

## BAB III

### ANALISIS DAN PERANCANGAN SISTEM

#### 3.1 Analisis Sistem

Analisis sistem adalah penguraian suatu sistem yang utuh ke dalam bagian-bagian komponennya yang bertujuan untuk mengidentifikasi dan menganalisa permasalahan, kesempatan, hambatan yang terjadi dan kebutuhan yang diharapkan sehingga dapat diusulkan perbaikannya.

Pengenalan gejala stroke pada setiap individu biasanya berbeda tergantung dari faktor-faktor yang mempengaruhi terjadinya stroke. Pengendalian stroke penting dilakukan agar angka kematian akibat stroke dapat dikurangi, selain itu bagi penderita stroke yang mengalami kecacatan tentu akan sangat merugikan bagi penderita, keluarga, maupun negara. Sehingga, pemerintah melalui Kementerian Kesehatan RI menerbitkan kegiatan-kegiatan pengendalian stroke. Salah satunya adalah dengan melakukan pengenalan sedini mungkin faktor resiko stroke. Deteksi dini pada penyakit stroke dapat dilakukan dengan mengenali faktor-faktor berupa data kesehatan yaitu Tekanan Darah, Kadar Gula Darah, Kolesterol Total, *Low Density Lipoprotein* (LDL), Umur, Jenis Kelamin, Asam Urat, *Blood Urea Nitrogen* (BUN) dan Kreatinin. Kemudian dokter akan menganalisis data dan mendiagnosa status dini risiko penyakit stroke pasien dalam kategori rendah, sedang dan tinggi. Setelah pasien mengetahui status deteksi dini stroke maka pasien akan diberikan informasi-informasi agar terhindar dari serangan stroke serta meminimalkan faktor-faktor risiko stroke.

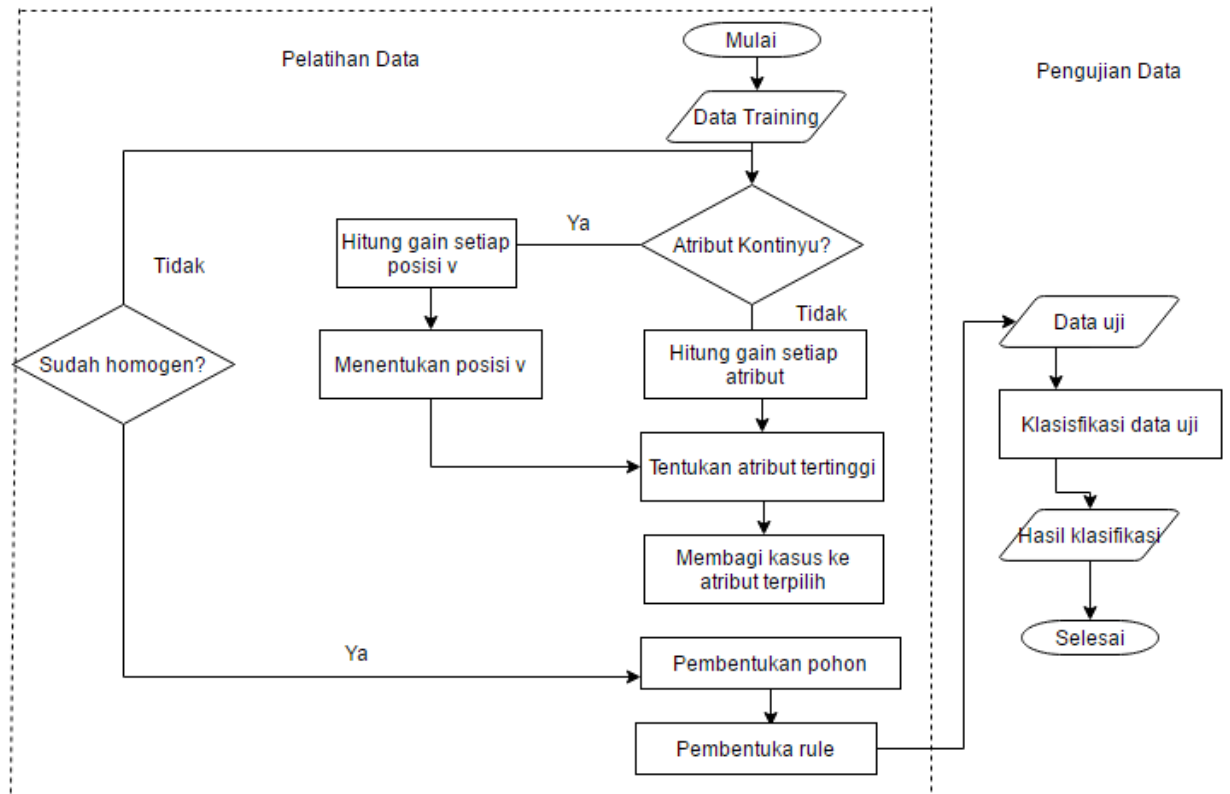
Proses pengenalan gejala serta faktor-faktor penyebab stroke yang berbeda-beda pula akan mempengaruhi dokter untuk proses diagnosa. Selain itu diagnosa yang dilakukan juga tidak mudah karena banyaknya faktor risiko yang beragam dan saling mempengaruhi, contohnya kolesterol (total dan LDL) dapat menyebabkan penyakit jantung juga dapat berpengaruh terhadap tekanan darah, jenis kelamin dapat memengaruhi nilai asam urat, asam urat

juga dapat memengaruhi tekanan darah dan kadar gula dapat memengaruhi tekanan darah. Oleh karena itu, dibutuhkan suatu sistem pendukung keputusan untuk membantu dokter dalam menentukan tingkat risiko stroke pasien dalam kategori rendah, sedang dan tinggi.

### 3.2 Hasil Analisis

Hasil analisis yang didapat adalah dengan menggunakan sistem ini dokter bisa mendapatkan informasi kategori resiko dini penyakit stroke berupa resiko rendah, resiko sedang dan resiko tinggi. Pembuatan aplikasi dengan memanfaatkan metode klasifikasi data mining dengan *Decision Tree C4.5* memerlukan data latih, data tersebut diperoleh dari data rekam medis pasien dengan diagnosa status risiko stroke oleh dokter, yang nantinya akan diolah dengan menggunakan metode *Decision Tree C4.5*. Hasil yang diperoleh dari perhitungan metode tersebut berupa hasil prediksi status deteksi dini risiko stroke pasien yang dapat membantu dokter dalam pengambilan keputusan.

Sistem yang dibangun merupakan aplikasi atau *tool* klasifikasi tingkat resiko penyakit stroke menggunakan teknik *data mining* dengan metode *Decision Tree C4.5*. Sistem ini akan menghasilkan nilai keluaran berupa kategori status resiko stroke yang tergolong ke dalam 3 kelas, yaitu kelas resiko rendah, resiko sedang dan resiko tinggi. Terdapat beberapa atribut yang dibutuhkan untuk mengklasifikasi tingkat resiko stroke pada pasien diantaranya: Tekanan Darah, Kadar Gula Darah, Kolesterol Total, *Low Density Lipoprotein* (LDL), Umur, Jenis Kelamin, Asam Urat, *Blood Urea Nitrogen* (BUN) dan Kreatinin. Diagram alir sistem pengklasifikasian untuk mendiagnosa awal resiko penyakit stroke dengan *Decision Tree C4.5* ditujukan pada **Gambar 3.1**



**Gambar 3.1** Flowchart System

**Penjelasan Gambar 3.1:**

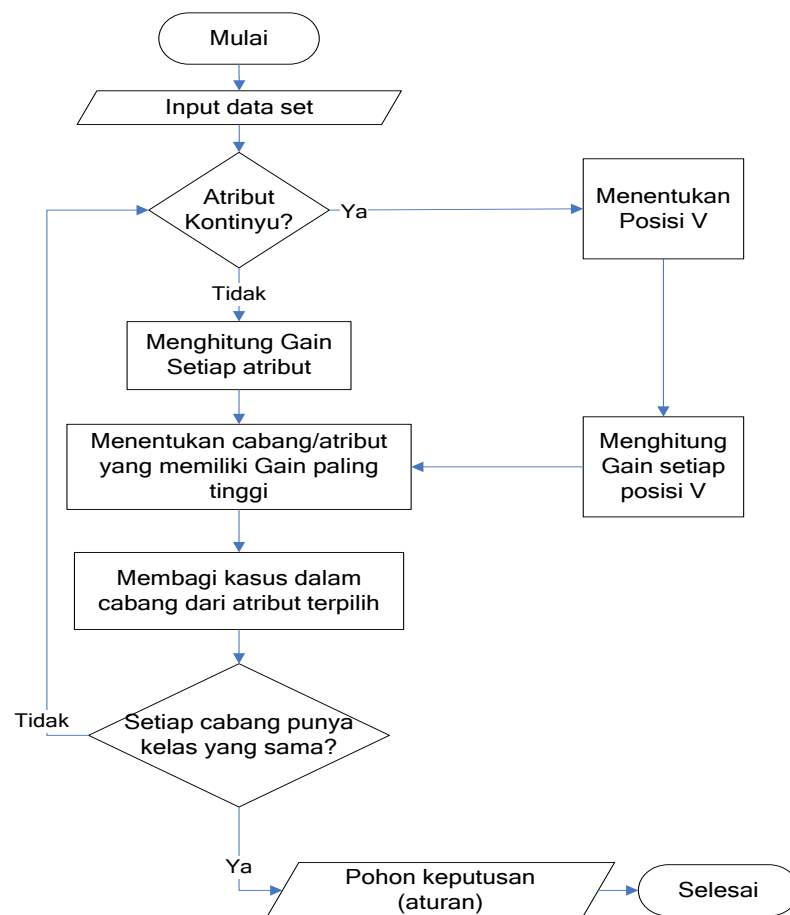
1. Pertama memasukkan data training (data pasien yang telah diperiksa) yang akan disimpan didalam *database*.
2. Lalu sistem akan memeriksa apakah atribut dari *data training* bertipe kontinyu (numerik)?
3. Jika atribut dari *data trining* bertipe kontinyu maka sistem akan menentukan posisi V, setelah itu akan dihitung *gain* setiap posisi V.
4. Jika atribut bertipe kategorikal (bukan kontinyu) maka akan dilakukan penghitungan *gain* untuk atribut tersebut.
5. Dari hasil penghitungan *gain* bertipe kontinyu maupun kategorikal tersebut selanjutnya akan dilakukan penentuan cabang/atribut yang memiliki nilai *gain* paling tinggi.
6. Pembagian kasus dalam cabang dari atribut terpilih.

