

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Sistem

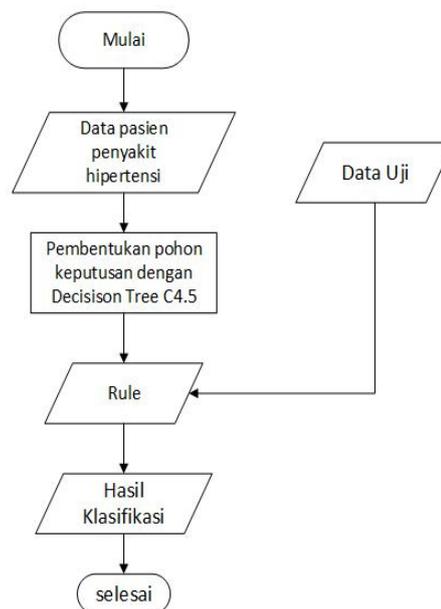
Hipertensi merupakan penyakit pembunuh senyap karena gejala awal yang tidak terlihat. Jika dibiarkan terlalu lama tanpa pengobatan yang tepat dapat menimbulkan komplikasi dan penyakit kardiovaskular bahkan bisa menimbulkan kematian. Dari data WHO menunjukkan hampir 50% penduduk Indonesia mengidap hipertensi. Badan kesehatan seperti WHO dan *British Hypertension Society* memiliki kepedulian yang besar terhadap masalah tekanan darah tinggi dan dampaknya terhadap kesehatan suatu negara. Jumlah penderita hipertensi di dunia terus meningkat setiap tahunnya, diperkirakan juga 9,4 juta orang meninggal akibat hipertensi dan komplikasi. Tingkat keparahan tekanan darah tinggi seseorang berbeda dengan orang lain, tingkat keparahan darah tinggi terbagi dalam beberapa status menurut WHO yaitu hipertensi grade 1 (ringan), hipertensi grade 2 (sedang), dan hipertensi grade 3 (berat). Petugas rumah sakit menentukan tingkat keparahan hipertensi pasien berdasarkan sistem sebelumnya yaitu ketentuan WHO yang ditunjukkan pada tabel 2.3. Tingkat keparahan tekanan darah tinggi (hipertensi) dapat diukur dari beberapa faktor, diantaranya usia, berat badan, tekanan darah sistolik, dan tekanan darah diastolik. Oleh karena itu, mengetahui tingkat keparahan hipertensi perlu dilakukan, untuk mengurangi angka kematian dan pencegahan resiko hipertensi dimasa yang akan datang.

3.2 Hasil Analisis

Hasil analisis yang didapat adalah sistem klasifikasi ini dapat membantu dokter untuk mengambil keputusan yang tepat mengenai kapan dan bagaimana penanganan penyakit tersebut untuk mencegah kemungkinan terburuk bagi penderita dengan mengklasifikasikan tingkat keparahan hipertensi apakah grade 1(ringan), grade 2(sedang), dan grade

3(berat). Pembuatan aplikasi data mining metode *Decision Tree C4.5* diperlukan data pembelajaran, data tersebut diperoleh dari rumah sakit muhammadiyah gresik. Data tersebut akan diolah menggunakan metode *Decision Tree C4.5*. Hasil yang diperoleh dari perhitungan metode *C4.5* yaitu berupa G1 (ringan), G2 (sedang), G3 (berat).

Sistem yang dibangun merupakan aplikasi atau *tool* klasifikasi pasien penderita hipertensi menggunakan teknik data mining dengan metode *Decision Tree C4.5*. Sistem ini akan menghasilkan nilai keluaran berupa kategori tingkat keparahan hipertensi yang tergolong dalam 3 kelas, yaitu kelas G1 (ringan), G2 (sedang), G3 (berat). Terdapat beberapa atribut yang dibutuhkan untuk mengklasifikasikan pasien penderita hipertensi diantaranya : usia, berat badan, tekanan sistolik, tekanan diastolik. Gambar 3.1 menjelaskan alur sistem pada aplikasi sistem klasifikasi penyakit hipertensi.



Gambar 3.1 *Flowcart* sistem *Decision Tree C4.5*

Penjelasan pada gambar 3.1 :

1. Pertama memasukkan data latih (data *training*) pasien penderita hipertensi yang sudah diperiksa dan akan disimpan didalam *database*.

2. Pembuatan pohon keputusan dengan metode *Decision Tree C4.5* berdasarkan data yang sudah disimpan didalam *database*.
3. Selanjutnya memasukkan data yang akan diklasifikasi (Data Uji).
4. Sistem melakukan klasifikasi data uji dengan menggunakan pohon keputusan yang sudah terbentuk pada proses sebelumnya.
5. Sistem mengeluarkan *output* klasifikasi.

Secara umum algoritma *Decision Tree C4.5* dapat dilihat pada gambar 2.6

3.3 Representasi Data

3.3.1 Sumber Data

Tahapan awal yang dilakukan dalam penelitian ini adalah menyiapkan data, dimana data diperoleh dari Rumah Sakit Muhammadiyah Gresik tahun 2016, 2017, 2018. Data yang digunakan adalah data pasien penyakit hipertensi tahun 2016, 2017, 2018. Data yang diperoleh akan digunakan dalam penelitian ini berupa data yang berkaitan dengan atribut hipertensi yaitu usia, berat badan, tekanan darah sistolik, dan tekanan darah diastolik. Jumlah data yang digunakan sebanyak 35 record dengan kelas “hipertensi G1 (ringan)”, “hipertensi G2 (sedang)”, dan “hipertensi G3 (berat)” masing-masing berjumlah 30 dan 5 yang akan dibagi menjadi data latih dan data uji. Data yang didapatkan tersebut dibagi menjadi 4 atribut seperti pada tabel 3.1:

Tabel 3.1 atribut data pasien penyakit hipertensi

No	Atribut	Ket
	Usia	Usia pasien penyakit hipertensi
	Berat badan	Berat badan pasien penyakit hipertensi
	Tekanan sistolik	Tekanan darah sistolik pasien
	Tekanan diastolik	Tekanan darah diastolik pasien

3.3.2 Persiapan Data

Data yang akan diproses untuk klasifikasi pasien penyakit hipertensi, diperoleh dari Rumah Sakit Muhammadiyah Gresik tahun 2016, 2017, 2018. Sebelum dilakukan proses klasifikasi maka data tersebut harus

melalui tahap *preprocessing*. Dari data-data tersebut yang dipilih untuk dijadikan sebagai atribut adalah usia, berat badan, tekanan diastolik, tekanan sistolik. Nilai atribut-atribut tersebut memiliki tipe numerik serta kelas bertipe kategorikal, rinciannya adalah sebagai berikut:

Tabel 3.2 Data Atribut

No	Atribut	keterangan	Tipe
1	Usia	Usia pasien hipertensi	Numerik
2	Berat Badan	Berat badan pasien	Numerik
3	Tekanan sistolik	Tekanan sistolik pasien	Numerik
4	Tekanan diastolik	Tekanan diastolik pasien	Numerik
5	Kelas	HT G1, HT G2, HT G3	Kategorikal

Terdapat dua macam data yang akan di gunakan yaitu data latih dan data uji. Data latih berfungsi untuk pembentukan pohon keputusan sedangkan data uji adalah data untuk pengujian sistem. Data di dapat dari Rumah Sakit Muhammadiyah Gresik.

