

Sistem Deteksi Dini Diabetes Melitus Dengan Teknik Klasifikasi Algoritma C4.5 Berdasarkan Rekam Medis di RS Tk. III Baladhika Husada Jember

Alviani Rodyatul Agustina¹, Mudafiq Riyan Pratama², Muhammad Yunus³, Ervina Rachmawati⁴

¹Manajemen Informasi Kesehatan, Politeknik Negeri Jember, alvianiagustina411@gmail.com

²Manajemen Informasi Kesehatan, Politeknik Negeri Jember, mudafiq.riyan@polije.ac.id

³Manajemen Informasi Kesehatan, Politeknik Negeri Jember, m.yunus@polije.ac.id

⁴Manajemen Informasi Kesehatan, Politeknik Negeri Jember, ervina_rachmawati@polije.ac.id

Keywords

Early Detection System,
C4.5 Algorithm,
Diabetes Mellitus,
Waterfall

ABSTRACT

Diabetes Mellitus (DM) is a chronic disease condition that occurs when the pancreas cannot produce insulin or when the body cannot use insulin effectively. At Baladhika Husada Jember Hospital, DM ranks among the top 10 diseases with the highest mortality rate of 6.99% in 2024. In efforts to prevent and control DM, a website-based early detection system was developed using the C4.5 algorithm classification technique with the Waterfall method. The research stages included creating C4.5 algorithm classification rules using RapidMiner tools, followed by development using the Waterfall method, which consists of the communication, planning, modeling, construction, and deployment stages. The classification rules were developed using preprocessed data from a total of 240 datasets, resulting in 172 clean datasets obtained from medical records at Baladhika Husada Jember Hospital. The training and testing data ratio was 50:50 using stratified sampling. Performance testing using the Confusion Matrix method yielded accuracy, precision, and recall values of 100% each, along with 8 classification rules that were subsequently implemented in the system. Based on the research results, random blood sugar is the most influential risk factor for DM, as it achieved the highest gain ratio. Recommendations for future researchers include increasing the amount of data and expanding the variety of data to help the system learn more complex patterns.

Kata Kunci

Sistem Deteksi Dini,
Algoritma C4.5,
Diabetes Mellitus,
Waterfall

ABSTRAK

Diabetes Melitus (DM) adalah kondisi penyakit kronis yang terjadi akibat pankreas tidak dapat memproduksi insulin atau ketika tubuh tidak dapat menggunakan insulin secara efektif. Pada RS Tk. III Baladhika Husada Jember, DM masuk ke dalam 10 besar penyakit dengan persentase kematian tertinggi sebesar 6,99% pada tahun 2024. Dalam upaya pencegahan dan pengendalian DM, maka dikembangkan sistem deteksi dini berbasis *website* dengan pendekatan teknik klasifikasi algoritma C4.5 menggunakan metode *Waterfall*. Tahapan penelitian ini meliputi pembuatan aturan klasifikasi algoritma C4.5 dengan *tools* RapidMiner, selanjutnya dilakukan pengembangan dengan metode *Waterfall* yaitu tahapan *communication*, *planning*, *modeling*, *construction* dan *deployment*. Penyusunan aturan klasifikasi menggunakan data hasil *preprocessing* dari total 240 *dataset* diperoleh data bersih sebanyak 172 *dataset* yang berasal dari rekam medis di RS Tk. III Baladhika Husada Jember. Perbandingan rasio *data training* dan *testing* adalah 50:50 dengan teknik *stratified sampling*. Proses pengujian kinerja dengan metode *Confusion Matrix* menghasilkan nilai *accuracy*, *precision*, dan *recall* masing-masing 100% serta aturan klasifikasi sebanyak 8 *rules* yang selanjutnya diimplementasikan dalam sistem. Berdasarkan hasil penelitian, gula darah acak merupakan faktor risiko yang paling berpengaruh terhadap penyakit DM karena memperoleh nilai *Gain Ratio* tertinggi. Saran untuk peneliti selanjutnya memperbanyak jumlah data dan memperluas ragam data agar dapat membantu sistem mempelajari pola yang lebih kompleks.

Korespondensi Penulis:

Alviani Rodyatul Agustina,
Politeknik Negeri Jember,
Jl. Mastrip PO BOX 164 Jember, Jawa Timur
Telepon : +6285895150185
Email: alvianiagustina411@gmail.com

Submitted : 02-01-2026; Accepted : 27-02-2026;
Published : 17-03-2026

Copyright (c) 2024 The Author (s) This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0)

1. PENDAHULUAN

Diabetes Melitus adalah kondisi penyakit kronis akibat pankreas tidak mampu memproduksi insulin yang cukup atau tubuh tidak dapat menggunakan insulin secara efektif dan ditandai dengan kadar glukosa darah yang meningkat [1]. Diabetes Melitus adalah penyakit tidak menular yang menjadi salah satu penyebab utama terkait angka morbiditas dan mortalitas. DM menempati urutan ke-9 pada tahun 2019 sebagai penyebab kematian di dunia [2]. Negara Indonesia pada tahun 2021 menempati peringkat ke tiga dengan angka *undiagnosed* diabetes tertinggi yaitu 14,3 juta kasus setelah China, dan India [3]. Berdasarkan Hasil Riset Kesehatan Dasar pada tahun 2018, prevalensi kasus Diabetes Melitus pada penduduk semua umur mencapai 2,6% dan menempati urutan ke lima [4].

Data pada Profil Dinas Kesehatan Provinsi Jawa Timur tahun 2021, Kabupaten Jember menempati urutan ke delapan dengan kasus Diabetes Melitus terbanyak di Provinsi Jawa Timur mencapai 37.786 dari total keseluruhan di Jawa Timur sebanyak 929.810 kasus [5]. Pada RS Tk. III Baladhika Husada Jember, Diabetes Melitus masuk ke dalam 10 besar penyakit dan memiliki persentase kematian tertinggi sebesar 6,99% pada tahun 2024.

Upaya dalam pencegahan penyakit dapat dilakukan dengan cara deteksi dini. Deteksi dini merupakan upaya pencegahan dan pengendalian penyakit tidak menular yang dilakukan terhadap individu atau kelompok secara rutin untuk menemukan faktor risiko sedini mungkin [6]. Salah satu cara pendeteksian dini DM dapat dilakukan dengan memanfaatkan klasifikasi *data mining* dan algoritma *machine learning*.