7. Sistem akan memeriksa apakah setiap cabang memiliki kelas yang sama?
8. Jika setiap cabang memiliki kelas yang berbeda maka dilakukan perhitungan kembali ke poin no.2.
9. Jika setiap cabang memiliki kelas yang sama maka menghasilkan pembentukan pohon keputusan dan dari pohon keputusan.
10. Dari pohon keputusan yang terbentuk maka akan dapat pula membentuk *rule* sesuai dengan pohon keputusan yang terbentuk.
11. Selanjutnya masukkan data uji yang akan dilakukan klasifikasi.
12. Sistem akan melakukan klasifikasi data uji dengan menggunakan pohon keputusan yang sudah terbentuk.
13. Sistem menghasilkan *output* hasil klasifikasi.

Algoritma *Decision Tree C4.5* dimulai dengan memasukkan data set berupa data latih, kemudian sistem akan melihat nilai untuk setiap atribut bernilai kontinyu atau tidak. Jika atribut bernilai kontinyu maka sistem akan melakukan perhitungann nilai  $V$ , jika atribut memiliki nilai kategorikal maka akan dilakukan perhitungan *gain* untuk setiap atribut, setelah memiliki nilai *gain* untuk semua atribut baik kategorikal maupun kontinyu maka sistem akan menentukan cabang/atribut mana yang memiliki nilai *gain* tertinggi. Selanjutnya dari *gain* tertinggi tersebut akan dibagi kedalam cabang dari atribut terpilih. Dari atribut yang terpilih tersebut akan dilihat apakah setiap cabang memiliki kelas yang sama atau tidak, jika setiap cabang sudah memiliki kelas yang sama maka akan terbentuk pohon keputusan, jika setiap cabang belum memiliki kelas yang sama maka akan dilakukan proses perhitungan lagi sampai semua cabang memiliki kelas yang sama.

Berikut ini adalah alur diagram dari algoritma *Decision Tree C4.5* digambarkan pada **Gambar 3.2**. Untuk memilih atribut sebagai simpul akar (*root node*) atau simpul dalam (*internal node*), didasarkan pada nilai *information gain* tertinggi dari atribut-atribut yang ada. Sebelum perhitungan *information gain*, akan dilakukan perhitungan *entropy*. *Entropy* merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma

C4.5 untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan data (*data set*). Semakin tinggi tingkat *entropy* dari sebuah data maka semakin homogen distribusi kelas pada data tersebut.



**Gambar 3.2** Algoritma Metode *Decision Tree C4.5*

### 3.3 Representasi Model *Decision Tree C4.5*

Dari 128 data pasien (data rekam medis pasien dan status stroke), diambil 90 data yang akan dijadikan data latih dan 38 sebagai data uji. Data latih disajikan pada tabel 3.1. Sedangkan data uji disajikan pada **Tabel 3.2**. Metode pengklasifikasian dengan menggunakan *Decision Tree C4.5* adalah pengembangan dari *ID3*, jika pada *ID3* pengklasifikasian hanya bisa dilakukan untuk atribut yang memiliki nilai kategorikal saja, maka *C4.5* dapat dilakukan untuk atribut yang memiliki nilai kategorikal maupun nilai kontinyu (numerik).

**Tabel 3.1** Data latih

<b>NO</b>	<b>TD</b>	<b>KG</b>	<b>KT</b>	<b>LDL</b>	<b>U</b>	<b>JK</b>	<b>AU</b>	<b>BUN</b>	<b>KRE</b>	<b>DIAGNOSA</b>
1	110	90	112	87	31	P	5,1	21,5	0,9	Rendah
2	90	100	152	81	43	P	5,2	26	0,9	Rendah
3	110	111	125	82	42	P	5,1	16,7	0,7	Rendah
4	100	110	115	86	52	P	4	31	0,8	Rendah
5	100	113	160	92	49	P	5	10,7	0,8	Rendah
6	90	106	117	95	51	L	5,6	32	0,8	Rendah
7	90	107	115	97	53	L	6	18	1	Rendah
8	100	96	147	77	54	P	4,2	20	1	Rendah
9	100	95	165	82	52	P	3,1	19,5	0,8	Rendah
10	90	97	117	86	49	P	3,1	27	0,7	Rendah
11	100	97	127	79	47	L	6,8	11,8	1	Rendah
12	100	107	140	77	45	L	6,4	25	0,9	Rendah
13	110	101	138	76	46	P	5,7	27	0,8	Rendah
14	100	107	105	82	30	P	4,8	20,4	0,9	Rendah
15	90	105	125	76	29	L	7,5	21,5	0,8	Rendah
16	90	105	125	74	29	L	6,1	25	0,8	Rendah
17	100	100	131	91	35	P	4,5	17	0,8	Rendah
18	110	115	125	90	35	L	5	12	0,9	Rendah
19	90	110	115	75	42	P	5	27	1	Rendah
20	110	113	135	75	40	L	4,8	11,8	1	Rendah
21	100	96	145	81	45	L	5,6	9,6	0,8	Rendah
22	100	95	156	81	37	L	5,7	36	0,9	Rendah
23	90	97	136	80	38	P	6	31,6	0,8	Rendah
24	100	101	135	81	46	P	3,7	21	0,8	Rendah
25	100	100	171	87	46	L	5	15,6	0,9	Rendah
26	110	125	115	90	52	P	5	25,7	0,5	Rendah
27	110	155	195	100	51	L	6,2	20	0,7	Rendah
28	90	108	187	123	43	L	6,7	25,8	0,6	Rendah
29	100	104	193	122	44	P	5,4	17,8	0,5	Rendah
30	110	109	139	127	55	L	6,8	20,7	0,8	Rendah
31	150	115	122	118	52	P	2	7,5	0,7	Sedang
32	140	132	167	115	75	P	4,8	39,2	1,5	Sedang
33	150	155	195	127	66	P	6,2	30,6	1	Sedang
34	135	175	219	162	48	P	6,2	19,8	1,1	Sedang
35	145	110	153	135	58	L	6,5	20	1,5	Sedang
36	150	87	167	107	50	L	6,7	20	1,5	Sedang
37	140	190	195	127	30	L	6,2	20	1,5	Sedang
38	150	186	260	168	55	P	3,1	6	1	Sedang
39	150	155	195	127	60	L	6,2	20	1,5	Sedang
40	160	110	194	108	60	L	10	20,4	1,5	Sedang
41	140	196	138	88	56	P	3,4	21,1	1,4	Sedang

**Tabel 3.1** Data latih (lanjutan)

<b>NO</b>	<b>TD</b>	<b>KG</b>	<b>KT</b>	<b>LDL</b>	<b>U</b>	<b>JK</b>	<b>AU</b>	<b>BUN</b>	<b>KRE</b>	<b>DIAGNOSA</b>
42	150	111	149	190	48	P	4,6	10,7	0,8	Sedang
43	145	188	208	170	42	L	6,4	14,3	1,3	Sedang
44	150	117	111	52	73	L	7,3	31,6	2,1	Sedang
45	140	136	143	120	71	L	8,9	23,4	1,9	Sedang
46	150	240	264	163	52	L	6,9	15,3	2,5	Sedang
47	150	181	186	137	57	P	3,5	6,2	0,5	Sedang
48	160	119	150	108	60	P	4,9	7,4	1	Sedang
49	130	140	221	148	52	L	8,4	11,1	0,8	Sedang
50	160	105	157	102	56	P	4,5	11,6	0,7	Sedang
51	130	95	210	153	62	L	5,2	7,7	1	Sedang
52	150	109	137	167	65	L	8,8	38,7	3,6	Sedang
53	145	196	200	160	62	P	2,8	13,5	0,6	Sedang
54	135	160	185	183	63	P	7,5	10,7	0,8	Sedang
55	140	191	176	86	66	P	6	20,4	1,2	Sedang
56	150	102	137	90	57	L	5,4	11,3	6,9	Sedang
57	170	111	207	168	80	L	7,9	31,7	2	Sedang
58	140	200	122	80	62	L	7,7	25,5	1,3	Sedang
59	150	128	237	154	50	P	4,5	12,4	0,7	Sedang
60	140	92	161	102	50	P	4	12,8	1	Sedang
61	200	167	263	158	78	L	8	15,8	1,4	Tinggi
62	180	421	277	165	56	L	4,7	14,9	1,5	Tinggi
63	200	136	294	195	60	P	4,8	16,2	0,8	Tinggi
64	170	150	190	130	65	L	6,2	20	1,5	Tinggi
65	190	127	133	165	65	P	3,8	18,2	0,5	Tinggi
66	180	136	177	136	65	L	7,3	13,5	1,2	Tinggi
67	170	112	189	150	67	P	5,2	11,1	0,8	Tinggi
68	200	174	187	125	65	L	6,2	12,9	0,9	Tinggi
69	200	186	195	127	64	P	6,2	13,6	0,9	Tinggi
70	180	247	189	130	55	P	6,2	20,6	1,1	Tinggi
71	210	204	229	162	49	P	7,2	23,1	0,7	Tinggi
72	170	146	208	109	61	P	5	8,6	1	Tinggi
73	200	305	232	167	65	L	12	25,8	2,4	Tinggi
74	190	119	195	127	63	L	6,2	13,2	0,7	Tinggi
75	180	146	229	162	45	P	8	5,4	0,8	Tinggi
76	180	117	191	117	66	P	4	6,1	0,8	Tinggi
77	190	156	154	125	55	L	7,8	14,6	1,3	Tinggi
78	190	260	172	128	69	P	3,4	10	0,5	Tinggi
79	170	192	209	150	59	L	8,2	34,7	1,9	Tinggi
80	190	100	190	124	62	P	6	14,8	1,2	Tinggi
81	200	243	282	182	54	P	6,2	11,5	0,7	Tinggi
82	200	138	267	188	66	P	6,4	13	0,9	Tinggi
83	170	107	138	196	65	L	6,8	9,5	1,1	Tinggi
84	180	107	215	150	65	P	5,4	14,6	0,9	Tinggi