Pada tabel 3.3 ditampilkan data keseluruhan yang akan digunakan. Sedangkan data latih pada tabel 3.4 dan data uji pada tabel 3.5.

Tabel 3.3 Data Seluruh pasien hipertensi

No	Usia	Berat badan	Tekanan sistolik	Tekanan diastolik	Kelas
1	60	59	160	90	G1
2	60	71	150	100	G1
3	75	49	170	100	G1
4	61	65	160	80	G1
5	71	48	170	90	G1
6	57	45	160	90	G1
7	45	60	160	80	G1
8	65	55	160	90	G1
9	71	48	140	80	G1
10	56	65	150	80	G1
11	70	52	160	100	G2
12	80	54	180	100	G2
13	65	60	160	100	G2
14	54	56	180	80	G2
15	65	56	160	80	G2
16	66	58	160	80	G2
17	42	60	160	90	G2
18	56	70	190	100	G2

19	50	75	180	100	G2
20	70	58	180	110	G2
21	65	40	180	120	G3
22	65	47	220	130	G3
23	51	45	200	120	G3
24	50	80	210	140	G3
25	68	52	180	100	G3
26	50	60	200	120	G3
27	49	64	200	100	G3
28	80	53	190	110	G3
29	47	60	210	100	G3
30	56	65	200	110	G3
31	56	65	150	80	G1
32	85	50	160	90	G2
33	55	65	170	80	G2
34	81	45	190	140	G3
35	80	50	210	140	G3

Tabel 3.4 Data Latih

No	Usia	Berat badan	Tekanan sistolik	Tekanan diastolik	Kelas
1	60	59	160	90	G1
2	60	71	150	100	G1
3	75	49	170	100	G1
4	61	65	160	80	G1
5	71	48	170	90	G1
6	57	45	160	90	G1
7	45	60	160	80	G1
8	65	55	160	90	G1
9	71	48	140	80	G1
10	56	65	150	80	G1
11	70	52	160	100	G2
12	80	54	180	100	G2
13	65	60	160	100	G2
14	54	56	180	80	G2
15	65	56	160	80	G2
16	66	58	160	80	G2
17	42	60	160	90	G2
18	56	70	190	100	G2
19	50	75	180	100	G2
20	70	58	180	110	G2
21	65	40	180	120	G3
22	65	47	220	130	G3
23	51	45	200	120	G3

24	50	80	210	140	G3
25	68	52	180	100	G3
26	50	60	200	120	G3
27	49	64	200	100	G3
28	80	53	190	110	G3
29	47	60	210	100	G3
30	56	65	200	110	G3

Tabel 3.5 Data Uji

No	Usia	Berat badan	Tekanan sistolik	Tekanan diastolik	Kelas
31	56	65	150	80	G1
32	85	50	160	90	G2
33	55	65	170	80	G2
34	81	45	190	140	G3
35	80	50	210	140	G3

3.3.3 Perhitungan *Decision Tree C4.5*

Perhitungan *Decision Tree C4.5* ini akan menggunakan data pada tabel 3.4 (data latih). Tabel tersebut akan diubah menjadi sebuah *Tree*. Sebelum melakukan perhitungan berikut, akan dijelaskan beberapa ketentuan dalam pembentukan *tree* pada kasus ini.

- Pemecahan cabang dilakukan secara biner yaitu pemecahan yang hanya mempunyai dua nilai yaitu \leq dan $>$ (kurang dari sama dengan dan lebih dari).
- Usia
Posisi V yang digunakan pada atribut usia adalah nilai antara (45, 50, 55, 60, 65, 70).
- Berat Badan
Posisi V yang digunakan pada atribut berat badan adalah nilai antara (40,50,60,70).
- Tekanan Sistolik
Posisi V yang digunakan pada atribut tekanan sistolik adalah nilai antara (140, 160, 180, 200, 220).
- Tekanan Diastolik

Posisi V yang digunakan pada atribut tekanan diastolik adalah nilai antara (85, 90, 95, 100, 105, 110, 115).

- Kelas

Pada variabel diagnosa kelas klasifikasi yang digunakan adalah hipertensi G1 (ringan), hipertensi G2 (sedang), hipertensi G3 (berat).

1. Langkah pertama yang harus dilakukan adalah menentukan atribut yang akan dijadikan akar (*root node*) dengan menghitung nilai gain yang paling tinggi. Sebelumnya yang akan dihitung yaitu *entropy* seluruh data. Berikut adalah perhitungan *entropy* seluruh data.

$$-\frac{10}{30} * \log_2 \left(\frac{10}{30} \right) - \frac{10}{30} * \log_2 \left(\frac{10}{30} \right) - \frac{10}{30} * \log_2 \left(\frac{10}{30} \right) = 1.5849625$$

Selanjutnya menghitung gain untuk setiap atribut, jika atribut memiliki nilai numerik akan dilakukan perhitungan untuk menentukan posisi gain tertinggi untuk masing-masing posisi.

2. Menghitung nilai numerik untuk atribut usia

Tabel 3.6 adalah hasil penentuan posisi v pada atribut usia, pada kolom yang diberi keterangan warna biru menunjukkan nilai gain tertinggi yang didapat pada posisi v = 55. Maka untuk atribut usia dilakukan diskretisasi pada v = 55 ketika menghitung *entropy* dan *gain* pada semua atribut.

3. Menghitung nilai numerik untuk atribut berat badan

Tabel 3.7 adalah hasil penentuan posisi v pada atribut berat badan, pada kolom yang diberi keterangan warna biru menunjukkan nilai gain tertinggi yang didapat pada posisi v = 50. Maka untuk atribut usia dilakukan diskretisasi pada v = 50 ketika menghitung *entropy* dan *gain* pada semua atribut.

4. Menghitung nilai numerik untuk atribut tekanan sistolik

Tabel 3.8 adalah hasil penentuan posisi v pada atribut tekanan sistolik, pada kolom yang diberi keterangan warna biru menunjukkan nilai gain tertinggi yang didapat pada posisi v = 180. Maka untuk atribut usia dilakukan diskretisasi pada v = 180 ketika menghitung *entropy* dan *gain* pada semua atribut.

5. Menghitung nilai numerik untuk atribut tekanan diastolik

Tabel 3.9 adalah hasil penentuan posisi v pada atribut tekanan diastolik, pada kolom yang diberi keterangan warna biru menunjukkan nilai *gain* tertinggi yang didapat pada posisi $v = 90$. Maka untuk atribut usia dilakukan diskretisasi pada $v = 90$.