Algoritma C4.5 adalah algoritma pada bidang *machine learning* yang mempunyai performa tinggi, baik dari segi *accuracy* maupun kecepatan perhitungan tinggi dalam pengklasifikasian [7]. Hasil pengklasifikasian berupa pohon keputusan, dapat menangani nilai atribut yang hilang, menghasilkan aturan (*rule*) yang mudah dimengerti dan proses perhitungan termasuk tercepat diantara algoritma lain karena dapat memotong variabel [8].

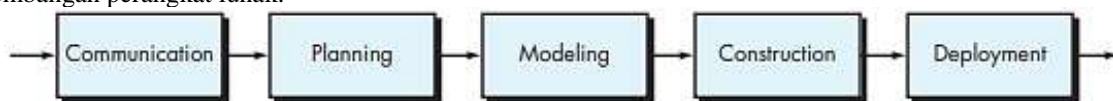
Dalam penelitian yang dilakukan oleh Pratama *et al.*, (2024) yang berjudul “*Diagnosis of Stroke and Diabetes Mellitus With Classification Techniques Using Decision Tree Method*”, hasil dari penelitian tersebut menunjukkan bahwa algoritma C4.5 memiliki *accuracy* 86.67%, *precision* 81.25%, dan *recall* 92.86% [9]. Penelitian lain juga dilakukan oleh Fitri *et al.*, (2023) yang berjudul “Perbandingan Klasifikasi Naïve Bayes dan C4.5 untuk Diagnosa Penyakit Stroke”, hasil dari penelitian tersebut menunjukkan bahwa algoritma C4.5 memiliki tingkat *accuracy* 92,22%, dan algoritma Naïve Bayes dengan tingkat *accuracy* sebesar 89,22% [10]. Maka, dapat diketahui bahwa algoritma C4.5 merupakan algoritma paling optimal dalam perbandingan klasifikasi diagnosis penyakit stroke.

Berdasarkan hasil dari penelitian terdahulu, algoritma C4.5 dalam pengklasifikasian terbukti efektif dengan tingkat *accuracy* yang tinggi, sehingga penulis tertarik untuk membuat suatu sistem deteksi dini penyakit Diabetes Melitus dengan menggunakan algoritma C4.5 berdasarkan rekam medis pasien dalam penentuannya menggunakan beberapa variabel yang meliputi, riwayat keluarga Diabetes Melitus, IMT, tekanan darah, dan gula darah acak [11]. Sistem deteksi ini dikembangkan dengan berbasis *website* dan adanya sistem ini dapat memudahkan dalam melakukan deteksi sedini mungkin, sehingga dapat dilakukan penanganan lebih awal bagi orang yang berisiko Diabetes Melitus.

2. METODE PENELITIAN

2.1 Jenis penelitian

Penelitian ini dilakukan untuk membuat sistem deteksi dini penyakit Diabetes Melitus berbasis *website* dengan metode kuantitatif menggunakan teknik klasifikasi dengan algoritma C4.5 dan metode *Waterfall* sebagai metode pengembangan perangkat lunak.



Gambar 1. Metode Waterfall [12]

Metode *Waterfall* adalah salah satu metode pengembangan sistem informasi yang termasuk bagian dari SDLC. *Waterfall* mengharuskan pengembangan sistem secara berurutan (sekuensial) sesuai dengan tahapannya [13]. Beberapa tahapan dari metode *Waterfall* dapat uraikan sebagai berikut:

- a. *Communication*
Pada tahap ini suatu komunikasi dilakukan dalam rangka menentukan tujuan yang akan menghasilkan inisiasi proyek mulai dari temuan permasalahan, pengumpulan data awal, analisis kebutuhan, fungsi serta tujuan dari pembuatan sistem.
- b. *Planning*
Pada tahap ini dilaksanakan perencanaan yaitu pembagian tugas, penjadwalan kegiatan, dan sumber daya yang dibutuhkan selama pengembangan sistem.
- c. *Modeling*
Pada tahap ini perancangan serta pemodelan sistem dilaksanakan yang berupa pembuatan struktur data, tampilan antarmuka sistem, maupun algoritma program. Bertujuan membuat gambaran sistem sehingga dapat lebih dipahami.
- d. *Construction*
Pada tahap ini proses pengkodean dan pengujian dari sistem dilaksanakan. Desain sistem pada tahap sebelumnya direalisasikan dengan diterjemahkan menjadi bahasa yang diproses oleh mesin berupa kumpulan kode.

Selanjutnya dilakukan pengujian untuk memeriksa kemungkinan kesalahan pada sistem sehingga dapat dilakukan perbaikan.

e. *Deployment*

Pada tahapan ini dilakukan proses penyebaran sistem dari lingkungan pengembang ke lingkungan produksi atau biasa disebut dengan *hosting*.

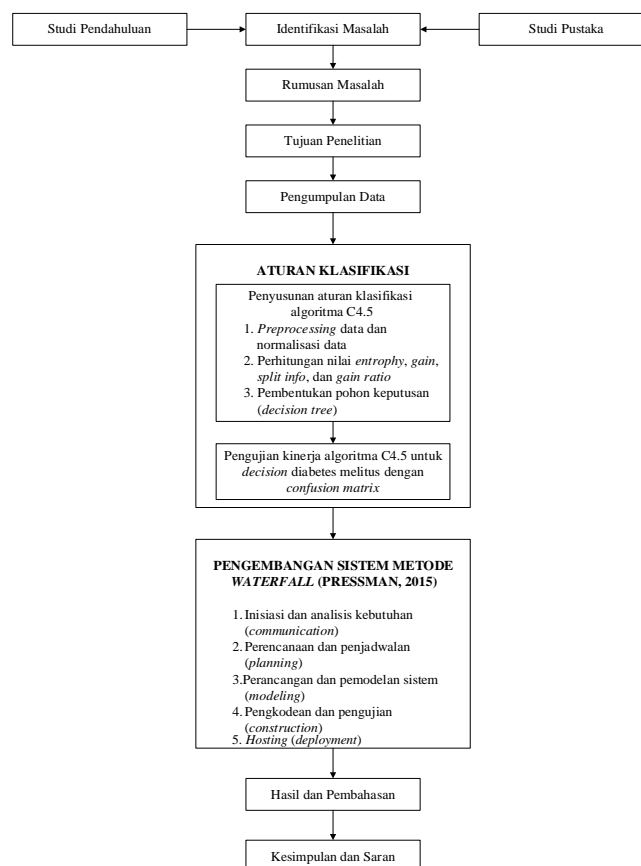
2.2 Objek Penelitian

Penelitian ini menggunakan objek berupa data sekunder yang diperoleh dari rekam medis pasien rawat inap periode 2020-2024 pada Rumah Sakit Tk. III Baladhika Husada Jember dengan diagnosis Diabetes Melitus (kode ICD X E11.-, E12.-, E13.-, E14.-) serta rekam medis lain dengan diagnosis *Non* Diabetes Melitus yaitu hipertensi (kode ICD X I10). Penggunaan diagnosis hipertensi ini didasarkan pada hasil konfirmasi dengan dokter spesialis penyakit dalam yang menyatakan bahwa hipertensi memiliki tingkat kemiripan faktor risiko dengan Diabetes Melitus.

