

## Evaluation of Machine Learning Models for Classifying Diabetes and Hypertension Using Patient Data from Public Health Center X

Febriyan Biopsa Minanda <sup>1\*</sup>, Anandha Widya Putri Rahmadhina <sup>2\*\*</sup>

\* Program Studi Informatika, Universitas Atma Jaya Yogyakarta

\*\* Program Studi Kesehatan Masyarakat, Universitas Mitra Indonesia  
[biopza27@gmail.com](mailto:biopza27@gmail.com) <sup>1</sup>, [anandawidya2811@gmail.com](mailto:anandawidya2811@gmail.com) <sup>2</sup>

### Article Info

#### Article history:

Received 2026-04-07

Revised 2026-05-13

Accepted 2026-05-25

#### Keyword:

*Diabetes Mellitus,  
Hypertension, Machine  
Learning, Ensemble learning,  
Classification*

### ABSTRACT

Major non-communicable diseases such as hypertension and diabetes mellitus continue to increase and require accurate and rapid early detection. This study aimed to evaluate the performance of several machine learning algorithms in classifying diabetes mellitus and hypertension using patient medical record data from a public health center in Bandar Lampung City. The dataset consisted of 662 patient records including age, gender, body mass index, blood glucose level, and blood pressure variables. The research stages included exploratory data analysis, data preprocessing, outlier handling using Interquartile Range (IQR), normalization using StandardScaler, and data splitting scenarios of 90:10, 80:20, and 70:30. A 10-fold cross validation approach was also applied to improve model validity. The algorithms tested were Random Forest, Support Vector Machine (SVM), Decision Tree, AdaBoost Classifier, and XGBoost Classifier. Evaluation results using accuracy, recall, precision, f1-score, and confusion matrix showed that Decision Tree achieved the best performance with 95% accuracy in the 80:20 scenario. The best model was implemented on the Streamlit platform and evaluated using the System Usability Scale (SUS) involving 23 respondents, obtaining an average score of 73.48 categorized as acceptable and good. The findings indicate that machine learning has the potential to support early screening and healthcare decision-making.



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

### I. PENDAHULUAN

Diabetes dan hipertensi termasuk dalam penyakit tidak menular yang prevalensinya meningkat baik secara global maupun di AS. Penyakit kronis yang dikenal sebagai hipertensi, ditandai dengan tekanan darah tinggi yang konsisten, seringkali salah didiagnosis dan karenanya disebut sebagai “pembunuh diam-diam”. Sebaliknya, peningkatan kadar glukosa darah akibat resistensi insulin merupakan ciri diabetes melitus, suatu penyakit metabolik [1, 2]. Komplikasi serius termasuk penyakit jantung, stroke, dan gagal ginjal dapat berkembang jika gangguan ini tidak dikelola dengan baik [4].

Diabetes dan hipertensi yang terjadi bersamaan dikaitkan dengan peningkatan beban penyakit dan risiko kematian bagi

pasien [4]. Karena situasi ini, kedua penyakit ini terutama menjadi target dalam upaya memerangi dan mengendalikan penyakit tidak menular [5]. Karenanya, diperlukan rencana yang efisien untuk meningkatkan manajemen penyakit dan deteksi dini [6].

Indonesia tidak kebal terhadap epidemi global (DM). berlandaskan (IDF) [7], Indonesia memiliki jumlah pasien diabetes tertinggi kelima di dunia pada tahun 2021. Dari 589 juta pada tahun 2024, jumlah orang yang hidup dengan diabetes diproyeksikan meningkat menjadi 783 juta pada tahun 2045. Perkiraan memperlihatkan kawasan Pasifik Barat akan mengalami peningkatan dari 215 juta kasus pada tahun 2024 menjadi 254 juta kasus pada 2050. Dari angka terendah sekitar 1,5% pada kelompok usia 20-44 tahun, prevalensi diabetes di Indonesia terus meningkat seiring bertambahnya

usia, mencapai > 20% pada kelompok usia 65+ tahun [1]. Sementara itu, prevalensi hipertensi di Indonesia mencapai 34,11% berlandaskan (SKI) tahun 2023, serta tren peningkatan juga mulai terjadi pada kelompok usia muda sehingga menjadi perhatian serius dalam kesehatan Masyarakat [2].

Meskipun prevalensi diabetes melitus dan hipertensi terus meningkat, proses deteksi dini pada pelayanan kesehatan primer masih menghadapi berbagai kendala. Keterbatasan kualitas dan kelengkapan data medis sering menyebabkan rendahnya akurasi dalam penentuan diagnosis penyakit [3]. Selain itu, banyaknya kasus tanpa gejala membuat penyakit tidak teridentifikasi sejak awal sehingga berisiko menimbulkan komplikasi yang lebih serius [4]. Kondisi tersebut menunjukkan bahwa pendekatan konvensional belum sepenuhnya mampu memberikan hasil diagnosis yang cepat dan tepat.

Sebagai solusi, pemanfaatan *machine learning* menjadi pendekatan yang efektif dalam meningkatkan akurasi klasifikasi dan prediksi penyakit berbasis data. Komputer ini dapat menyaring kumpulan data yang sangat besar dengan cara yang tidak dapat dilakukan manusia [11]. Salah satu cara untuk meningkatkan sistem pendukung keputusan klinis adalah dengan menggunakan algoritma klasifikasi pada data kesehatan [12]. Akibatnya, pembelajaran mesin suatu hari nanti dapat membantu dokter membuat keputusan yang lebih baik dan mendeteksi penyakit lebih awal.

Banyak penelitian sebelumnya telah menunjukkan bahwa pembelajaran mesin memiliki potensi besar untuk meningkatkan ketepatan kategorisasi penyakit kronis. Diabetes dan tekanan darah tinggi termasuk di antaranya. *Support Vector Machine* (SVM) mengungguli *Naive Bayes* dengan tingkat akurasi 93,37% dalam klasifikasi hipertensi, menurut penelitian yang menggunakan data pasien dari klinik kesehatan masyarakat [13]. Melalui penggunaan 532 kumpulan data, penelitian lebih lanjut mengungkapkan bahwa beberapa algoritma klasifikasi mengungguli yang lain. Sebagai contoh, Regresi Logistik Biner mencapai sensitivitas 93,33%, berbeda dengan *Naive Bayes* yang sebesar 63,64% [14]. Dengan menggunakan 768 rekam medis pasien, algoritma SVM menunjukkan efisiensi yang luar biasa dalam kategorisasi diabetes (tingkat akurasi 71%, dikutip dalam [15]). Akurasi dan presisi model dalam klasifikasi gangguan keduanya ditingkatkan ketika dilatih dengan *Deep Neural Network* (DNN) [16]. Meskipun strategi pemilihan fitur memengaruhi akurasi, penelitian lebih lanjut menunjukkan bahwa SVM melampaui sejumlah algoritma lain, termasuk *Naive Bayes*, *Random Forest*, serta *AdaBoost* [17].

Dengan pertimbangan tersebut, studi mempergunakan data pasien yang dikumpulkan dari biro kesehatan masyarakat untuk membandingkan akurasi beberapa algoritma pembelajaran mesin dalam mengidentifikasi kasus diabetes melitus serta hipertensi. Visi studi guna menetapkan pendekatan terbaik dengan membandingkan kinerja algoritma. *Recall*, akurasi, presisi, *f1-score*, dan matriks kebingungan adalah beberapa metrik dipergunakan penilaian

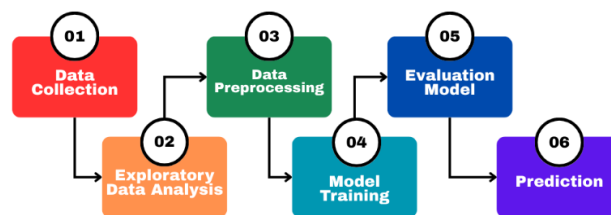
untuk menemukan akurasi klasifikasi. Temuan evaluasi harus mengungkapkan kelebihan serta kekurangan relatif algoritma. Penilaian kinerja yang ketat mengarah pada pemilihan model terbaik.

Studi diharapkan berkontribusi dalam pengembangan sistem pendukung keputusan di pelayanan kesehatan. Model yang dihasilkan akan diimplementasikan dalam aplikasi prediksi berbasis *web* menggunakan *platform Streamlit*. Pendekatan tersebut diharapkan mampu meningkatkan deteksi dini serta kualitas pengambilan keputusan dalam penanganan penyakit tidak menular. Sistem diperkembangkannya harapannya bisa dipakai jadi instrumen bantu dalam skrining awal pasien. Selain itu, temuan studi bisa jadi referensi dalam pengembangan teknologi kesehatan berbasis *machine learning* di masa mendatang.

## II. METODE

### A. Tahap Penelitian

Penelitian yang dilakukan disusun secara sistematis untuk memastikan proses analisis berjalan terstruktur dan terukur. Setiap tahapan saling berkaitan dan membentuk alur kerja yang berkesinambungan. Pendekatan tersebut digunakan guna hasilkan model prediksi akurat serta dapat dipertanggungjawabkan.