Tabel 3.6 Hasil perhitungan *gain* atribut usia

USIA	45	50	55	60	
Diagnosa <=	>	<=	<=	<=	>
G1	1 0.5	9 0.52631676	1 0.401050703	9 0.529684	1 0.352214
G2	1 0.5	9 0.52631676	2 0.516387121	8 0.529935	3 0.528321
G3	0 0	10 0.53050958	4 0.46134567	6 0.505722	5 0.471109
sum	2	28	7	23	9
entropi	1	1.5831	1.3788	1.5653	1.3516
gain	0.0407	0.0632	0.0978	1.5452	1.5656
					0.0194

Tabel 3.7 Hasil perhitungan *gain* atribut berat badan

BB	40	50	60	70	
Diagnosa <=	>	<=	<=	<=	>
G1	0 0	10 0.52967341	6 0.505722	7 0.525661	3 0.530639
G2	0 0	10 0.52967341	0 0	8 0.530702	2 0.5
G3	1 0	9 0.52387945	3 0.523882466	7 0.522324	3 0.530639
sum	1	29	7	22	8
entropi	0	1.5832	0.9852	1.5505	1.5613
gain	0.0545	0.1664	0.0085	1.585	1.585
					0

Tabel 3.8 Hasil perhitungan *gain* atribut tekanan sistolik

TS	140	160	180	200		
Diagnosa	<=	<=	<=	<=	>	>
G1	9 0.52387945	8 0.431039827	2 0.363231	10 0.509709	0 0	10 0.530726
G2	10 0.52967341	5 0.530196778	5 0.519275	9 0.523882	1 0.352214	10 0.530726
G3	10 0.52967341	0 0	10 0.450315	2 0.323078	8 0.151044	7 0.504916
sum	29	13	17	21	9	27
entropi	1.5832	0.9612	1.3328	1.3567	0.5033	1.5664
gain	0.0545	0.4132			0.4843	0.1752

Tabel 3.9 Hasil perhitungan *gain* atribut tekanan diastolik

TD	85	90	95	100		
Diagnosa	<=	<=	<=	<=	>	>
G1	4 0.461346	6 0.5057216	8 0.389975	2 0.352214	8 0.389975	2 0.352214
G2	3 0.461346	7 0.52232388	4 0.528320834	6 0.528321	4 0.528321	6 0.528321
G3	0 0.461346	10 0.5224495	10 0.471109	0 0	10 0.471109	3 0.391973
sum	7	23	12	18	12	18
entropi	1.384	1.5505	0.9183	1.3516	0.9183	1.3516
gain	0.0733	0.4067	0.4067	0.4067	1.4365	0.5436
						0.3866

6. Perhitungan *gain* untuk menentukan *node* akar

Setelah semua atribut diperoleh *entropy* dan *gain* tertinggi maka ditentukan *gain* yang tertinggi diantara semua atribut. Atribut yang terpilih dengan *gain* tertinggi akan dijadikan *root node* atau akar dan hasilnya disajikan pada tabel 3.10.

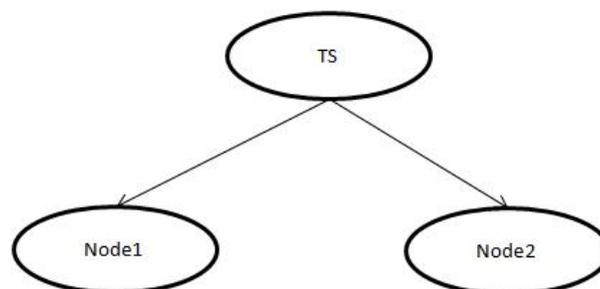
Tabel 3.10 perhitungan *gain* untuk mencari *node* akar

Perhitungan entropy dan gain						NODE 1		
		jml	G1	G2	G3	Entropy	Gain	
Total		30	10	10	10	1.584963		
USIA	<=55	9	1	3	5	1.3516	0.0978	
	>55	21	9	7	5	1.5452		
BB	<=50	7	4	0	3	0.9852	0.1664	
	>50	23	6	10	7	1.5505		
TS	<=180	21	10	9	2	1.3567	0.4843	
	>180	9	0	1	8	0.5033		
TD	<=90	12	8	4	0	0.9183	0.4067	
	>90	18	2	6	10	1.3516		

Hasil yang didapat di Tabel 3.10 menunjukkan bahwa *gain* tertinggi ada pada atribut TS (Tekanan Sistolik) keterangan warna kuning. Maka tekanan sistolik dijadikan *node* akar.

7. Pembentukan pohon keputusan setelah *node* akar terbentuk

Pada atribut TS ≤ 180 dan > 180 memiliki kelas yang berbeda, sehingga keduanya akan menjadi *node* seperti ditunjukkan pada gambar 3.2.



Gambar 3.2 pohon keputusan yang terbentuk pada *node* 1

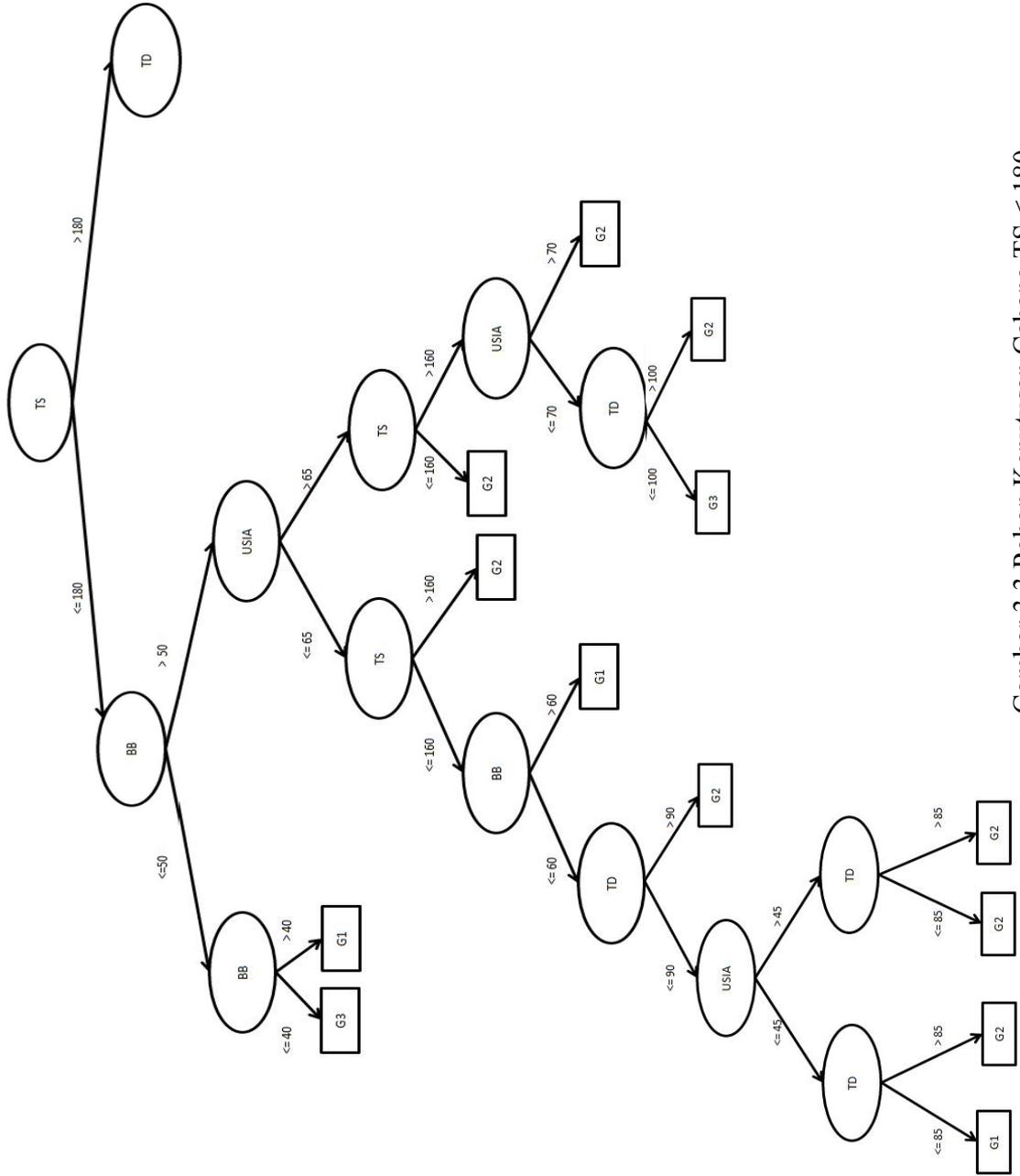
8. Langkah selanjutnya yaitu penyajian data pada cabang $TS \leq 180$