**Tabel 3.1** Data latih (lanjutan)

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
85	180	206	270	184	70	P	3,8	19,5	1,3	Tinggi
86	190	120	197	117	65	P	7,8	13,3	0,9	Tinggi
87	200	253	249	187	49	P	5,2	14,1	1	Tinggi
88	200	169	193	128	65	L	6,2	19,9	1,6	Tinggi
89	170	130	207	160	52	P	5,7	10,4	1,2	Tinggi
90	180	328	269	108	50	P	8,6	9,5	0,6	Tinggi

**Tabel 3.2** Data uji

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
1	110	110	162	89	45	L	6	27	1,1	Rendah
2	106	121	161	56	49	P	4,7	29	0,8	Rendah
3	100	129	141	78	43	L	5,1	17,5	0,7	Rendah
4	100	155	173	133	56	L	7,1	31,6	0,9	Rendah
5	110	110	195	127	45	L	6,2	27,4	0,8	Rendah
6	90	100	125	121	40	P	5,1	21,5	0,6	Rendah
7	110	106	190	130	45	P	4,4	20,7	0,9	Rendah
8	100	100	195	127	28	L	6,5	16,2	0,6	Rendah
9	90	163	132	75	41	P	4,5	25	0,7	Rendah
10	110	115	170	91	32	P	3,5	9,3	1	Rendah
11	110	120	162	96	37	P	3	17,9	1	Rendah
12	100	109	137	92	36	L	8	38	3,6	Rendah
13	110	115	229	173	50	L	5,5	12,3	1,1	Rendah
14	90	92	221	164	48	P	4,1	16,9	0,6	Rendah
15	100	135	113	63	50	L	6,1	27,1	0,8	Rendah
16	100	102	137	90	47	L	5,4	11,3	0,7	Rendah
17	110	111	117	52	43	L	7,3	31,6	0,9	Rendah
18	100	109	137	92	60	L	6,5	25,7	0,6	Rendah
19	145	135	113	63	80	L	8,1	39,4	2,9	Sedang
20	145	115	229	173	71	L	5,5	12,3	1,1	Sedang
21	150	117	198	113	56	L	9	9,8	1,1	Sedang
22	150	98	147	95	60	L	2,8	21,5	1,5	Sedang
23	140	108	222	146	40	P	3,3	11,6	0,5	Sedang
24	140	78	243	182	59	P	4,9	11,7	0,6	Sedang
25	150	91	162	125	55	P	4,8	20,2	1	Sedang
26	140	98	176	106	52	L	10	14,4	1,1	Sedang
27	150	220	168	106	47	P	3	7,4	0,5	Sedang
28	145	104	192	112	52	L	8,2	15,1	1,1	Sedang
29	150	116	159	92	55	P	4,2	12	0,8	Sedang
30	140	128	155	102	42	L	8	10,5	0,9	Sedang
31	180	257	195	140	56	P	5	14,7	0,7	Tinggi



**Tabel 3.2** Data Uji (lanjutan)

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
32	180	140	194	134	69	L	8,6	28,7	1,7	Tinggi
33	190	173	147	147	70	L	6,6	21,5	0,9	Tinggi
34	170	86	214	106	80	L	7,8	9,6	0,8	Tinggi
35	180	82	130	185	72	P	4,7	12,7	0,5	Tinggi
36	170	152	195	127	55	L	6,2	14,9	1,2	Tinggi
37	170	223	375	205	60	P	6,2	79,2	3,8	Tinggi
38	190	231	183	134	76	L	9,8	17,9	1,2	Tinggi

Sedangkan atribut yang digunakan dalam penelitian ini adalah sebagai

berikut:

**Tabel 3.3** Data Atribut

No	Atribut	Keterangan	Tipe
1	TD	Tekanan darah pasien	Numerik
2	KG	Kadar gula darah pasien	Numerik
3	KT	Kolestrol total pasien	Numerik
4	LDL	Kolesterol <i>Low Density Lipoprotein</i>	Numerik
5	U	Usia	Numerik
6	JK	Jenis kelamin	Kategorikal
7	AU	Asam urat	Numerik
8	BUN	<i>Blood Urea Nitrogen (BUN)</i>	Numerik
9	KRE	Kandungan kreatinin pasien	Numerik

Perhitungan *decision tree C4.5* ini akan menggunakan data pada **Tabel 3.1** (*data training*). Tabel tersebut akan diubah menjadi sebuah pohon keputusan. Sebelum melakukan perhitungan, berikut akan dijelaskan ketentuan dalam pembentukan pohon keputusan pada kasus ini:

a) Pemecahan cabang dilakukan secara biner yaitu pemecahan yang hanya mempunyai dua nilai yakni  $\leq$  dan  $>$  (kurang dari sama dengan dan lebih dari)

b) Tekanan darah (TD)

Posisi  $v$  yang digunakan pada atribut TD adalah nilai antara (90, 105, 120, 135, 150, 165, 180, 195)

## c) Kadar gula darah (KG)

Posisi v yang digunakan pada atribut KG adalah nilai antara (90, 110, 130, 150, 170, 190, 210, 230, 250)

## d) Kolesterol total (KT)

Posisi v yang digunakan pada atribut KG adalah nilai antara (115, 130, 145, 160, 175, 190, 205, 220, 235)

e) Kolesterol *Low Density Lipoprotein* (LDL)

Posisi v yang digunakan pada atribut LDL adalah nilai antara (80, 90, 100, 110, 120, 130, 140, 150, 160)

## f) Usia (U)

Posisi v yang digunakan pada atribut U adalah nilai antara (30, 35, 40, 45, 50, 55, 60, 65)

## g) Jenis kelamin (JK)

Jenis kelamin adalah atribut yang bernilai kategorikal sehingga tidak perlu ditentukan nilai posisi v.

## h) Asam urat (AU)

Posisi v yang digunakan pada atribut AU adalah nilai antara (3, 4, 5, 6, 7, 8)

## i) Blood urea nitrogen (BUN)

Posisi v yang digunakan pada atribut BUN adalah nilai antara (10, 15, 20, 25, 30)

## j) Kreatinin (KRE)

Posisi v yang digunakan pada atribut KRE adalah nilai antara ((0,5), (0,6), (0,7), (0,8), (0,9), (1), (1,1), (1,2), (1,3))

## k) Kelas

Pada variabel diagnosa yang kelas klasifikasi yang digunakan adalah Rendah, Sedang dan Tinggi

Langkah pertama adalah memilih atribut yang akan dijadikan akar (*root node*) dengan menghitung nilai gain yang paling tinggi. Sebelumnya yang akan dihitung adalah nilai *entropy* semua data. Berikut adalah perhitungan *entropy* semua data.

$$\begin{aligned} Entropy(S) &= -\frac{30}{90} * \log_2\left(\frac{30}{90}\right) - \frac{30}{90} * \log_2\left(\frac{30}{90}\right) - \frac{30}{90} * \log_2\left(\frac{30}{90}\right) \\ &= 0,52832 + 0,52832 + 0,52832 = 1.58496 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap atribut, jika atribut memiliki nilai numerik maka akan dilakukan perhitungan untuk menentukan posisi  $v$ . Berikut adalah perhitungan untuk menentukan posisi *gain* tertinggi untuk masing-masing posisi.

**Tabel 3.4** adalah hasil penentuan posisi  $v$  pada atribut tekanan darah, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 120$ . Maka untuk atribut tekanan darah dilakukan diskretisasi pada  $v = 120$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.5** adalah hasil penentuan posisi  $v$  pada atribut kadar gula, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 110$ . Maka untuk atribut kadar gula dilakukan diskretisasi pada  $v = 110$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.6** adalah hasil penentuan posisi  $v$  pada atribut kolestrol total, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 175$ . Maka untuk atribut kadar gula dilakukan diskretisasi pada  $v = 175$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.7** adalah hasil penentuan posisi  $v$  pada atribut LDL (*Low Density Lippoprotein*), pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 100$ . Maka untuk atribut LDL (*Low Density Lippoprotein*) dilakukan diskretisasi pada  $v = 100$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.4** Posisi v untuk Pemecahan Atribut Tekanan Darah

TD	90		105		120		135		150		165		180		195	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	9	21	22	8	30	0	30	0	30	0	30	0	30	0	30	0
Sedang	0	30	0	30	0	30	4	26	26	4	29	1	30	0	30	0
Tinggi	0	30	0	30	0	30	0	30	0	30	0	30	14	16	20	10
Jumlah	9	81	22	68	30	60	34	56	56	34	59	31	74	16	80	10
Entropy	<b>0</b>	<b>1,5664</b>	<b>0</b>	<b>1,4049</b>	<b>0</b>	<b>1</b>	<b>0,5226</b>	<b>0,9963</b>	<b>0,9963</b>	<b>0,5226</b>	<b>0,9998</b>	<b>0,2056</b>	<b>1,5106</b>	<b>0</b>	<b>1,5613</b>	<b>0</b>
Gain	<b>0.1752</b>		<b>0.5235</b>		<b>0.9183</b>		<b>0,7676</b>		<b>0.7676</b>		<b>0.8587</b>		<b>0.3429</b>		<b>0.1971</b>	

