



Implementasi Algoritma K-Nearest Neighbors (KNN) dalam Deteksi Dini Hipertensi berdasarkan Analisis Tekanan Darah

Implementation of K-Nearest Neighbors Algorithm for Early Detection of Hypertension Based on Blood Pressure Analysis

Ary Prandika Siregar* & Said Iskandar Al Idrus

Program Studi Ilmu Komputer, Universitas Negeri Medan, Indonesia

Diterima: 16 April 2025; Direview: 20 April 2025; Disetujui: 24 April 2025

*Corresponding Email: aryprandika0902@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem deteksi hipertensi berbasis web menggunakan algoritma K-Nearest Neighbors (KNN) serta menganalisis tingkat keakuratannya dalam mengklasifikasikan status hipertensi. Data penelitian diperoleh dari 447 catatan medis pasien di RSKG Rasyida, yang mencakup delapan variabel: jenis kelamin, usia, tekanan darah sistolik, tekanan darah diastolik, tinggi badan, berat badan, indeks massa tubuh (IMT), dan status hipertensi. Tahap preprocessing dilakukan melalui tiga langkah utama, yaitu pemilihan fitur (usia, tekanan darah sistolik, tekanan darah diastolik, dan IMT), penyeimbangan data menggunakan teknik undersampling, serta normalisasi data dengan metode Min-Max, sehingga diperoleh 425 data seimbang dengan lima kategori hipertensi. Sistem web yang dikembangkan memiliki modul login, dashboard, input data, hasil deteksi, dan riwayat deteksi, serta telah diuji menggunakan metode black box. Hasil terbaik diperoleh pada nilai $k = 13$ dengan akurasi 92,94%, precision 94%, recall 93%, dan F1-score 93%. Temuan ini menunjukkan bahwa sistem yang dikembangkan mampu mendeteksi hipertensi secara akurat dan dapat menjadi alat bantu skrining awal berbasis data bagi tenaga medis.

Kata Kunci: Hipertensi; Deteksi Dini; Klasifikasi; K-Nearest Neighbors (KNN); Sistem Berbasis Website

Abstract

This study aims to develop a web-based hypertension detection system using the K-Nearest Neighbors (KNN) algorithm and to analyze its accuracy in classifying hypertension status. The dataset was obtained from 447 patient medical records at RSKG Rasyida, consisting of eight variables: gender, age, systolic blood pressure, diastolic blood pressure, height, weight, body mass index (BMI), and hypertension status. The preprocessing stage involved three main steps—feature selection (age, systolic and diastolic blood pressure, BMI), data balancing using undersampling, and data normalization through the Min-Max method—resulting in 425 balanced data samples with five hypertension categories. The web application includes modules for login, dashboard, data input, detection results, and detection history, and has been evaluated using black box testing. The best KNN performance was achieved at $k = 13$ with 92.94% accuracy, 94% precision, 93% recall, and 93% F1-score. These results indicate that the proposed system can accurately classify hypertension and serve as an effective, data-driven screening tool for healthcare professionals.

Keywords: Hypertension; Early Detection; Classification; K-Nearest Neighbors; Web-Based System

PENDAHULUAN

Hipertensi merupakan salah satu penyebab kematian terbesar di dunia, baik di negara maju maupun negara berkembang, termasuk Indonesia. World Health Organization melaporkan lebih dari 1,28 miliar orang dewasa menderita hipertensi, dan diperkirakan pada tahun 2025 jumlahnya mencapai 1,5 miliar orang. Selain itu, diperkirakan setiap tahunnya sebanyak 10,44 juta orang meninggal karena hipertensi dan komplikasinya [1]. Kondisi ini berbahaya karena hipertensi sering kali tidak menimbulkan gejala hingga menyebabkan komplikasi serius seperti stroke, gagal ginjal, dan penyakit jantung [2].

Berdasarkan data dari Badan Penelitian dan Pengembangan Kesehatan, prevalensi hipertensi di Indonesia pada tahun 2013 tercatat 25,8% dan mengalami peningkatan menjadi 34,1% pada tahun 2018. Pada tahun 2023, prevalensi kejadian hipertensi mencapai 30,8%. Sumatera Utara adalah salah satu provinsi di Indonesia dengan prevalensi kejadian hipertensi yang mengalami fluktuasi. Pada tahun 2013, prevalensi hipertensi yaitu sebesar 24,7% dan mengalami peningkatan pada tahun 2018 yaitu 29,19%. Berdasarkan Survei Kesehatan Indonesia, prevalensi hipertensi di Sumatera Utara mencapai 25,4% pada tahun 2023. Jika dilihat berdasarkan jenis kelamin, perempuan menjadi kelompok tertinggi mengalami hipertensi yaitu sebesar 34,7%, sedangkan pada laki-laki sebesar 26,9% [3]. Angka ini menunjukkan bahwa hipertensi merupakan penyakit yang cukup banyak diderita masyarakat di daerah tersebut.