Gambar 1. Tahapan Penelitian

Gambar 1 gambarkan alur fase studi dilakukan secara berkesinambungan. Setiap tahapan merepresentasikan proses metodologis dalam pengolahan data hingga pembentukan model prediktif. Proses studi diawali pengumpulan data serta akhirnya tahap prediksi hasil:

- 1) *Data Collection* : Tahap pengumpulan data yang relevan dan representatif dari berbagai sumber sesuai dengan kebutuhan penelitian.
- 2) *Exploratory Data Analysis* : Tahap analisis eksploratif untuk mengidentifikasi karakteristik data, pola distribusi, serta anomali yang mungkin muncul.
- 3) *Data Preprocessing* : Tahap praproses data yang meliputi pembersihan, normalisasi, dan transformasi data agar layak digunakan dalam pemodelan.
- 4) *Model Training* : Tahap pelatihan model dengan memanfaatkan data yang telah diproses guna mempelajari pola dan hubungan antarvariabel.
- 5) *Evaluation Model* : Tahap evaluasi performa model menggunakan metrik pengukuran tertentu untuk menilai tingkat akurasi dan keandalan model.

- 6) *Prediction* : Tahap penerapan model terpilih untuk menghasilkan prediksi berdasarkan data baru belum pernah dipakai.

### B. Data Collection

Data dipergunakan studi ialah data primer didapat langsung dari salah satu puskesmas Kota Bandar Lampung. Pengumpulan data dilaksanakan atas catatan RM pasien berkorelasi beserta pemeriksaan penyakit tidak menular, khususnya diabetes melitus dan hipertensi. Data yang diperoleh mencakup informasi identitas dasar pasien serta hasil pemeriksaan klinis yang relevan untuk proses klasifikasi penyakit.

Dalam penggunaan data medis, penelitian memperhatikan aspek etika dan privasi dengan memperoleh izin dari pihak puskesmas, menjaga kerahasiaan informasi pasien, serta tidak mencantumkan identitas pribadi seperti nama, alamat, dan nomor rekam medis pada *dataset* yang dianalisis. Seluruh data digunakan hanya untuk kepentingan penelitian dan disajikan dalam bentuk anonim agar kerahasiaan pasien tetap terlindungi. Penggunaan data tersebut bertujuan untuk memastikan kesesuaian data dengan kondisi nyata di lapangan sehingga dapat meningkatkan validitas hasil penelitian. Adapun deskripsi *dataset* yang digunakan adalah sebagai berikut:

TABEL I  
DESKRIPSI DATASET

Nama	Deskripsi	Nilai
Tanggal Lahir	Tanggal lahir pasien yang digunakan untuk mengetahui usia pasien	Datetime
Jenis Kelamin	Jenis kelamin pasien	Lelaki atau Perempuan
TB	Tinggi badan pasien dalam satuan centimeter	Numerik (cm)
BB	Berat badan pasien dalam satuan kilogram	Numerik (kg)
GDP	Gula Darah Puasa pasien	Numerik (mg/dL)
GDPP	Gula Darah Post Prandial (setelah makan) pasien	Numerik (mg/dL)
Sitole	Tekanan darah sistolik pasien	Numerik (mmHg)
Ditole	Tekanan darah diastolik pasien	Numerik (mmHg)
Diagnosa	Hasil diagnosis pasien	Diabetes, Hipertensi, atau Komplikasi (Diabetes dan Hipertensi)

Berlandaskan Tabel 1, *dataset* yang digunakan dalam penelitian memuat sejumlah variabel yang merepresentasikan karakteristik identitas pasien serta hasil pemeriksaan klinis yang berkaitan dengan diabetes melitus dan hipertensi. Variabel Tanggal Lahir digunakan untuk mengetahui usia pasien dan disajikan dalam format *datetime*, sedangkan

variabel Jenis Kelamin menunjukkan kategori pasien yaitu laki-laki atau perempuan. Variabel antropometri terdiri atas TB dan BB, yang masing-masing menunjukkan tinggi badan pasien dalam satuan *centimeter* (cm) dan berat badan pasien dalam satuan kilogram (kg) dengan tipe data numerik.

Selanjutnya, variabel klinis meliputi GDP serta GDPP sebagai, keduanya disajikan dalam bentuk numerik dengan satuan mg/dL. Selain itu, variabel tekanan darah terdiri atas *Sitole* perhatikan tekanan darah sistolik serta *Ditole* yang menunjukkan tekanan darah diastolik, dengan tipe data numerik dalam satuan mmHg. Adapun variabel Diagnosa berperan sebagai label atau target klasifikasi yang menunjukkan hasil akhir pemeriksaan pasien.

Dalam penelitian ini, diagnosis tidak diklasifikasikan secara terpisah sebagai *multi-label*, melainkan menggunakan pendekatan *multi-class classification* dengan tiga kelas diagnosis yang saling eksklusif, yaitu diabetes melitus, hipertensi, dan komplikasi. Kelas komplikasi merupakan kategori gabungan bagi pasien yang mengalami diabetes melitus dan hipertensi secara bersamaan. Dengan demikian, setiap pasien hanya memiliki satu label diagnosis akhir, sehingga keseluruhan variabel pada Tabel 1 dapat digunakan sebagai dasar dalam proses klasifikasi menggunakan algoritma *machine learning*. Contoh data yang digunakan dalam penelitian disajikan pada Gambar 2:

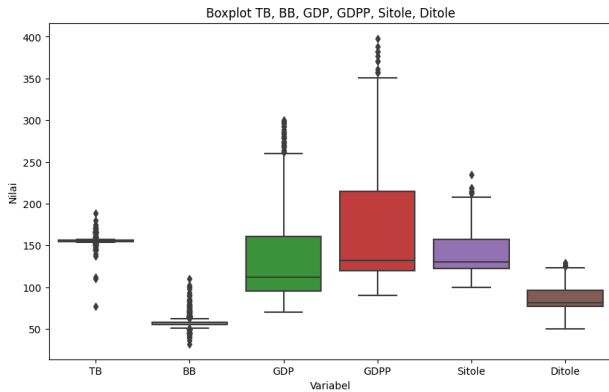
No	Tanggal Lahir	Jenis Kelamin	TB	BB	GDP	GDPP	Sitole	Ditole	Diagnosa
1	11/05/1962	Laki-Laki	156	59	157	182	110	73	Diabetes
2	09/12/1955	Laki-Laki	165	60	106	107	130	90	Hipertensi
3	12/11/1962	Laki-Laki	156	59	110	120	195	109	Hipertensi
4	27/10/1968	Perempuan	150	50	140	155	110	70	Diabetes
5	24/08/1957	Perempuan	156	59	270	287	189	100	Diabetes dan Hipertensi
8	05/03/1990	Perempuan	155	55	116	123	122	77	Hipertensi
9	20/04/1964	Laki-Laki	160	60	216	223	120	80	Diabetes dan Hipertensi
10	23/03/1963	Laki-Laki	144	44	229	250	128	86	Diabetes dan Hipertensi
...	...	...	...	...	...	...	...	...	...
663	08/12/1981	Perempuan	156	59	100	123	164	101	Hipertensi

Gambar 2. Sampel Data Mentah

Berdasarkan Gambar 2 *dataset* dipergunakan studi meliputi 662 data pasien beserta 9 kolom variabel meliputi informasi identitas serta hasil pemeriksaan klinis. Variabel yang tersedia meliputi Tanggal Lahir, Jenis Kelamin, TB, BB, GDP, GDPP, *Sitole*, *Ditole*, serta Diagnosa sebagai label klasifikasi. Data menunjukkan variasi karakteristik pasien dengan nilai numerik pada parameter antropometri dan klinis, serta kategori diagnosis yang terdiri dari diabetes, hipertensi, dan komplikasi (diabetes dan hipertensi). Struktur dataset tersebut memungkinkan penggunaan sebagai data latih serta uji saat klasifikasi memakai algoritma *machine learning* untuk mengidentifikasi pola penyakit secara lebih akurat.

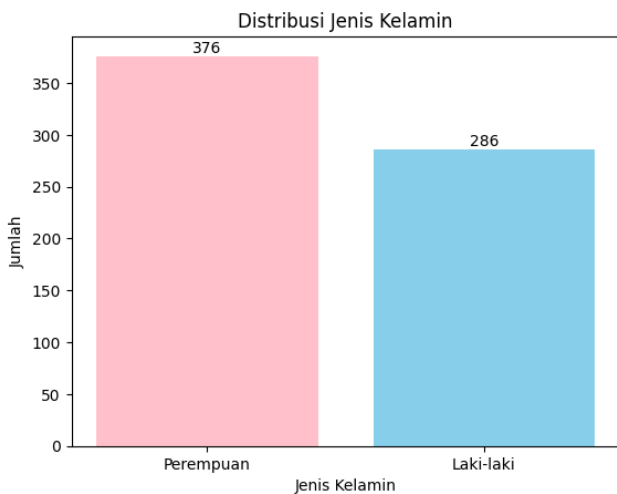
### C. Exploratory Data Analysis

Eksplorasi data dilaksanakan guna mendapati gambaran awalnya terkait ciri *dataset* yang digunakan dalam penelitian. Proses berikut bertujuan untuk melihat pola distribusi, kecenderungan nilai, serta potensi adanya anomali dalam data. Selain itu, tahapan EDA menjadi dasar dalam menentukan langkah pengolahan data sebelum proses pemodelan dilakukan.



