

Klasifikasi Penyakit Hipertensi Dan Diabetes Berbasis Web Pada Klinik Pratama Rumkitban 01.08.03 Batam

Mira Chandra Kirana ^{1*}, Michel K ^{2**}

* Teknik Informatika, Politeknik Negeri Batam

** Teknik Multimedia Jaringan, Politeknik Negeri Batam

mira@polibatam.ac.id¹, michaeloke14@gmail.com²

Article Info

Article history:

Received 09-06-2021

Revised 20-07-2021

Accepted 22-07-2021

Keyword:

*Algoritma C45,
Classification,
Diabetes,
Hypertension,
Tree Decision.*

ABSTRACT

The management of outpatient medical records at the Rumkitban Primary Clinic 01.08.03 Batam is still manual and causes many limitations and problems. This problem resulted in the inability of the clinic to run the Chronic Disease Treatment Program (PROLANIS) organized by BPJS-Health. The purpose of the study is to facilitate data processing and then from that data it can be used to classify hypertension and diabetes then the results of the classification are displayed in graphical form. This study discusses 2 diseases, namely hypertension and diabetes. The system uses the C45 Tree Decision Algorithm for automatic data processing. The attributes used are glucose, diastolic, systolic, insulin, and age to support the decision-making system. The system can make a decision whether the patient has hypertension, diabetes or not. The results of this study are the accuracy of classification accuracy in the system for hypertension shows 16.667% accuracy and 83.333% accuracy is not correct, then the calculation of diabetes classification accuracy shows 96.667% accuracy, and 3.333% accuracy is incorrect. This system is integrated with Mysql database to store the results.



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

I. PENDAHULUAN

Klinik Pratama Rumkitban 01.08.03 Batam merupakan instansi militer angkatan darat yang bergerak di bidang kesehatan, dimana fungsi utamanya adalah memberikan pelayanan kesehatan dan balai pengobatan kepada TNI, PNS dan keluarga TNI dan PNS. Klinik Pratama Rumkitban 01.08.03 Batam sebagai salah satu institusi pelayanan kesehatan umum, juga menerima pasien umum berobat dari masyarakat umum dan pasien BPJS Kesehatan, sehingga membutuhkan sebuah sistem informasi yang tingkat akurat dan h1 tinggi, serta memadai untuk meningkatkan pelayanan Klinik Pratama Rumkitban kepada pasien dan lingkungan terkait. Unit yang menyediakan informasi penting untuk pimpinan dalam mengambil suatu keputusan dalam meningkatkan pelayanan terhadap pasien adalah unit rawat jalan. Unit rawat jalan setiap harinya melakukan aktifitas pelayanan terhadap pasien yang mendapatkan pelayanan pengobatan dan diagnosa oleh dokter.

Pengelolaan penyakit kronis (Prolanis) [1] yaitu hipertensi dan diabetes dari BPJS-Kesehatan ini belum berjalan sepenuhnya dikarenakan kurangnya SDM, pendataan peserta prolanis yang belum jelas, dan manajemen rekam medis pasien yang kurang baik. Berdasarkan masalah yang timbul sehigga muncul ide untuk membuat sebuah aplikasi agar berkurangnya masalah yang ada, maka dari penelitian ini akan segera dibuat sistem untuk mengklasifikasi pasien penyakit kronis berbasis *web* agar nantinya dapat mendata peserta prolanis dan melakukann pemantauan peserta prolanis yang terdapat pada Rumkitban 01.08.03 Batam. Untuk menangani permasalahan tersebut perlu adanya pengembangan sistem informasi rekam medis rawat jalan di Klinik Pratama Rumkitban 01.08.03 Batam.

Penelitian sebelumnya yang di lakukan oleh [2] membuktikan seluruh variabel MMUST berpengaruh terhadap kesuksesan implementasi RME dengan nilai R2 kepuasan informasi 0,394; harapan kinerja 0,292; kepuasan keseluruhan 0,602; manfaat keseluruhan 0,444; dan sikap 0,655. Nilai Goodness of Fit (GoF) sebesar 0.5777, sehingga

dapat disimpulkan model penelitian ini secara substansial merepresentasikan hasil penelitian. Dengan dilakukan penelitian ini dapat membantu klinik dalam pengolahan data secara akurat dan tepat. Dengan memiliki data yang akurat dan tepat klinik mampu menjalankan program pengolahan penyakit kronis yang diselenggarakan oleh BPJS-Kesehatan.

Penelitian sebelumnya yang dilakukan oleh [3] menunjukkan bahwa dengan mengubah rekam medis dan banyaknya data medis yang tersedia kedalam digitalisasi akan menjadi area potensi penelitian yang luas dan berguna bagi bidang kesehatan maupun bidang lainnya. Maka Penelitian yang dilakukan menggunakan web agar memudahkan dalam mengolah data karena sistem untuk mengakses, memanipulasi, dan mengunduh dokumen yang terdapat dalam komputer [4] sangat terorganisir dengan baik dan tercatat dengan rapi dalam sistem. Oleh sebab itu dibutuhkan sebuah web server pada penelitian ini agar berguna untuk pengolahan dan pengujian data yang sangat banyak.

Rekam medis ini juga di didukung dengan fitur untuk mengklasifikasi penyakit hipertensi dan diabetes sesuai dengan program dari BPJS-Kesehatan mengolah data pasien hipertensi dan diabetes. Penelitian sebelumnya yang dilakukan oleh [5] menunjukkan bahwa sebuah sistem dapat memberikan informasi tentang diagnosa awal penyakit diabetes kepada masyarakat agar dapat melakukan tindakan awal atau pencegahan terhadap penyakit diabetes militus. Dalam mengolah data tersebut sistem di didukung dengan beberapa atribut yang di inputkan oleh dokter.