**Tabel 3.5** Posisi v untuk Pemecahan Atribut Kadar Gula

KG	90		110		130		150		170		190		210		230		250	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	1	29	24	6	29	1	29	1	30	0	30	0	30	0	30	0	30	0
Sedang	1	29	8	22	14	16	17	13	20	10	25	5	29	1	29	1	30	0
Tinggi	0	30	3	27	9	21	15	15	18	12	20	10	23	7	23	7	25	5
Jumlah	2	88	35	55	52	38	61	29	68	22	75	15	82	8	82	8	85	5
Entropy	<b>1</b>	<b>1,5848</b>	<b>1,1637</b>	<b>1,3814</b>	<b>1,4175</b>	<b>1,1364</b>	<b>1,5214</b>	<b>1,1784</b>	<b>1,5477</b>	<b>0,994</b>	<b>1,5656</b>	<b>0,9183</b>	<b>1,5755</b>	<b>0,5436</b>	<b>1,5755</b>	<b>0,5436</b>	<b>1,5799</b>	<b>0</b>
Gain	<b>0.0132</b>		<b>0.2882</b>		<b>0.2861</b>		<b>0.1741</b>		<b>0.1728</b>		<b>0.1272</b>		<b>0.2861</b>		<b>0.1012</b>		<b>0.0928</b>	

**Tabel 3.6** Posisi v untuk Pemecahan Atribut Kolestrol Total

KT	115		130		145		160		175		190		205		220		235	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	6	24	13	17	21	9	25	5	27	3	28	2	30	0	30	0	30	0
Sedang	1	29	3	27	7	23	11	19	14	16	17	13	22	8	26	4	27	3
Tinggi	0	30	0	30	2	28	3	27	4	26	10	20	15	15	19	11	22	8
Jumlah	7	83	16	74	30	60	39	51	45	45	55	35	67	23	75	15	79	11
Entropy	<b>0,5917</b>	<b>1,5783</b>	<b>0,6962</b>	<b>1,5463</b>	<b>1.1106</b>	<b>1,4539</b>	<b>1.2109</b>	<b>1,3449</b>	<b>1.2766</b>	<b>1,2482</b>	<b>1.4666</b>	<b>1,228</b>	<b>1.53</b>	<b>0,9321</b>	<b>1.5604</b>	<b>0,8366</b>	<b>1.5734</b>	<b>0,8454</b>
Gain	<b>0.0834</b>		<b>0.1898</b>		<b>0.2981</b>		<b>0,2981</b>		<b>0.3226</b>		<b>0.2112</b>		<b>0.2078</b>		<b>0.1452</b>		<b>0.1005</b>	

**Tabel 3.7** Posisi v untuk Pemecahan Atribut LDL

LDL	80		90		100		110		120		130		140		150		160	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	9	21	22	8	27	3	27	3	27	3	30	0	30	0	30	0	30	0
Sedang	2	28	5	25	5	25	10	20	13	17	16	14	18	12	19	11	22	8
Tinggi	0	30	0	30	0	30	2	28	4	26	13	17	14	16	17	13	19	11
Jumlah	11	79	27	63	32	58	39	51	44	46	59	31	62	28	66	224	71	19
Entropy	<b>0,684</b>	<b>1,569</b>	<b>0,6913</b>	<b>1,4169</b>	<b>0,6253</b>	<b>1,2363</b>	<b>1,0905</b>	<b>1,245</b>	<b>1,2665</b>	<b>1,2528</b>	<b>1,4875</b>	<b>0,9932</b>	<b>1,5095</b>	<b>0,9852</b>	<b>1,5383</b>	<b>0,995</b>	<b>1,5578</b>	<b>0,9819</b>
Gain	<b>0,1241</b>		<b>0,3857</b>		<b>0,5659</b>		<b>0,4069</b>		<b>0,3255</b>		<b>0,2677</b>		<b>0,2386</b>		<b>0,1915</b>		<b>0,1487</b>	

**Tabel 3.8** adalah hasil penentuan posisi  $v$  pada atribut usia, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 55$ . Maka untuk atribut usia dilakukan diskretisasi pada  $v = 55$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.9** adalah hasil penentuan posisi  $v$  pada atribut asam urat, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 6$ . Maka untuk atribut asam urat dilakukan diskretisasi pada  $v = 6$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.10** adalah hasil penentuan posisi  $v$  pada atribut Blood Urea Nitrogen (BUN), pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 15$ . Maka untuk atribut BUN dilakukan diskretisasi pada  $v = 15$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.11** adalah hasil penentuan posisi  $v$  pada atribut kreatinin, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 1$ . Maka untuk atribut kreatinin dilakukan diskretisasi pada  $v = 1$  ketika menghitung *entropy* dan *gain* pada semua atribut.

Selanjutnya menghitung *entropy* dan *gain* untuk atribut jenis kelamin, perhitungan untuk atribut ini tidak perlu menggunakan posisi  $v$  karena atribut memiliki nilai kategorikal. Berikut adalah perhitungan untuk mencari *gain* jenis kelamin.

$$\begin{aligned} \text{Gain} &= 1.5849 - \left( \frac{49}{90} * \left( -\frac{30}{49} * \log_2 \left( \frac{30}{49} \right) - \frac{11}{49} * \log_2 \left( \frac{11}{49} \right) - \frac{8}{49} * \log_2 \left( \frac{8}{49} \right) \right) \right) \\ &\quad + \frac{41}{90} * \left( -\frac{0}{41} * \log_2 \left( \frac{0}{41} \right) - \frac{19}{41} * \log_2 \left( \frac{19}{41} \right) - \frac{22}{41} * \log_2 \left( \frac{22}{41} \right) \right) = 0,0095 \end{aligned}$$

Setelah melakukan perhitungan *gain* untuk atribut jenis kelamin maka dilanjutkan untuk menghitung nilai  $v$  untuk masing-masing atribut selanjutnya. Sehingga diperoleh perhitungan pada **Tabel 3.8** sampai **Tabel 3.11**.

**Tabel 3.8** Posisi v untuk Pemecahan Atribut Usia

U	30		35		40		45		50		55		60		65	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	3	27	6	24	9	21	16	14	22	8	30	0	30	0	30	0
Sedang	1	29	1	29	1	29	2	28	7	23	11	19	19	11	24	6
Tinggi	0	30	0	30	0	30	1	29	4	26	8	22	11	19	24	6
Jumlah	4	86	7	83	10	80	19	71	33	57	49	41	60	30	78	12
Entropy	<b>0,8113</b>	<b>1,5836</b>	<b>0,5917</b>	<b>1,5783</b>	<b>0,469</b>	<b>1,5678</b>	<b>0,7742</b>	<b>1,5189</b>	<b>1,2335</b>	<b>1,4425</b>	<b>1,3441</b>	<b>0,9961</b>	<b>1,474</b>	<b>0,9481</b>	<b>1,5766</b>	<b>1</b>
Gain	<b>0,0357</b>		<b>0,0834</b>		<b>0,1393</b>		<b>0,2233</b>		<b>0,2191</b>		<b>0,3994</b>		<b>0,2863</b>		<b>0,0852</b>	

**Tabel 3.9** Posisi v untuk Pemecahan Atribut Asam Urat

AU	3		4		5		6		7		8	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	0	30	4	26	13	17	23	7	29	1	30	0
Sedang	2	28	6	24	11	19	14	16	22	8	26	4
Tinggi	0	30	4	26	7	23	12	18	21	9	27	3
Jumlah	2	88	14	76	31	59	49	41	72	18	83	7
Entropy	<b>0</b>	<b>1,5842</b>	<b>1,5567</b>	<b>1,584</b>	<b>1,5409</b>	<b>1,5735</b>	<b>1,5256</b>	<b>1,4866</b>	<b>1,5695</b>	<b>1,2516</b>	<b>1,5823</b>	<b>0,9852</b>
Gain	<b>0,036</b>		<b>0,0052</b>		<b>0,0227</b>		<b>0,0771</b>		<b>0,079</b>		<b>0,0491</b>	

**Tabel 3.10** Posisi v untuk Pemecahan Atribut BUN

BUN	10		15		20		25		30	
	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	1	29	5	25	13	17	20	10	26	4
Sedang	5	25	14	16	20	10	24	6	25	5
Tinggi	6	24	20	10	26	4	28	2	29	1
Jumlah	12	78	39	51	59	31	72	18	80	10
Entropy	<b>1,325</b>	<b>1,5801</b>	<b>1,4046</b>	<b>1,4898</b>	<b>1,5309</b>	<b>1,383</b>	<b>1,5715</b>	<b>1,3516</b>	<b>1,5821</b>	<b>1,361</b>
Gain	<b>0,036</b>		<b>0,1321</b>		<b>0,105</b>		<b>0,0574</b>		<b>0,0274</b>	

**Tabel 3.11** Posisi v untuk Pemecahan Atribut Kreatinin

KRE	0,5		0,6		0,7		0,8		0,9		1		1,1		1,2		1,3	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>	<=	>
Rendah	2	28	3	27	6	24	18	12	25	5	30	0	30	0	30	0	30	0
Sedang	1	29	2	28	5	25	8	22	8	22	13	17	14	16	15	15	17	13
Tinggi	2	28	3	27	6	24	10	20	15	15	17	13	19	11	22	8	24	6
Jumlah	5	85	8	82	17	73	36	54	48	42	60	30	63	27	67	23	71	19
Entropy	<b>1,5219</b>	<b>1,5848</b>	<b>1,5613</b>	<b>1,5847</b>	<b>1,5799</b>	<b>1,5847</b>	<b>1,4955</b>	<b>1,5407</b>	<b>1,4454</b>	<b>1,3847</b>	<b>1,4936</b>	<b>0,9871</b>	<b>1,5135</b>	<b>0,9751</b>	<b>1,53</b>	<b>0,9321</b>	<b>1,5479</b>	<b>0,8997</b>
Gain	<b>0,0037</b>		<b>0,0023</b>		<b>0,0012</b>		<b>0,0623</b>		<b>0,1679</b>		<b>0,2602</b>		<b>0,233</b>		<b>0,2078</b>		<b>0,1739</b>	



Setelah semua atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula, pada tabel keterangan warna kuning adalah nilai *gain* terbesar untuk atribut yang terpilih. Hasilnya disajikan pada **Tabel 3.12**.