Berdasarkan data yang diperoleh dari RSKG Rasyida pada tahun 2025 menunjukkan distribusi status hipertensi. Terdapat 447 pasien, di antaranya 95 orang dalam kategori Hipertensi Tingkat 2, 92 orang dalam kategori Hipertensi Tingkat 1, dan 88 orang mengalami Krisis Hipertensi. Sementara itu, pasien dengan status Prehipertensi sebanyak 87 orang dan pasien dengan status Normal sebanyak 85 orang. Pola ini menunjukkan bahwa mayoritas pasien sudah mengalami hipertensi di tahap yang lebih berat, sehingga membutuhkan perawatan medis yang lebih intensif.

Dilihat dari segi demografi, distribusi usia berdasarkan jenis kelamin menunjukkan bahwa kasus hipertensi lebih sering dialami oleh kelompok usia 45–60 tahun, baik pria maupun wanita. Pada usia 45–55 tahun, jumlah pasien wanita mencapai 43 orang, sedangkan pria sebanyak 57 orang. Di kelompok usia 56–60 tahun, jumlah pasien pria (57 orang) masih lebih banyak dibandingkan wanita (44 orang). Hal ini menunjukkan

bahwa risiko terkena hipertensi semakin tinggi seiring bertambahnya usia, terutama pada usia dewasa paruh baya hingga lanjut usia.

Hipertensi sering kali tidak menunjukkan gejala, sehingga banyak penderita tidak menyadari kondisinya. Oleh karena itu, deteksi dini sangat penting untuk mencegah kerusakan organ dan komplikasi serius [4]. Penelitian ini menawarkan solusi berupa sistem deteksi dini berbasis algoritma K-Nearest Neighbors (KNN). Algoritma ini bekerja dengan membandingkan data kesehatan pasien seperti usia, tekanan darah sistolik, tekanan darah diastolik, dan indeks massa tubuh (IMT) dengan data pasien sebelumnya yang sudah diketahui status hipertensinya. Hasil klasifikasi ditentukan berdasarkan tingkat kesamaan antar data[5]. Sistem deteksi dini hipertensi yang dikembangkan dalam penelitian ini ditujukan sebagai alat bantu bagi tenaga medis di rumah sakit maupun klinik untuk melakukan pemeriksaan awal pasien sebagai dukungan dalam mempercepat proses skrining dan membantu tenaga medis dalam pengambilan keputusan[6].

Di era transformasi digital, teknologi machine learning semakin banyak digunakan di bidang kesehatan, termasuk untuk analisis medis. Salah satu algoritma yang sering digunakan adalah K-Nearest Neighbors (KNN). Algoritma KNN dipilih karena mudah dipahami, tidak memerlukan asumsi distribusi data, dan cukup efisien untuk dataset medis dengan jumlah variabel terbatas [7]. Algoritma ini bekerja dengan membandingkan data baru terhadap data historis pasien berdasarkan tingkat kesamaan, sehingga hasil prediksi diperoleh dari pola data sebelumnya[8].Kelebihan ini menjadikan KNN lebih fleksibel dalam menangani berbagai jenis data pasien dan mampu memberikan hasil prediksi yang akurat apabila nilai k ditentukan secara tepat.

Meskipun algoritma K-Nearest Neighbors (KNN) banyak digunakan dan terbukti efektif dalam klasifikasi medis, penerapannya pada kasus hipertensi masih menghadapi beberapa masalah [9]. Beberapa di antaranya adalah ketidakseimbangan jumlah data antara pasien normal dan pasien hipertensi. Kondisi ini dapat membuat model cenderung bias terhadap kelas mayoritas sehingga meningkatkan risiko kesalahan klasifikasi. Selain itu, performa KNN sangat dipengaruhi oleh pemilihan parameter, terutama jumlah tetangga terdekat (k) dan jenis metrik jarak yang digunakan[10]. Untuk mengatasi masalah tersebut, penelitian ini menggunakan metode undersampling guna menyeimbangkan distribusi data tanpa menambah data sintesis, serta melakukan pencarian nilai k optimal agar model lebih stabil dan akurat. Dengan langkah ini, KNN

diharapkan lebih efektif dalam mendeteksi kondisi hipertensi berdasarkan data klinis dari rumah sakit maupun layanan kesehatan lainnya[10].