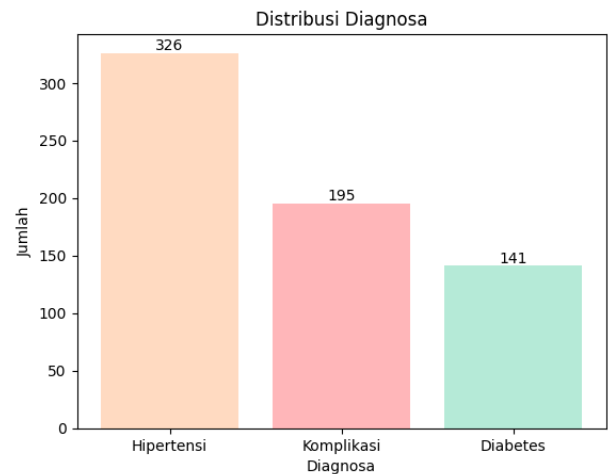
Gambar 3. Boxplot TB, BB, GDP, GDPP, Sitole, Ditole

Visualisasi pada Gambar 3 memperlihatkan penyebaran nilai dari variabel TB, BB, GDP, GDPP, Sitole, dan Ditole yang menunjukkan variasi antar data. Nilai GDP dan GDPP tampak memiliki rentang yang lebih lebar dibandingkan variabel lainnya, yang mengindikasikan perbedaan kondisi kadar gula darah pada pasien. Variabel tekanan darah, baik sistolik maupun diastolik, juga menunjukkan distribusi yang cukup menyebar dengan nilai tengah berada pada kisaran tertentu. Selain itu, masih terdapat *outlier* pada beberapa variabel, terutama pada TB, BB, GDP dan GDPP, yang menunjukkan adanya nilai ekstrem dalam *dataset*.



Gambar 4. Distribusi Jenis Kelamin

Pada Gambar 4, terlihat sebaran gender tak seimbang lelaki serta perempuannya. Jumlah pasien perempuan tercatat lebih banyak dibandingkan laki-laki, yaitu sebanyak 376 data, sedangkan laki-laki berjumlah 286 data. Kondisi tersebut menunjukkan bahwa *dataset* didominasi oleh pasien perempuan.



Gambar 5. Distribusi Diagnosa

Gambar 5 menunjukkan bahwa distribusi diagnosis didominasi oleh kasus hipertensi dibandingkan kategori lainnya. Jumlah pasien dengan diagnosis hipertensi mencapai 326 data, sementara kategori komplikasi sebanyak 195 data dan diabetes sebanyak 141 data. Hal ini mengindikasikan bahwa hipertensi menjadi kondisi yang paling banyak ditemukan dalam *dataset* penelitian.

**D. Data Preprocessing**

Tahap *preprocessing* data guna menyiapkan *dataset* supaya lebih siap dipakai saat permodelan *machine learning*. Dilaksanakannya transformasi terhadap beberapa variabel guna meningkatkan relevansi fitur yang digunakan. Kolom tanggal lahir diubah menjadi variabel Umur untuk merepresentasikan usia pasien secara langsung tanpa perlu perhitungan ulang, sedangkan variabel TB serta BB (dikonversi jadi IMT) jadi indikator yang lebih representatif dalam menilai kondisi fisik pasien. Proses transformasi tersebut bertujuan untuk menyederhanakan data sekaligus meningkatkan kualitas fitur yang digunakan dalam analisis. Sampel hasil *preprocessing* data ditampilkan pada Tabel 2.

TABEL II  
SAMPLE DATA PREPROCESSING

Umur	JK	IMT	GDP	GDPP	Sitole	Ditole	Diagnosa
63	L	24,14	157	182	110	73	DM
70	L	22,04	106	107	130	90	HT
63	L	24,14	110	120	195	109	HT
57	P	22,22	140	155	110	70	SM
68	P	24,14	270	287	189	100	KP
36	P	22,89	116	123	122	77	HT
61	L	23,44	216	223	120	80	KP
63	L	21,22	229	250	128	86	KP
...	...	...	...	...	...	...	...
44	P	24,14	100	123	164	101	HT

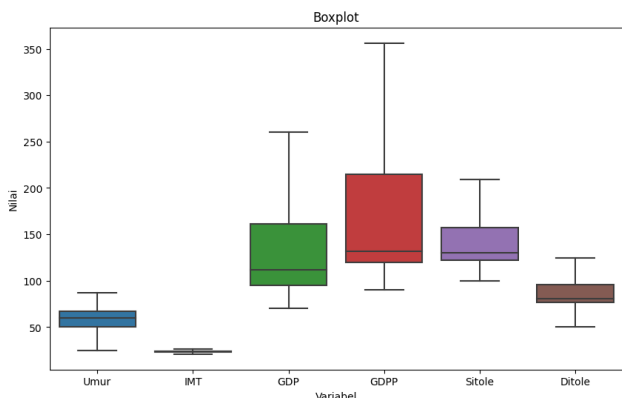
JK: Jenis Kelamin, L: Lelaki, P: Perempuan

IMT: Indeks Massa Tubuh

DM: Diabetes, HT: Hipertensi, KP: Komplikasi (Diabetes dan Hipertensi)

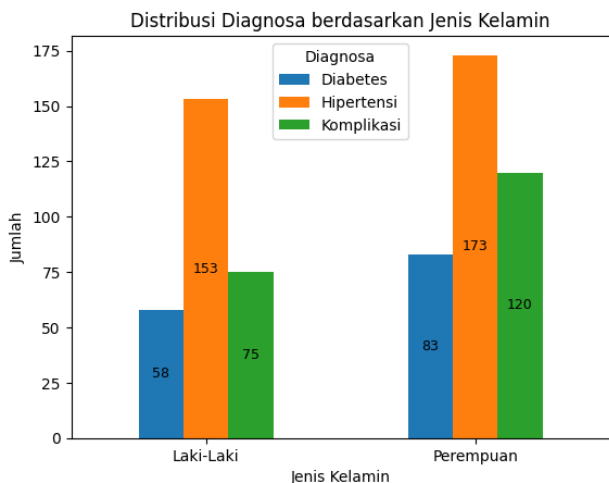
Penanganan *outlier* dilakukan untuk mengurangi pengaruh nilai ekstrem bisa jadi keberpengaruhannya temuan analisis serta performanya *machine learning*. Metode digunakan adalah

pendekatan berbasis *Interquartile Range* (IQR) dengan membatasi nilai data agar tetap berada dalam kisaran batas bawah serta atas sudah ditetapkan. Proses tersebut diterapkan pada beberapa variabel numerik seperti umur, IMT, GDP, GDPP, tekanan darah *sitole*, dan *ditole* sehingga distribusi data menjadi lebih stabil dan representatif. Pendekatan yang dilakukan tidak menghapus data, melainkan menyesuaikan nilai ekstrem agar tetap berada dalam batas yang wajar sehingga informasi penting dalam *dataset* tetap terjaga. Hasil visualisasi *boxplot* setelah proses penanganan *outlier* ditampilkan pada Gambar 6.



Gambar 6. Boxplot Umur, IMT, GDP, GDPP, Sitole, Ditole

Visualisasi pada Gambar 6 menunjukkan bahwa distribusi data pada setiap variabel menjadi lebih stabil dibandingkan sebelumnya. Rentang nilai antar variabel terlihat lebih terkontrol tanpa adanya penyimpangan yang ekstrem. Selain itu, data sudah tidak memiliki *outlier* sehingga sebaran nilai menjadi lebih representatif untuk proses selanjutnya.



Gambar 7. Distribusi Diagnosa berdasarkan Jenis Kelamin

Pada Gambar 7, terlihat bahwa distribusi diagnosis berbeda antara laki-laki dan perempuan. Jumlah kasus hipertensi menjadi yang paling dominan pada kedua kelompok, yaitu sebanyak 153 kasus pada laki-laki dan 173 kasus pada perempuan. Selain itu, kasus komplikasi juga menunjukkan

jumlah yang cukup tinggi, yaitu 75 kasus pada laki-laki dan 120 kasus pada perempuan. Sementara itu, kasus diabetes memiliki jumlah paling sedikit, yaitu 58 kasus pada laki-laki dan 83 kasus pada perempuan, meskipun tetap menunjukkan perbedaan distribusi antar jenis kelamin.

*E. Model Training*

Tahap model *training* diawali dengan proses *preprocessing* data yang mencakup *encoding* variabel kategorikal dan normalisasi data numerik. Variabel kategorikal seperti jenis kelamin terlebih dahulu dikonversi ke bentuk numerik menggunakan teknik label *encoding* agar dapat diproses oleh algoritma *machine learning*. Selanjutnya, data distandarisasi menggunakan teknik *StandardScaler* sebelum memulai fase pelatihan model untuk memastikan bahwa semua fitur berada pada skala yang sama. Kinerja optimal model pembelajaran mesin dipastikan di seluruh rentang nilai yang mungkin variabel dengan mengikuti proses ini. Faktor yang penting klasifikasi meliputi usia, jenis kelamin, BMI, PDB, sitole, dan ditole, fitur-fitur ini digunakan untuk penskalaan. Data dianggap dinormalisasi ketika distribusinya diselarskan membuat deviasi standar jadi satu serta rerata mendekati nol.

Guna mulai pelatihan model, dataset dipisahkan menjadi set pelatihan dan pengujian sesuai dengan tiga skenario perbandingan yang mungkin: 90:10, 80:20, dan 70:30. Dengan mengevaluasi ketahanan model terhadap berbagai proporsi data, bagian bermaksud memberikan hasil yang lebih konsisten serta optimal. Setiap skenario dirancang untuk melatih model serta menguji kemampuannya untuk mengklasifikasikan data baru. Selain itu, untuk meningkatkan validitas dan mengurangi potensi bias hasil pengujian, penelitian ini juga menerapkan metode *cross validation* menggunakan *10-fold cross validation*. Pembagian *dataset* tiap contoh diperlihatkan Tabel 3.