**Tabel 3.12** Hasil Perhitungan *Entropy* dan *Gain* Untuk Node Akar

Perhitungan entropy dan gain		NODE 1					
Total		jml	Rendah	Sedang	Tinggi	Entropy	Gain
		90	30	30	30	1,584963	
Tekanan Darah	$\leq 120$	30	30	0	0	<b>0</b>	<b>0,9183</b>
	$> 120$	60	0	30	30	1	
Kadar Gula	$\leq 110$	35	24	8	3	1,1637	0,2882
	$> 110$	55	6	22	27	1,3814	
Kolestrol Total	$\leq 175$	45	27	14	4	1,2766	0,3226
	$> 175$	45	3	16	26	1,2482	
LDL	$\leq 100$	32	27	5	0	0,6253	0,5659
	$> 100$	58	3	25	30	1,2363	
Usia	$\leq 55$	49	30	11	8	1,3441	0,3994
	$> 55$	41	0	19	22	0,9961	
Jenis Kelamin	L	40	14	15	11	1,5729	0,0095
	P	50	16	15	19	1,5776	
Asam Urat	$\leq 7$	72	29	22	21	1,5695	0,0790
	$> 7$	18	1	8	9	1,2516	
BUN	$\leq 15$	39	5	14	20	1,4046	0,1321
	$> 15$	51	25	16	10	1,4898	
Kreatinin	$\leq 1$	60	30	13	17	1,4936	0,2602
	$> 1$	30	0	17	13	0,9871	

Hasil yang didapat di **Tabel 3.12** menunjukkan bahwa *gain* tertinggi ada di atribut Tekanan Darah, maka Tekanan Darah dijadikan sebagai *node* akar. Pada atribut Tekanan Darah *entropy* yang dihasilkan pada nilai  $\leq 120$  memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki tekanan darah  $\leq 120$  masuk dalam kelas diagnosa Rendah. Sehingga untuk pemangkasan data yang akan digunakan pada perhitungan selanjutnya hanya memiliki dua kelas saja yaitu Sedang dan Tinggi karena data latih dengan kelas Rendah sudah habis (*homogen*). Pohon yang terbentuk untuk proses pada node 1 adalah sebagai berikut.



**Gambar 3.3** Pohon Keputusan yang Terbentuk pada Node 1

**Tabel 3.13** Data Kasus pada Tekanan Darah  $\leq 120$

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
1	110	90	112	87	31	P	5,1	21,5	0,9	Rendah
2	90	100	152	81	43	P	5,2	26	0,9	Rendah
3	110	111	125	82	42	P	5,1	16,7	0,7	Rendah
4	100	110	115	86	52	P	4	31	0,8	Rendah
5	100	113	160	92	49	P	5	10,7	0,8	Rendah
6	90	106	117	95	51	L	5,6	32	0,8	Rendah
7	90	107	115	97	53	L	6	18	1	Rendah
8	100	96	147	77	54	P	4,2	20	1	Rendah
9	100	95	165	82	52	P	3,1	19,5	0,8	Rendah
10	90	97	117	86	49	P	3,1	27	0,7	Rendah
11	100	97	127	79	47	L	6,8	11,8	1	Rendah
12	100	107	140	77	45	L	6,4	25	0,9	Rendah
13	110	101	138	76	46	P	5,7	27	0,8	Rendah
14	100	107	105	82	30	P	4,8	20,4	0,9	Rendah
15	90	105	125	76	29	L	7,5	21,5	0,8	Rendah
16	90	105	125	74	29	L	6,1	25	0,8	Rendah
17	100	100	131	91	35	P	4,5	17	0,8	Rendah
18	110	115	125	90	35	L	5	12	0,9	Rendah
19	90	110	115	75	42	P	5	27	1	Rendah
20	110	113	135	75	40	L	4,8	11,8	1	Rendah
21	100	96	145	81	45	L	5,6	9,6	0,8	Rendah
22	100	95	156	81	37	L	5,7	36	0,9	Rendah
23	90	97	136	80	38	P	6	31,6	0,8	Rendah
24	100	101	135	81	46	P	3,7	21	0,8	Rendah
25	100	100	171	87	46	L	5	15,6	0,9	Rendah
26	110	125	115	90	52	P	5	25,7	0,5	Rendah
27	110	155	195	100	51	L	6,2	20	0,7	Rendah
28	90	108	187	123	43	L	6,7	25,8	0,6	Rendah
29	100	104	193	122	44	P	5,4	17,8	0,5	Rendah
30	110	109	139	127	55	L	6,8	20,7	0,8	Rendah

**Tabel 3.14** Data Kasus pada Tekanan Darah >120

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
31	150	115	122	118	52	P	2	7.5	0.7	Sedang
32	140	132	167	115	75	P	4.8	39.2	1.5	Sedang
33	150	155	195	127	66	P	6.2	30.6	1	Sedang
34	135	175	219	162	48	P	6.2	19.8	1.1	Sedang
35	145	110	153	135	58	L	6.5	20	1.5	Sedang
36	150	87	167	107	50	L	6.7	20	1.5	Sedang
37	140	190	195	127	30	L	6.2	20	1.5	Sedang
38	150	186	260	168	55	P	3.1	6	1	Sedang
39	150	155	195	127	60	L	6.2	20	1.5	Sedang
40	160	110	194	108	60	L	10	20.4	1.5	Sedang
41	140	196	138	88	56	P	3.4	21.1	1.4	Sedang
42	150	111	149	190	48	P	4.6	10.7	0.8	Sedang
43	145	188	208	170	42	L	6.4	14.3	1.3	Sedang
44	150	117	111	52	73	L	7.3	31.6	2.1	Sedang
45	140	136	143	120	71	L	8.9	23.4	1.9	Sedang
46	150	240	264	163	52	L	6.9	15.3	2.5	Sedang
47	150	181	186	137	57	P	3.5	6.2	0.5	Sedang
48	160	119	150	108	60	P	4.9	7.4	1	Sedang
49	130	140	221	148	52	L	8.4	11.1	0.8	Sedang
50	160	105	157	102	56	P	4.5	11.6	0.7	Sedang
51	130	95	210	153	62	L	5.2	7.7	1	Sedang
52	150	109	137	167	65	L	8.8	38.7	3.6	Sedang
53	145	196	200	160	62	P	2.8	13.5	0.6	Sedang
54	135	160	185	183	63	P	7.5	10.7	0.8	Sedang
55	140	191	176	86	66	P	6	20.4	1.2	Sedang
56	150	102	137	90	57	L	5.4	11.3	6.9	Sedang
57	170	111	207	168	80	L	7.9	31.7	2	Sedang
58	140	200	122	80	62	L	7.7	25.5	1.3	Sedang
59	150	128	237	154	50	P	4.5	12.4	0.7	Sedang
60	140	92	161	102	50	P	4	12.8	1	Sedang
61	200	167	263	158	78	L	8	15.8	1.4	Tinggi
62	180	421	277	165	56	L	4.7	14.9	1.5	Tinggi
63	200	136	294	195	60	P	4.8	16.2	0.8	Tinggi
64	170	150	190	130	65	L	6.2	20	1.5	Tinggi
65	190	127	133	165	65	P	3.8	18.2	0.5	Tinggi
66	180	136	177	136	65	L	7.3	13.5	1.2	Tinggi
67	170	112	189	150	67	P	5.2	11.1	0.8	Tinggi
68	200	174	187	125	65	L	6.2	12.9	0.9	Tinggi
69	200	186	195	127	64	P	6.2	13.6	0.9	Tinggi

**Tabel 3.14** Data Kasus pada Tekanan Darah >120 (lanjutan)

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA
70	180	247	189	130	55	P	6.2	20.6	1.1	Tinggi
71	210	204	229	162	49	P	7.2	23.1	0.7	Tinggi
72	170	146	208	109	61	P	5	8.6	1	Tinggi
73	200	305	232	167	65	L	12.2	25.8	2.4	Tinggi
74	190	119	195	127	63	L	6.2	13.2	0.7	Tinggi
75	180	146	229	162	45	P	8	5.4	0.8	Tinggi
76	180	117	191	117	66	P	4	6.1	0.8	Tinggi
77	190	156	154	125	55	L	7.8	14.6	1.3	Tinggi
78	190	260	172	128	69	P	3.4	10	0.5	Tinggi
79	170	192	209	150	59	L	8.2	34.7	1.9	Tinggi
80	190	100	190	124	62	P	6	14.8	1.2	Tinggi
81	200	243	282	182	54	P	6.2	11.5	0.7	Tinggi
82	200	138	267	188	66	P	6.4	13	0.9	Tinggi
83	170	107	138	196	65	L	6.8	9.5	1.1	Tinggi
84	180	107	215	150	65	P	5.4	14.6	0.9	Tinggi
85	180	206	270	184	70	P	3.8	19.5	1.3	Tinggi
86	190	120	197	117	65	P	7.8	13.3	0.9	Tinggi
87	200	253	249	187	49	P	5.2	14.1	1	Tinggi
88	200	169	193	128	65	L	6.2	19.9	1.6	Tinggi
89	170	130	207	160	52	P	5.7	10.4	1.2	Tinggi
90	180	328	269	108	50	P	8.6	9.5	0.6	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Tekanan Darah yang berada >120 saja. Hasil pemecahan tersebut menyisakan 60 data yang terdiri dari 30 data dengan kelas diagnosa Sedang dan 30 data dengan kelas diagnosa Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk *entropy* keseluruhan

$$\begin{aligned}
 Entropy(S) &= -\frac{30}{60} * \log_2 \left( \frac{30}{60} \right) - \frac{30}{60} * \log_2 \left( \frac{30}{9060} \right) \\
 &= 0,5 + 0,5 = 1
 \end{aligned}$$

Setelah mengetahui nilai *entropy* keseluruhan maka dilakukan perhitungan dengan menentukan posisi *v* untuk setiap atribut.