Ini merujuk pada beberapa penelitian sebelumnya yang membahas penggunaan algoritma KNN diantaranya yang dilakukan oleh Penelitian yang dilakukan Sari (2024), *Application of the K-Nearest Neighbor Method for Classification of Hypertension Diseases (Case Study: Stabat Health Center)*. Menerapkan KNN pada data pasien Puskesmas Stabat. Hasilnya menunjukkan akurasi sebesar 97,12%, yang menunjukkan bahwa metode ini efektif dalam mendeteksi hipertensi. Penelitian yang dilakukan Rodiana (2024), *Classification of Hypertension Patients in Palembang by K-Nearest Neighbor and Local Mean K-Nearest Neighbor*, mengombinasikan KNN dengan Local Mean K-Nearest Neighbor untuk mengklasifikasikan pasien hipertensi di Palembang, dan\ mendapatkan akurasi 96,97%. Penelitian yang dilakukan Resky (2025), *PCA and t-SNE Implementation for KNN Hypertension Classification Visualization*, membandingkan KNN murni dengan KNN yang menggunakan reduksi dimensi melalui PCA dan t-SNE. Hasilnya, KNN tanpa reduksi dimensi mencapai akurasi 99%, KNN dengan PCA mencapai 100% dan KNN dengan t-SNE mencapai 99%, menunjukkan bahwa optimasi fitur bisa meningkatkan kinerja model. Penelitian Rahim (2021) menerapkan metode Modified K-Nearest Neighbor (MKNN) pada sistem pakar untuk klasifikasi tingkat risiko hipertensi, dengan akurasi sebesar 86,67%.

Selain KNN, beberapa penelitian juga menggunakan metode lain untuk mengklasifikasikan hipertensi. Penelitian yang dilakukan oleh Novianti (2024) berjudul "Klasifikasi Penyakit Hipertensi Menggunakan Metode Random Forest" membahas penerapan algoritma Random Forest dalam klasifikasi tingkat hipertensi. Hasilnya menunjukkan akurasi sebesar 98% pada data training dan 95% pada data testing. Penelitian ini membuktikan bahwa metode Random Forest cukup efektif dalam mengklasifikasikan berbagai tingkat hipertensi.

Meskipun algoritma K-Nearest Neighbors (KNN) dikenal mampu memberikan akurasi yang baik dalam klasifikasi hipertensi, sebagian besar penelitian terdahulu yang menggunakan KNN untuk kasus hipertensi masih terbatas pada simulasi model atau pengujian terbatas, tanpa implementasi nyata dalam bentuk sistem yang dapat langsung digunakan oleh tenaga medis. Selain itu, beberapa kendala teknis seperti ketidakseimbangan data antar kelas dan sensitivitas terhadap pemilihan parameter k jarang diperhatikan, padahal hal ini sangat berpengaruh terhadap akurasi klasifikasi. Di



sisi lain, pada fasilitas kesehatan di Indonesia, belum banyak tersedia sistem berbasis web yang mampu mengklasifikasikan status hipertensi pasien. Kondisi ini menunjukkan adanya celah penelitian yang perlu diisi.

Pada praktik layanan Kesehatan di RSKG, dokter atau perawat memang sudah dapat menentukan apakah pasien mengalami hipertensi atau tidak berdasarkan hasil pengukuran tekanan darah. Namun, klasifikasi yang digunakan masih terbatas pada kategori sederhana dan belum dilengkapi dengan dukungan sistem digital yang mampu memberikan penentuan status hipertensi secara lebih rinci, seperti Normal, Prehipertensi, Hipertensi Tingkat 1, Hipertensi Tingkat 2, atau Krisis Hipertensi. Hal ini sangat penting dalam pelayanan kesehatan, karena klasifikasi yang jelas membantu menentukan tingkat risiko serta tindakan medis yang tepat.

Penelitian ini difokuskan untuk mendukung tenaga medis di fasilitas kesehatan melalui sistem digital berbasis KNN yang mampu memberikan hasil klasifikasi cepat, objektif, dan sesuai standar [11]. Dengan demikian, proses skrining hipertensi dapat dilakukan lebih efisien, sehingga diharapkan membantu deteksi dini, mencegah keterlambatan diagnosis, serta mengurangi risiko komplikasi serius seperti gagal ginjal, stroke, atau kematian akibat hipertensi[12].

METODE PENELITIAN

Penelitian ini termasuk dalam kategori penelitian terapan (*applied research*) yang bertujuan untuk mengimplementasikan algoritma K-Nearest Neighbors (KNN) dalam membangun sistem deteksi dini hipertensi berdasarkan analisis data tekanan darah [13]. Populasi penelitian mencakup pasien yang melakukan pemeriksaan tekanan darah di Rumah Sakit Khusus Ginjal (RSKG) Rasyida Medan, dengan sampel berjumlah 447 pasien berusia ≥ 18 tahun yang dipilih berdasarkan kriteria inklusi dan telah divalidasi oleh tenaga medis untuk menjamin keakuratan data klinis. Data diperoleh melalui metode dokumentasi rekam medis dan divalidasi oleh dokter spesialis serta kepala bidang keperawatan guna memastikan kesesuaian antara tekanan darah, usia, dan status hipertensi. Tujuh variabel digunakan sebagai input (*independen*), yaitu jenis kelamin, usia, tekanan darah sistolik, tekanan darah diastolik, berat badan, tinggi badan, dan indeks massa tubuh (IMT), sedangkan variabel *output* (*dependen*) adalah status hipertensi yang diklasifikasikan menjadi lima kategori: Normal, Prehipertensi, Hipertensi Tingkat 1, Hipertensi Tingkat 2, dan Krisis Hipertensi. Data yang terkumpul melalui