2.3 Teknik Pengumpulan Data

Teknik observasi digunakan dalam penelitian ini dimana dilakukan pada rekam medis pasien berdasarkan variabel-variabel yang sudah ditentukan. Observasi dilakukan untuk pengumpulan data yang dibutuhkan dalam pembuatan aturan dengan algoritma C4.5. Selain itu, juga dilakukan dokumentasi terhadap formulir rekam medis.

2.4 Tahapan Penelitian



Gambar 2. Tahapan Penelitian

Berikut uraian dari tahapan penelitian:

- a. Studi Pendahuluan, diawali dengan upaya pencarian permasalahan dan pengumpulan data-data yang berkaitan melalui laporan global Diabetes Melitus dari WHO, *International Diabetes Federation* (IDF), profil kesehatan Jawa Timur, profil kesehatan Kabupaten Jember, dan data laporan RS Tk.III Baladhika Husada Jember.
- b. Studi Pustaka, pengumpulan informasi yang diperoleh dari berbagai sumber seperti artikel penelitian terdahulu, buku elektronik, laporan, dan regulasi yang berkaitan dengan penelitian.
- c. Identifikasi Masalah, pencarian titik masalah yang berkaitan dengan kejadian penyakit Diabetes Melitus.
- d. Rumusan Masalah, pertanyaan singkat mengenai permasalahan dan solusi yang akan dibahas dalam penelitian.
- e. Tujuan Penelitian, menggambarkan langkah-langkah yang dilalui oleh peneliti dalam mencapai hasil penelitian.
- f. Pengumpulan Data, mengobservasi rekam medis pasien rawat inap di RS Tk. III Baladhika Husada Jember.

- g. Aturan Klasifikasi, dilakukan penyusunan aturan C4.5 meliputi *Preprocessing* data, menormalisasikan data, menghitung nilai *Entropy*, *Gain*, *Split Info*, dan *Gain Ratio* serta membuat *decision tree*. Selanjutnya pengujian kinerja dengan *Confusion Matrix*.
- h. Pembuatan Sistem, menggunakan metode *Waterfall* dimulai dari tahap *communication* untuk analisis kebutuhan. Tahap *planning* yaitu perencanaan dan penjadwalan. Tahap *modeling* yaitu perancangan dan pemodelan dengan membuat gambaran sistem berupa desain *interface*, *Flowchart System*, *data flow diagram* (DFD), dan *Entity Relationship Diagram* (ERD). Tahap *Construction* meliputi pengkodean program dan pengujian sistem yang telah dibangun. Tahapan terakhir adalah *Deployment* dengan menempatkan sistem deteksi dini DM pada *Server Hosting*.
- i. Hasil dan Pembahasan, penjabaran pembahasan dari hasil pengembangan sistem deteksi dini DM.
- j. Kesimpulan dan Saran, peneliti merumuskan kesimpulan berdasarkan hasil penelitian dan menyusun saran untuk penelitian kedepannya yang bersifat membangun.

3. HASIL DAN ANALISIS

Penelitian ini dilaksanakan melalui dua tahapan utama, yaitu analisis data dan pengembangan sistem. Pada tahap analisis data dilakukan serangkaian proses meliputi pengumpulan data, *preprocessing* data, normalisasi data, *evaluation*, dan *presentation knowledge*. Hasil analisis data selanjutnya digunakan untuk pengembangan sistem deteksi dini diabetes melitus dengan metode *Waterfall*.

3.1 Analisis Data

Analisis data dilakukan sesuai dengan proses *Knowledge Discovery in Database* (KDD), dimulai dengan pengumpulan data rekam medis rawat inap dengan diagnosis Diabetes Melitus dan *Non Diabetes Melitus*, selanjutnya *preprocessing* data meliputi *data selection* dan *data cleaning*, normalisasi data, proses *mining*, *evaluation*, *presentation knowledge*, dan pengujian kinerja.

3.1.1 Pengumpulan Data

Dataset diperoleh dari rekam medis pasien rawat inap tahun 2020 sampai dengan 2024 di RS Tk. III Baladhika Husada Jember. Data yang dikumpulkan sebanyak 240 data, dengan diagnosis Diabetes Melitus sebanyak 140 data dan *Non Diabetes Melitus* sebanyak 100 data. Lembar observasi digunakan sebagai instrumen pengumpulan data yang dibuat dalam format *excel*.

3.1.2 Data Selection

Data yang terkumpul selanjutnya dilakukan pemilihan data sesuai dengan variabel faktor risiko DM yang telah ditentukan. Berdasarkan Pedoman Pengendalian Diabetes Melitus dan Penyakit Metabolik yang diterbitkan oleh Depkes RI Tahun 2008 [11], serta hasil konfirmasi dengan dokter spesialis penyakit dalam di RS Tk. III Baladhika Husada Jember dan menyesuaikan dengan data rekam medis yang tersedia didukung dengan referensi pada *e-book* yang berjudul *Clinical Reasoning : Endokrin/Metabolisme/Nutrisi* oleh dr. Muhammad Agita Hutomo, M.M.R Tahun 2024 [14], maka penentuan risiko penyakit Diabetes Melitus menggunakan 4 variabel, yaitu riwayat keluarga DM, IMT, tekanan darah dan gula darah acak.