TABEL III  
SKENARIO PEMBAGIAN DATASET

Skenario	Data Latih	Data Uji
90:10	595 data	67 data
80:20	592 data	133 data
70:30	463 data	199 data

Proses *tuning* model dilakukan melalui skenario *hyperparameter* pada sebagian algoritma *machine learning*, seperti *Random Forest*, *SVM*, *Decision Tree*, *AdaBoost Classifier*, serta *XGBoost Classifier*. Pengujian dilakukan dengan mencoba berbagai kombinasi parameter untuk memperoleh konfigurasi terbaik yang mampu meningkatkan performa model dalam klasifikasi nyadata. Tiap algoritma ada parameter beda sehingga proses pencarian kombinasi optimal dilakukan secara sistematis untuk menghasilkan model dengan akurasi terbaik. Skenario *hyperparameter* diperlukan studi diTabel 4.

Beragam pendekatan pembelajaran mesin, tiapnya dengan serangkaian parameter diselarskan kinerja model, termasuk dalam skenario hiperparameter studi ini (Tabel 4). *Random*

*Forest* ialah metode mempertimbangkan banyak variabel. Di antara pengaturan ini ialah kedalaman pohon maksimum, jumlah estimator, distribusi sampel node, serta ukuran sampel *split* serta *leaf* minimum. Pendekatan SVM menggunakan kernel fungsi basis radial (*rbf*), menyesuaikan kompleksitas model beserta penyesuaian pada parameter C, serta menilai pengaruh setiap titik data memakai  $\gamma$ .

TABEL IV  
SKENARIO HYPERPARAMETER MODEL

Model	Parameter	Nilai
Random Forest	<i>n_estimators</i>	100, 200, 300
	<i>max_depth</i>	None, 5, 10, 20
	<i>min_samples_split</i>	2, 5, 10
	<i>min_samples_leaf</i>	1, 2, 4
Support Vector Machine	C	0.1, 1, 10
	<i>gamma</i>	scale, 0.1, 0.01, 0.001
Decision Tree	<i>kernel</i>	Rbf
	<i>criterion</i>	<i>gini</i> , <i>entropy</i> , <i>log_loss</i>
	<i>max_depth</i>	None, 5, 10, 20
	<i>min_samples_split</i>	2, 5, 10
AdaBoost Classifier	<i>min_samples_leaf</i>	1, 2, 4
	<i>max_features</i>	None, <i>sqrt</i> , <i>log2</i>
	<i>n_estimators</i>	50, 100, 200
	<i>learning_rate</i>	0.01, 0.1, 1, 10
XGBoost Classifier	<i>algorithm</i>	SAMME, SAMME.R
	<i>n_estimators</i>	50, 100, 200
	<i>learning_rate</i>	0.01, 0.1, 3
	<i>max_depth</i>	3, 5, 7
XGBoost Classifier	<i>subsample</i>	0.8, 1.0
	<i>colsample_bytree</i>	0.8, 1.0

Jika melaksanakan eksperimen dengan metode Pohon Keputusan, memakai beragam parameter guna atur jumlah fitur yang digunakan tiap cabang, seperti fitur maksimum, split sampel minimum, kedalaman maksimum, dan leaf sampel minimum. Sementara itu, *AdaBoost Classifier* menggunakan parameter *n\_estimators*, *learning\_rate*, dan *algorithm* untuk mengatur proses *boosting* dalam meningkatkan performa model. Adapun *XGBoost Classifier* menggunakan parameter *n\_estimators*, *learning\_rate*, *max\_depth*, serta *subsample* dan *colsample\_bytree*. Variasi parameter tersebut bertujuan untuk menemukan kombinasi terbaik yang mampu menghasilkan kinerja klasifikasi yang paling optimal.

### III. HASIL DAN PEMBAHASAN

#### A. Evaluation Model

Dilaksanakan guna mengukur kinerja algoritma dalam mengklasifikasikan data berdasarkan parameter sudah ditetapkan Tabel 3. Tes dilakukan beberapa skenario pembagian data guna mengetahui kestabilan performa model yang juga telah ditentukan pada Tabel 4. Temuan evaluasinya

ditampilkan bentukan metrik akurasi, *recall*, *precision*, serta *f1-score* yang disajikan Tabel 5.

TABEL V  
HASIL EVALUASI MODEL

Skenario	Model	Best Param	Evaluation			
			Acc	Rc	Pc	F1
90:10	RF	<i>max_depth</i> : none, <i>min_samples_leaf</i> : 1, <i>min_samples_split</i> : 5, <i>n_estimators</i> : 100	0.84	0.83	0.82	0.82
	SVM	C: 10, <i>gamma</i> : scale, <i>kernel</i> : rbf	0.82	0.82	0.81	0.81
	DC	<i>criterion</i> : <i>gini</i> , <i>max_depth</i> : none, <i>max_features</i> : none, <i>min_samples_leaf</i> : 1, <i>min_samples_split</i> : 2	0.87	0.86	0.86	0.86
	ADA	<i>algorithm</i> : SAMME, <i>learning_rate</i> : 1, <i>n_estimators</i> : 50	0.84	0.83	0.82	0.82
	XGB	<i>colsample_bytree</i> : 0.8, <i>learning_rate</i> : 0.01, <i>max_depth</i> : 3, <i>n_estimators</i> : 100, <i>subsample</i> : 0.8	0.79	0.81	0.78	0.79
	80:20	RF	<i>max_depth</i> : none, <i>min_samples_leaf</i> : 1, <i>min_samples_split</i> : 2, <i>n_estimators</i> : 200	0.94	0.94	0.93
SVM		C: 0.1, <i>gamma</i> : scale, <i>kernel</i> : rbf	0.94	0.93	0.93	0.93
DC		<i>criterion</i> : <i>gini</i> , <i>max_depth</i> : none, <i>max_features</i> : none, <i>min_samples_leaf</i> : 1, <i>min_samples_split</i> : 2	0.95	0.95	0.95	0.95

	ADA	<i>algorithm:</i> SAMME, <i>learning rate:</i> 0.1, <i>n estimators:</i> 200	0.94	0.94	0.93	0.93
	XGB	<i>colsample bytree:</i> 0.8, <i>learning rate:</i> 0.3, <i>max depth:</i> 3, <i>n estimators:</i> 50, <i>subsample:</i> 0.8	0.94	0.94	0.93	0.93
70:30	RF	<i>max depth:</i> none, <i>min samples leaf:</i> 1, <i>min samples split:</i> 2, <i>n estimators:</i> 200	0.92	0.91	0.90	0.90
	SVM	<i>C:</i> 1, <i>gamma:</i> scale, <i>kernel:</i> rbf	0.91	0.90	0.89	0.90
	DC	<i>criterion:</i> gini, <i>max depth:</i> none, <i>max features:</i> none, <i>min samples leaf:</i> 1, <i>min samples split:</i> 2	0.93	0.92	0.92	0.92
	ADA	<i>algorithm:</i> SAMME, <i>learning rate:</i> 0.1, <i>n estimators:</i> 200	0.92	0.91	0.90	0.90
	XGB	<i>colsample bytree:</i> 0.8, <i>learning rate:</i> 0.3, <i>max depth:</i> 3, <i>n estimators:</i> 50, <i>subsample:</i> 0.8	0.91	0.91	0.90	0.90

**RF:** Random Forest, **SVM:** Support Vector Machine, **DC:** Decision Tree, **ADA:** AdaBoost Classifier, **XGB:** XGBoost Classifier  
**Acc:** Accuracy, **Rc:** Recall, **Pc:** Precision, **F1:** F1-Score

Berdasarkan Tabel 5, terlihat bahwa performa setiap model menunjukkan perbedaan pada masing-masing skenario pembagian data. Pada skenario 90:10, model dengan performa terbaik adalah *Decision Tree* dengan nilai akurasi sebesar 0,87, *recall* 0,86, *precision* 0,86, dan *f1-score* 0,86. *Random Forest* dan *AdaBoost* menunjukkan hasil yang sama dengan akurasi sebesar 0,84, sedangkan *SVM* memperoleh akurasi 0,82. Sementara itu, *XGBoost* menghasilkan akurasi paling rendah pada skenario ini, yaitu sebesar 0,79. Hasil

tersebut menunjukkan bahwa pada jumlah data uji yang lebih kecil, *Decision Tree* mampu memberikan performa terbaik dibandingkan model lainnya.

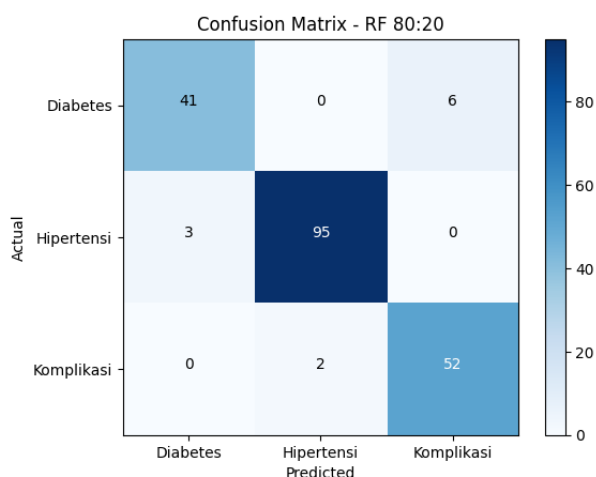
Pada skenario 80:20, seluruh model mengalami peningkatan performa yang cukup signifikan. *Decision Tree* kembali menunjukkan hasil terbaik dengan nilai akurasi, *recall*, *precision*, dan *f1-score* masing-masing sebesar 0,95. *Random Forest*, *SVM*, *AdaBoost*, dan *XGBoost* juga menunjukkan performa yang baik dengan akurasi sebesar 0,94. Nilai *recall*, *precision*, dan *f1-score* pada keempat model tersebut berada pada rentang 0,93 hingga 0,94. Hal ini menunjukkan bahwa skenario 80:20 mampu memberikan keseimbangan yang baik antara jumlah data latih dan data uji, sehingga model dapat mempelajari pola data secara optimal serta tetap mampu diuji dengan jumlah data yang cukup representatif.