**Tabel 3.15** adalah hasil penentuan posisi *v* pada atribut tekanan darah, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain*

tertinggi didapat pada posisi  $v = 165$ . Maka untuk atribut tekanan darah dilakukan diskretisasi pada  $v = 165$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.16** adalah hasil penentuan posisi  $v$  pada atribut kadar gula, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 250$ . Maka untuk atribut kadar gula dilakukan diskretisasi pada  $v = 250$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.17** adalah hasil penentuan posisi  $v$  pada atribut kolestrol total, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 175$ . Maka untuk atribut kolestrol total dilakukan diskretisasi pada  $v = 175$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.18** adalah hasil penentuan posisi  $v$  pada atribut LDL, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 90$  dan  $v = 100$  sehingga dalam hal ini nilai  $v$  yang dipilih adalah yang paling kiri yaitu pada posisi  $v = 90$ . Maka untuk atribut LDL dilakukan diskretisasi pada  $v = 90$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.19** adalah hasil penentuan posisi  $v$  pada atribut usia, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 60$ . Maka untuk atribut usia dilakukan diskretisasi pada  $v = 60$  ketika menghitung *entropy* dan *gain* pada semua atribut.

Selanjutnya menghitung *entropy* dan *gain* untuk atribut jenis kelamin, perhitungan untuk atribut ini tidak perlu menggunakan posisi  $v$  karena atribut memiliki nilai kategorikal. Berikut adalah perhitungannya.

$$\begin{aligned} Gain = 1 - & \left( \frac{26}{60} * \left( -\frac{15}{26} * \log_2 \left( \frac{15}{26} \right) - \frac{11}{26} * \log_2 \left( \frac{11}{26} \right) \right) + \frac{34}{60} \right. \\ & \left. * \left( -\frac{15}{34} * \log_2 \left( \frac{15}{34} \right) - \frac{19}{34} * \log_2 \left( \frac{19}{34} \right) \right) \right) = 0,0131 \end{aligned}$$

**Tabel 3.15** Posisi v untuk Pemecahan Atribut Tekanan Darah

TD	120		135		150		165		180		195	
	<=	>	<=	<=	>	>	<=	>	<=	>	<=	>
Sedang	0	30	4	26	26	4	29	1	30	0	30	0
Tinggi	0	30	0	30	0	30	0	30	14	16	20	10
Jumlah	0	60	4	56	26	34	29	31	44	16	50	10
Entropy	<b>0</b>	<b>1</b>	<b>0</b>	<b>0,9963</b>	<b>0</b>	<b>0,5226</b>	<b>0</b>	<b>0,2056</b>	<b>0,9024</b>	<b>0</b>	<b>0,971</b>	<b>0</b>
Gain	<b>0</b>		<b>0,0701</b>		<b>0,7039</b>		<b>0,8938</b>		<b>0,3382</b>		<b>0,1908</b>	

**Tabel 3.16** Posisi v untuk Pemecahan Atribut Kadar Gula

KG	90		110		130		150		170		190		210		230		250	
	<=	>	<=	>	<=	>	<=	>	<=	>	<=	<=	>	>	<=	>	<=	>
Sedang	1	29	8	22	14	16	17	13	20	10	25	5	29	1	29	1	30	0
Tinggi	0	30	3	27	9	21	15	15	18	12	20	10	23	7	23	7	25	5
Jumlah	1	59	11	49	23	37	32	28	68	22	45	15	52	8	52	8	55	5
Entropy	<b>0</b>	<b>0,9998</b>	<b>0,8454</b>	<b>0,9925</b>	<b>0,9656</b>	<b>0,9868</b>	<b>0,9972</b>	<b>0,9963</b>	<b>1,5477</b>	<b>0,994</b>	<b>0,9911</b>	<b>0,9183</b>	<b>0,9904</b>	<b>0,5436</b>	<b>0,9904</b>	<b>0,5436</b>	<b>0,994</b>	<b>0</b>
Gain	<b>0,0169</b>		<b>0,0345</b>		<b>0,0213</b>		<b>0,0032</b>		<b>0,1728</b>		<b>0,0271</b>		<b>0,0692</b>		<b>0,0692</b>		<b>0,0888</b>	

**Tabel 3.17** Posisi v untuk Pemecahan Atribut Kolestrol Total

KT	115		130		145		160		175		190		205		220		235	
	<=	>	<=	<=	>	>	<=	>	<=	>	<=	<=	>	>	<=	>	<=	>
Sedang	1	29	3	27	7	23	11	19	14	16	17	13	22	8	26	4	27	3
Tinggi	0	30	0	30	2	28	3	27	4	26	10	20	15	15	19	11	22	8
Jumlah	7	59	3	57	9	51	14	46	18	42	27	33	37	23	45	15	49	11
Entropy	<b>0</b>	<b>0,9998</b>	<b>0</b>	<b>0,998</b>	<b>0,7642</b>	<b>0,9931</b>	<b>0,0796</b>	<b>0,9781</b>	<b>0,7642</b>	<b>0,9587</b>	<b>0,951</b>	<b>0,9673</b>	<b>0,974</b>	<b>0,9321</b>	<b>0,9825</b>	<b>0,8366</b>	<b>0,9925</b>	<b>0,8454</b>
Gain	<b>0,0169</b>		<b>0,0519</b>		<b>0,0412</b>		<b>0,0752</b>		<b>0,0997</b>		<b>0,04</b>		<b>0,0421</b>		<b>0,054</b>		<b>0,0345</b>	

**Tabel 3.18** Posisi v untuk Pemecahan Atribut LDL

LDL	80		90		100		110		120		130		140		150		160	
	<=	>	<=	<=	>	>	<=	>	<=	>	<=	>	<=	<=	>	>	<=	>
Sedang	2	28	5	25	5	25	10	20	13	17	16	14	18	12	19	11	22	8
Tinggi	0	30	0	30	0	30	2	28	4	26	13	17	14	16	17	13	19	11
Jumlah	2	58	5	55	5	55	12	48	17	43	29	31	32	28	36	224	41	19
Entropy	<b>0</b>	<b>0,9991</b>	<b>0</b>	<b>0,994</b>	<b>0</b>	<b>0,994</b>	<b>0,65</b>	<b>0,9799</b>	<b>0,7871</b>	<b>0,9682</b>	<b>0,9923</b>	<b>0,9932</b>	<b>0,9887</b>	<b>0,9852</b>	<b>0,9978</b>	<b>0,995</b>	<b>0,9961</b>	<b>0,9819</b>
Gain	<b>0,0342</b>		<b>0,0888</b>		<b>0,0888</b>		<b>0,0861</b>		<b>0,0831</b>		<b>0,0072</b>		<b>0,0129</b>		<b>0,0033</b>		<b>0,0084</b>	

**Tabel 3.19** Posisi v untuk Pemecahan Atribut Usia

U	30		35		40		45		50		55		60		65	
	<=	>	<=	<=	>	>	<=	>	<=	>	<=	>	<=	<=	>	>
Sedang	1	29	1	29	1	29	2	28	7	23	11	19	19	11	24	6
Tinggi	0	30	0	30	0	30	1	29	4	26	8	22	11	19	24	6
Jumlah	1	59	1	59	1	59	3	57	11	49	19	41	30	30	48	12
Entropy	<b>0</b>	<b>0,9998</b>	<b>0</b>	<b>0,9998</b>	<b>0</b>	<b>0,9998</b>	<b>0,9183</b>	<b>0,9998</b>	<b>0,9457</b>	<b>0,9973</b>	<b>0,9819</b>	<b>0,9961</b>	<b>0,9481</b>	<b>0,9481</b>	<b>1</b>	<b>1</b>
Gain	<b>0,0169</b>		<b>0,0169</b>		<b>0,0169</b>		<b>0,0043</b>		<b>0,0122</b>		<b>0,0084</b>		<b>0,0519</b>		<b>0</b>	

**Tabel 3.20** Posisi v untuk Pemecahan Atribut Asam Urat

AU	3		4		5		6		7		8	
	<=	>	<=	<=	<=	>	>	>	<=	>	<=	>
Sedang	2	28	6	24	11	19	14	16	22	8	26	4
Tinggi	0	30	4	26	7	23	12	18	21	9	27	3
Jumlah	2	58	10	50	18	42	26	34	43	17	53	7
Entropy	<b>0</b>	<b>0,9991</b>	<b>0,971</b>	<b>0,9988</b>	<b>0,9641</b>	<b>0,9934</b>	<b>0,9957</b>	<b>0,9975</b>	<b>0,9996</b>	<b>0,9975</b>	<b>0,9997</b>	<b>0,9852</b>
Gain	<b>0,0342</b>		<b>0,0058</b>		<b>0,0154</b>		<b>0,0033</b>		<b>0,079</b>		<b>0,002</b>	



**Tabel 3.21** Posisi v untuk Pemecahan Atribut BUN

BUN	10		15		20		25		30	
	<=	>	<=	<=	<=	>	>	>	<=	>
Sedang	5	25	14	16	20	10	24	6	25	5
Tinggi	6	24	20	10	26	4	28	2	29	1
Jumlah	11	49	34	26	46	14	52	8	54	6
Entropy	<b>0,994</b>	<b>0,9997</b>	<b>0,9774</b>	<b>0,9612</b>	<b>0,9877</b>	<b>0,8631</b>	<b>0,9957</b>	<b>0,8113</b>	<b>0,996</b>	<b>0,65</b>
Gain	<b>0,0013</b>		<b>0,0296</b>		<b>0,0414</b>		<b>0,0289</b>		<b>0,0386</b>	

**Tabel 3.22** Posisi v untuk Pemecahan Atribut Kreatinin

KRE	0,5		0,6		0,7		0,8		0,9		1		1,1		1,2		1,3	
	<=	>	<=	<=	<=	>	>	>	<=	>	<=	>	<=	>	<=	>	<=	>
Sedang	1	29	2	28	5	25	8	22	8	22	13	17	14	16	15	15	17	13
Tinggi	2	28	3	27	6	24	10	20	15	15	17	13	19	11	22	8	24	6
Jumlah	3	57	5	55	11	49	18	42	23	37	30	30	33	27	37	23	41	19
Entropy	<b>0,9183</b>	<b>0,9998</b>	<b>0,971</b>	<b>0,9998</b>	<b>0,994</b>	<b>0,9997</b>	<b>0,9911</b>	<b>0,9984</b>	<b>0,9321</b>	<b>0,974</b>	<b>0,9871</b>	<b>0,9871</b>	<b>0,9834</b>	<b>0,9751</b>	<b>0,974</b>	<b>0,9321</b>	<b>0,9789</b>	<b>0,8997</b>
Gain	<b>0,0043</b>		<b>0,0026</b>		<b>0,0013</b>		<b>0,0038</b>		<b>0,0421</b>		<b>0,0129</b>		<b>0,0203</b>		<b>0,0421</b>		<b>0,0462</b>	

**Tabel 3.20** adalah hasil penentuan posisi  $v$  pada atribut asam urat, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 7$ . Maka untuk atribut asam urat dilakukan diskretisasi pada  $v = 7$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.21** adalah hasil penentuan posisi  $v$  pada atribut BUN, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 20$ . Maka untuk atribut BUN dilakukan diskretisasi pada  $v = 20$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.22** adalah hasil penentuan posisi  $v$  pada atribut kreatinin, pada kolom yang diberi keterangan warna kuning menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 1,3$ . Maka untuk atribut kreatinin dilakukan diskretisasi pada  $v = 1,3$  ketika menghitung *entropy* dan *gain* pada semua atribut. Setelah melakukan semua perhitungan untuk masing-masing atribut maka akan dipilih *gain* tertinggi dari semua atribut. **Tabel 3.23** menunjukkan bahwa atribut tekanan darah memiliki nilai *gain* tertinggi (keterangan warna kuning) sehingga atribut tekanan darah dipilih sebagai *node* cabang.