Berikut data dengan atribut tekanan sistoliknya $TS \leq 180$ akan disajikan dalam tabel 3.11.

Tabel 3.11 data pada atribut $TS \leq 180$

No	Nama	Usia	Berat Badan	T.sistolik	T.diastolik	status
1	p1	60	59	160	90	G1
2	p2	60	71	150	100	G1
3	p3	75	49	170	100	G1
4	p4	61	65	160	80	G1
5	p5	71	48	170	90	G1
6	p6	57	45	160	90	G1
7	p7	45	60	160	80	G1
8	p8	65	55	160	90	G1
9	p9	71	48	140	80	G1
10	p10	56	65	150	80	G1
11	p11	70	52	160	100	G2
12	p12	80	54	180	100	G2
13	p13	65	60	160	100	G2
14	p14	54	56	180	80	G2
15	p15	65	56	160	80	G2
16	p16	66	58	160	80	G2
17	p17	42	60	160	90	G2
18	p19	50	75	180	100	G2
19	p20	70	58	180	110	G2
20	p21	65	40	180	120	G3
21	p25	68	52	180	100	G3

Data dengan atribut tekanan sistolik $TS \leq 180$ berjumlah 21 data dengan kelas yang berbeda yaitu G1 dan G2. Sehingga perlu dilakukan perhitungan *entropy* dan *gain* untuk menentukan *node* selanjutnya. Langkah berikutnya dikerjakan seperti langkah ke 2 sampai langkah 8, sehingga diperoleh hasil cabang $TS \leq 180$ ditunjukkan pada gambar 3.3.



Gambar 3.3 Pohon Keputusan Cabang TS ≤ 180

9. Selanjutnya penentuan untuk cabang $TS > 180$

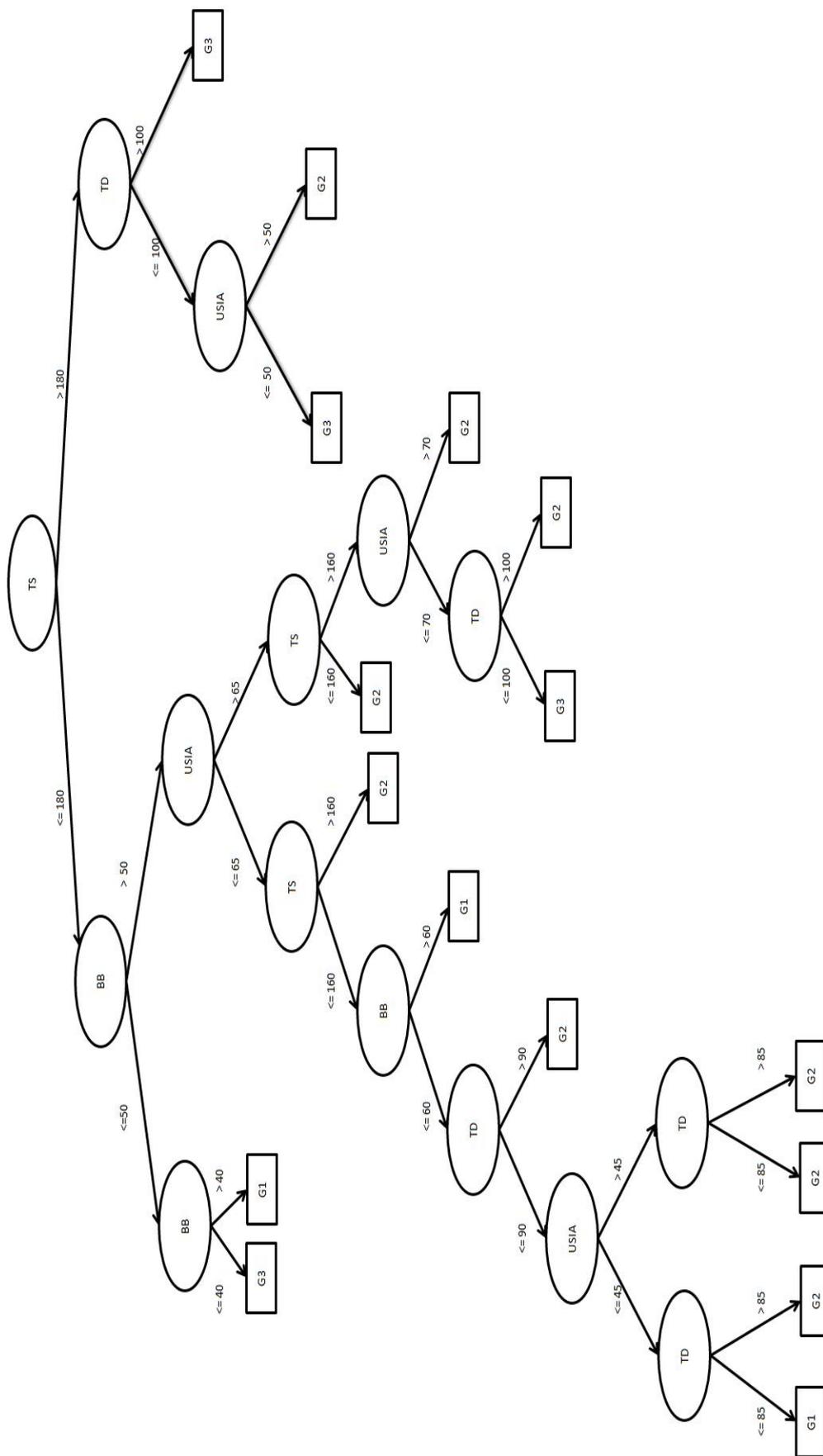
Data yang termasuk dalam kriteria $TS > 180$ berjumlah 9 data yang ditunjukkan pada tabel 3.12. perhitungan dilakukan sesuai dengan langkah 1 sampai langkah 8.