tahap *preprocessing*, mencakup pemeriksaan kelengkapan data, *balancing* menggunakan *undersampling* untuk menghindari bias kelas, *feature selection* untuk memilih empat fitur paling berpengaruh (usia, sistolik, diastolik, IMT), dan normalisasi dengan Min-Max Scaling agar setiap variabel berada pada rentang 0–1. Dataset kemudian dibagi menjadi data latih (80%) dan data uji (20%), tanpa cross-validation karena ukuran data relatif kecil. Nilai *k* optimal ditentukan melalui pengujian nilai ganjil 1–21 dengan memilih akurasi tertinggi pada data uji. Proses klasifikasi dilakukan menggunakan metode jarak Euclidean dengan pembobotan jarak (*weights = 'distance'*) agar tetangga terdekat memiliki pengaruh lebih besar terhadap hasil prediksi. Evaluasi model dilakukan menggunakan *confusion matrix*, dengan metrik akurasi, presisi, *recall*, dan *F1-score* untuk mengukur performa klasifikasi. Setelah model KNN mencapai performa terbaik, dilakukan perancangan dan implementasi sistem berbasis web menggunakan Visual Studio Code, XAMPP, MySQL, dan browser web sebagai media antarmuka untuk deteksi dini hipertensi.

HASIL DAN PEMBAHASAN

Dataset yang digunakan mencakup 8 Variabel, yaitu: jenis kelamin, usia, tekanan darah sistolik, tekanan darah diastolik, tinggi badan, berat badan, IMT (indeks massa tubuh) serta status hipertensi. Hasil yang diperoleh sebanyak 447 data pasien status normal dan pasien yang mengalami hipertensi.

1) Variabel Jenis Kelamin

Distribusi jenis kelamin menunjukkan bahwa penelitian lebih banyak berjenis kelamin laki-laki (296 orang) dibandingkan dengan perempuan (191 orang). Hal ini mengindikasikan bahwa dominasi data penelitian lebih banyak berasal dari laki-laki, sehingga dapat memberikan gambaran prevalensi hipertensi yang lebih kuat pada kelompok tersebut.

2) Variabel Usia

Distribusi Sebaran usia pasien berkisar antara 20–80 tahun, dengan mayoritas berada pada rentang 40–60 tahun. Puncak distribusi tercatat pada usia sekitar 50 tahun dengan 68 pasien, menegaskan bahwa risiko hipertensi lebih banyak muncul pada kelompok usia dewasa hingga lanjut.

3) Variabel Tekanan Darah Sistolik

Distribusi Tekanan Sistolik pasien berada pada rentang 90–220 mmHg. Sebagian besar pasien memiliki tekanan sistolik 120–140 mmHg dengan frekuensi tertinggi



lebih dari 50 pasien, yang menunjukkan dominasi kategori normal hingga prehipertensi.

4) Variabel Tekanan Darah Diastolik

Distribusi Tekanan Diastolik berkisar antara 60–120 mmHg, dengan puncak pada kisaran 80–90 mmHg sebanyak 75 pasien. Hal ini menandakan bahwa sebagian besar pasien berada pada kategori batas normal hingga tinggi.

5) Variabel Tinggi Badan

Distribusi Tinggi badan pasien berkisar 145–185 cm, dengan mayoritas berada pada 160–170 cm. Frekuensi tertinggi adalah 90 pasien pada tinggi sekitar 160 cm, yang sesuai dengan rata-rata tinggi badan orang dewasa di Indonesia.

6) Variabel Berat Badan

Distribusi Berat Badan pasien berkisar 40–120 kg. Distribusi terbanyak terdapat pada kisaran 55–70 kg, dengan jumlah pasien terbanyak (126 pasien) berada pada berat 55–60 kg. Sebagian besar pasien berada pada kategori berat badan normal, meskipun ada yang mengalami obesitas.

7) Variabel IMT (Indeks Massa Tubuh)

Distribusi Nilai IMT pasien bervariasi antara 15–38, dengan dominasi pada kisaran 20–25. Frekuensi tertinggi terdapat pada nilai 22–23 sebanyak 111 pasien. Berdasarkan standar WHO, sebagian besar pasien termasuk kategori berat badan normal, meskipun terdapat kelompok dengan obesitas.

