *Area Under the Curve* (AUC) sebesar 0,9874 secara langsung menjawab rumusan masalah pertama mengenai reliabilitas model. Capaian kuantitatif tersebut mengonfirmasi bahwa mesin prediksi ini memiliki kemampuan diskriminasi kelas yang sangat superior untuk digunakan sebagai instrumen diagnosis.

Keberhasilan performa tersebut dipengaruhi oleh karakteristik arsitektur *Random Forest* yang sangat adaptif terhadap struktur data klinis yang kompleks. Algoritma berbasis ansambel ini mampu menangkap hubungan non-linear antar variabel tanpa memerlukan asumsi distribusi data yang ketat. Karakteristik ini sangat krusial mengingat data rekam medis di puskesmas sering kali memiliki kombinasi fitur numerik dan kategorikal yang heterogen.

Optimalnya klasifikasi juga tidak terlepas dari intervensi strategis melalui tahapan optimasi *hyperparameter* menggunakan metode *GridSearchCV*. Penetapan konfigurasi terbaik dengan seratus pohon keputusan dan batasan kedalaman sepuluh level terbukti efektif dalam menjinakkan varians model. Pembatasan kedalaman arsitektur ini secara sistematis mencegah terjadinya *overfitting* sehingga model tetap tangguh saat dihadapkan pada data pasien baru.

Analisis metrik *Feature Importance* menunjukkan bahwa algoritma mampu memberikan bobot analitik yang objektif dan selaras dengan literatur medis. Penempatan variabel riwayat tekanan darah, indeks massa tubuh (BMI), dan usia sebagai prediktor teratas membuktikan adanya penalaran logis terhadap patologi kardiovaskular. Keselarasan antara logika pembelahan *node* dan fakta empiris kedokteran memberikan jaminan bahwa keputusan model didasarkan pada kausalitas klinis yang sah.

Selain metrik komputasional, ketangguhan model ini juga terbukti dalam menangani dinamika input pengguna yang kompleks dan non-linear selama pengujian. Fenomena di mana pasien dengan riwayat hipertensi namun berskor stres sedang diprediksi "Normal" terjadi karena model membaca kontribusi faktor risiko lain (seperti BMI dan asupan garam) yang terkelola dengan baik. Sebaliknya, pengguna tanpa riwayat hipertensi namun memiliki skor stres yang ekstrem dapat langsung terdeteksi "Beresiko", karena algoritma *Random Forest* bekerja mengevaluasi pembobotan risiko secara multifaktorial dari sepuluh variabel sekaligus.

Dinamika prediksi ini mendapatkan validasi langsung dari pakar klinis Puskesmas Melintang, Ibu Kartini, AMK., yang mengonfirmasi bahwa diagnosis hipertensi tidak dapat ditentukan hanya oleh satu parameter. Beliau memaparkan bahwa status risiko pasien merupakan akumulasi dari kombinasi faktor seperti tekanan darah, tingkat stres, usia, dan pola hidup. Justifikasi klinis ini menegaskan bahwa aplikasi web ini beroperasi secara cerdas dan objektif sebagai alat bantu prediksi yang meniru logika diagnosis medis secara holistik di dunia nyata.



Temuan impresif dalam riset ini secara eksplisit mengafirmasi kesimpulan dari studi terdahulu yang dilakukan oleh Aditya et al. (2024) serta Novianti et al. (2023). Serupa dengan studi-studi rujukan tersebut, algoritma *Random Forest* kembali menunjukkan stabilitas klasifikasi yang luar biasa saat membedah variabel fisik mendasar pasien. Konsistensi hasil lintas penelitian ini memperkuat legitimasi algoritma tersebut sebagai standar emas dalam ranah klasifikasi risiko penyakit tidak menular.

Implementasi optimasi parameter dalam penelitian ini juga berhasil mengatasi batasan akurasi yang sebelumnya sempat dilaporkan oleh Pramesti Dewi et al. (2022). Meskipun beberapa literatur lain merekomendasikan algoritma *XGBoost* untuk klasifikasi skala masif, arsitektur *Random Forest* dinilai jauh lebih rasional untuk skenario Puskesmas Melintang. Hal ini dikarenakan *Random Forest* menawarkan keseimbangan beban komputasi yang ringan dengan tingkat interpretasi yang lebih mudah dipahami oleh tenaga medis.

Terkait rumusan masalah kedua, migrasi arsitektur sistem ke lingkungan *Google Cloud Platform* (GCP) terbukti menjadi solusi implementasi yang efisien. Pemanfaatan infrastruktur awan tersebut berhasil mendobrak keterbatasan spesifikasi perangkat keras fisik yang sering dihadapi oleh fasilitas kesehatan primer. Dengan pendekatan ini, aplikasi deteksi dini dapat beroperasi secara persisten tanpa membebani anggaran pemeliharaan perangkat lokal di puskesmas.

Penggunaan kerangka kerja Flask pada sisi *backend* dan Streamlit pada sisi *frontend* menciptakan ekosistem perangkat lunak yang responsif namun bertenaga. Flask berfungsi mengisolasi beban komputasi pemrosesan algoritma di peladen, sementara Streamlit menyederhanakan interaksi melalui desain formulir yang bersih. Sinergi teknologi ini memastikan tenaga medis dapat memasukkan data klinis dan melihat hasil prediksi secara instan tanpa hambatan operasional.

Kehadiran modul generator dokumen PDF otomatis merupakan terobosan penting untuk menjembatani inovasi digital dengan birokrasi pengarsipan fisik. Otomatisasi pencetakan laporan rekam medis ini secara signifikan memangkas durasi pelayanan administratif bagi petugas kesehatan di lapangan. Integrasi ini membuktikan bahwa luaran model prediksi medis dapat diadaptasikan secara praktis untuk mendukung efisiensi alur kerja layanan kesehatan masyarakat.