3.1.3 Data Cleaning

Proses *cleaning* dilakukan dengan membuang data yang bersifat *missing value*, inkonsistensi, dan duplikasi. *Missing value* banyak ditemukan pada variabel IMT, tekanan darah, dan riwayat keluarga DM. Inkonsistensi dan duplikasi data tidak ditemukan. Data-data tersebut dilakukan pembersihan dengan dihilangkan agar data relevan dengan diagnosis Diabetes Melitus maupun *Non Diabetes Melitus* dan tersisa sebanyak 172 data, dimana diagnosis Diabetes Melitus sebanyak 112 data dan *Non Diabetes Melitus* sebanyak 60 data.

Tabel 1. Data Penelitian

No.	IMT	Tekanan Darah	Gula Darah Acak	Riwayat Keluarga	Class
1	22	120/70 mmHg	309 mg/dL	Tidak Ada	Diabetes Melitus
2	24	110/70 mmHg	158 mg/dL	Tidak Ada	Diabetes Melitus
3	21,3	110/70 mmHg	273 mg/dL	Tidak Ada	Diabetes Melitus
4	20	120/80 mmHg	300 mg/dL	Tidak Ada	Diabetes Melitus
5	23	180/100 mm/Hg	479 mg/dL	Tidak Ada	Diabetes Melitus
6	33,3	200/110 mmHg	110 mg/dL	Tidak Ada	<i>Non Diabetes Melitus</i>
7	20,8	170/110 mmHg	97,4mg/dL	Tidak Ada	<i>Non Diabetes Melitus</i>

No.	IMT	Tekanan Darah	Gula Darah Acak	Riwayat Keluarga	Class
8	21,6	160/100 mmHg	108 mg/dL	Tidak Ada	Non Diabetes Melitus
9	21	220/100 mmHg	99 mg/dL	Tidak Ada	Non Diabetes Melitus
...
172	22	170/80 mmHg	65 mg/dL	Tidak Ada	Non Diabetes Melitus

3.1.4 Normalisasi Data

Normalisasi data merupakan proses dalam *data mining* untuk menetapkan rentang nilai atribut atau variabel dari data yang telah dikumpulkan [15]. Normalisasi disini dilakukan dengan mengkategorikan nilai pada masing-masing variabel dengan kategori yang telah ditetapkan. Berikut pengkategorianya:

- a. Riwayat Keluarga DM : Ada dan Tidak Ada [11].
- b. IMT : Underweight (< 18,5 g/m2), Normal (18,5 - 24,9 g/m2), Obesitas I (25 – 29,9 g/m2), dan Obesitas II (\geq 30 g/m2) [14].
- c. Tekanan Darah : Normal (< 120 dan 80 mmHg), Pra-hipertensi (120 - 139 atau 80 - 89 mmHg), Hipertensi Tingkat I (140 - 159 atau 90 - 99 mmHg), Hipertensi Tingkat II (> 160 atau > 100 mmHg)[16].
- d. Gula Darah Acak : Normal (< 140 mg/dL), Prediabetes (140 – 199 mg/dL), dan Diabetes (\geq 200 mg/dL) [14].

Tabel 2. Hasil Normalisasi Data

No.	IMT	Tekanan Darah	Gula Darah Acak	Riwayat Keluarga	Class
1	Normal	Pra-hipertensi	Diabetes	Tidak Ada	Diabetes Melitus
2	Normal	Normal	Prediabetes	Tidak Ada	Diabetes Melitus
3	Normal	Normal	Diabetes	Tidak Ada	Diabetes Melitus
4	Normal	Pra-hipertensi	Diabetes	Tidak Ada	Diabetes Melitus
5	Normal	Hipertensi Tingkat II	Diabetes	Tidak Ada	Diabetes Melitus
6	Obesitas II	Hipertensi Tingkat II	Normal	Tidak Ada	Non Diabetes Melitus
7	Normal	Hipertensi Tingkat II	Normal	Tidak Ada	Non Diabetes Melitus
8	Normal	Hipertensi Tingkat II	Normal	Tidak Ada	Non Diabetes Melitus
9	Normal	Hipertensi Tingkat II	Normal	Tidak Ada	Non Diabetes Melitus
...
172	Normal	Hipertensi Tingkat II	Normal	Tidak Ada	Non Diabetes Melitus

Tabel 2 merupakan data yang telah dinormalisasikan sesuai dengan kategori masing-masing variabel, selanjutnya dapat dilakukan proses *mining*.

3.1.5 Proses Mining

Proses *mining* adalah tahapan pengujian dengan pembagian rasio antara *data training* dan *data testing* dengan *tools* RapidMiner. Teknik pemilihan sampel menggunakan *stratified sampling* yaitu pemilihan *data testing* dilakukan secara acak namun tetap memperhatikan komposisi dari masing-masing kelas. *Stratified sampling* dapat memastikan setiap kelompok atau sub-kelompok terdistribusi dengan proporsional, sehingga dapat mencegah adanya bias ketika terdapat suatu kelas atau atribut dominan dalam sampel [17]. Perbandingan sampel yang digunakan adalah 95:5, 90:10, 85:15, 80:20, 75:25, 70:30, 65:35, 60:40, 55:45, dan 50:50. Misalnya pada penelitian ini terdapat 172 data yang terdiri dari 112 data kelas Diabetes Melitus dan 60 data kelas Non Diabetes Melitus, perbandingan yang digunakan dalam pembagian yaitu 50:50. Artinya *data testing* dengan kelas Diabetes Melitus adalah 50% dari jumlah data dengan kelas Diabetes Melitus, yaitu 56 data. Sama halnya dengan *data testing* dengan kelas Non Diabetes Melitus adalah sebanyak 50% dari jumlah data dengan kelas Non Diabetes Melitus, yaitu 30 data. Proses pengujian dilakukan dengan beberapa perbandingan yang berbeda bertujuan untuk mencari jumlah komposisi terbaik dengan *accuracy* tertinggi. Hal ini sejalan dengan penelitian yang dilakukan oleh Musu et al., (2021) bahwa tingkat *accuracy* algoritma C4.5 bersifat fluktuatif, sehingga dalam pengujian dilakukan dengan menguji perbandingan antara *data training* dan *data testing* untuk mendapatkan *accuracy* yang terbaik. Berikut merupakan hasil pengujian tingkat *accuracy* dari beberapa perbandingan data yang digunakan [18].