8) Variabel Status Hipertensi

Distribusi Status Hipertensi relatif seimbang. Pasien terbanyak berada pada kategori Hipertensi Tingkat 2 (95 pasien), diikuti oleh Hipertensi Tingkat 1 (92 pasien), Krisis Hipertensi (88 pasien), Prehipertensi (87 pasien), dan Normal (85 pasien). Hasil ini menunjukkan bahwa mayoritas pasien berada pada kategori hipertensi, dengan kecenderungan ke arah derajat yang lebih berat.

Setelah data diperoleh, dilakukan tahap preprocessing untuk meningkatkan kualitas dan kesiapan data sebelum digunakan dalam proses klasifikasi menggunakan algoritma KNN [14]. Tahap ini mencakup tiga langkah utama, yaitu *feature selection*, *data balancing*, dan normalisasi data. Pada tahap *feature selection*, dipilih empat variabel yang paling berpengaruh terhadap status hipertensi, yaitu usia, tekanan darah sistolik, tekanan darah diastolik, dan indeks massa tubuh (IMT), karena faktor-faktor tersebut secara klinis merupakan indikator utama hipertensi dan secara teknis menghasilkan performa model



yang lebih stabil [15]. Variabel lain seperti jenis kelamin tidak digunakan karena menyebabkan penurunan akurasi akibat perbedaan skala data. Selanjutnya dilakukan penyeimbangan data (balancing) menggunakan teknik undersampling untuk mengatasi ketimpangan jumlah data antar kelas, di mana sebagian data pada kelas mayoritas dikurangi hingga setiap kelas memiliki jumlah yang seimbang, menghasilkan 425 data akhir yang siap digunakan untuk pelatihan model. Setelah itu diterapkan normalisasi Min-Max untuk menyamakan skala antar variabel numerik agar seluruh fitur memiliki rentang nilai yang setara, sehingga tidak ada satu variabel pun yang mendominasi dalam perhitungan jarak pada algoritma KNN. Tahapan ini memastikan bahwa data yang digunakan sudah bersih, seimbang, dan terstandarisasi, sehingga mendukung terbentuknya model klasifikasi hipertensi yang akurat dan stabil. Adapun rumus normalisasi min-max yang digunakan adalah sebagai berikut:

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4.1)$$

Keterangan:

x_{new} : nilai setelah normalisasi

x : nilai sebelum normalisasi

x_{min} : nilai minimal dari data

x_{max} : nilai maksimal dari data

Adapun variabel-variabel yang dilakukan normalisasi dalam penelitian ini meliputi Usia, Tekanan Darah Sistolik, Tekanan Darah Diastolik, dan Indeks Massa Tubuh:

1) Usia

Hasil normalisasi variabel usia menggunakan metode *Min-Max* menunjukkan bahwa nilai usia dapat ditransformasikan ke dalam rentang 0 hingga 1. Pada kategori Dewasa Muda dengan usia 18 tahun, diperoleh nilai normalisasi sebesar 0, karena merupakan nilai usia terendah dalam data. Selanjutnya, kategori Dewasa dengan usia 40 tahun memiliki nilai normalisasi 0,333, selanjutnya kategori Paruh Baya dengan usia 55 tahun menghasilkan nilai 0,583. Adapun kategori Lansia pada usia 65 tahun memperoleh nilai

0,750. Terakhir, kategori Lansia Akhir dengan usia 80 tahun menghasilkan nilai normalisasi 1, yang merupakan nilai tertinggi.

2) Tekanan Darah Sistolik

Hasil normalisasi variabel tekanan darah sistolik menggunakan metode *Min-Max* menunjukkan bahwa nilai dapat ditransformasikan ke dalam rentang 0 hingga 1. Pada kategori Normal dengan nilai sistolik 90 mmHg, diperoleh nilai normalisasi 0, karena merupakan nilai terendah dalam data. Selanjutnya, kategori Prehipertensi dengan nilai sistolik 130 mmHg memiliki nilai normalisasi 0,308, selanjutnya kategori Hipertensi Tingkat 1 dengan nilai 150 mmHg menghasilkan nilai normalisasi 0,462. Pada kategori Hipertensi Tingkat 2 dengan nilai 170 mmHg, diperoleh nilai normalisasi 0,615. Adapun kategori Krisis Hipertensi pada nilai sistolik 220 mmHg menghasilkan nilai normalisasi 1, yang merupakan nilai tertinggi dalam data.

3) Tekanan Darah Diastolik

Hasil normalisasi variabel tekanan darah diastolik menggunakan metode *Min-Max* menunjukkan bahwa nilai diastolik berhasil ditransformasikan ke dalam rentang 0 hingga 1. Pada kategori Normal dengan nilai sistolik 60 mmHg, diperoleh nilai normalisasi 0, karena merupakan nilai terendah dalam data. Selanjutnya, pada kategori Prehipertensi dengan nilai diastolik 85 mmHg diperoleh nilai normalisasi 0,417, selanjutnya kategori Hipertensi Tingkat 1 dengan nilai 95 mmHg menghasilkan nilai normalisasi 0,583. Pada kategori Hipertensi Tingkat 2 dengan nilai 105 mmHg diperoleh nilai normalisasi 0,750. Terakhir, kategori Krisis Hipertensi dengan nilai diastolik 120 mmHg memperoleh nilai normalisasi 1, yang merupakan nilai tertinggi dalam data.