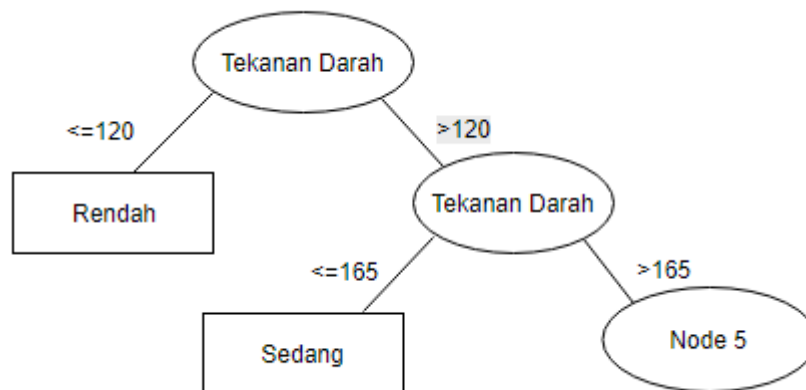
**Tabel 3.23** Hasil Perhitungan *Entropy* dan *Gain* untuk *Node* Akar 3

Perhitungan entropy dan gain			NODE 3			
Total		jml	Sedang	Tinggi	Entropy	Gain
		60	30	30	1	
Tekanan Darah	<=165	29	29	0	0	0.8938
	>165	31	1	30	0.2056	
Kadar Gula	<=250	55	30	25	0.994	0.0888
	>250	5	0	5	0	
Kolestrol Total	<=175	18	14	4	0.7642	0.0997
	>175	42	16	26	0.9587	
LDL	<=90	5	5	0	0	0.0888
	>90	55	25	30	0.994	
Usia	<=60	30	19	11	0.9481	0.0519
	>60	30	11	19	0.9481	
Jenis Kelamin	L	26	15	11	0.9829	0.0131
	P	34	15	19	0.9900	
Asam Urat	<=3	2	2	0	0	0.0342
	>3	58	28	30	0.9991	
BUN	<=20	46	20	26	0.9877	0.0414
	>20	14	10	4	0.8631	

**Tabel 3.23** Hasil Perhitungan *Entropy* dan *Gain* untuk *Node* Akar 3(lanjutan)

		jml	Sedang	Tinggi	Entropy	Gain
Kreatinin	$\leq 1,3$	41	17	24	0.9789	0.0462
	$> 1,3$	19	13	6	0.8997	

Dari perhitungan tersebut diperoleh atribut tekanan darah memiliki nilai *gain* lebih tinggi dibanding dengan atribut lainnya sehingga diperoleh pohon keputusan seperti berikut.

**Gambar 3.4** Pohon keputusan node 3

Perhitungan selanjutnya dilakukan pemangkasan data lagi sesuai dengan hasil pada **Tabel 3.22**, dari table tersebut menyisakan 31 data yang terdiri dari 1 data dengan diagnosa Sedang dan 30 data dengan diagnosa Tinggi. Data tersebut kemudian dilakukan proses perhitungan lagi sehingga diperoleh hasil seperti pada **Tabel 3.23**.

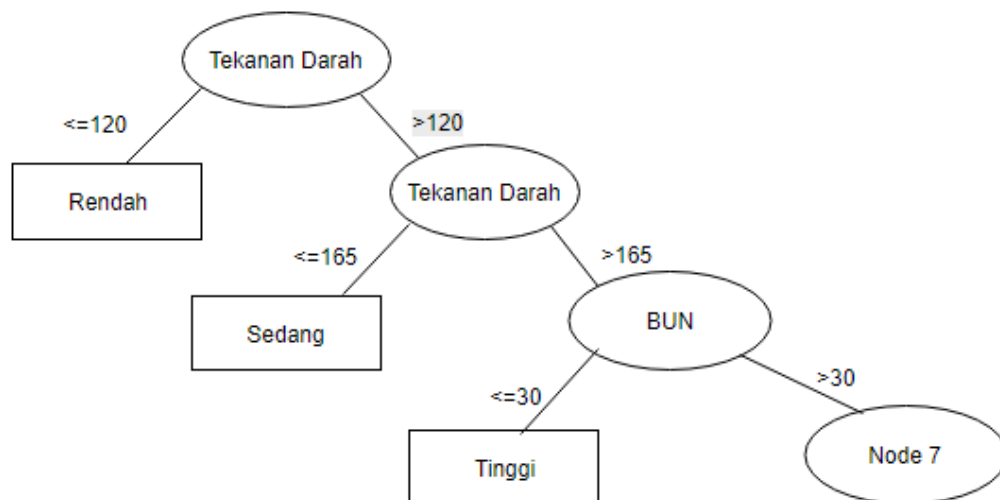
**Tabel 3.24** Hasil Perhitungan *Entropy* dan *Gain* untuk *Node* Akar 5

Perhitungan entropy dan gain		NODE 5				
		jml	Sedang	Tinggi	Entropy	Gain
<b>Total</b>		31	1	30	0.205593	
<b>Tekanan Darah</b>	$\leq 180$	15	1	14	0.3534	0.0346
	$> 180$	16	0	16	0	
<b>Kadar Gula</b>	$\leq 130$	10	1	9	0.469	0.0543
	$> 130$	21	0	21	0	
<b>Kolestrol Total</b>	$\leq 205$	15	0	15	0	0.0315
	$> 205$	16	1	15	0.3373	
<b>LDL</b>	$\leq 160$	19	0	19	0	0.0454
	$> 160$	12	1	11	0.4138	

**Tabel 3.24** Hasil Perhitungan *Entropy* dan *Gain* untuk Node Akar 5 (lanjutan)

		jml	Sedang	Tinggi	Entropy	Gain
Usia	<=65	24	0	24	0	0.0720
	>65	7	1	6	0.5917	
Jenis Kelamin	L	12	1	11	0.4138	0.0454
	P	19	0	19	0.0000	
Asam Urat	<=7	21	0	21	0	0.0543
	>7	10	1	9	0.469	
BUN	<=30	29	0	29	0	0.1411
	>30	2	1	1	1	
Kreatinin	<=1,3	24	0	24	0	0.0720
	>1,3	7	1	6	0.5917	

Dari **Tabel 2.24** menunjukkan atribut BUN memiliki nilai gain paling tinggi dibanding dengan atribut lainnya (keterangan warna kuning) sehingga atribut tersebut dipilih sebagai *node* cabang. Dari hasil tersebut dapat diperoleh pohon keputusan seperti pada gambar 3.5

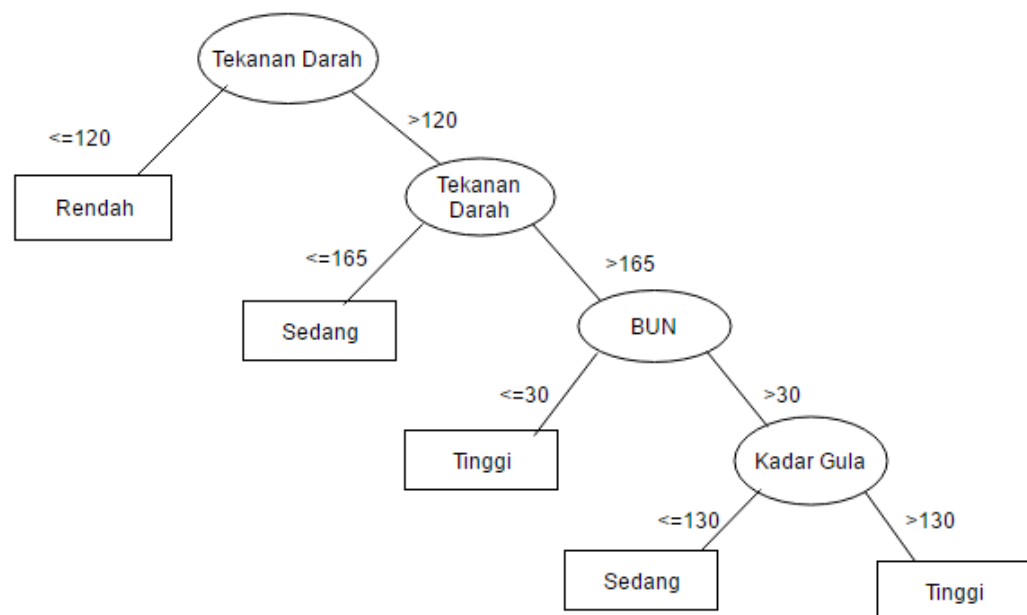
**Gambar 3.5** Pohon Keputusan Node 5

Dari **Tabel 3.23** dapat diketahui hanya tersisa 2 data saja yaitu 1 untuk diagnosa Sedang dan 1 untuk diagnosa Tinggi, maka diteruskan ke proses selanjutnya sehingga diperoleh hasil sebagai berikut.