Tabel 3.12 data pada atribut $TS > 180$

No	Nama	Usia	Berat Badan	T.sistolik	T.diastolik	status
18	p18	56	70	190	100	G2
22	p22	65	47	220	130	G3
23	p23	51	45	200	120	G3
24	p24	50	80	210	140	G3
26	p26	50	60	200	120	G3
27	p27	49	64	200	100	G3
28	p28	80	53	190	110	G3
29	p29	47	60	210	100	G3
30	p30	56	65	200	110	G3

10. Pohon keputusan pada cabang $TS > 180$

Perolehan pohon keputusan pada cabang $TS > 180$ melewati beberapa perhitungan, yang mana dikerjakan sesuai langkah 1 sampai langkah ke 8. Sehingga diperoleh hasil pohon keputusan yang ditunjukkan pada gambar 3.4.



Gambar 3.4 Hasil pohon keputusan klasifikasi penyakit hipertensi

Dari pohon keputusan tersebut akan dijadikan dalam bentuk aturan IF THEN sebagai berikut :

1. TS \leq 180 AND BB \leq 50 AND BB \leq 40 THEN G3
2. TS \leq 180 AND BB \leq 50 AND BB $>$ 40 THEN G1
3. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB \leq 60 AND TD \leq 90 AND USIA \leq 45 AND TD \leq 85 THEN G1
4. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB \leq 60 AND TD \leq 90 AND USIA \leq 45 AND TD $>$ 85 THEN G2
5. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS $>$ 160 THEN G2
6. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB $>$ 60 THEN G1
7. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB \leq 60 AND TD $>$ 90 THEN G2
8. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB \leq 60 AND TD \leq 90 AND USIA $>$ 45 AND TD \leq 85 THEN G2
9. TS \leq 180 AND BB $>$ 50 AND USIA \leq 65 AND TS \leq 160 AND BB \leq 60 AND TD \leq 90 AND USIA $>$ 45 AND TD $>$ 85 THEN G2
10. TS \leq 180 AND BB $>$ 50 AND USIA $>$ 65 AND TS \leq 160 THEN G2
11. TS \leq 180 AND BB $>$ 50 AND USIA $>$ 65 AND TS $>$ 160 AND USIA $>$ 70 THEN G2
12. TS \leq 180 AND BB $>$ 50 AND USIA $>$ 65 AND TS $>$ 160 AND USIA \leq 70 AND TD \leq 100 THEN G3
13. TS \leq 180 AND BB $>$ 50 AND USIA $>$ 65 AND TS $>$ 160 AND USIA \leq 70 AND TD $>$ 100 THEN G2
14. TS $>$ 180 AND TD $>$ 100 G3
15. TS $>$ 180 AND TD \leq 100 AND USIA \leq 50 THEN G3
16. TS $>$ 180 AND TD \leq 100 AND USIA $>$ 50 THEN G2

Dari perhitungan data latih digunakan untuk membentuk *rule*, sehingga dai *rule* tersebut dapat dijadikan aturan sebagai pengklasifikasian untuk data uji. Dari *rule* yang sudah terbentuk dilakukan pengujian terhadap data latih dan data uji, untuk mengetahui apakah *rule* yang

terbentuk sudah maksimal atau belum. Jika hasil pengujian data latih dan data uji menunjukkan nilai akurasi yang tinggi dan nilai *error* yang rendah maka sudah dapat disimpulkan bahwa *rule* yang terbentuk sudah maksimal. Berikut hasil pengujian data latih dan data uji masing-masing di tunjukkan pada Tabel 3.13 dan Tabel 3.14.

Tabel 3.13 Pengujian Data Latih

No	Usia	Berat badan	Tekanan sistolik	Tekanan diastolik	Kelas	Hasil
1	60	59	160	90	G1	Benar
2	60	71	150	100	G1	Benar
3	75	49	170	100	G1	Benar
4	61	65	160	80	G1	Benar
5	71	48	170	90	G1	Benar
6	57	45	160	90	G1	Benar
7	45	60	160	80	G1	Benar
8	65	55	160	90	G1	Benar
9	71	48	140	80	G1	Benar
10	56	65	150	80	G1	Benar
11	70	52	160	100	G2	Benar
12	80	54	180	100	G2	Benar
13	65	60	160	100	G2	Benar
14	54	56	180	80	G2	Benar
15	65	56	160	80	G2	Benar
16	66	58	160	80	G2	Benar
17	42	60	160	90	G2	Benar
18	56	70	190	100	G2	Benar
19	50	75	180	100	G2	Benar
20	70	58	180	110	G2	Benar
21	65	40	180	120	G3	Benar
22	65	47	220	130	G3	Benar
23	51	45	200	120	G3	Benar
24	50	80	210	140	G3	Benar
25	68	52	180	100	G3	Benar
26	50	60	200	120	G3	Benar
27	49	64	200	100	G3	Benar
28	80	53	190	110	G3	Benar
29	47	60	210	100	G3	Benar
30	56	65	200	110	G3	Benar

Tabel 3.14 Pengujian Data Uji

No	Usia	Berat badan	Tekanan sistolik	Tekanan diastolik	Kelas	Hasil
31	56	65	150	80	G1	Benar
32	85	50	160	90	G2	Salah
33	55	65	170	80	G2	Benar
34	81	45	190	140	G3	Benar
35	80	50	210	140	G3	Benar

3.4 Kebutuhan Pembuatan Sistem

Kebutuhan-kebutuhan untuk pembangunan sistem sebagai berikut :

A. Perangkat Lunak

Adapun perangkat lunak yang dibutuhkan dalam pembangunan aplikasi tersebut adalah sebagai berikut :

1. Aplikasi browser *Firefox Mozilla* atau *Google Chrome* yang berfungsi untuk mengakses klasifikasi penyakit hipertensi dengan *Decision Tree C4.5* yang berbasis *web*.
2. *Mysql* adalah sebuah perangkat lunak yang berfungsi sebagai penyimpanan data atau *database*.
3. XAMPP adalah aplikasi berisi semua aplikasi *server* yang berfungsi untuk membuat *server* lokal atau *localhost*.

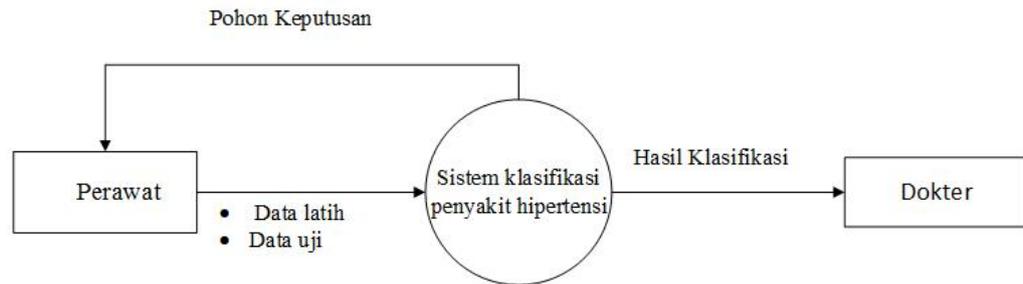
B. Kebutuhan Perangkat Keras

1. Prosesor Intel Pentium
2. RAM 4 GB
3. HDD 500 GB
4. Monitor
5. *Keyboard dan mouse*
6. Printer

3.5 Perancangan sistem

Bagian ini akan menjelaskan rancangan sistem seperti *flowcart*, *context diagram*, diagram berjenjang dan *data flow diagram* (DFD).