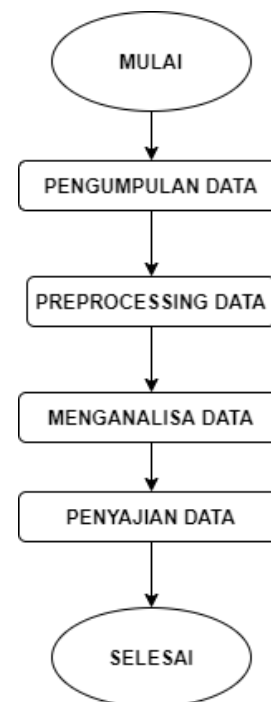
Pada penelitian ini menggunakan beberapa atribut seperti glukosa, diastole, sistol, insulin, dan usia. Menggunakan algoritma c45 sebagai pengolahan datanya untuk menciptakan sebuah pohon keputusan yang akan di gunakan oleh sistem untuk mengklasifikasi penyakit hipertensi dan diabetes. Karena pada penelitian sebelumnya yang dilakukan oleh [6] diperoleh bahwa performa tertinggi dalam mengklasifikasi dicapai ketika algoritma IDE3 menggunakan 5 atribut yaitu gpost, glun, upost, urin dan actn. Dimana kelima atribut tersebut diperoleh menggunakan algoritma *Correlation based Feature Selection* (CFS) dengan nilai rata-rata akurasi sebesar 84.77, nilai rata-rata sensitifity sebesar 87.18, dan nilai rata-rata specificity sebesar 82.37. dengan menggunakan 5 atribut juga sehingga di penelitian ini dapat menghasilkan tingkat ketepatan klasifikasi penyakit yang akurasinya tinggi juga. dasar algoritma C4.5 adalah pohon keputusan. Pada dasarnya Algoritma data mining C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokan yang bersifat prediktif. [7]

Maka dari Penelitian ini diharapkan Menyediakan sebuah solusi pencatatan rekam medis secara tepat dan akurat, kemudian data yang dihasilkan dapat digunakan untuk diklasifikasi menjadi penyakit hipertensi dan diabetes dan ditampilkan hasil dalam bentuk grafik. Dalam perancangan dan pengembangan serta dalam penerapannya menyesuaikan dengan kebutuhan dan perkembangan teknologi terkini yang relevan.

II. METODE PENELITIAN

A. Alur Penelitian

Untuk mencapai tujuan yang diharapkan dan sesuai dengan latar belakang, dilakukan beberapa tahapan yaitu antara lain.



Gambar 1. Alur penelitian

Gambar 1 menggambarkan seluruh tahap yang ada dalam proses penelitian. Penelitian yang dilakukan terbagi menjadi 3 tahapan yaitu mengumpulkan data, preprocessing data, dan menganalisa data. pada tahap awal ini menggumpulkan data langkah yang dilakukan yaitu mengambil data pasien penyakit hipertensi dan diabetes dari rekam medis Klinik Pratama Rumkitban 01.08.03 Batam kemudian dilanjutkan dengan tahap *preprocessing* data yang dilakukan adalah data *cleaning*, data *integration*, data *transformation*, dan data *reduction*. Pada data *cleaning* digunakan untuk melengkapi, menghilangkan data yang tidak lengkap, dan menghaluskan atau meniadakan data *noisy*, ada juga memperbaiki data yang tidak konsisten (berubah-berubah).

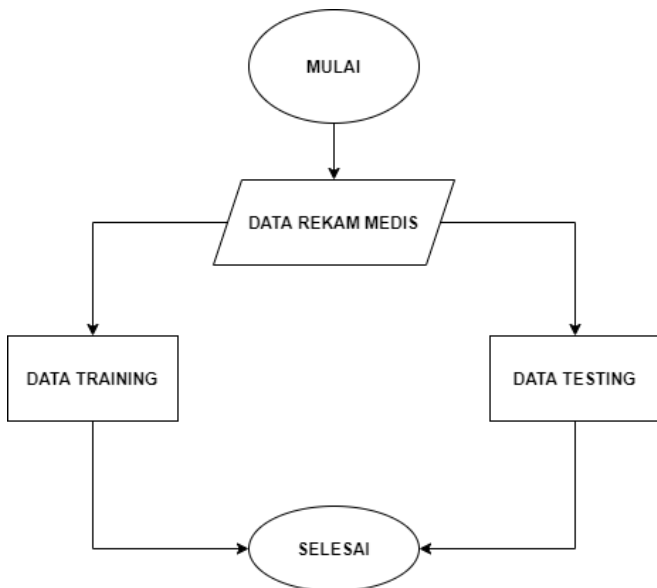
Pada data *integration* berguna untuk mengabungkan data dari beberapa database kedalam satu database baru. Pada data *transformation* berguna untuk mengubah data agar berbentuk yang sesuai proses mining. Pada data *reduction* berguna untuk mengurangi ukuran data kemudian menghasilkan analisis yang sama.

Dimana pada tahap ini menghasilkan data yang memiliki kualitas baik atau data akurat untuk dilakukan proses mining pada sistem, petugas mengelompokkan data menjadi 2 yaitu data training dan testing. Pada data *training* akan diproses terlebih dahulu oleh algoritma C4.5 untuk menghasilkan sebuah pohon keputusan/ rule untuk sistemnya. Yang nantinya akan digunakan untuk prediksi kelas data baru yang

belum pernah ada. Kemudian untuk data testing digunakan untuk mengukur sejauh mana tingkat keberhasilan sistem melakukan klasifikasi dengan akurat dan tepat.

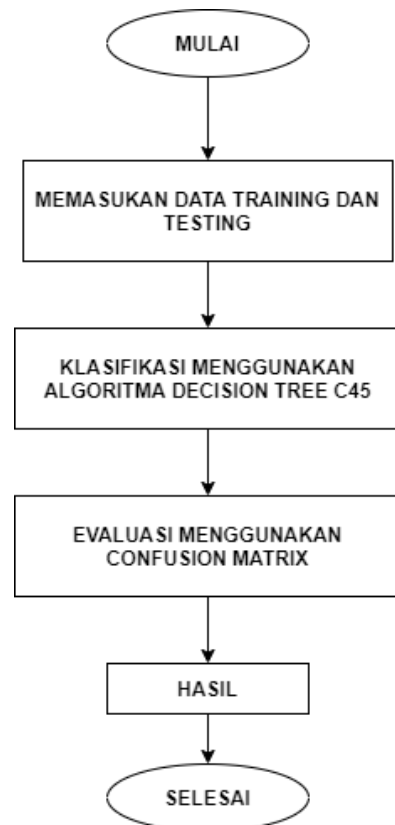
Lalu dilanjutkan tahapan menganalisa merupakan tahapan akhir yang dimana data akan dimasukkan kedalam aplikasi kemudian sistem melakukan klasifikasi dengan sendirinya dari pohon keputusan yang sudah di dapatkan dari hasil mining menggunakan algoritma C4.5, sehingga sistem menghasilkan klasifikasi penyakit.

- 1) *Pengumpulan Data*: data yang digunakan berasal dari data rekam medis Klinik Pratama Rumkitban 01.08.03 Batam. Atribut-atribut dalam sistem ini menggunakan diantaranya: usia, kadar glukosa darah, tekanan darah, insulin.
- 2) *Preprocessing Data*: pada gambar 2 menggambarkan tahapan *preprocessing* data yaitu dengan melakukan *preprocessing* terhadap data hasil rekam medis dan kemudian dilanjutkan dengan mengelompokkan data menjadi 2 bagian yaitu data *training* dan data *testing*. Pada Data training digunakan oleh algoritma *decision tree* C4.5 untuk membentuk sebuah pohon keputusan / rule untuk system nantinya. Pada data *testing* untuk mengukur sejauh mana tingkat keberhasilan sistem melakukan klasifikasi dengan akurat dan tepat.