Tabel 3. Hasil Accuracy Perbandingan Jumlah Dataset

No	Perbandingan Jumlah Dataset	Accuracy (%)
1	95:5 / 163:8	87,50%
2	90:10 / 154:17	94,12%
3	85:15 / 146:25	96%
4	80:20 / 137:34	97,06%
5	75:25 / 129:43	97,67%
6	70:30 / 120:51	98,04%
7	65:35 / 111:60	98,33%
8	60:40 / 103:68	98,53%
9	55:45 / 94:77	98,70%
10	50:50 / 86:86	100%

Berdasarkan tabel di atas nilai *accuracy* tertinggi diperoleh dari hasil perbandingan 50:50 antara *data training* dan *data testing* sebesar 100%. Oleh karena itu, aturan (*rule*) klasifikasi dibuat dengan menggunakan perbandingan 50:50 yang terdiri atas 86 *data training* dan 86 *data testing*.

3.1.6 Evaluation

Evaluasi terhadap perbandingan *class* hasil klasifikasi dengan *class* aktual dilakukan dengan *Confusion Matrix*.

Tabel 4. *Confusion Matrix* Hasil Algoritma C4.5 Menggunakan RapidMiner

		Prediksi	
		DM	Non DM
Kelas Asli	DM	TP = 56	FN = 0
	Non DM	FP = 0	TN = 30

Berdasarkan Tabel 4, hasil klasifikasi dari 86 data antara *class* aktual dengan *class* prediksi telah sesuai. Selanjutnya menghitung nilai *accuracy*, *precision* dan *recall* dengan perhitungan sebagai berikut.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 = \frac{56 + 30}{56 + 30 + 0 + 0} \times 100 = 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\% = \frac{56}{56 + 0} \times 100\% = 100\%$$

$$Recall = \frac{TP}{TP + FN} \times 100\% = \frac{56}{56 + 0} \times 100\% = 100\%$$

Nilai *accuracy* menunjukkan banyaknya data yang berhasil diprediksi dengan benar baik positif maupun negatif dari keseluruhan data, artinya data yang diprediksi benar Diabetes Melitus dan Non Diabetes Melitus dari keseluruhan kelas sebesar 100%. Nilai *precision* menunjukkan jumlah hasil prediksi benar positif dari keseluruhan hasil data yang diprediksi positif, artinya data yang diprediksi Diabetes Melitus yang benar dari keseluruhan data yang diprediksi Diabetes Melitus sebesar 100%. Nilai *recall* menunjukkan hasil prediksi benar positif dari keseluruhan data yang benar positif, artinya data yang diprediksi benar Diabetes Melitus dari keseluruhan data yang diprediksi benar Diabetes Melitus sebesar 100%. Nilai *accuracy* menunjukkan seberapa sering sistem deteksi ini benar secara keseluruhan, *precision* menunjukkan keakuratan data prediksi seberapa tepat sistem deteksi dalam memprediksi benar DM, *recall* menunjukkan keakuratan data aktual seberapa banyak data yang benar DM dengan data hasil prediksi.

Setelah dilakukan analisis terhadap *data training* dan *data testing* yang digunakan, hasil nilai *accuracy*, *precision*, dan *recall* mencapai nilai yang sempurna disebabkan karena kurang beragamnya sampel data yang digunakan. Selain itu, pemilihan parameter gula darah acak juga menjadi salah satu penyebab, karena model menjadikan gula darah acak sebagai acuan utama dalam proses klasifikasi, ketika gula darah acak bernilai >140 mg/dL maka akan terklasifikasi menjadi DM begitupun sebaliknya. Kadar glukosa darah yang tinggi melebihi batas normal menjadi karakteristik utama penyakit Diabetes Melitus [19].

3.1.7 Presentation Knowledge

Data training yang telah dilakukan proses perhitungan klasifikasi menggunakan algoritma C4.5 akan menghasilkan pohon keputusan. Pohon keputusan tersebut membentuk aturan klasifikasi sebanyak 8 *rules* dalam klasifikasi faktor risiko penyakit Diabetes Melitus. Berikut merupakan pohon keputusan dan aturan dari algoritma C4.5 yang dihasilkan.



Gambar 3. Pohon Keputusan Algoritma C4.5

3.2 Pengembangan Sistem

Pengembangan sistem deteksi dini Diabetes Melitus dilakukan dengan metode pengembangan *Waterfall*, berikut penjabaran setiap tahapannya:

3.2.1 Communication

Pada tahapan ini dilakukan analisis kebutuhan fungsional dan non fungsional. Kebutuhan fungsional diperoleh dari wawancara dengan dokter spesialis penyakit dalam di RS Tk. III Baladhika Husada Jember. Untuk penentuan faktor risiko DM dan fitur-fitur yang akan dikembangkan dalam sistem deteksi dini DM. Kebutuhan non fungsional mencakup alat-alat dan bahan yang diperlukan dalam pengembangan.

3.2.2 Planning

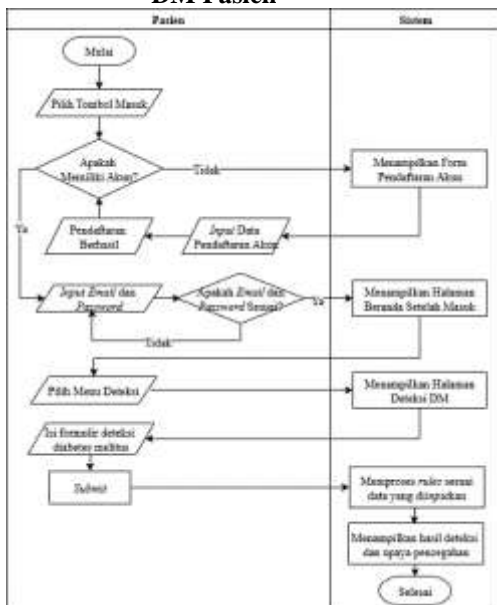
Perencanaan ini dilakukan dengan penjadwalan kegiatan dan disajikan dalam *Gantt Chart* yang digunakan untuk menggambarkan jadwal dan tugas.