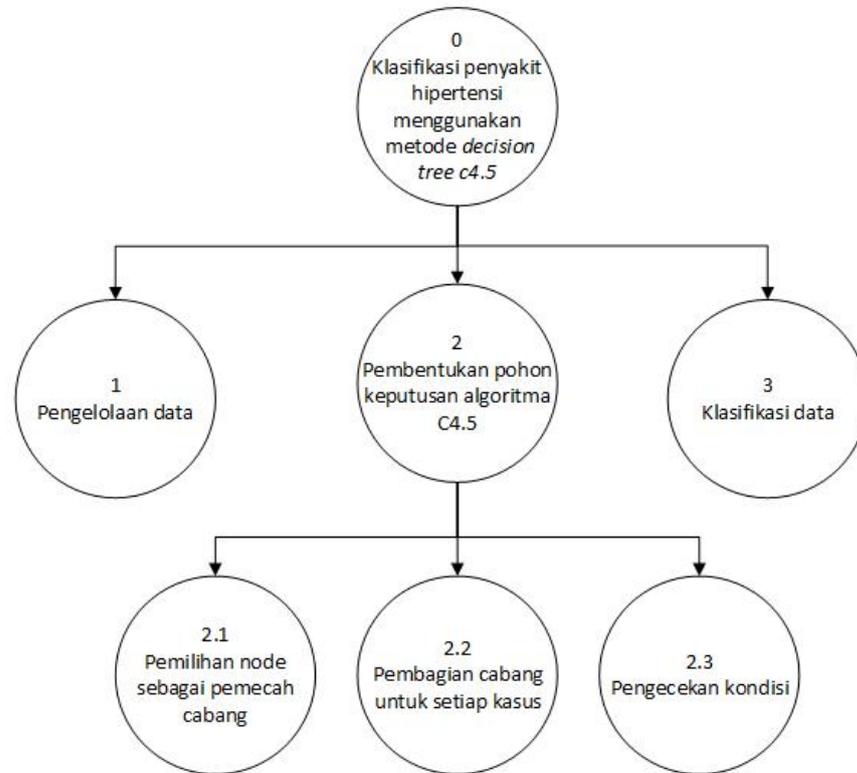
3.5.1 Context Diagram



Gambar 3.5 *Context Diagram*

Penjelasan dari gambar 3.5, terlihat bahwa yang terlibat atau *entity* dalam sistem ini adalah perawat dan dokter. Perawat memasukkan data pasien penyakit hipertensi sebagai data latih yang terdiri dari usia, berat badan, tekanan sistolik, dan tekanan diastolik. Data tersebut digunakan sebagai data latih atau *data training* yang akan diproses untuk pembentukan pohon keputusan. Keluaran dari sistem untuk perawat adalah rule atau pohon keputusan pasien penyakit hipertensi berdasarkan data yang telah dimasukkan. Setelah pohon keputusan terbentuk, perawat memasukkan data uji pasien hipertensi untuk diklasifikasikan dan dokter akan menerima hasil klasifikasi pasien.

3.5.2 Diagram Berjenjang



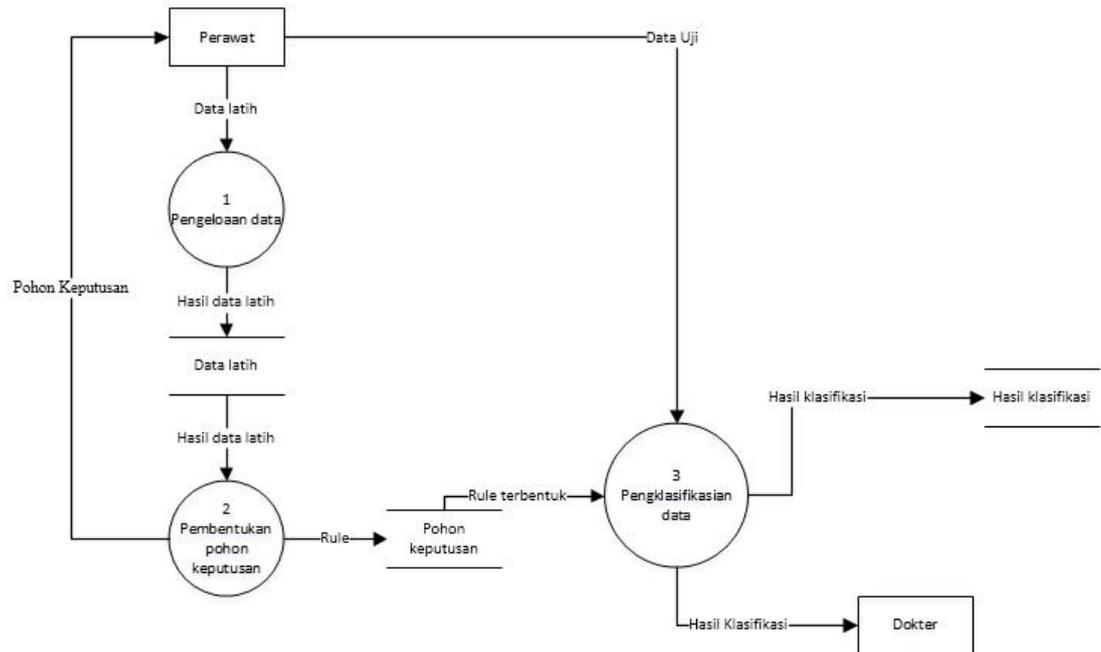
Gambar 3.6 Diagram Berjenjang

Diagram berjenjang disajikan pada gambar 3.6 berikut penjelasannya:

- 1) *Top level* : Sistem klasifikasi penyakit hipertensi di Rumah Sakit Muhammadiyah Gresik.
- 2) *Level 0* :
 1. Pengelolaan data, merupakan proses pengelolaan data latih training atau data yang akan digunakan dalam pembentukan pohon keputusan.
 2. Pembentukan aturan (pohon keputusan) dengan metode C4.5, yang didalamnya terdapat tiga proses.
 3. Pengklasifikasian data uji menggunakan aturan yang sudah terbentuk.
- 3) *Level 1* :
 - 2.1 Pemilihan *node* sebagai pemecah cabang.
 - 2.2 Pembagian cabang pada setiap kasus.