Gambar 2. Tahap Preprocessing

3) *Analisa Data*



Gambar 3. Tahap Analisa Data

Gambar 3 menggambarkan tahap menganalisa yang dimulai dari tahap memasukan data *training* dan data *testing* ke dalam aplikasi, tahap selanjutnya adalah tahap untuk melakukan klasifikasi yang diproses oleh algoritma *Decision Tree* C4.5 [8].

$$\text{Entropi (S)} = \sum_{i=1}^k -p_i \log_2 p_i$$

Keterangan:

- S: Himpunan (dataset) kasus
- k: Banyaknya partisi S
- Pi: Probabilitaas yang didapat dari Sum(Ya) atau Sum(Tidak) dibagi total kasus

Setelah mendapatkan entropi dari keseluruhan kasus, lakukan analisis pada setiap atribut dan nilai-nilainya dan hitung entropinya. Langkah berikutnya yaitu dengan menghitung *Gain*, rumus daripada *Gain* adalah sebagai berikut: [8]

$$\text{Gain(A)} = \text{Entropi (S)} - \sum_{i=1}^k \frac{|S_i|}{|S|} \times \text{Entropi (S}_i\text{)}$$

Keterangan:

- S: himpunan kasus
- A: atribut
- n: jumlah partisi atribut A

- |Si|: jumlah kasus pada partisi ke-i
- |S|: jumlah kasus dalam S

Setelah itu akan system akan menghasilkan sebuah pohon keputusan yang akan digunakan untuk menghasilkan klasifikasi penyakit hipertensi dan diabetes. Kemudian dilanjutkan Tahapan untuk mengevaluasi hasil klasifikasi data yang dilakukan sistem untuk memperkirakan hasil klasifikasi yang dilakukan benar atau salah. Sistem juga dapat menghitung tingkat akurasi ketepatan klasifikasi dengan menggunakan *confusion matrix*, yang dimana contohnya terdapat pada table 1.

TABLE 1.
MODEL CONFUSION MATRIX

Kelas	Terklasifikasi Positif	Terklasifikasi Negatif
Positif	TP (True Positive)	FN (False Negative)
Negatif	FP (False Positive)	TN (True Negative)

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \quad 1$$

$$Presisi = \frac{TP}{FP+TP} * 100\% \quad 2$$

$$Recall = \frac{TP}{FN+TP} * 100\% \quad 3$$

Keterangan:

- TP adalah *True Positive*, yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem.
- TN adalah *True Negative*, yaitu jumlah data negatif yang terklasifikasi dengan benar oleh sistem.
- FN adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah oleh sistem.
- FP adalah *False Positive*, yaitu jumlah data positif namun terklasifikasi salah oleh sistem

Tahap yang terakhir yaitu mendapatkan prediksi penyakit diabetes dan hipertensi dari hasil klasifikasi tersebut.

4) *Penyajian Data*

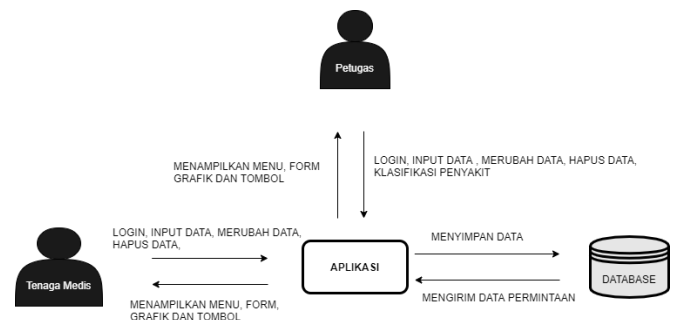
Tabel 2 merupakan data yang dikumpulkan, setelah itu diletakkan di dalam sebuah database yang nantinya akan dilakukan *preprocessing* agar menghasilkan data yang sesuai untuk proses *training* dan *testing*.

TABLE 2.
PENYAJIAN DATA

Kadar Gula	DIASTOL	SISTOL	Kadar Insulin	Usia	Kelas
148	72	35	0	50	Diabetes
85	66	29	0	31	Tidak Diabetes dan tidak Hipertensi
183	64	0	0	32	Diabetes dan tidak hipertensi
89	66	23	94	21	Tidak Diabetes dan tidak Hipertensi
137	40	35	168	33	Diabetes dan tidak Hipertensi
116	74	0	0	30	Tidak Tidak Diabetes dan tidak Hipertensi
78	50	32	88	26	Diabetes dan tidak Hipertensi
115	0	0	0	29	Tidak Diabetes dan tidak Hipertensi
197	70	45	543	53	Diabetes dan tidak Hipertensi
125	96	0	0	54	Diabetes dan Hipertensi

Sumber: Rekam Medis Klinik Pratama Rumkitban 01.08.03 Batam)

B. *Deskripsi Umum Sistem*



Gambar 2. Deskripsi umum sistem

Gambar 4 deskripsi umum sistem yang akan dirancang,

Sistem aplikasi diatas meliputi:

- 1) Pasien melakukan pendaftaran berobat melalui Petugas
- 2) Petugas mendaftarkan pasien dengan *login* ke dalam aplikasi dan menginput data pasien, Petugas memiliki hak ases untuk menginput, mengupdate, mendelete data pasien, jika sudah selesai di daftarkan petugas memberikan nomor antrian.
- 3) Tenaga medis memberikan pelayanan kesehatan atau pun tindakan medis terhadap pasien, setelah itu menginputkan tindakan, diagnosa, dan obat kedalam aplikasi. Tenaga medis hanya di berikan hak ases menginput data berobat, mengubah data berobat, memeriksa data pasien, memeriksa hasil klasifikasi penyakit hipertensi dan diabetes kemudian dapat memeriksa riwayat berobat pasien.

- 4) Petugas memiliki hak ases mengelola user, data kunjungan sakit, daftar kunjungan peserta prolanis, dan dapat mencetak laporan. Kemudian petugas dapat mengklasifikasi pasien penyakit kronis.