3.2.3 Modeling

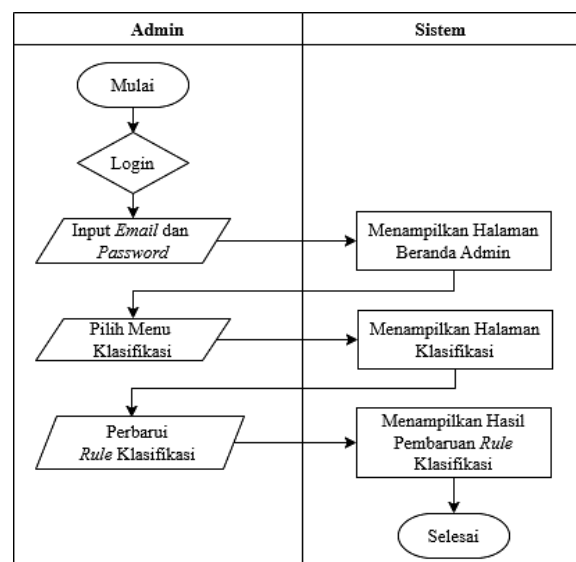
Pada tahapan ini dilakukan perancangan *Flowchart System*, *Data Flow Diagram (DFD)*, *Entity Relationship Diagram (ERD)*, dan *Desain Interface*.

a. Flowchart System

Flowchart System Untuk Deteksi Dini DM Pasien



Flowchart System Admin Untuk Memperbarui Rule Klasifikasi



Gambar 4. Flowchart System

Flowchart System Untuk Deteksi Dini DM Pasien merupakan gambaran alur ketika pasien akan melakukan deteksi, pasien login dengan memilih tombol masuk kemudian memasukkan email dan password. Apabila belum memiliki akun maka harus melakukan pendaftaran baru kemudian dapat login pada sistem. Ketika melakukan pendaftaran akun, pasien menginput data pendaftaran, setelah selesai dapat melakukan login. Setelah login akan tampil

halaman beranda setelah *login* kemudian pilih fitur ‘Deteksi’, didalamnya terdapat formulir deteksi berisi data-data yang dibutuhkan untuk deteksi. Setelah itu, pasien mengisi formulir deteksi, kemudian pilih tombol ‘Submit’ ketika sudah selesai maka sistem akan memproses *rules* atau aturan klasifikasi dan sistem akan menampilkan hasil deteksi.

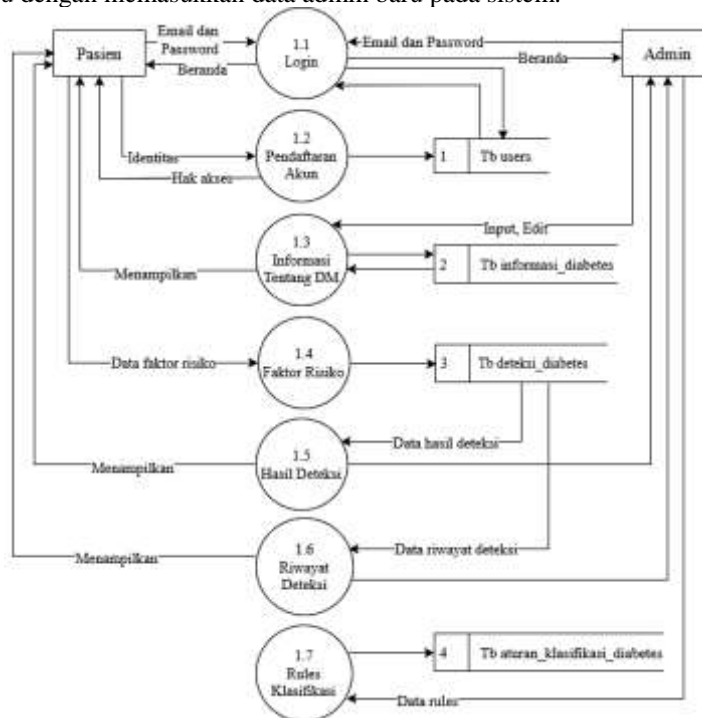
Flowchart System Admin Untuk Memperbarui Rule Klasifikasi merupakan gambaran admin untuk memperbarui *rule* klasifikasi pada sistem. Admin *login* akun terlebih dahulu dengan memasukkan *email* dan *password*. Setelah itu, akan masuk pada beranda admin. Pada beranda admin pilih menu ‘Klasifikasi’ kemudian admin dapat melakukan pembaruan *rules* setelah selesai, maka sistem akan menampilkan hasil *rule* klasifikasi yang telah diperbarui.

b. *Data Flow Diagram (DFD)*



Gambar 5. DFD Level 0

DFD *level 0* dari sistem deteksi dimana terdiri atas dua entitas yaitu pasien dan admin. Pasien dalam sistem ini dapat memasukkan data diri dan data faktor risiko dari Diabetes Melitus yang dimiliki kemudian akan diproses oleh sistem deteksi dan pasien akan mendapatkan hasil dari deteksi berdasarkan data faktor. Admin yang telah memiliki akses dapat melihat, mengedit, menghapus informasi tentang DM, mengelola *rules* algoritma C4.5 dan dapat menambah admin yang baru dengan memasukkan data admin baru pada sistem.

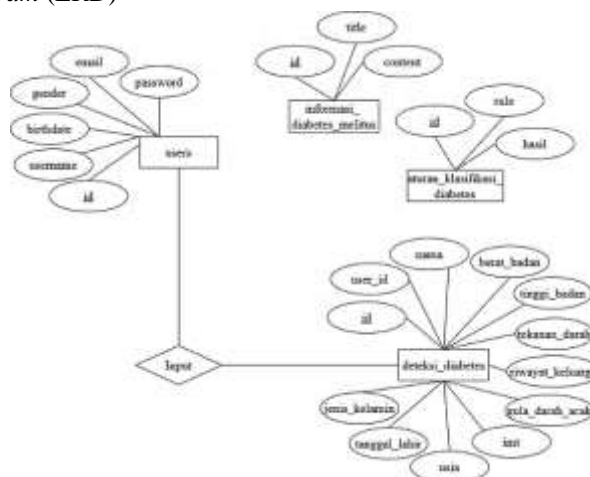


Gambar 6. DFD Level 1

DFD *level 1* dimana terdapat dua entitas yaitu pasien dan admin. Pasien dapat melakukan *login* dan pendaftaran akun jika belum memiliki akun. Data pasien akan tersimpan pada tabel users dalam *database*. Pasien juga dapat melihat informasi tentang DM yang memuat informasi seputar DM dimana data tersebut tersimpan pada *tb informasi diabetes* dalam *database*. Selain itu, pasien dapat melakukan deteksi dengan input faktor risiko dimana sistem akan melakukan proses klasifikasi dengan *rules* algoritma C4.5 dan hasilnya disimpan pada tabel deteksi diabetes dalam *database*. Hasil deteksi tersebut juga dapat diakses kembali oleh pasien pada menu riwayat.