Pada skenario 70:30, performa model mengalami sedikit penurunan dibandingkan skenario 80:20, namun hasilnya masih tergolong stabil. *Decision Tree* tetap memperoleh performa tertinggi dengan akurasi sebesar 0,93, *recall* 0,92, *precision* 0,92, dan *f1-score* 0,92. *Random Forest* dan *AdaBoost* memperoleh akurasi sebesar 0,92, sedangkan *SVM* dan *XGBoost* memperoleh akurasi sebesar 0,91. Penurunan performa pada skenario ini dapat terjadi karena proporsi data latih menjadi lebih kecil dibandingkan skenario 80:20, sehingga model memiliki lebih sedikit data untuk mempelajari pola klasifikasi.

Secara keseluruhan, evaluasi model pada tiga skenario split menunjukkan bahwa *Decision Tree* memiliki performa paling baik dan stabil dibandingkan algoritma lainnya. Hal ini terlihat dari nilai akurasi *Decision Tree* yang selalu berada pada posisi tertinggi di setiap skenario, yaitu 0,87 pada 90:10, 0,95 pada 80:20, dan 0,93 pada 70:30. Skenario 80:20 menjadi pembagian data yang menghasilkan performa paling optimal karena hampir seluruh model memperoleh nilai evaluasi yang tinggi dan seimbang. Selain itu, nilai *recall*, *precision*, dan *f1-score* yang relatif konsisten menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi yang akurat, tetapi juga memiliki kemampuan klasifikasi yang cukup baik pada setiap kelas. Hasil *confusion matrix* dari masing-masing model ditampilkan pada Gambar 8 hingga Gambar 12.

#### 1) Random Forest

Model *Random Forest* menunjukkan kinerja paling optimal pada skenario pembagian data 80:20 dengan nilai akurasi sebesar 0,94. Hasil tersebut mengindikasikan bahwa *Random Forest* mampu mengenali pola data klinis secara efektif, terutama pada variabel yang berkaitan dengan kadar gula darah dan tekanan darah. Stabilitas performa yang ditunjukkan pada skenario 80:20 juga memperlihatkan bahwa model cukup baik dalam menangani variasi data serta mampu mempertahankan kinerja klasifikasi pada data uji. Visualisasi *confusion matrix* untuk *Random Forest* ditampilkan pada Gambar 8.

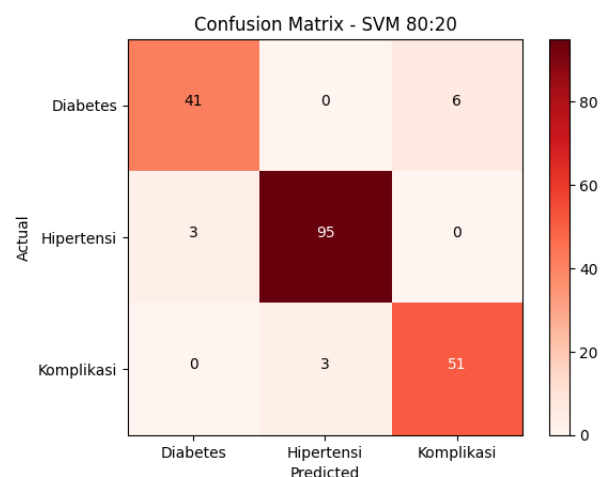


Gambar 8. Confusion Matrix Random Forest Ratio 80:20

*Confusion matrix* pada Gambar 8 menunjukkan bahwa model *Random Forest* mampu mengklasifikasikan sebagian besar data dengan tepat. Pada kelas diabetes, sebanyak 41 data berhasil diprediksi dengan benar, sedangkan 6 data masih salah diklasifikasikan sebagai komplikasi. Pada kelas hipertensi, model menunjukkan kemampuan deteksi yang sangat baik dengan 95 data berhasil diklasifikasikan dengan benar dan hanya 3 data yang salah diprediksi sebagai diabetes. Sementara itu, pada kelas komplikasi, sebanyak 52 data berhasil diprediksi dengan benar dan 2 data salah diklasifikasikan sebagai hipertensi. Hasil ini menunjukkan bahwa *Random Forest* memiliki kemampuan yang kuat dalam mendeteksi kelas hipertensi dengan tingkat kesalahan yang rendah, sekaligus tetap mampu membedakan kelas diabetes dan komplikasi secara cukup baik. Secara keseluruhan, kesalahan klasifikasi yang relatif kecil menunjukkan bahwa model memiliki performa yang baik dalam membedakan ketiga kategori diagnosis.

## 2) Support Vector Machine SVM

Pada algoritma *Support Vector Machine* (SVM), performa terbaik diperoleh pada skenario pembagian data 80:20 dengan nilai akurasi sebesar 0,94. Hasil tersebut memperlihatkan bahwa SVM memiliki kemampuan yang baik dalam membentuk batas pemisah antar kelas diagnosis berdasarkan pola data klinis yang digunakan. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi pada kelas tertentu, terutama antara kelas diabetes dan komplikasi yang memiliki karakteristik klinis saling beririsan. Visualisasi *confusion matrix* untuk SVM ditampilkan pada Gambar 9.

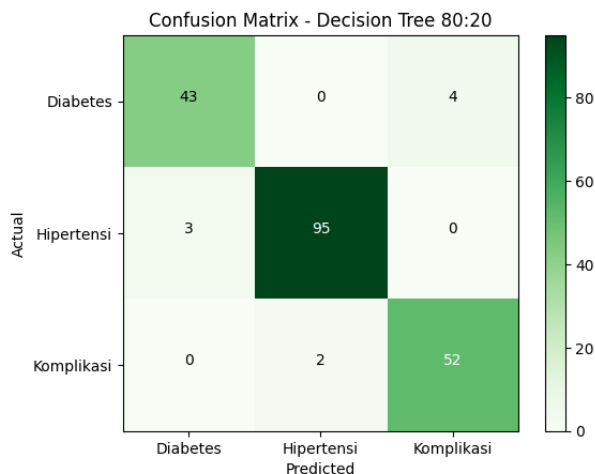


Gambar 9. Confusion Matrix Support Vector Machine Ratio 80:20

Berdasarkan Gambar 9, *confusion matrix* pada model SVM menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan tepat. Pada kelas diabetes, sebanyak 41 data berhasil diprediksi dengan benar, sedangkan 6 data masih salah diklasifikasikan sebagai komplikasi. Pada kelas hipertensi, model menunjukkan kemampuan deteksi yang sangat baik dengan 95 data berhasil diklasifikasikan secara benar dan hanya 3 data yang salah diprediksi sebagai diabetes. Sementara itu, pada kelas komplikasi, sebanyak 51 data berhasil diklasifikasikan dengan benar, sedangkan 3 data salah diprediksi sebagai hipertensi. Hasil ini menunjukkan bahwa SVM memiliki performa yang kuat dalam mendeteksi kelas hipertensi dengan tingkat kesalahan yang rendah. Namun, model masih mengalami sedikit kesulitan dalam membedakan kelas diabetes dan komplikasi, karena kedua kelas tersebut dapat memiliki kemiripan pola pada variabel gula darah dan tekanan darah.

## 3) Decision Tree

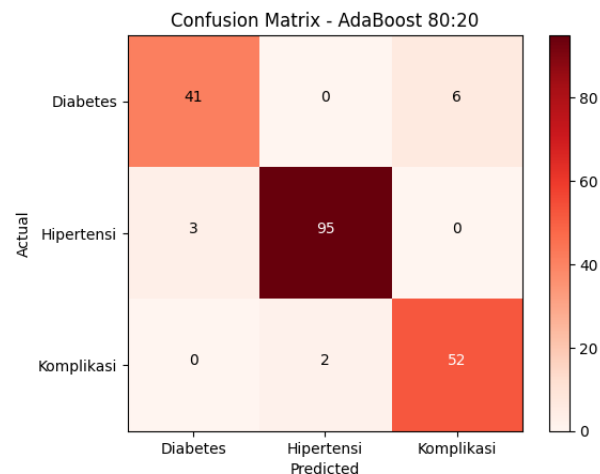
Algoritma *Decision Tree* mencapai performa terbaik pada skenario pembagian data 80:20 dengan nilai akurasi sebesar 0,95. Hasil tersebut memperlihatkan bahwa *Decision Tree* mampu membentuk aturan keputusan yang efektif berdasarkan pola variabel klinis, seperti kadar gula darah, tekanan darah, umur, dan IMT. Meskipun model *Decision Tree* berpotensi mengalami *overfitting*, hasil evaluasi pada skenario 80:20 menunjukkan bahwa model mampu memberikan performa yang optimal pada data uji. Visualisasi *confusion matrix* untuk *Decision Tree* ditampilkan pada Gambar 10.