C. Use Case Diagram

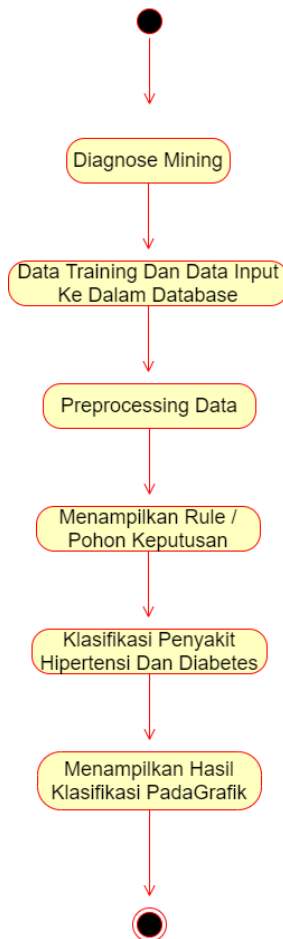
Pada gambar 5 *Use Case Diagram* menjelaskan bahwa user melakukan login terlebih dahulu untuk dapat mengelola data user, mengelola data pasien, mengelola data dokter, mengelola data rekam medis, memeriksa riwayat, memeriksa laporan dan mengklasifikasi penyakit hipertensi dan diabetes.

Lampiran I

Gambar 3. Use case diagram sistem

D. Activity Diagram

Gambar 6 menjelaskan tentang *Activity Diagram* mengklasifikasi penyakit hipertensi dan diabetes:



Gambar 4. Activity Diagram klasifikasi penyakit

E. ER-Diagram

Pada gambar 7 menjelaskan tentang hubungan antar entitas pada sistem dan uji data mining.

Lampiran II

Gambar 5. ER-Diagram sistem

III. HASIL DAN PEMBAHASAN

Metode yang dipakai dalam penelitian ini adalah algoritma klasifikasi dengan Algoritma Decision Tree yang digunakan untuk melakukan klasifikasi penyakit hipertensi dan diabetes pada data yang ada di sistem. Untuk mendapatkan kinerja algoritma klasifikasi ini menggunakan dataset tersedia pada rekam medis klinik pratama rumkitban 01.08.03 Batam. Dataset yang digunakan memiliki 5 atribut terdiri dari data nominal dan numerik. Kelas data sasaran adalah hipertensi, tidak hipertensi, diabetes dan tidak diabetes. Sistem ini terdiri dari beberapa tahap yaitu preprocessing untuk menghilangkan data – data yang noisy ataupun redudansi, selanjutnya melakukan klasifikasi dengan Algoritma *Decision tree C45*. Setelah itu dilakukan evaluasi dengan *confusion matrix*.

A. Preprocessing Data

Pada proses ini adalah mempersiapkan data yang berkualitas baik untuk diproses data mining yang akan diolah oleh algoritma *C45*, *Tree desicion* dan proses testing menggunakan *confusion matrix*. *Preprocessing* data ini dilakukan oleh sistem karena pada sistem telah di tanamkan algoritma *c45* sehingga sistem otomatis melakukan preprocessing dengan sendirinya. Pada Tabel 3 dan Tabel 4 merupakan tabel data training yang telah melalui proses preprocessing dari data rekam medis klinik pratama Rumkitban 01.08.03 Batam. Data training ini diproses menggunakan algoritma *C45* dan *Tree Decision*. Proses *preprocessing* data yang dilakukan pada sistem ini adalah merubah data kosong dengan rata-rata data per *variable* dan mengelompokkan data-data kedalam beberapa kelas.

Adapun pengelompokan data untuk masing-masing *variable* ini berdasarkan Kemenkes RI tahun 2014 tentang hipertensi [9] dan Kemenkes RI tahun 2019 tentang diabetes [10] sebagai berikut:

- Glukosa = ≤ 100 , > 100 , ≤ 120 , > 120 , ≤ 140 , > 140
- Diastol = ≤ 70 , > 70 , ≤ 80 , > 80 , ≤ 90 , > 90 ,
- Sistol = ≤ 110 , > 110 , ≤ 120 , > 120 , ≤ 130 , > 130
- Insulin = ≤ 200 , > 200 , ≤ 300 , > 300 , ≤ 400 , > 400
- Usia = ≤ 17 , > 17 , ≤ 25 , > 25 , ≤ 36 , > 36 , ≤ 56 , > 56
- Kelas = Hipertensi, tidak hipertensi, Diabetes, tidak diabetes

Tabel 3 merupakan sebuah data training yang berisi data pasien yang menderita penyakit hipertensi yang didapat dari Klinik Pratama Rumkitban 01.08.03 Batam dan telah melalui proses preocesing sehingga data dapat di proses oleh sistem dengan algoritma *C45*.

TABEL 3
DATA PENYAKIT HIPERTENSI SETELAH PREPROCESSING

Nama Pasien	Glukosa	Diastol	Sistol	Insulin	Usia	Kelas
1	100	80	130	200	64	hipertensi
2	100	80	120	200	57	tidak
3	100	90	120	200	59	tidak
4	100	90	130	200	66	hipertensi
5	100	80	120	200	53	tidak
6	100	85	121	200	19	tidak
7	100	70	115	200	31	tidak
8	100	80	110	200	52	tidak
9	100	73	122	200	44	tidak
10	100	88	118	200	34	tidak
11	100	86	123	200	35	tidak
12	100	80	109	200	49	tidak
13	100	70	135	200	62	hipertensi
14	100	80	133	200	41	hipertensi
15	100	90	132	200	52	hipertensi
16	115	85	130	200	64	hipertensi
17	123	77	113	240	65	tidak
18	135	85	90	231	39	tidak
19	100	90	130	210	62	hipertensi
20	125	80	125	213	53	tidak
26	100	85	121	200	19	tidak
27	100	70	115	200	31	tidak
28	100	80	110	200	52	tidak
29	100	73	122	200	44	tidak
30	100	88	118	200	34	tidak
31	100	86	123	200	35	tidak
32	100	80	109	200	49	tidak
33	100	70	135	200	62	hipertensi
34	100	80	133	200	41	hipertensi
35	100	90	132	200	52	hipertensi
36	115	85	130	200	64	hipertensi
37	123	77	113	240	65	tidak
38	135	85	90	231	39	tidak
39	100	90	130	210	62	hipertensi
40	125	80	125	213	53	tidak
41	139	94	133	230	39	hipertensi
42	80	75	108	200	28	tidak
43	80	80	120	200	20	tidak
44	80	90	135	200	40	hipertensi
45	85	91	133	200	67	hipertensi
46	80	91	132	200	52	hipertensi
47	125	88	125	240	61	tidak
48	80	95	134	200	62	hipertensi
49	88	95	135	200	46	hipertensi
50	100	94	134	200	35	hipertensi
51	103	99	136	200	38	hipertensi
52	135	100	136	245	61	hipertensi
53	132	102	133	210	51	hipertensi
54	140	95	100	241	64	tidak
55	130	80	133	222	60	tidak
56	132	80	132	210	58	hipertensi
57	135	95	135	200	55	hipertensi
58	145	85	115	200	62	tidak
59	80	96	136	200	45	hipertensi
60	136	85	96	200	57	tidak
61	85	95	133	200	50	hipertensi
62	125	85	115	200	65	hipertensi
63	100	102	140	200	52	hipertensi
64	102	98	135	200	43	hipertensi
65	105	95	133	200	51	hipertensi
66	100	95	131	200	66	hipertensi
67	90	90	131	200	45	hipertensi