Admin dapat *login* dan masuk pada beranda admin dimana datanya tersimpan pada tabel users dalam *database*. Admin memiliki akses untuk melihat, menginput, mengedit, dan menghapus informasi tentang DM yang nantinya data tentang DM akan tersimpan pada tabel informasi diabetes dalam *database*. Selain itu, admin juga memiliki akses untuk melihat hasil dan riwayat deteksi seluruh pasien dimana data tersebut tersimpan pada tabel deteksi diabetes dalam *database*. Admin juga dapat memperbarui *rules* algoritma C4.5 pada menu klasifikasi dimana tersimpan pada tabel aturan klasifikasi diabetes dalam *database*.

c. Entity Relationship Diagram (ERD)



Gambar 7. Entity Relationship Diagram (ERD)

Gambar 7 merupakan *entity relationship diagram* sistem deteksi dini Diabetes Melitus yang terdiri atas 4 entitas yaitu *users*, *deteksi_diabetes*, *informasi_diabetes*, dan *aturan_klasifikasi_diabetes*. Terdapat dua entitas yang berhubungan yaitu antara entitas *users* dengan *deteksi_diabetes* relasinya *One to Many* adalah satu *user* dapat menginputkan beberapa data untuk melakukan deteksi dan data tersebut akan masuk dalam tabel *deteksi_diabetes* yang nantinya akan diproses oleh sistem dan muncul hasil deteksi.



Gambar 8. Physical Data Model (PDM)

Pada Gambar 8 merupakan *physical data model* yang menunjukkan isi pada tabel dalam *database* meliputi nama kolom, tipe data, ukuran, *primary key*, dan lain sebagainya.

d. Desain Interface



Gambar 9. Desain Interface Halaman Deteksi

Gambar 9 merupakan rancangan tampilan halaman formulir deteksi dimana pengguna yang telah memiliki akun dapat menggunakan fitur ini. Setelah pengguna masuk maka dapat melakukan deteksi dini secara mandiri dengan mengisi variabel-variabel yang terdapat pada sistem kemudian pilih 'Submit' jika telah selesai. Kemudian hasil deteksi akan muncul berisiko DM atau tidak berisiko DM.

3.2.4 Construction

Pada tahapan ini dilakukan pengkodean dan pengujian sistem. Pengkodean sistem dilakukan dengan aplikasi *Visual Studio Code* dalam bahasa pemrograman PHP, *Framework Codeigniter 3*, dan *Database MySQL*. Pengujian sistem dilakukan melalui 2 tahapan yaitu pengujian fungsionalitas dengan *Black-Box Testing* pada 2 admin (PMIK) serta 2 pasien dimana fungsi dari masing-masing menu *user* telah dinyatakan berhasil dijalankan dan pengujian kinerja

sistem dilakukan dengan memasukkan 86 *data testing* ke dalam sistem dan terklasifikasi sesuai dengan *class* aktualnya.

Tabel 5. Hasil Pengujian Kinerja Sistem

No	RapidMiner		Sistem
	Class Aktual	Class Klasifikasi	
1	Diabetes Melitus	Diabetes Melitus	Sesuai
2	Diabetes Melitus	Diabetes Melitus	Sesuai
3	Diabetes Melitus	Diabetes Melitus	Sesuai
4	Diabetes Melitus	Diabetes Melitus	Sesuai
5	Diabetes Melitus	Diabetes Melitus	Sesuai
6	Non Diabetes Melitus	Non Diabetes Melitus	Sesuai
7	Non Diabetes Melitus	Non Diabetes Melitus	Sesuai
8	Non Diabetes Melitus	Non Diabetes Melitus	Sesuai
9	Non Diabetes Melitus	Non Diabetes Melitus	Sesuai
...
86	Non Diabetes Melitus	Non Diabetes Melitus	Sesuai

3.2.5 Deployment

Pada tahapan ini dilakukan dengan menempatkan sistem pada *server hosting* yang terhubung dengan jaringan internet, sistem deteksi dini DM dapat diakses melalui alamat <https://diabetologix.mikpolije.com> secara *online* oleh pengguna.

4. KESIMPULAN

- Penyusunan aturan klasifikasi menggunakan data hasil *preprocessing* dari total 240 dataset diperoleh data bersih sebanyak 172 *dataset* yang berasal dari rekam medis di RS Tk. III Baladhika Husada Jember, menggunakan perbandingan rasio 50:50 dengan teknik *stratified sampling* dimana masing-masing data *training* dan testing berjumlah 86 data. Aturan klasifikasi yang terbentuk sebanyak 8 *rules* dimana variabel gula darah acak menjadi faktor risiko yang paling berpengaruh terhadap penyakit diabetes melitus karena memperoleh nilai *Gain Ratio* tertinggi sehingga variabel gula darah acak menjadi *node* akar pada pohon keputusan yang terbentuk.
- Pengujian kinerja algoritma C4.5 dengan metode *Confusion Matrix* dari hasil perhitungan menggunakan *tools* RapidMiner menghasilkan nilai *accuracy* sebesar 100%, *precision* 100%, dan *recall* 100%. Nilai *accuracy* menunjukkan seberapa sering sistem deteksi ini benar secara keseluruhan, *precision* menunjukkan keakuratan data prediksi seberapa tepat sistem deteksi dalam memprediksi benar DM, *recall* menunjukkan keakuratan data aktual seberapa banyak data yang benar DM dengan data hasil prediksi.
- Tahapan *communication* dilakukan dengan menganalisis kebutuhan dalam pengembangan sistem, meliputi variabel faktor risiko yang digunakan untuk deteksi dini DM dan fitur-fitur yang diperlukan.
- Tahapan *planning* dilakukan dengan pembuatan jadwal kegiatan pengembangan sistem deteksi dini Diabetes Melitus dalam bentuk *Gantt Chart* yang digunakan untuk menggambarkan jadwal dan tugas.
- Tahapan *modeling* dilakukan dengan pembuatan *flowchart system*, *data flow diagram*, *entity relationship diagram*, dan desain *interface* dalam perancangan sistem deteksi dini Diabetes Melitus.
- Tahapan *construction* meliputi pengkodean dan pengujian sistem. Pengkodean sistem dilakukan dengan aplikasi *Visual Studio Code* dalam bahasa pemrograman PHP, *framework CodeIgniter 3*, dan *database MySQL*. Pengujian sistem dilakukan melalui 2 tahapan yaitu pengujian fungsionalitas dengan *black-box testing* pada 2 admin (PMIK) serta 2 pasien dimana fungsi dari masing-masing menu *user* telah dinyatakan berhasil dijalankan dan pengujian kinerja sistem dilakukan dengan memasukkan 86 data *testing* ke dalam sistem dan terklasifikasi sesuai dengan *class* aktualnya.
- Tahapan *deployment* dilakukan dengan menempatkan sistem pada *server hosting* yang terhubung dengan jaringan internet, sistem deteksi dini DM dapat diakses melalui alamat <https://diabetologix.mikpolije.com> secara *online* oleh pengguna.