4) IMT (Indeks Massa Tubuh)

Hasil normalisasi variabel Indeks Massa Tubuh (IMT) menggunakan metode *Min-Max* menunjukkan bahwa nilai IMT berhasil ditransformasikan ke dalam rentang 0 hingga 1. Pada kategori Kurus dengan nilai 15.0, diperoleh nilai normalisasi 0, karena merupakan nilai terendah dalam data. Selanjutnya, pada kategori Normal dengan nilai IMT 22.0 diperoleh nilai normalisasi 0,304, sedangkan kategori Overweight dengan nilai 28.0 menghasilkan nilai normalisasi 0,565. Pada kategori Obesitas dengan nilai 35.0 diperoleh nilai normalisasi 0,870. Terakhir kategori Obesitas Parah dengan nilai 38.0 memperoleh nilai normalisasi 1, yang merupakan nilai tertinggi dalam data.

Model KNN pada penelitian ini menghasilkan akurasi sebesar 92,94%, dengan nilai *precision* 94%, *recall* 93%, dan *F1-score* 93%. Hasil ini tergolong tinggi, meskipun



sedikit lebih rendah dibandingkan beberapa penelitian lain yang mencapai akurasi di atas 97%. Perbedaan ini terjadi karena penelitian ini menggunakan data asli pasien dari Rumah Sakit Khusus Ginjal Rasyida, bukan data simulasi atau dataset publik yang sudah disesuaikan. Data klinis nyata memiliki variasi tekanan darah, usia, dan kondisi fisik yang lebih beragam, sehingga model bekerja pada kondisi yang lebih kompleks. Hal ini membuat akurasi sedikit lebih rendah, tetapi hasilnya lebih realistis dan sesuai dengan kondisi sebenarnya di lapangan medis. Dengan hasil tersebut, model KNN dapat dianggap stabil dan dapat diandalkan untuk membantu tenaga medis dalam mendeteksi hipertensi secara dini, karena mampu mempertahankan keseimbangan antara akurasi tinggi dan kemampuan mengenali pasien secara benar.

Gambar 1. Hasil *Classification Report* penerapan model algoritma KNN.

Classification Report:				
	precision	recall	f1-score	support
Hipertensi Tingkat 1	0.81	1.00	0.89	17
Hipertensi Tingkat 2	1.00	0.82	0.90	17
Krisis Hipertensi	0.94	1.00	0.97	17
Normal	1.00	0.94	0.97	17
Prehipertensi	0.94	0.88	0.91	17
accuracy			0.93	85
macro avg	0.94	0.93	0.93	85
weighted avg	0.94	0.93	0.93	85

Berikut ini adalah contoh perhitungan manual algoritma KNN dengan menggunakan variabel dan atribut dari 13 data yang ada pada dataset, yaitu Umur (x_1), Tekanan Systolik (x_2), Tekanan Diastolik (x_3), Indeks Massa Tubuh yang dihitung dari Tinggi dan Berat badan (x_4). Seluruh data digunakan dalam perhitungan ini telah melalui tahapan preprocessing, yang mencakup transformasi data menggunakan feature selection, penyeimbangan data menggunakan teknik undersampling, dan normalisasi menggunakan metode min-max normalization.

Adapun data yang digunakan dalam proses ini terdiri dari data pelatihan (*train data*) dan data pengujian (*test data*), yang disajikan sebagai berikut:

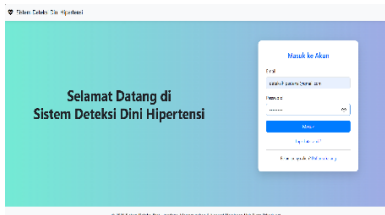
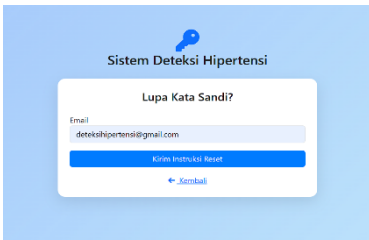
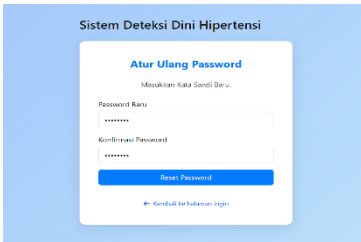
Tabel 1. Data Training