68	135	85	120	235	52	tidak
69	135	91	135	200	63	hipertensi
70	139	94	133	230	39	hipertensi
71	100	90	125	200	38	hipertensi
72	100	85	115	200	61	hipertensi
73	120	81	130	199	65	hipertensi
74	120	90	129	200	52	hipertensi

Table 4 merupakan sebuah data training yang berisi data pasien yang menderita penyakit diabetes militus yang didapat dari Klinik Pratama Rumkitban 01.08.03 Batam dan telah melalui proses preproceing sehingga data dapat di proses oleh sistem dengan algoritma c45.

TABEL 4
DATA PENYAKIT DIABETES SETELAH PREPROCESSING

Nama Pasien	Glukosa	Diastol	Sistol	Insulin	Usia	Kelas
1	100	80	130	200	64	tidak
2	100	80	120	200	57	tidak
3	100	90	120	200	59	tidak
4	100	90	130	200	66	tidak
5	100	80	120	200	53	tidak
6	100	85	121	200	19	tidak
7	100	70	115	200	31	tidak
8	100	80	110	200	52	tidak
9	100	73	122	200	44	tidak
10	100	88	118	200	34	tidak
11	100	86	123	200	35	tidak
12	100	80	109	200	49	tidak
13	100	70	135	200	62	tidak
14	100	80	133	200	41	tidak
15	100	90	132	200	52	tidak
16	115	85	130	200	64	tidak
17	123	77	113	240	65	diabetes
18	135	85	90	231	39	diabetes
19	100	90	130	210	62	tidak
20	125	80	125	213	53	diabetes
26	100	85	121	200	19	tidak
27	100	70	115	200	31	tidak
28	100	80	110	200	52	tidak
29	100	73	122	200	44	tidak
30	100	88	118	200	34	tidak
31	100	86	123	200	35	tidak
32	100	80	109	200	49	tidak
33	100	70	135	200	62	tidak
34	100	80	133	200	41	tidak
35	100	90	132	200	52	tidak
36	115	85	130	200	64	tidak
37	123	77	113	240	65	tidak
38	135	85	90	231	39	tidak
39	100	90	130	210	62	tidak
40	125	80	125	213	53	tidak
41	139	94	133	230	39	diabetes
42	80	75	108	200	28	tidak
43	80	80	120	200	20	tidak
44	80	90	135	200	40	tidak
45	85	91	133	200	67	tidak
46	80	91	132	200	52	tidak
47	125	88	125	240	61	diabetes
48	80	95	134	200	62	tidak
49	88	95	135	200	46	tidak
50	100	94	134	200	35	tidak
51	103	99	136	200	38	tidak
52	135	100	136	245	61	tidak
53	132	102	133	210	51	tidak
54	140	95	100	241	64	tidak
55	130	80	133	222	60	tidak
56	132	80	132	210	58	tidak
57	135	95	135	200	55	tidak
58	145	85	115	200	62	tidak
59	80	96	136	200	45	tidak
60	136	85	96	200	57	tidak
61	85	95	133	200	50	tidak
62	125	85	115	200	65	tidak
63	100	102	140	200	52	tidak
64	102	98	135	200	43	tidak
65	105	95	133	200	51	tidak
66	100	95	131	200	66	tidak
67	90	90	131	200	45	tidak

52	135	100	136	245	61	diabetes
53	132	102	133	210	51	diabetes
54	140	95	100	241	64	diabetes
55	130	80	133	222	60	diabetes
56	132	80	132	210	58	diabetes
57	135	95	135	200	55	diabetes
58	145	85	115	200	62	diabetes
59	80	96	136	200	45	tidak
60	136	85	96	200	57	diabetes
61	85	95	133	200	50	tidak
62	125	85	115	200	65	tidak
63	100	102	140	200	52	diabetes
64	102	98	135	200	43	tidak
65	105	95	133	200	51	tidak
66	100	95	131	200	66	tidak
67	90	90	131	200	45	tidak
68	135	85	120	235	52	diabetes
69	135	91	135	200	63	tidak
70	139	94	133	230	39	diabetes
74	100	75	100	210	60	diabetes
75	120	20	110	200	49	diabetes
76	135	80	110	200	68	diabetes
77	110	70	115	200	80	diabetes
78	120	80	110	210	62	diabetes
79	120	81	130	199	65	diabetes

B. Perhitungan dengan Algoritma C45

Setelah proses *preprocessing* data akan dilakukan perhitungan menggunakan algoritma C45. Tahap pertama yang dilakukan adalah melakukan perhitungan jumlah kelas, kemudian melakukan perhitungan jumlah kasus per kelas, setelah itu mencari entropy setiap atribut lalu setelah selesai mencari entropy dilanjutkan mencari nilai gain. Nilai gain dengan nilai tertinggi merupakan akarnya kemudian cabang dan seterusnya dari keputusan kemudian di hitung kembali sampai nilai entropy seluruhnya atribut mencapai 0 maka sistem akan berhenti menghitung.

1. Perhitungan Algoritma C45 Penyakit Hipertensi

Pada table 4 menunjukkan hasil hasil perhitungan algoritma c45 yang dilakukan oleh sistem. Sehingga nantinya akan menunjukkan hasil keputusan yang berbentuk *Tree Decision* penyakit Hipertensi.