Validasi akhir melalui *User Acceptance Testing* (UAT) menunjukkan bahwa kerumitan model matematis berhasil disembunyikan dengan baik di balik antarmuka yang ramah pengguna. Dominasi skor "Sangat Setuju" dari berbagai demografi responden membuktikan bahwa aplikasi ini sangat inklusif bagi berbagai tingkat literasi teknologi. Secara keseluruhan, sistem ini telah bertransformasi menjadi perangkat intervensi digital yang aplikatif dan siap pakai untuk mengakselerasi prosedur skrining di Puskesmas Melintang.

KESIMPULAN

Penerapan algoritma *Random Forest* dalam penelitian ini terbukti sangat andal dalam memprediksi risiko penyakit hipertensi dengan tingkat akurasi mencapai 95,21% dan nilai *Area Under the Curve* (AUC) sebesar 0,9874. Pencapaian performa yang luar biasa ini diperoleh melalui penerapan teknik optimasi *hyperparameter* menggunakan *GridSearchCV* yang



menetapkan konfigurasi seratus pohon keputusan dan kedalaman maksimal sepuluh level. Hasil tersebut secara langsung membuktikan bahwa model mampu mengenali pola medis pasien secara presisi tanpa terjebak dalam kondisi *overfitting*.

Model klasifikasi tersebut juga telah berhasil diimplementasikan ke dalam sebuah sistem aplikasi berbasis web fungsional menggunakan arsitektur *backend* Flask dan *frontend* Streamlit. Penyebaran sistem dilakukan melalui infrastruktur *Google Cloud Platform* (GCP) yang dilengkapi dengan fitur generator dokumen PDF untuk mencetak laporan rekam medis secara otomatis. Integrasi teknologi ini memberikan solusi digital konkret bagi Puskesmas Melintang dalam mempercepat proses skrining kardiovaskular sekaligus mengurangi beban administrasi pencatatan manual.

Kelayakan operasional dari aplikasi ini telah divalidasi secara komprehensif melalui uji penerimaan pengguna yang melibatkan dua belas responden dari berbagai latar belakang profesi. Hasil tabulasi kuesioner menunjukkan tingkat penerimaan yang sangat positif pada metrik kemudahan penggunaan, fungsionalitas, kualitas keluaran, serta utilitas sistem. Dengan demikian, arsitektur sistem deteksi dini hipertensi ini dinyatakan sangat layak dan siap dioperasikan oleh tenaga medis maupun masyarakat awam untuk mendukung transformasi digital layanan kesehatan.

DAFTAR PUSTAKA

- Aditya, M. F. R., Azizah, N. L., & Indahyanti, U. (2024). Prediksi penyakit hipertensi menggunakan metode Decision Tree dan Random Forest. *Jurnal Komputasi*, Universitas Nusa Mandiri.
- Alpaydin, E. (2024). *Introduction to machine learning* (4th ed.). MIT Press.
- Chicco, D., Tötsch, N., & Jurman, G. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 14, 56706. <https://doi.org/10.1038/s41598-024-56706-x>.
- Harahap et al. (2021). Hubungan Peran Keluarga Mengontrol Gaya Hidup Penderita Hipertensi.
- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An Introduction to Statistical Learning: with Applications in Python*. Springer.
- Kementerian Kesehatan Republik Indonesia. (2023). *Profil kesehatan Indonesia 2023*. Jakarta: Kementerian Kesehatan RI.
- Liu, Z., Chen, T., & Wang, X. (2021). Data preprocessing in healthcare machine learning: Challenges and solutions. *Computational Intelligence in Medicine*, 18(3), 101–118.
- Novianti, S., Alkadri, S. P. A., & Fakhruzi, I. (2023). Klasifikasi penyakit hipertensi menggunakan metode Random Forest. *Jurnal Progresif*, STMIK Banjarbaru.



- Nurjanah, A., Sari, L., Lanaya, D., & Djuria, R. F. (2022). Efektivitas Media Leaflet terhadap Pengetahuan Penggunaan Obat Antihipertensi di Puskesmas Melintang. *Jurnal Ilmu dan Teknologi Kesehatan Terpadu*, 3(1), 6–14. <https://doi.org/10.53579/jitkt.v3i1.75>.
- Pramesti Dewi, P., Purwono, & Kurniawan, S. D. (2022). Pemanfaatan teknologi machine learning pada klasifikasi jenis hipertensi berdasarkan fitur pribadi. *Jurnal SMARTCOMP*, Politeknik Harapan Bersama.
- Radzi, N. M., Harun, M. H., & Othman, M. (2022). User Acceptance Testing (UAT) Framework for web-based applications in the healthcare sector. *IEEE Access*, 10, 45123-45135.
- Septian, E., Khaefi, M. R., Athoillah, A., Aisyah, D. N., Hardhantyo, M., & Rahman, F. M. (2025). Prediction of Personalised Hypertension Using Machine Learning in Indonesian Population. *PubMed Central (PMC)*.
- Singh, P. (2021). *Deploy machine learning models to production: With Flask, Streamlit, Docker, and Kubernetes on Google Cloud Platform*. Apress.
- Tharwat, A. (2021). Classification assessment methods: a detailed tutorial. *Applied Computing and Informatics*, 17(1), 168-192.
- World Health Organization. (2023). *Hypertension fact sheet 2023*. Geneva: WHO Press.
- Yusup, R. M., & Rijanto, E. (2025). Analisis komparatif model pembelajaran mesin untuk memprediksi hipertensi ke dalam empat kelas berdasarkan JNC 8. *Jurnal JTK3TI*, Universitas Komputer Indonesia.