Gambar 10. *Confusion Matrix Decision Tree Ratio 80:20*

*Confusion matrix* pada Gambar 10 memperlihatkan bahwa model *Decision Tree* mampu mengklasifikasikan sebagian besar data dengan tepat. Pada kelas diabetes, sebanyak 43 data berhasil diprediksi dengan benar, sedangkan 4 data salah diklasifikasikan sebagai komplikasi. Pada kelas hipertensi, model menunjukkan kemampuan deteksi yang sangat baik dengan 95 data berhasil diklasifikasikan secara benar dan hanya 3 data yang salah diprediksi sebagai diabetes. Sementara itu, pada kelas komplikasi, sebanyak 52 data berhasil diklasifikasikan dengan benar dan 2 data salah diprediksi sebagai hipertensi. Hasil ini menunjukkan bahwa *Decision Tree* memiliki kemampuan yang kuat dalam mendeteksi kelas hipertensi dengan tingkat kesalahan yang rendah. Selain itu, model juga cukup baik dalam membedakan kelas diabetes dan komplikasi, meskipun masih terdapat sedikit kesalahan pada kedua kelas tersebut.

#### 4) *AdaBoost Classifier*

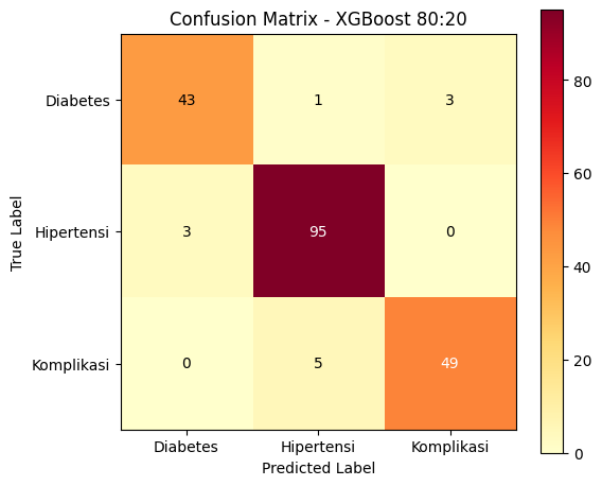
Model *AdaBoost Classifier* menunjukkan performa terbaik pada skenario pembagian data 80:20 dengan nilai akurasi sebesar 0,94. Hasil tersebut memperlihatkan bahwa metode boosting pada *AdaBoost* mampu meningkatkan kemampuan model dalam mengenali pola diagnosis berdasarkan variabel klinis yang digunakan. Meskipun performanya masih sedikit berada di bawah *Decision Tree*, hasil evaluasi menunjukkan bahwa *AdaBoost* mampu memberikan klasifikasi yang stabil pada data uji. Visualisasi *confusion matrix* untuk *AdaBoost* ditampilkan pada Gambar 11.

Gambar 11. *Confusion Matrix AdaBoost Classifier Ratio 80:20*

Berdasarkan Gambar 11, *confusion matrix* pada model *AdaBoost* menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan tepat. Pada kelas diabetes, sebanyak 41 data berhasil diprediksi dengan benar, sedangkan 6 data salah diklasifikasikan sebagai komplikasi. Pada kelas hipertensi, model menunjukkan kemampuan deteksi yang sangat baik dengan 95 data berhasil diklasifikasikan secara benar dan hanya 3 data yang salah diprediksi sebagai diabetes. Sementara itu, pada kelas komplikasi, sebanyak 52 data berhasil diklasifikasikan dengan benar dan 2 data salah diprediksi sebagai hipertensi. Hasil ini menunjukkan bahwa *AdaBoost* memiliki kemampuan yang kuat dalam mendeteksi kelas hipertensi dengan tingkat kesalahan yang rendah. Namun, model masih menunjukkan sedikit kesalahan dalam membedakan kelas diabetes dan komplikasi, yang kemungkinan disebabkan oleh kemiripan karakteristik klinis antara pasien diabetes dan pasien komplikasi.

#### 5) *XGBoost Classifier*

Pada skenario pembagian data 80:20, *XGBoost Classifier* memperoleh performa terbaik dengan nilai akurasi sebesar 0,94. Melalui mekanisme *boosting* yang dilakukan secara bertahap, *XGBoost* dapat mempelajari pola dari variabel klinis secara efektif sehingga menghasilkan prediksi yang cukup akurat. Meskipun hasilnya masih sedikit lebih rendah dibandingkan *Decision Tree*, *XGBoost* tetap menunjukkan kinerja yang baik pada data uji. Visualisasi *confusion matrix* untuk *XGBoost Classifier* disajikan pada Gambar 12.



Gambar 12. Confusion Matrix XGBoost Classifier Ratio 80:20

Berdasarkan Gambar 12, model *XGBoost* mampu mengklasifikasikan sebagian besar data secara tepat. Pada kelas diabetes, terdapat 43 data yang berhasil diprediksi dengan benar, sedangkan 1 data salah diklasifikasikan sebagai hipertensi dan 3 data sebagai komplikasi. Kemampuan deteksi terbaik terlihat pada kelas hipertensi, dengan 95 data berhasil dikenali secara benar dan hanya 3 data yang keliru diprediksi sebagai diabetes. Sementara itu, pada kelas komplikasi, model berhasil mengklasifikasikan 49 data dengan benar, namun masih terdapat 5 data yang salah diprediksi sebagai hipertensi. Hasil tersebut menunjukkan bahwa *XGBoost* sangat baik dalam mengenali kelas hipertensi dengan tingkat kesalahan yang rendah, meskipun masih mengalami beberapa kesalahan pada kelas komplikasi, terutama ketika data komplikasi terklasifikasi sebagai hipertensi.

### B. Prediction

Tahap prediksi dilakukan dengan menggunakan *Decision Tree* sebagai model terbaik berdasarkan hasil evaluasi sebelumnya. Pemilihan model tersebut didasarkan pada performa yang secara deskriptif paling unggul dibandingkan algoritma lain. Pada skenario pembagian data 80:20, *Decision Tree* memperoleh nilai *accuracy*, *recall*, *precision*, dan *f1-score* tertinggi, yaitu masing-masing sebesar 0,95. Nilai tersebut lebih tinggi dibandingkan *Random Forest*, *SVM*, *AdaBoost*, dan *XGBoost* yang sama-sama memperoleh akurasi sebesar 0,94 dengan *f1-score* sebesar 0,93. Selain itu, *Decision Tree* juga menunjukkan performa terbaik pada skenario 90:10 dan 70:30, sehingga dapat dinilai memiliki hasil yang relatif lebih konsisten dibandingkan model lainnya. Proses prediksi selanjutnya dilakukan dengan menggunakan data baru yang ditampilkan pada Tabel 6.

TABEL VI  
DATA PREDIKSI

Umur	JK	TB	BB	GDP	GDPP	Sitole	Ditole
56	P	160	68	105	125	155	98
45	L	165	70	180	220	140	90
60	L	167	76	185	230	160	100
29	L	168	75	110	140	110	75
65	P	153	58	160	205	158	98
50	P	157	62	108	130	150	95
38	L	170	80	200	250	150	95
58	P	154	59	112	138	170	105
54	P	159	66	170	215	155	96
42	L	172	80	115	140	160	100
50	P	160	65	150	190	130	95
57	L	169	79	178	225	170	106
63	L	166	72	110	135	165	102
60	P	155	60	130	160	120	80
49	L	171	82	195	240	165	104

JK: Jenis Kelamin, L: Laki-laki, P: Perempuan

Berdasarkan Tabel 6, data prediksi yang digunakan merupakan data baru yang belum pernah dilibatkan dalam proses pelatihan model sebelumnya. Data tersebut mencakup variabel umur, jenis kelamin, tinggi badan, berat badan, kadar gula darah, serta tekanan darah yang digunakan sebagai input dalam proses klasifikasi. Variasi nilai pada setiap variabel menunjukkan perbedaan karakteristik pasien yang akan diuji menggunakan model terbaik. Penggunaan data baru ini bertujuan untuk menguji kemampuan generalisasi model dalam menghasilkan prediksi yang akurat, dimana data prediksi tersebut akan diimplementasikan pada *platform Streamlit*.



Gambar 13. Tampilan awal pada Platform Streamlit

Gambar 13 menampilkan tampilan awal (*dashboard*) dari *platform Streamlit* yang digunakan sebagai media implementasi sistem prediksi diabetes melitus dan hipertensi. Pada tampilan tersebut, pengguna dapat memasukkan data pasien sesuai dengan variabel yang dibutuhkan oleh model, seperti umur, jenis kelamin, IMT, kadar gula darah, serta tekanan darah. Antarmuka dirancang sederhana dan interaktif sehingga memudahkan pengguna dalam melakukan proses prediksi secara langsung. Sistem tersebut berfungsi sebagai alat bantu dalam mendukung deteksi dini penyakit berbasis *machine learning*, dan hasil prediksi menggunakan data baru ditampilkan pada Tabel 7.