- Jumlah data = 68
- Jumlah Hipertensi = 37
- Jumlah Aman = 31
- Entropy All = 0.994

TABEL 5
PERHITUNGAN SELURUH DATA TRAINING HIPERTENSI

Nilai Atribut	Jumlah data	Jumlah Hipertensi	Jumlah Aman	Entropy	Gain
glukosa<=100	42	23	19	0.993	
glukosa>100	26	14	12	0.996	
					0

glukosa<=120	49	30	19	0.963	
glukosa>120	19	7	12	0.949	
					0.035
glukosa<=140	67	37	30	0.992	
glukosa>140	1	0	1	0	
	0			0	0.017
diastol<=70	4	2	2	1	
diastol>70	64	35	29	0.994	
	0			0	-0
diastol<=80	22	5	17	0.773	
diastol>80	46	32	14	0.887	
	0			0	0.144
diastol<=90	49	19	30	0.963	
diastol>90	19	18	1	0.297	
	0			0	0.217
sistol<=110	9	0	9	0	
sistol>110	59	37	22	0.953	
	0			0	0.167
sistol<=120	23	2	21	0.426	
sistol>120	45	35	10	0.764	
	0			0	0.344
sistol<=130	41	11	30	0.839	
sistol>130	27	26	1	0.229	
	0			0	0.397
insulin<=200	52	31	21	0.973	
insulin>200	16	6	10	0.954	
	0			0	0.025
insulin<=300	68	37	31	0.994	
insulin>300	0	0	0	0	
	0			0	0

insulin<=400	68	37	31	0.994	
insulin>400	0	0	0	0	
	0			0	0
usia<=17	0	0	0	0	
usia>17	68	37	31	0.994	
	0			0	0
usia<=25	3	0	3	0	
usia>25	65	37	28	0.986	
	0			0	0.052
usia<=36	11	1	10	0.439	
usia>36	57	36	21	0.949	
	0			0	0.128
usia<=56	43	21	22	1	
usia>56	25	16	9	0.943	
	0			0	0.015

Table 5 memiliki data sebanyak 68, dari data tersebut terdapat 37 yang menderita penyakit hipertensi dan sebanyak 31 orang tidak menderita penyakit hipertensi, setelah melalui proses algoritma c45 didapat lah entropy seluruh data sebesar 0.994, kemudian di lanjutkan dengan mencari gain setiap atribut seperti yang terlihat pada table 5 hasil gain di setiap atribut selanjut akan di pilih atribut dengan gain yang tertinggi dan Atribut terpilih = Sistol <= 130 & > 130, dengan nilai gain = 0.397 kemudian terbentuk cabang 1 dengan atribut (sistol<=130). Kemudian sistem lanjut menghitung sampai nilai pada atribut menjadi nol atau sama. Sehingga mendapat pohon keputusan seperti gambar 8

Lampiran III

Gambar 6. Pohon keputusan hipertensi

2. Perhitungan Algoritma C45 Penyakit Diabetes

Pada table 4.8 menunjukkan hasil hasil perhitungan Entropy dan gain yang dilakukan oleh sistem. Sehingga nantinya akan menunjukkan hasil keputusan yang berbentuk Tree Decision untuk penyakit diabetes.

- Jumlah data = 71
- Jumlah Diabetes = 22
- Jumlah Aman = 49
- Entropy All = 0.893

TABEL 5
PERHITUNGAN SELURUH DATA TRAINING DIABETES

Nilai Atribut	Jumlah data	Jumlah Diabetes	Jumlah Aman	Entropy	Gain
glukosa<=100	41	2	39	0.281	
glukosa>100	30	20	10	0.918	
	0			0	0.343
glukosa<=120	50	6	44	0.529	

glukosa>120	21	16	5	0.792	
	0			0	0.286
glukosa<=140	70	21	49	0.881	
glukosa>140	1	1	0	0	
	0			0	0.024
diastol<=70	6	2	4	0.918	
diastol>70	65	20	45	0.89	
	0			0	0.001
diastol<=80	28	9	19	0.906	
diastol>80	43	13	30	0.884	
	0			0	0
diastol<=90	52	15	37	0.867	
diastol>90	19	7	12	0.949	
	0			0	0.004
sistol<=110	13	7	6	0.996	
sistol>110	58	15	43	0.825	
	0			0	0.037
sistol<=120	27	11	16	0.975	
sistol>120	44	11	33	0.811	
	0			0	0.02
sistol<=130	43	14	29	0.91	
sistol>130	28	8	20	0.863	
	0			0	0.002
insulin<=200	52	8	44	0.619	
insulin>200	19	14	5	0.831	
	0			0	0.217
insulin<=300	71	22	49	0.893	
insulin>300	0	0	0	0	
	0			0	0
insulin<=400	71	22	49	0.893	
insulin>400	0	0	0	0	
	0			0	0
usia<=17	0	0	0	0	
usia>17	71	22	49	0.893	
	0			0	0
usia<=25	3	0	3	0	
usia>25	68	22	46	0.908	
	0			0	0.023
usia<=36	11	0	11	0	
usia>36	60	22	38	0.948	
	0			0	0.092
usia<=56	42	9	33	0.75	
usia>56	29	13	16	0.992	
	0			0	0.044

Table 6 memiliki data sebanyak 71, dari data tersebut terdapat 22 yang menderita penyakit diabetes militus dan sebanyak 49 orang tidak menderita penyakit diabetes militus, setelah melalui proses algoritma c45 didapat lah entropy seluruh data sebesar 0.893, kemudian di lanjutkan dengan mencari gain setiap atribut seperti yang terlihat pada table 6 hasil gain di setiap atribut selanjut akan di pilih atribut dengan gain yang tertinggi dan Atribut terpilih adalah Glukosa <= 100, dengan nilai gain = 0.343, kemudian terbentuk cabang 1 dari atribut (glukosa<=100) dan seterusnya hingga nilai setiap atribut menjadi nol atau sama. Hasil pohon keputusan seperti gambar 9 berikut.

Lampiran IV

Gambar 7. Pohon keputusan diabetes

C. Hasil Klasifikasi

Setelah mendapat rule / pohon keputusan kedua penyakit tersebut petugas dapat menginput data berobat pasien pada sistem data didapatkan dari pasien yang berobat pada klinik dan data pasien itu di berikan oleh dokter yang melakukan pemeriksaan. sistem menghasilkan klasifikasi seperti yang ditunjukkan oleh gambar 10. Dari 30 data pasien diantaranya ada yang menderita hipertensi, diabetes dan tidak ada sama sekali menderita penyakit keduanya. Setelah diinputkan kedalam sistem petugas dapat melakukan klasifikasi pada menu diagnose mining dan didapatkan 30 menderita hipertensi dan 4 pasien menderita penyakit diabetes. tapi hasil perhitungan akurasi ketepatan klasifikasi pada sistem untuk penyakit hipertensi menunjukkan 16.667% akurasi tepat dan 83.333% akurasi tidak tepat, kemudian pada perhitungan akurasi klasifikasi diabetes menunjukkan 96.667% akurasi tepat dan 3.333 % akurasi tidak tepat. Hasil perhitungan tersebut langsung di tampilkan pada grafik dihalaman dashboard, dapat dilihat pada gambar 10.