2.3 Pengecekan kondisi, yaitu jika masih ada kasus yang memiliki kelas yang berbeda maka mengulangi.

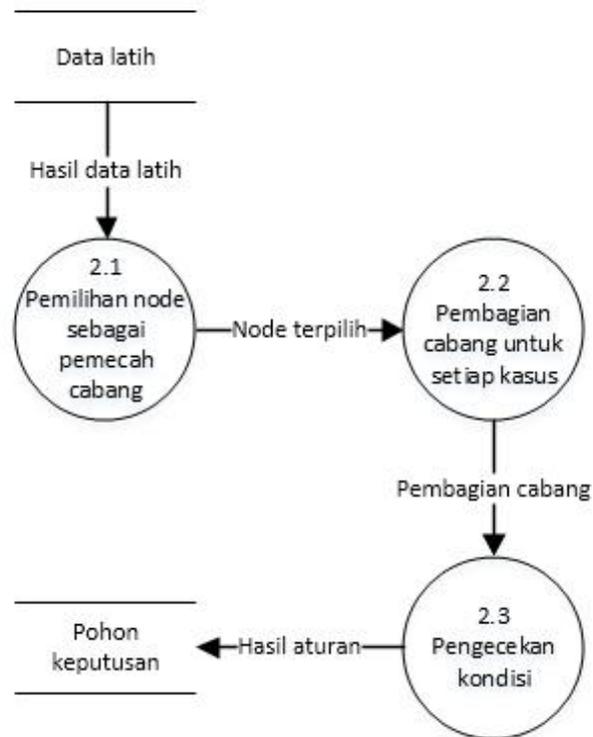
3.5.3 Data Flow Diagram Level 0



Gambar 3.7 DFD Level 0 Aliran data dalam sistem

DFD level 0 pada gambar 3.7 menjelaskan aliran data pada sistem. Terdapat empat proses didalam sistem tersebut. Proses satu adalah perawat memasukkan data latih yang akan digunakan untuk proses pembentukan aturan (*rule*). *Rule* yang sudah terbentuk akan dijadikan untuk proses pengklasifikasian data uji. Hasilnya akan disimpan dalam hasil klasifikasi.

3.5.4 Data Flow Diagram Level 1



Gambar 3.8 DFD Level 1 Proses pembentukan aturan (pohon keputusan)

Proses pembentukan aturan menggunakan metode *decision tree* c4.5 ini memiliki tiga proses didalamnya yaitu, proses pemilihan *node* yang akan dijadikan sebagai pemecah cabang, membagi cabang pada setiap kasus, dan proses pengecekan kondisi. Jika ada kasus yang memiliki kelas berbeda. Maka akan mengulangi pada proses pemilihan *node*. Hasil dari proses ini adalah aturan atau pohon keputusan yang akan disimpan pada *database*.

3.5.5 Struktur Tabel

Struktur tabel ini menjelaskan tabel atau tempat penyimpanan data yang digunakan untuk keperluan sistem yang akan dibangun. Berikut adalah struktur dari tabel-tabel yang akan digunakan.

a. Data latih

Tabel ini digunakan untuk menyimpan data latih atau data yang akan diproses pada pembentukan pohon keputusan. Tabel ini berisi rekam

medis dan diagnosa yang didapat dari Rumah Sakit Muhammadiyah Gresik. Struktur tabel ini dapat dilihat pada Tabel 3.15.

Tabel 3.15 struktur tabel data latihan

Name	Type	Length	Key
Id	int	30	Primery key
Nama	varchar	30	
Usia	int	30	
Berat badan	int	30	
Tsistolik	int	30	
Tdiastolik	int	30	
Status	varchar	30	

b. Data uji

Tabel ini digunakan untuk menyimpan data pengujian, yaitu untuk menguji tingkat akurasi dari pohon keputusan yang terbentuk. Struktur tabelnya ditunjukkan pada Tabel 3.16.

Tabel 3.16 struktur tabel data uji

Name	Type	Length	Key
Id	int	30	Primery key
Usia	int	30	
Berat badan	int	30	
t_sistolik	int	30	
t_diastolik	int	30	
Status_asli	varchar	30	
Status_klasifikasi	varchar	30	

c. Hasil klasifikasi

Tabel ini digunakan untuk menyimpan data hasil klasifikasi. Strukturnya ditunjukkan pada tabel 3.17.

Tabel 3.17 struktur tabel hasil klasifikasi

Name	Type	Length	Key
Id	int	30	Primery key
Usia	int	30	
Berat badan	int	30	
t_sistolik	int	30	
t_diastolik	int	30	
hasil	varchar	30	

d. *Gain*

Tabel ini merupakan *temporary* yang digunakan untuk menampung hasil perhitungan *gain* seperti pada tabel 3.18.

Tabel 3.18 struktur tabel *gain*

Name	Type	Length	Key
id	int	30	Primery key
Node id	Int	30	
atribut	varchar	30	
gain	double		

e. User

Tabel *user* ini dibuat supaya bisa mengakses aplikasi ini. Data dari *user* tersebut tersimpan dalam tabel *user*. Struktur tabel ini dapat dilihat pada tabel 3.19.

Tabel 3.19 struktur tabel user

Name	Type	Length	Key
User id	int	30	Primery key
Nama	varchar	30	
username	varchar	30	
password	varchar	30	
type	Char	30	

f. Keputusan

Tabel ini menampung hasil dari proses pembentukan pohon keputusan, yaitu menampung aturan-aturan yang telah terbentuk diunjukkan pada tabel 3.20.

Tabel 3.20 struktur tabel keputusan

Name	Type	Length	Key
id	int	30	Primery key
Parent	Text		
Akar	Text		
Keputusan	Varchar	30	

3.5.6 Desain Antar Muka

a) Halaman login

Halaman login digunakan user untuk masuk kedalam sistem.

Login
Username <input type="text"/>
Password <input type="password"/>
<input type="button" value="Login"/>

Gambar 3.9 Rancangan halaman login

b) Halaman Utama

Menu yang ditampilkan untuk pengguna sebagai admin yaitu daftar dashboard, data Latih, proses, rule, dan klasifikasi. Terdapat tombol logout yang digunakan admin untuk keluar aplikasi.

RSMG	Logout
Dashboard	Klasifikasi Penyakit Hipertensi menggunakan metode Decision Tree C4.5
Data Latih	
Proses	
Rule	
Klasifikasi	
	Footer

Gambar 3.10 Rancangan halaman utama

c) Halaman data latih

Halaman data latih digunakan admin untuk melihat data latih, serta untuk memasukkan data latih yang baru dengan meng-klik tombol

tambah data latih, admin juga dapat meng-edit serta menghapus data latih.

RSMG		Logout																									
Dashboard	<div style="text-align: center; margin-bottom: 10px;"> <input type="button" value="Tambah Data"/> </div> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>No</th> <th>Nama</th> <th>Usia</th> <th>Berat Badan</th> <th>Tsistolik</th> <th>Tdiastolik</th> <th>Aksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Hapus Edit</td> </tr> <tr> <td>2</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Hapus Edit</td> </tr> </tbody> </table>						No	Nama	Usia	Berat Badan	Tsistolik	Tdiastolik	Aksi	1						Hapus Edit	2						Hapus Edit
No							Nama	Usia	Berat Badan	Tsistolik	Tdiastolik	Aksi															
1												Hapus Edit															
2												Hapus Edit															
Data Latih																											
Proses																											
Rule																											
Klasifikasi																											
Footer																											

Gambar 3.11 Rancangan halaman data latih

d) Halaman tambah data latih

Halaman tambah data latih nantinya akan digunakan admin untuk menambah data latih, dengan meng-inputkan atribut yang sudah ditentukan.

RSMG		Logout	
Dashboard	<div style="margin-bottom: 10px;">Usia <input style="width: 150px;" type="text"/></div> <div style="margin-bottom: 10px;">BB <input style="width: 150px;" type="text"/></div> <div style="margin-bottom: 10px;">T.sistol <input style="width: 150px;" type="text"/></div> <div style="margin-bottom: 10px;">T.diastol <input style="width: 150px;" type="text"/></div> <div style="text-align: center; margin-bottom: 10px;"> <input type="button" value="Submit"/> </div>		
Data Latih			
Proses			
Rule			
Klasifikasi			
Footer			

Gambar 3.12 Rancangan halaman tambah data latih

e) Halaman proses C4.5

Tampilan halaman ini akan digunakan admin untuk melakukan proses mining dengan algoritma C4.5.