TABEL VII  
HASIL PREDIKSI

Umur	JK	TB	BB	GDP	GDPP	Sitole	Ditole	Diagnosa
56	P	160	68	105	125	155	98	Hipertensi
45	L	165	70	180	220	140	90	Diabetes
60	L	167	76	185	230	160	100	Komplikasi
29	L	168	75	110	140	110	75	Hipertensi
65	P	153	58	160	205	158	98	Komplikasi
50	P	157	62	108	130	150	95	Hipertensi
38	L	170	80	200	250	150	95	Diabetes
58	P	154	59	112	138	170	105	Hipertensi
54	P	159	66	170	215	155	96	Komplikasi
42	L	172	80	115	140	160	100	Hipertensi
50	P	160	65	150	190	130	95	Diabetes
57	L	169	79	178	225	170	106	Komplikasi
63	L	166	72	110	135	165	102	Hipertensi
60	P	155	60	130	160	120	80	Diabetes
49	L	171	82	195	240	165	104	Hipertensi

JK: Jenis Kelamin, L: Laki-laki, P: Perempuan

Pada Tabel 7, hasil prediksi terhadap 15 data baru menunjukkan bahwa model mengklasifikasikan pasien ke dalam tiga kategori diagnosis, yaitu hipertensi, diabetes melitus, dan komplikasi. Sebanyak 6 data diprediksi sebagai hipertensi, 4 data sebagai diabetes melitus, dan 5 data sebagai komplikasi. Pasien yang diprediksi mengalami hipertensi umumnya memiliki nilai tekanan darah yang lebih dominan, sedangkan pasien yang diprediksi mengalami diabetes melitus cenderung menunjukkan kadar GDP dan GDPP yang lebih tinggi. Sementara itu, kategori komplikasi ditemukan pada pasien dengan kombinasi kadar gula darah dan tekanan darah yang sama-sama meningkat. Hasil tersebut menunjukkan bahwa model *Decision Tree* mampu mengidentifikasi pola diagnosis berdasarkan kombinasi variabel klinis yang digunakan dalam proses prediksi.



Gambar 14. Tampilan Hasil Prediksi pada Platform Streamlit

Gambar 14 menunjukkan hasil prediksi yang ditampilkan pada platform Streamlit setelah data pasien diproses oleh sistem. Informasi yang disajikan berupa diagnosis seperti diabetes melitus, hipertensi, atau komplikasi (diabetes dan hipertensi). Tampilan dibuat sederhana agar hasil prediksi mudah dipahami pengguna.

Setelah sistem prediksi berhasil diimplementasikan pada platform Streamlit, tahap selanjutnya dilakukan pengujian kegunaan atau *usability testing* menggunakan metode *System Usability Scale* (SUS). Pengujian ini bertujuan untuk menilai tingkat kemudahan penggunaan, kejelasan tampilan, serta

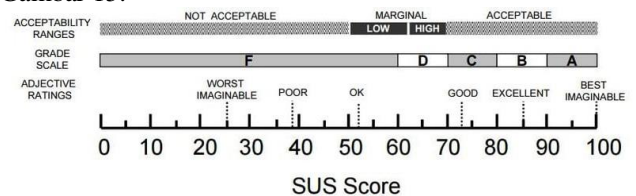
kenyamanan pengguna saat berinteraksi dengan sistem yang dikembangkan. Pertanyaan yang digunakan dalam pengujian *usability* disajikan pada Tabel 8.

TABEL VIII  
PERTANYAAN PENGUJIAN USABILITY

No	Pertanyaan	Skala
1	Saya merasa akan menggunakan aplikasi prediksi ini apabila diperlukan.	1 - 5
2	Saya merasa aplikasi prediksi ini terlalu rumit untuk digunakan.	1 - 5
3	Saya merasa aplikasi prediksi ini mudah digunakan.	1 - 5
4	Saya merasa membutuhkan bantuan teknis dari orang lain untuk dapat menggunakan aplikasi prediksi ini.	1 - 5
5	Saya merasa fitur-fitur pada aplikasi prediksi ini sudah terintegrasi dengan baik.	1 - 5
6	Saya merasa tampilan dan fungsi pada aplikasi prediksi ini belum konsisten.	1 - 5
7	Saya merasa sebagian besar pengguna dapat mempelajari cara menggunakan aplikasi prediksi ini dengan cepat.	1 - 5
8	Saya merasa aplikasi prediksi ini merepotkan untuk digunakan.	1 - 5
9	Saya merasa percaya diri saat menggunakan aplikasi prediksi ini.	1 - 5
10	Saya merasa perlu mempelajari banyak hal terlebih dahulu sebelum dapat menggunakan aplikasi prediksi ini.	1 - 5

Sumber: diadaptasi dari [18]

Berdasarkan Tabel 8, pertanyaan pengujian *usability* menggunakan metode *System Usability Scale* (SUS) terdiri atas 10 pertanyaan yang dinilai dengan skala Likert 1 sampai 5, di mana skor 1 menunjukkan tingkat persetujuan paling rendah dan skor 5 menunjukkan tingkat persetujuan paling tinggi. Hasil skor SUS kemudian diinterpretasikan berdasarkan *SUS Score Percentile Rank*, yaitu *Grade A* untuk skor  $\geq 80,3$ , *Grade B* untuk skor  $\geq 74$  dan  $< 80,3$ , *Grade C* untuk skor  $\geq 68$  dan  $< 74$ , *Grade D* untuk skor  $\geq 51$  dan  $< 68$ , serta *Grade F* untuk skor  $< 51$ , sebagaimana ditampilkan pada Gambar 15.



Gambar 15. System Usability Scale (SUS) scores

Pengujian *usability* dilakukan dengan menyebarkan kuesioner *System Usability Scale* (SUS) kepada pengguna aplikasi prediksi. Dari proses penyebaran tersebut, diperoleh sebanyak 23 responden yang memberikan penilaian terhadap kegunaan sistem. Hasil penilaian dari 23 responden tersebut dapat dilihat pada Tabel 9.

TABEL IX  
HASIL PENILAIAN USABILITY

Responden	Jumlah Skor	Nilai SUS
1	29	72,5
2	29	72,5
3	30	75
4	28	70
5	31	77,5
6	30	75
7	27	67,5
8	33	82,5
9	29	72,5
10	28	70
11	30	75
12	30	75
13	25	62,5
14	33	82,5
15	29	72,5
16	30	75
17	28	70
18	30	75
19	34	85
20	26	65
21	31	77,5
22	29	72,5
23	27	67,5
<b>Rata-rata</b>	<b>29,39</b>	<b>73,48</b>

Berdasarkan Tabel 9, hasil perhitungan *System Usability Scale* (SUS) dari 23 responden memperoleh nilai rata-rata sebesar 73,48. Mengacu pada Gambar 15, nilai tersebut berada pada kategori *acceptable*, termasuk dalam *Grade C*, dan berada pada rentang penilaian *good*. Hasil ini menunjukkan bahwa aplikasi prediksi memiliki tingkat kegunaan yang baik dan dapat diterima oleh pengguna.

### C. Pembahasan

Hasil penelitian menunjukkan bahwa *Decision Tree* memiliki performa tertinggi dibandingkan model lainnya dalam klasifikasi diabetes melitus dan hipertensi. Pada skenario pembagian data 80:20, *Decision Tree* memperoleh nilai *accuracy*, *recall*, *precision*, dan *f1-score* masing-masing sebesar 0,95, lebih tinggi dibandingkan *Random Forest*, *SVM*, *AdaBoost*, dan *XGBoost* yang memperoleh akurasi sebesar 0,94. Nilai tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik dalam mengenali pola hubungan antarvariabel klinis, seperti kadar gula darah dan tekanan darah. Meskipun akurasi 0,95 tergolong tinggi, masih terdapat kemungkinan kesalahan klasifikasi sehingga dalam konteks medis model tetap perlu digunakan secara hati-hati dan tidak dapat dijadikan satu-satunya dasar dalam penentuan diagnosis pasien.

Pada tahap prediksi, model *Decision Tree* mampu mengklasifikasikan 15 data baru ke dalam tiga kategori diagnosis, yaitu 6 data hipertensi, 4 data diabetes melitus, dan 5 data komplikasi. Hasil tersebut menunjukkan bahwa model dapat mengenali pola diagnosis berdasarkan kombinasi nilai

GDP, GDPP, serta tekanan darah yang digunakan sebagai variabel *input*. Implementasi pada *platform Streamlit* memperlihatkan bahwa model dapat diterapkan secara interaktif sebagai alat bantu skrining awal, namun penggunaannya untuk mendukung pengambilan keputusan klinis masih memerlukan validasi lebih lanjut pada lingkungan medis nyata dengan data yang lebih luas dan evaluasi langsung oleh tenaga kesehatan.

Konsisten dengan penelitian lain, hasil memperlihatkan bahwa pendekatan *Decision Tree* berkinerja baik ketika digunakan untuk klasifikasi data medis. Sementara Hadi dan Sirodj mencapai tingkat akurasi 74,78% dalam klasifikasi diabetes mempertimbangkan parameter kritis seperti usia, BMI, serta glukosa, Junus dkk. mencapai akurasi yang lebih tinggi yaitu 98,08% serta mengalahkan *SVM* [20]. Prediksi hipertensi memakai algoritma ensemble seperti *Random Forest*, *AdaBoost*, serta *Gradient Boosting* efektif, seperti yang ditunjukkan oleh Muftisany dkk. [21], dan oleh Khalim dkk. [22], *Random Forest* melampaui *Naive Bayes*. Temuan ini memperlihatkan akurasi serta stabilitas *Random Forest* dalam hal klasifikasi data kesehatan.

## IV. KESIMPULAN

Berdasarkan hasil penelitian, penerapan metode *machine learning* dalam klasifikasi diabetes melitus dan hipertensi menunjukkan performa yang baik. Dari lima algoritma yang diuji, *Decision Tree* memberikan hasil terbaik dengan nilai *accuracy*, *recall*, *precision*, dan *f1-score* masing-masing sebesar 0,95 pada skenario pembagian data 80:20. Hasil evaluasi pada tiga skenario pembagian data serta penerapan *10-fold cross validation* menunjukkan bahwa *Decision Tree* memiliki performa yang relatif baik dan konsisten dalam mengenali pola hubungan antarvariabel klinis, seperti kadar gula darah, tekanan darah, usia, dan indeks massa tubuh. Pengujian menggunakan 15 data baru juga menunjukkan bahwa model mampu mengklasifikasikan data ke dalam tiga kategori diagnosis, yaitu 6 data hipertensi, 4 data diabetes melitus, dan 5 data komplikasi, sehingga memperlihatkan kemampuan model dalam melakukan prediksi terhadap data di luar proses pelatihan.