Lampiran V

Gambar 8. Grafik hasil klasifikasi penyakit hipertensi dan diabetes

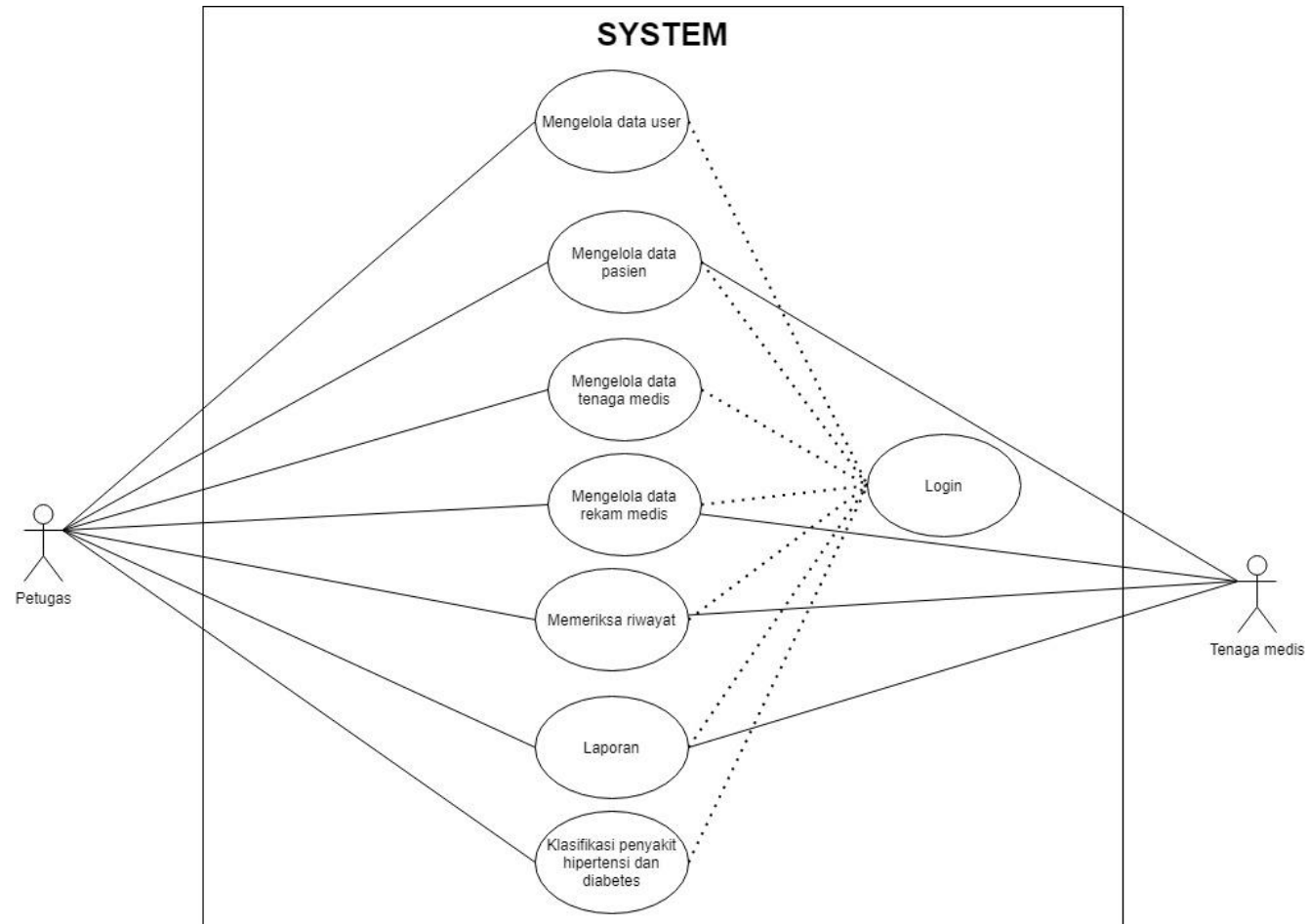
V. KESIMPULAN

Berdasarkan dari tujuan penelitian ini dapat disimpulkan sistem rekam medis dalam bentuk digital sangat membantu dalam mengelola data yang sangat banyak, membantu proses pengambilan keputusan dengan cepat, dan memudahkan klinik dalam menjalankan program pengelolaan penyakit kronis (Prolanis) dari BPJS-Kesehatan. Hasil klasifikasi data mining bahwa algoritma C4.5 dapat digunakan untuk mengklasifikasi penyakit hipertensi, diabetes mellitus, tidak hipertensi, dan tidak diabetes melitus. Hasil yang ditampilkan oleh Grafik hanya menampilkan data klasifikasi penyakit hipertensi dan diabetes melitus pada bulan yang berjalan. Saran untuk penelitian selanjutnya adalah dapat menggunakan metode klasifikasi yang lain seperti *Support Vector Machine (SVM)*, *K-Nearest Neighbor (KNN)*, *Algoritma CART (Classification and Regreesion Trees)*, *Neural Network (NN)*, maupun metode klasifikasi lainnya.