Saran untuk peneliti selanjutnya dapat memperbanyak jumlah data dan memperluas ragam data agar data lebih bervariasi sehingga dapat membantu sistem mempelajari pola yang lebih kompleks. Selain itu, dapat dilakukan penambahan variabel faktor risiko yang lebih lengkap sehingga aturan klasifikasi yang terbentuk lebih detail serta dapat dikembangkan dengan pengelompokan sesuai jenis-jenis atau tipe dari penyakit Diabetes Melitus.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan selama proses penelitian ini. Ucapan terima kasih disampaikan kepada RS Tk. III Baladhika Husada Jember sebagai instansi yang telah bersedia menjadi objek penelitian. Penulis juga menghargai kontribusi dari dosen pembimbing, rekan-rekan, serta pihak yang turut membantu dalam penyusunan dan publikasi artikel ini.

REFERENSI

- [1] WHO, "Diabetes," World Health Organization. Accessed: Sep. 16, 2023. [Online]. Available: https://www-who-int.translate.goog/news-room/fact-sheets/detail/diabetes?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc
- [2] WHO, "10 Penyebab Kematian Teratas," World Health Organization. Accessed: Sep. 16, 2023. [Online]. Available: https://www-who-int.translate.goog/news-room/fact-sheets/detail/the-top-10-causes-of-death?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc
- [3] IDF, *IDF Diabetes Atlas: 10TH Edition 2021*, 10th ed. Belgia: International Diabetes Federation, 2021.
- [4] Kemenkes RI, "Hasil Utama Rischesdas 2018," 2018. doi: 10.12691/jfmr-2-12-26.
- [5] Dinkes Jatim, *Profil Kesehatan Provinsi Jawa Timur 2021*, 2021st ed. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2022.
- [6] Kemenkes RI, "Peraturan Menteri Kesehatan RI No 71 Tahun 2015 Tentang Penanggulangan Penyakit Tidak Menular," *Peratur. Menteri Kesehat. RI No 71 Tahun 2015 Tentang Penanggulangan Penyakit Tidak Menular*, no. 1775, p. 32, 2015.
- [7] M. Burhan Hanif and G. Guntoro Setiaji, "Meningkatkan Kinerja Decision Tree C4.5 dengan Seleksi Fitur Korelasi Pearson pada Deteksi Penyakit Diabetes," *Indones. J. Comput. Sci.*, vol. 11, no. 2, pp. 685–695, 2022, doi: 10.33022/ijcs.v11i2.3087.
- [8] R. Haqmanullah Pambudi, B. Darma Setiawan, and Indriati, "Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] M. R. Pratama, A. L. Suryana, G. Alfiansyah, Z. Olivia, I. Nurawati, and P. Destarianto, "Diagnosis of Stroke and Diabetes Mellitus With Classification Techniques Using Decision Tree Method," *Int. J. Heal. Inf. Syst.*, vol. 2, no. 1, pp. 1–8, 2024, doi: 10.47134/ijhis.v2i1.36.
- [10] N. Gusrialni Fitri, S. Adilya, and F. Azizi, "Comparison of the Naive Bayes Classification System and C4.5 for the Diagnosis of Stroke Perbandingan Sistem Klasifikasi Naive Bayes dan C4.5 Untuk Diagnosa Penyakit stoke," pp. 49–55, 2023.
- [11] Depkes RI, "Pedoman Pengendalian Diabetes Melitus dan Penyakit Metabolik.pdf." 2008.
- [12] P. R. S and B. R. Maxim, *Software Engineering*, 8th ed., vol. 11, no. 1. New York: McGraw-Hill Education, 2015. [Online]. Available: http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI
- [13] J. A. Ramadhan, D. T. Haniva, and A. Suharso, "Systematic Literature Review Penggunaan Metodologi Pengembangan Sistem Informasi Waterfall, Agile, dan Hybrid," *JIEET J. Inf. Eng. Educ. Technol.*, vol. 07, no. 01, pp. 36–42, 2023, [Online]. Available: <https://journal.unesa.ac.id/index.php/jieet/article/view/21941>
- [14] M. A. Hutomo, *Clinical Reasoning: Endokrin/Metabolisme/Nutrisi*. Yogyakarta: Fakultas Kedokteran Universitas Ahmad Dahlan, 2024.
- [15] Noviandi, "Data Mining," Jakarta Barat: Universitas Esa Unggul, 2018, pp. 1–6.
- [16] S. J. Adrian and Tommy, "Hipertensi Esensial: Diagnosis dan Tatalaksana Terbaru pada Dewasa," *Cdk-274*, vol. 46, no. 3, pp. 172–178, 2019, [Online]. Available: <http://www.cdkjournal.com/index.php/CDK/article/view/503%0A> diakses pada tanggal 28 oktober 2020
- [17] A. B. Alfianah, "Analisis Perbandingan Kinerja Algoritma K-Nearest Neighbor (KNN) dan Algoritma C4.5 Untuk Klasifikasi Teknik Persalinan," 2024.
- [18] W. Musu, A. Ibrahim, and Heriadi, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5," *Pros. Semin. Ilm. Sist. Inf. Dan Teknol. Inf.*, vol. X, no. 1, pp. 186–195, 2021.
- [19] Menkes, "Keputusan Menteri Kesehatan Republik Indonesia Nomor HK.01.07/MENKES/603/2020 Tentang Pedoman Nasional Pelayanan Kedokteran Tata Laksana Diabetes Melitus Tipe 2 Dewasa," 2020.