Implementasi model ke dalam *platform* berbasis *Streamlit* menghasilkan aplikasi prediksi yang bersifat interaktif dan mudah digunakan. Hasil pengujian *usability* menggunakan *System Usability Scale* (SUS) terhadap 23 responden memperoleh skor rata-rata 73,48, yang termasuk dalam kategori *acceptable*, *Grade C*, dan berada pada rentang penilaian *good*. Penggunaan sistem yang dikembangkan diharapkan dapat dimanfaatkan sebagai alat bantu dalam proses skrining awal diabetes melitus dan hipertensi. Namun, hasil prediksi yang diberikan oleh sistem tidak digunakan sebagai penetapan diagnosis akhir, sehingga pemeriksaan lebih lanjut dan kepastian kondisi pasien tetap perlu dilakukan secara langsung oleh tenaga medis.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada institusi yang telah memberikan dukungan dalam pelaksanaan penelitian, khususnya pihak Puskesmas di Kota Bandar Lampung yang telah memfasilitasi pengumpulan data. Apresiasi juga disampaikan kepada seluruh pihak yang telah berkontribusi baik secara langsung maupun tidak langsung, sehingga penelitian dapat terlaksana dengan baik dan menghasilkan luaran yang bermanfaat.

### DAFTAR PUSTAKA

- [1] Kemenkes, "Diabetes Melitus Tipe 2," Kementerian Kesehatan Republik Indonesia. Diakses: 20 Januari 2026. [Daring]. Tersedia pada: <https://ayosehat.kemkes.go.id/topik-penyakit/diabetes-penyakit-ginjal/diabetes-melitus-tipe-2>
- [2] WHO, "Diabetes," World Health Organization. [Daring]. Tersedia pada: <https://www.who.int/news-room/fact-sheets/detail/diabetes>
- [3] WHO, "Hypertension," World Health Organization. [Daring]. Tersedia pada: <https://www.who.int/news-room/fact-sheets/detail/hypertension>
- [4] A. A. M. Hezam, H. B. M. Shaghdar, dan L. Chen, "The connection between hypertension and diabetes and their role in heart and kidney disease development," *Journal of Research in Medical Sciences*, vol. 29, no. 1, Apr 2024, doi: 10.4103/jrms.jrms\_470\_23.
- [5] Anggie Margaretha Kubalang, Vera D. Tombokan, dan Jonesius Eden Manoppo, "Gambaran Hasil Screening Diabetes Melitus pada Keluarga Anggota Prolanis Maleosan," *Jurnal Riset Ilmu Kesehatan Umum dan Farmasi (JRIKUF)*, vol. 3, no. 4, hlm. 53–64, Okt 2025, doi: 10.57213/jrikuf.v3i4.878.
- [6] N. R. Shafriani dan Y. Rahmawati, "Pemeriksaan Kesehatan sebagai Upaya Pencegahan Penyakit Tidak Menular bagi Ibu-Ibu PKK Dusun Nglarang RT 06 Tlogoadi, Mlati, Sleman, Yogyakarta," *Jurnal PEDAMAS (Pengabdian Kepada Masyarakat)*, vol. 2, no. 3, hlm. 840–846, 2024, [Daring]. Tersedia pada: <https://pekatpk.my.id/index.php/JP/article/view/323>
- [7] A. Gavriila, "Prevalensi Diabetes Melitus di Indonesia: Angka Resmi Riskesdas dan IDF 2025," Good Stats. [Daring]. Tersedia pada: <https://goodstats.id/article/prevalensi-diabetes-melitus-di-indonesia-angka-resmi-riskesdas-dan-idf-2025-mhkJA>
- [8] H. UMS, "Indonesia Peringkat Lima Dunia, Hipertensi Ancam Remaja Usia Produktif," UMS News. [Daring]. Tersedia pada: <https://news.ums.ac.id/id/05/2025/indonesia-peringkat-lima-dunia-hipertensi-ancam-remaja-usia-produktif/>
- [9] R. G. Wardhana, G. Wang, dan F. Sibuea, "Penerapan Machine Learning Dalam Prediksi Tingkat Kasus Penyakit di Indonesia," *Journal of Information System Management (JOISM) e-ISSN*, vol. 5, no. 1, hlm. 40–45, 2023.
- [10] A. Pinandito, S. A. Wicaksono, dan S. H. Wijoyo, "Implementasi Machine Learning dalam Deteksi Risiko Tinggi Diabetes Melitus pada Kehamilan," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 4, hlm. 739–746, Agu 2023, doi: 10.25126/jtiik.2023107005.
- [11] W. Van Woensel dkk., "A community-of-practice-based evaluation methodology for knowledge intensive computational methods and its application to multimorbidity decision support," *J. Biomed. Inform.*, vol. 142, hlm. 104395, Jun 2023, doi: 10.1016/j.jbi.2023.104395.
- [12] A. S. Biyantoro dan B. Prasetyo, "Penerapan Decision Tree untuk Klasifikasi Status Kesehatan dengan perbandingan KNN dan Naive Bayes," *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)*, vol. 4, no. 1, hlm. 47–55, Mar 2024, doi: 10.57152/ijirse.v4i1.1342.
- [13] M. D. Jannah, A. Fauzi, C. Emilia, dan H. Hikmayanti, "Klasifikasi Penyakit Hipertensi Menggunakan Support Vector Machine dan Naive Bayes," *Jurnal Algoritma*, vol. 22, no. 1, hlm. 905–913, Jun 2025, doi: 10.33364/algoritma/v.22-1.2316.
- [14] R. Sahila, T. Widiharih, dan I. T. Utami, "Analisis Klasifikasi Menggunakan Regresi Logistik Biner dan Algoritma Naive Bayes Classifier Pada Penyakit Hipertensi," *Jurnal Gaussian*, vol. 13, no. 2, hlm. 319–327, Nov 2024, doi: 10.14710/j.gauss.13.2.319-327.
- [15] P. R. Putri dan R. Alit, "Klasifikasi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine (SVM)," *Journal of Informatics and Computer Science*, vol. 06, no. 3, hlm. 740–746, 2024, doi: 10.26740/jinacs.v6n03.p740-746.
- [16] M. Rizky, A. Pramuntadi, W. D. Prastowo, dan D. H. Gutama, "Implementasi Metode Deep Neural Network pada Klasifikasi Penyakit Diabetes Melitus Tipe 2," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 3, hlm. 1043–1050, Jun 2024, doi: 10.57152/malcom.v4i3.1279.
- [17] A. Wantoro, Zulkifli, A. Fi. Yulia, D. Y. Ayu, dan S. Mustofa, "Evaluasi Kinerja Algoritma Machine Learning (ML) Menggunakan Seleksi Fitur Pada Klasifikasi Diabetes," *JIP (Jurnal Informatika Polinema)*, vol. 11, no. 3, hlm. 311–316, Jun 2025, [Daring]. Tersedia pada: [www.kaggle.com/uciml/pima-Feishal-Azriel-Arya-Putra,Bibit-Waluyo,Risqi-Faturahman,Wahyu-Dwi-Purwoprasetyo,dan-Ito-Setiawan](http://www.kaggle.com/uciml/pima-Feishal-Azriel-Arya-Putra,Bibit-Waluyo,Risqi-Faturahman,Wahyu-Dwi-Purwoprasetyo,dan-Ito-Setiawan)
- [18] Feishal Azriel Arya Putra, Bibit Waluyo, Risqi Faturahman, Wahyu Dwi Purwoprasetyo, dan Ito Setiawan, "Analisis Usability Testing Menggunakan Metode System Usability Scale terhadap Kepuasan Pengguna Website Kemahasiswaan Universitas Amikom Purwokerto," *Uranus: Jurnal Ilmiah Teknik Elektro, Sains dan Informatika*, vol. 3, no. 1, hlm. 121–130, Jan 2025, doi: 10.61132/uranus.v3i1.673.
- [19] D. A. Hadi dan D. A. N. Sirodj, "Metode Random Forest untuk Klasifikasi Penyakit Diabetes," *Bandung Conference Series: Statistics*, vol. 3, no. 2, hlm. 428–435, Agu 2023, doi: 10.29313/bcss.v3i2.8354.
- [20] C. Z. V. Junus, T. Tarno, dan P. Kartikasari, "Klasifikasi Menggunakan Metode Support Vector Machine dan Random Forest untuk Deteksi Awal Risiko Diabetes Melitus," *Jurnal Gaussian*, vol. 11, no. 3, hlm. 386–396, Jan 2023, doi: 10.14710/j.gauss.11.3.386-396.
- [21] H. Muftisany, T. F. Efendi, dan N. A. Rozaq Rais, "Perbandingan Kinerja Algoritma Random Forest, AdaBoost, dan Gradient Boosting dalam Memprediksi Risiko Penyakit Hipertensi," *Faktor Exacta*, vol. 18, no. 2, hlm. 161, Okt 2025, doi: 10.30998/faktorexacta.v18i2.28959.
- [22] K. A. Khalim, U. Hayati, dan A. Bahtiar, "Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest dan Naive Bayes," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, hlm. 498–504, 2023, doi: 10.36040/jati.v7i1.6376.