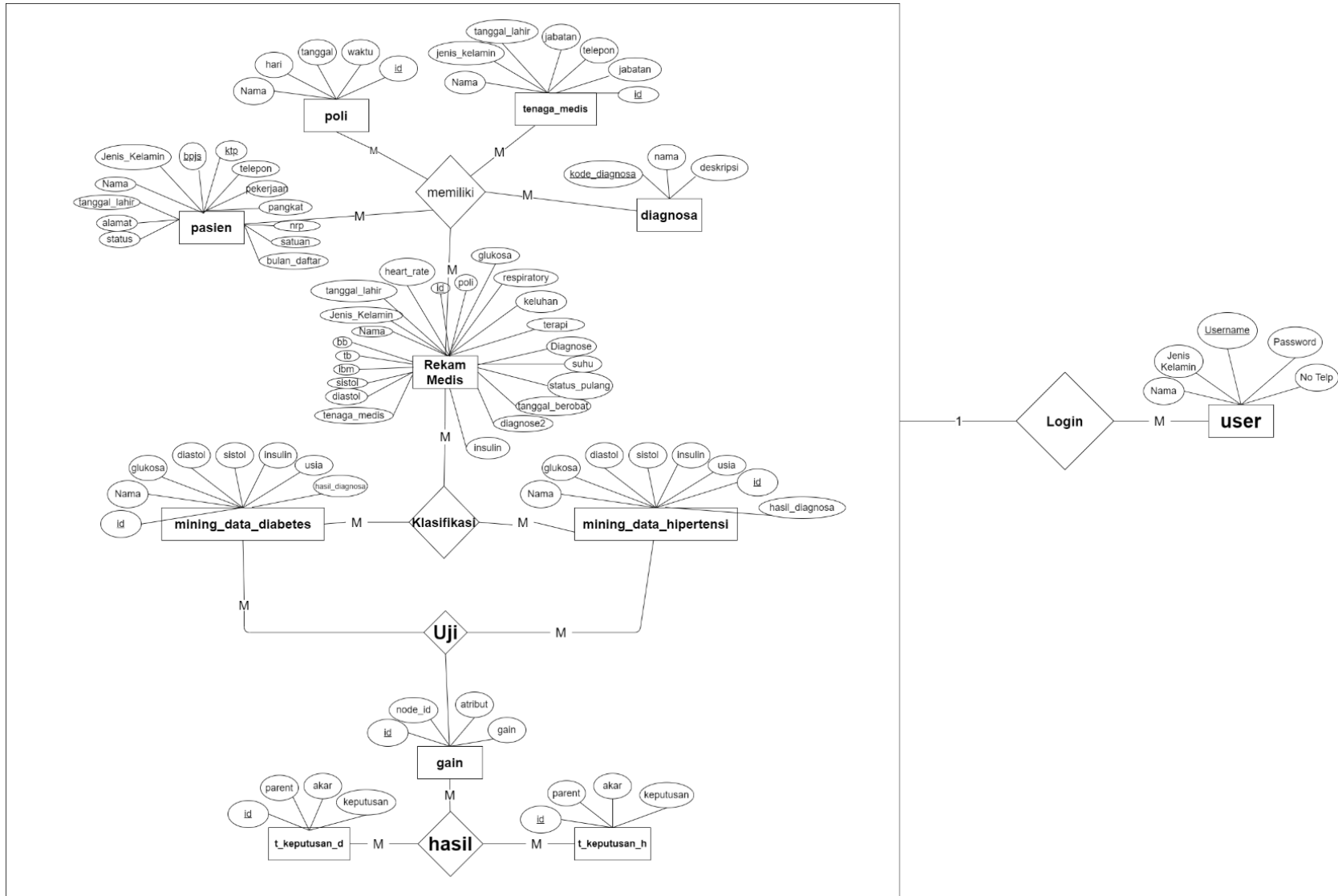
DAFTAR PUSTAKA

- [1] B. Kesehatan, "Panduan Praktis Prolanis," 2014. [Online]. Available: <https://bpjs-kesehatan.go.id/bpjs/dmdocuments/06-PROLANIS.pdf>. [Accessed 20 Juli 2021].
- [2] A. Rika, H. Kusnanto and W. Istiono, "Analisis Kesuksesan Implementasi Rekam Medis Elektronik Di RS Universitas Gadjah Mada," *Jurnal Sistem Informasi*, 2017.
- [3] L. Pereira, R. Rijo, C. Silva and R. Martinho, "Text Mining Applied to Electronic Medical Records: A Literature Review," *International Journal of E-Health and Medical Communications*, pp. 1-18, 2015.
- [4] KBBI, "KBBI Daring," Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2016. [Online]. Available: <https://kbbi.kemdikbud.go.id/entri/web>. [Accessed 08 Februari 2021].
- [5] A. J. N. Susanto, *Klasifikasi Penyakit Diabetes Mellitus Menggunakan Algoritma Decision Tree C4.5*, Batam: Politeknik Negeri Batam, 2018.
- [6] M. S. Efendi and H. A. Wibawa, *Prediksi Penyakit Diabetes Menggunakan Algoritma ID3 dengan Pemilihan Atribut Terbaik*, vol. VI, Purwokerto: JUITA, 2018, pp. 1-8.
- [7] F. Hermawati, *Data Mining*, Yogyakarta: Andi, 2013.
- [8] F. S. Sulaeman and M. A. Rilmansyah, "Aplikasi Penerapan Algoritma C45 untuk Memprediksi Kelulusan," *Jurnal Media Teknik & Sistem Industri*, vol. 5, pp. 41-54, 2021.
- [9] K. RI, *Pusdatin Hari Hipertensi Sedunia*, Jakarta Selatan: Kementerian Kesehatan RI, 2014, pp. 1-8.
- [10] K. RI, *Pusdatin Hari Diabetes Sedunia*, Jakarta Selatan: Kementerian Kesehatan RI, 2018.

Lampiran I. Use Case Diagram



Lampiran II. ER-Diagram



Lampiran III. Rule Hipertensi

Rule Hipertensi

HIDE TABLE

NO	PARENT	ROOT	DECISION
1	(sistol>130) AND (diastol<=80)	(glukosa<=100)	hipertensi
2	(sistol>130) AND (diastol<=80)	(glukosa>100)	tidak
3	(sistol>130)	(diastol>80)	hipertensi

Lampiran IV. Rule Diabetes

Rule Diabetes

HIDE TABLE

NO	PARENT	ROOT	DECISION
1	(glukosa<=100) AND (insulin<=200)	(diastol<=90)	tidak
2	(glukosa<=100) AND (insulin>200)	(diastol<=80)	diabetes
3	(glukosa<=100) AND (insulin>200)	(diastol>80)	tidak
4	(glukosa>100) AND (insulin<=200) AND (sistol<=120)	(diastol<=80)	diabetes
5	(glukosa>100) AND (insulin<=200) AND (sistol<=120) AND (diastol>80) AND (glukosa<=140)	(sistol<=110)	diabetes
6	(glukosa>100) AND (insulin<=200) AND (sistol<=120) AND (diastol>80) AND (glukosa<=140)	(sistol>110)	tidak
7	(glukosa>100) AND (insulin<=200) AND (sistol<=120) AND (diastol>80)	(glukosa>140)	diabetes
8	(glukosa>100) AND (insulin<=200) AND (sistol>120) AND (glukosa<=120)	(diastol<=90)	tidak
9	(glukosa>100) AND (insulin<=200) AND (sistol>120) AND (glukosa<=120)	(diastol>90)	tidak
10	(glukosa>100) AND (insulin>200) AND (sistol<=130)	(glukosa<=120)	diabetes
11	(glukosa>100) AND (insulin>200) AND (sistol<=130) AND (glukosa>120)	(diastol<=80)	diabetes
12	(glukosa>100) AND (insulin>200) AND (sistol<=130) AND (glukosa>120) AND (diastol>80) AND (sistol<=110)	(diastol<=90)	diabetes
13	(glukosa>100) AND (insulin>200) AND (sistol<=130) AND (glukosa>120) AND (diastol>80) AND (sistol<=110)	(diastol>90)	diabetes
14	(glukosa>100) AND (insulin>200) AND (sistol<=130) AND (glukosa>120) AND (diastol>80)	(sistol>110)	diabetes
15	(glukosa>100) AND (insulin>200)	(sistol>130)	diabetes

Lampiran V: Hasil Klasifikasi