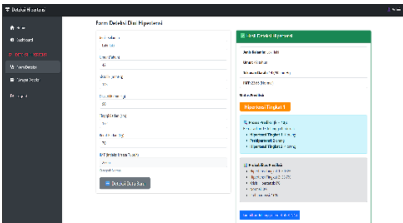
Titik	x_1	x_2	x_3	x_4	y
P1	0,516	0,315	0,466	0,306	Prehipertensi
P2	0,300	0,361	0,483	0,383	Prehipertensi

P3	0,383	0,423	0,583	0,358	Hipertensi Tingkat 1
P4	0,433	0,369	0,533	0,349	Hipertensi Tingkat 1
P5	0,226	0,346	0,500	0,326	Hipertensi Tingkat 1
P6	0,466	0,446	0,550	0,341	Hipertensi Tingkat 1
P7	0,466	0,384	0,500	0,306	Hipertensi Tingkat 1
P8	0,416	0,446	0,383	0,337	Hipertensi Tingkat 1
P9	0,223	0,476	0,483	0,401	Hipertensi Tingkat 1
P10	0,383	0,538	0,550	0,430	Hipertensi Tingkat 2
P11	0,400	0,569	0,483	0,283	Hipertensi Tingkat 2
P12	0,333	0,600	0,500	0,401	Hipertensi Tingkat 2
P13	0,333	0,515	0,666	0,402	Hipertensi Tingkat 2

Berdasarkan data tabel diatas menampilkan data pasien dari dataset, dengan label Pasien 1 hingga Pasien 13. Masing-masing data yang sudah memiliki kelas (variabel y) yang diketahui dan digunakan sebagai data training dalam proses perhitungan manual algoritma KNN. Seluruh data telah melalui proses normalisasi menggunakan min-max, sesuai dengan hasil yang diperoleh pada tahap preprocessing sebelumnya.

Tabel 2. Black Box Testing

No	Aktivitas Pengujian	Hasil yang Diharapkan	Hasil Pengujian	Keterangan
1.	Halaman Login Awal	Menampilkan halaman awal sekaligus halaman login sebelum masuk kedalam sistem		[✓] Valid
2.	Halaman Lupa Password	Menampilkan halaman ketika user/admin mengalami lupa password		[✓] Valid
3.	Halaman Reset Password	Menampilkan halaman untuk mengganti password dengan memasukkan password baru		[✓] Valid
				[] Invalid

4. Halaman Dashboard	Menampilkan hasil distribusi jumlah total deteksi yang sudah dilakukan berdasarkan 5 status hipertensi		[✓] Valid
[] Invalid			
5. Halaman Input	Menampilkan halaman berisi form input untuk melakukan proses deteksi		[✓] Valid
[] Invalid			
6. Halaman Hasil	Menampilkan hasil inputan data yang ada kemudian menampilkan hasil deteksi hipertensi		[✓] Valid
[] Invalid			
7. Halaman Riwayat	Menampilkan Halaman riwayat hasil deteksi yang sudah dilakukan sebelumnya.		[✓] Valid
[] Invalid			

Hasil pengujian dengan *black box testing* menunjukkan bahwa seluruh fungsi utama sistem berjalan sesuai kebutuhan. Hal ini menegaskan bahwa penggunaan metode *black box testing* tepat untuk memastikan keandalan fungsionalitas sistem dari sisi pengguna.

SIMPULAN

Sistem deteksi dini hipertensi berbasis web menggunakan algoritma K-Nearest Neighbors (KNN) telah berhasil dirancang dan dibangun sebagai alat bantu skrining awal bagi tenaga medis di RSKG Rasyida. Sistem ini mampu memproses data klinis pasien

(usia, tekanan darah sistolik/diastolik, tinggi badan, berat badan, dan IMT) secara terstruktur melalui interface web yang mencakup fitur login, input data, hasil deteksi, dan riwayat deteksi. Seluruh fungsi sistem telah diuji secara fungsional menggunakan black box testing dan terbukti bekerja sesuai kebutuhan, sehingga siap digunakan untuk mendukung proses deteksi hipertensi secara cepat, akurat, dan terstandar.

Algoritma K-Nearest Neighbors (KNN) menunjukkan tingkat akurasi yang sangat baik dalam klasifikasi status hipertensi setelah melalui proses optimasi data. Dengan penerapan feature selection (pemilihan usia, Tekanan sistolik, Tekanan diastolik, dan IMT), penyeimbangan data menggunakan undersampling, serta normalisasi Min-Max, model mencapai akurasi tertinggi sebesar 92,94% pada nilai $k = 13$. Evaluasi lebih lanjut menunjukkan precision 94%, recall 93%, dan F1-score 93%, membuktikan bahwa algoritma KNN efektif dan andal dalam mengklasifikasikan lima kategori hipertensi (Normal, Prehipertensi, Hipertensi Tingkat 1, Hipertensi Tingkat 2, dan Krisis Hipertensi).