RSMG		Logout																			
Dashboard	<p>DATA LATIH</p> <table border="1"> <thead> <tr> <th>No</th> <th>Nama</th> <th>Usia</th> <th>Berat Badan</th> <th>Tsistolik</th> <th>Tdiastolik</th> </tr> </thead> <tbody> <tr> <td>1</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>2</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table> <p style="text-align: right;">PROSES C45</p> <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px auto; width: 80%;">Tabel data proses C4.5</div>			No	Nama	Usia	Berat Badan	Tsistolik	Tdiastolik	1						2					
No				Nama	Usia	Berat Badan	Tsistolik	Tdiastolik													
1																					
2																					
Data Latih																					
Proses																					
Rule																					
Klasifikasi																					
	Footer																				

Gambar 3.13 Rancangan halaman proses C4.5

f) Halaman rule

Halaman ini merupakan halaman dimana rule yang terbentuk dari proses mining ditampilkan. Terdapat dua pilihan bagi admin untuk menghapus rule dan untuk menguji rule. Hasil dari uji rule akan ditampilkan di halaman uji rule.

RSMG		Logout	
Dashboard	<p>RULE</p> <p style="text-align: right;"> <input type="button" value="Hapus rule"/> <input type="button" value="Uji rule"/> </p> <div style="border: 1px solid black; padding: 5px; margin: 10px auto;"> <pre> 1.IF TS <= 160 AND BB > 60 THEN Label = G1 2.IF TS <= 160 AND BB <= 60 AND TD > 90 THEN Label = G2 3.IF TS <= 160 AND BB <= 60 AND TD <= 90 AND BB <= 50 THEN Label = G1 4.IF TS <= 160 AND BB <= 60 AND TD <= 90 AND BB > 50 AND USIA > 65 THEN Label = G2 </pre> </div>		
Data Latih			
Proses			
Rule			
Klasifikasi			

Gambar 3.14 halaman rule

g) Halaman uji rule

Halaman ini digunakan admin untuk menguji rule, dengan memasukkan data uji yang akan di uji dengan rule yang sudah terbentuk. Uji rule akan menampilkan beberapa informasi yang diperlukan seperti akurasi dan laju error.

RSMG		Logout																						
Dashboard	<p><u>UJI RULE</u></p> <p>Tambah Data Uji</p> <table border="1"> <thead> <tr> <th>No</th> <th>Nama</th> <th>Usia</th> <th>Berat Badan</th> <th>Tsistolik</th> <th>Tdiastolik</th> <th>Aksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Hapus Edit</td> </tr> <tr> <td>2</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Hapus Edit</td> </tr> </tbody> </table> <p>Hitung Akurasi</p> <p>Jumlah data yang diprediksi : Jumlah data prediksi tepat : Jumlah data prediksi tidak tepat : Akurasi : Laju error :</p>			No	Nama	Usia	Berat Badan	Tsistolik	Tdiastolik	Aksi	1						Hapus Edit	2						Hapus Edit
No				Nama	Usia	Berat Badan	Tsistolik	Tdiastolik	Aksi															
1									Hapus Edit															
2									Hapus Edit															
Data Latih																								
Proses																								
Rule																								
Klasifikasi																								
	Footer																							

Gambar 3.15 Uji rule

h) Halaman klasifikasi

Halaman klasifikasi digunakan admin untuk memasukkan data pasien baru yang akan diklasifikasikan. Setelah itu akan muncul hasil klasifikasi dari data pasien yang baru masukkan.

RSMG		Logout									
Dashboard	<p>KLASIFIKASI</p> <table border="1"> <tbody> <tr> <td>Usia</td> <td><input type="text"/></td> </tr> <tr> <td>BB</td> <td><input type="text"/></td> </tr> <tr> <td>T.sistolik</td> <td><input type="text"/></td> </tr> <tr> <td>T.diastolik</td> <td><input type="text"/></td> </tr> </tbody> </table> <p>Klasifikasi</p> <p>Hasil klasifikasi anda :</p>			Usia	<input type="text"/>	BB	<input type="text"/>	T.sistolik	<input type="text"/>	T.diastolik	<input type="text"/>
Usia				<input type="text"/>							
BB				<input type="text"/>							
T.sistolik				<input type="text"/>							
T.diastolik				<input type="text"/>							
Data Latih											
Proses											
Rule											
Klasifikasi											
	Footer										

Gambar 3.16 Halaman Klasifikasi

3.6 Evaluasi Sistem

Sistem prediksi (klasifikasi) tidak bisa bekerja 100% benar, maka pada bagian ini akan mengevaluasi hasil perhitungan prediksi. Evaluasi ini menggunakan *Confusion Matrik* yaitu tabel yang digunakan untuk menentukan kinerja suatu model klasifikasi. Untuk mengukur nilai akurasi yang didapat dari hasil pengujian, menggunakan rumus 2.5, dan rumus 2.6. Berikut ini merupakan contoh perhitungan *Counfusion Matrix* dengan menggunakan nilai dari hasil yang diperoleh data uji dapat dilihat pada tabel 3.21.

Tabel 3.21 Matrik Confusion data uji

Evaluasi	Hasil Prediksi		
	G1	G2	G3
G1	1	0	0
G2	1	1	0
G3	0	0	2

Berikut ini merupakan perhitungan akurasi dan laju error dengan menggunakan nilai dari hasil yang diperoleh dari pengujian data latih. Dari hasil prediksi diketahui :

Jumlah data dengan prediksi sesuai = 4

Jumlah data dengan prediksi tidak sesuai = 1

Jumlah prediksi yang dilakukan = 5

$$Akurasi = \frac{4}{5} = 1 \times 100\% = 80\%$$

$$Laju\ error = \frac{1}{5} = 0 \times 100\% = 20\%$$

Sebelum membuat aplikasi klasifikasi pasien penyakit hipertensi dengan metode decision tree C4.5 ini, perlu dilakukan beberapa skenario pengujian sistem terlebih dahulu, agar sistem dapat berjalan sesuai dengan tujuan pembuatannya.

- a) Data tahun 2016, 2017, 2018 disediakan dua macam data yaitu data latih sebanyak 90 data, dan data uji sebanyak 9 data. Data latih di gunakan untuk membentuk pohon keputusan dan data uji digunakan untuk menguji

akurasi. Selanjutnya membentuk pohon keputusan dari data latih yang sudah di sediakan. Untuk data uji, diklasifikasikan berdasarkan pohon keputusan yang terbentuk. Kemudian akan dihitung akurasi baik atau tidaknya pohon keputusan yang sudah terbentuk.

- b) Pada uji hasil klasifikasi maka sistem dapat menentukan grade pasien penderita hipertensi apakah G1 (ringan), G2 (sedang) , atau G3 (berat).
- c) Akurasi sistem diperoleh dari hasil presentase akurasi ketepatan sistem dan hasil dari presentase akurasi kesalahan sistem.