DAFTAR PUSTAKA

- [1] M. Maulia and H. K. Hengky, "Analisis kejadian penyakit hipertensi di Kabupaten Pinrang," *J. Ilm. Mns. Dan Kesehatan*, vol. 4, no. 3, pp. 324–331, 2021.
- [2] S. Islamy, W. R. Gusti, and M. Zakariyah, "Penerapan IoT Pada Prototipe Pengukur Tekanan Darah Non-Invasive Berbasis ESP8266," *JST (Jurnal Sains dan Teknol.)*, vol. 12, no. 3, 2023.
- [3] D. Sianipar and R. Lubis, "Hubungan Aktivitas Fisik dan Kualitas Tidur dengan Kejadian Hipertensi pada Lansia di Wilayah Kerja Puskesmas Matiti Kecamatan Doloksanggul Tahun 2025: Penelitian," *J. Pengabd. Masy. dan Ris. Pendidik.*, vol. 4, no. 1, pp. 3599–3605, 2025.
- [4] D. P. K. RI, "Hari Hipertensi Dunia 2019: 'Know Your Number, Kendalikan Tekanan Darahmu dengan CERDIK,'" *Andrew's Disease of the Skin Clinical Dermatology.*, 2019. <https://p2ptm.kemkes.go.id/tag/hari-hipertensi-dunia-2019-know-your-number-kendalikan-tekanan-darahmu-dengan-cerdik%0Ahttps://p2ptm.kemkes.go.id/kegiatan-p2ptm/pusat-/hari-hipertensi-dunia-2019-know-your-number-kendalikan-tekanan-darahmu-dengan-cerdik>
- [5] R. Rodiana, "Classification of Hypertension Patients in Palembang by K-Nearest Neighbor and Local Mean K-Nearest Neighbor," *J. Stat. Data Sci.*, vol. 3, no. 1, pp. 27–35, Aug. 2024, doi: 10.33369/jsds.v3i1.32381.
- [6] D. Kurniawan, M. Wahyudi, L. Pujiastuti, and S. Sumanto, "Deteksi dan Prediksi Cerdas Penyakit Paru-Paru dengan Algoritma Random Fores," *Indones. J. Comput. Sci.*, vol. 3, no. 1, pp. 51–56, 2024.
- [7] F. L. ABDILLAH, "PREDIKSI HIPERTENSI DENGAN ALGORITMA MACHINE LEARNING MENGGUNAKAN K-NEAREST NEIGHBORS (K-NN)." Universitas Mercu Buana Jakarta, 2024.
- [8] R. R. Ismail, H. Wijaya, J. Siregar, and N. Nugroho, "Sistem Pakar Untuk Mendiagnosa Gangguan Somatisasi Menggunakan Metode K-Nearest Neighbors (KNN)".
- [9] M. A. Purba, "Implementasi Algoritma Nearest Neighbor dalam Mengukur Tingkat Kepuasan Masyarakat pada Pelayanan Samsat Medan Selatan," *Pelita Inform. Inf. dan Inform.*, vol. 8, no. 2, pp. 212–217, 2019.
- [10] P. P. Allorerung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 3, pp. 178–191, 2024.
- [11] J. Siregar and D. Wijayanti, "INTEGRASI MODEL NEURAL NETWORK DENGAN SVM DAN KNN UNTUK DETEKSI DINI PENYAKIT JANTUNG PADA PENDERITA HIPERTENSI," *Technol. J. Ilm.*, vol.



- 16, no. 1, pp. 104–109, 2025.
- [12] E. Erjon *et al.*, “Edukasi dan Deteksi Dini Pemeriksaan Tekanan Darah dalam Mencegah Risiko Komplikasi Hipertensi,” *Cakrawala J. Pengabd. Masy. Glob.*, vol. 4, no. 1, pp. 41–46, 2025.
- [13] F. D. FADILLAH, “Penerapan Metode K-Nearest Neighbor Untuk Klasifikasi Penyakit Hipertensi.” Universitas Malikussaleh, 2023.
- [14] S. Tondang, R. R. Prasetyo, R. Fulvian, Y. G. Sitorus, and G. Chrisnawati, “Analisis Perbandingan Algoritma K-Nearest Neighbor dan Ensemble Learning dalam Klasifikasi Penyakit Obesitas,” *RIGGSJ. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 4536–4548, 2025.
- [15] P. B. Utomo, M. Faruqziddan, E. H. S. Aulia, and S. D. Azzahra, “Perbandingan Skenario Balancing Oversampling dan Undersampling dalam Klasifikasi Resiko Kambuh Kanker Tiroid menggunakan Algoritma SVM Linear,” *JAMIJ. Ahli Muda Indones.*, vol. 5, no. 2, pp. 172–182, 2024.