**Tabel 3.25** Hasil Perhitungan *Entropy* dan *Gain* untuk Node Akar 7

Perhitungan entropy dan gain		NODE 7				
Total		jml	Sedang	Tinggi	Entropy	Gain
		2	1	1	1	
Kadar Gula	$\leq 130$	1	1	0	0	1,0000
	$> 130$	1	0	1	0	
LDL	$\leq 150$	1	0	1	0	1,0000
	$> 150$	1	1	0	0	
Usia	$\leq 60$	1	0	1	0	1,0000
	$> 60$	1	1	0	0	
Asam Urat	$\leq 8$	1	1	0	0	1,0000
	$> 8$	1	0	1	0	

Pada **Tabel 3.25** semua nilai *gain* pada setiap atribut memiliki nilai yang sama yaitu 1, sehingga semua atribut yang tersisa yaitu Kadar Gula, LDL, usia, dan Asam Urat dapat dijadikan *node* akar.



**Gambar 3.6** pohon keputusan node 7

Dari pohon keputuan tersebut terdiri dari 5 *node leaf* (daun) yang dijadikan sebagai kelas keputusan yang dipilih. Pada node awal atribut tekanan darah terpilih sebagai node akar dengan pemecah nilai  $v$  pada nilai 120. Jika tekanan darah berada pada  $\leq 120$  maka diagnosa yang diberikan adalah pada

status “Rendah”, sedangkan pada nilai  $>120$  maka akan dilakukan perhitungan lebih lanjut untuk memperoleh node cabang lagi. Pada node selanjutnya diketahui atribut tekanan darah terpilih sebagai node cabang dengan nilai  $v$  pada 165. Jika tekanan darah berada pada nilai  $>120$  dan berada  $\leq 165$  maka diagnose yang diberikan adalah “Sedang”, sedangkan pada nilai  $>165$  akan dilakukan perhitungan lagi.

*Node* selanjutnya diketahui atribut BUN terpilih sebagai *node* cabang dengan posisi  $v$  pada nilai 30, jika tekanan darah berada pada nilai  $>165$  dengan nilai BUN  $\leq 30$  maka diagnosa yang diberikan adalah “Tinggi” sedangkan untuk nilai BUN  $>30$  akan dilakukan perhitungan untuk memperoleh node cabang lagi. Pada perhitungan terakhir diketahui atribut kadar gula menjadi node cabang dengan nilai  $v$  pada nilai 130. Jika tekanan darah berada pada nilai  $>165$ , nilai BUN  $>30$  dan kadar gula  $\leq 130$  maka diagnosa yang diberikan adalah pada status resiko “Sedang”, sedangkan jika tekanan darah berada pada nilai  $>165$ , nilai BUN  $>30$  dan kadar gula  $>130$  maka diagnosa yang diberikan adalah pada status resiko “Tinggi.”

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

RULE 1: Jika TEKANAN DARAH  $\leq 120$  maka diagnosa RENDAH

RULE 2 : Jika TEKANAN DARAH  $>120$  dan  $\leq 165$  maka diagnosa SEDANG

RULE 3 : Jika TEKANAN DARAH  $>165$  dan BUN  $\leq 30$  maka diagnosa TINGGI

RULE 4 : Jika TEKANAN DARAH  $>165$  dan BUN  $>30$  dan KADAR GULA  $\leq 130$  maka diagnosa SEDANG

RULE 5 : Jika TEKANAN DARAH  $>165$  dan BUN  $>30$  dan KADAR GULA  $>130$  maka diagnosa TINGGI

*Rule* yang terbentuk digunakan untuk melakukan klasifikasi status resiko dini penyakit stroke. Pada pengklasifikasian data uji, *rule* digunakan untuk melihat kelas status diagnosa awal yang diperoleh dari perhitungan berdasarkan data latih. Perhitungan data latih digunakan untuk membentuk *rule*

sehingga dari *rule* tersebut dapat dijadikan aturan sebagai pengklasifikasi untuk data uji.

Dari *rule* yang terbentuk maka akan dilakukan pengujian baik untuk data latih serta data uji. Pengujian ini dilakukan untuk mengetahui apakah *rule* yang terbentuk sudah maksimal (mendekati benar) atau belum. Jika pengujian data latih dan data uji dengan memperhatikan *rule* yang terbentuk dari perhitungan menunjukkan nilai akurasi yang tinggi dan memiliki nilai *error* yang rendah maka dapat disimpulkan bahwa *rule* yang terbentuk sudah maksimal. Berikut adalah hasil pengujian untuk data latih dapat dilihat pada **Tabel 3.26**.

**Tabel 3.26** Hasil prediksi data latih

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA	SISTEM	HASIL
1	110	90	112	87	31	P	5,1	21,5	0,9	Rendah	Rendah	Benar
2	90	100	152	81	43	P	5,2	26	0,9	Rendah	Rendah	Benar
3	110	111	125	82	42	P	5,1	16,7	0,7	Rendah	Rendah	Benar
4	100	110	115	86	52	P	4	31	0,8	Rendah	Rendah	Benar
5	100	113	160	92	49	P	5	10,7	0,8	Rendah	Rendah	Benar
6	90	106	117	95	51	L	5,6	32	0,8	Rendah	Rendah	Benar
7	90	107	115	97	53	L	6	18	1	Rendah	Rendah	Benar
8	100	96	147	77	54	P	4,2	20	1	Rendah	Rendah	Benar
9	100	95	165	82	52	P	3,1	19,5	0,8	Rendah	Rendah	Benar
10	90	97	117	86	49	P	3,1	27	0,7	Rendah	Rendah	Benar
11	100	97	127	79	47	L	6,8	11,8	1	Rendah	Rendah	Benar
12	100	107	140	77	45	L	6,4	25	0,9	Rendah	Rendah	Benar
13	110	101	138	76	46	P	5,7	27	0,8	Rendah	Rendah	Benar
14	100	107	105	82	30	P	4,8	20,4	0,9	Rendah	Rendah	Benar
15	90	105	125	76	29	L	7,5	21,5	0,8	Rendah	Rendah	Benar
16	90	105	125	74	29	L	6,1	25	0,8	Rendah	Rendah	Benar
17	100	100	131	91	35	P	4,5	17	0,8	Rendah	Rendah	Benar
18	110	115	125	90	35	L	5	12	0,9	Rendah	Rendah	Benar
19	90	110	115	75	42	P	5	27	1	Rendah	Rendah	Benar
20	110	113	135	75	40	L	4,8	11,8	1	Rendah	Rendah	Benar
21	100	96	145	81	45	L	5,6	9,6	0,8	Rendah	Rendah	Benar
22	100	95	156	81	37	L	5,7	36	0,9	Rendah	Rendah	Benar
23	90	97	136	80	38	P	6	31,6	0,8	Rendah	Rendah	Benar
24	100	101	135	81	46	P	3,7	21	0,8	Rendah	Rendah	Benar
25	100	100	171	87	46	L	5	15,6	0,9	Rendah	Rendah	Benar
31	110	125	115	90	52	P	5	25,7	0,5	Rendah	Rendah	Benar

**Tabel 3.26** Hasil prediksi data latih (lanjutan)

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA	SISTEM	HASIL
32	110	155	195	100	51	L	6,2	20	0,7	Rendah	Rendah	Benar
33	90	108	187	123	43	L	6,7	25,8	0,6	Rendah	Rendah	Benar
34	100	104	193	122	44	P	5,4	17,8	0,5	Rendah	Rendah	Benar
35	110	109	139	127	55	L	6,8	20,7	0,8	Rendah	Rendah	Benar
50	150	115	122	118	52	P	2	7,5	0,7	Sedang	Sedang	Benar
51	140	132	167	115	75	P	4,8	39,2	1,5	Sedang	Sedang	Benar
52	150	155	195	127	66	P	6,2	30,6	1	Sedang	Sedang	Benar
53	135	175	219	162	48	P	6,2	19,8	1,1	Sedang	Sedang	Benar
54	145	110	153	135	58	L	6,5	20	1,5	Sedang	Sedang	Benar
55	150	87	167	107	50	L	6,7	20	1,5	Sedang	Sedang	Benar
56	140	190	195	127	30	L	6,2	20	1,5	Sedang	Sedang	Benar
57	150	186	260	168	55	P	3,1	6	1	Sedang	Sedang	Benar
58	150	155	195	127	60	L	6,2	20	1,5	Sedang	Sedang	Benar
59	160	110	194	108	60	L	10	20,4	1,5	Sedang	Sedang	Benar
60	140	196	138	88	56	P	3,4	21,1	1,4	Sedang	Sedang	Benar
61	150	111	149	190	48	P	4,6	10,7	0,8	Sedang	Sedang	Benar
62	145	188	208	170	42	L	6,4	14,3	1,3	Sedang	Sedang	Benar
63	150	117	111	52	73	L	7,3	31,6	2,1	Sedang	Sedang	Benar
64	140	136	143	120	71	L	8,9	23,4	1,9	Sedang	Sedang	Benar
65	150	240	264	163	52	L	6,9	15,3	2,5	Sedang	Sedang	Benar
66	150	181	186	137	57	P	3,5	6,2	0,5	Sedang	Sedang	Benar
67	160	119	150	108	60	P	4,9	7,4	1	Sedang	Sedang	Benar
68	130	140	221	148	52	L	8,4	11,1	0,8	Sedang	Sedang	Benar
69	160	105	157	102	56	P	4,5	11,6	0,7	Sedang	Sedang	Benar
70	130	95	210	153	62	L	5,2	7,7	1	Sedang	Sedang	Benar
71	150	109	137	167	65	L	8,8	38,7	3,6	Sedang	Sedang	Benar
72	145	196	200	160	62	P	2,8	13,5	0,6	Sedang	Sedang	Benar
73	135	160	185	183	63	P	7,5	10,7	0,8	Sedang	Sedang	Benar
74	140	191	176	86	66	P	6	20,4	1,2	Sedang	Sedang	Benar
49	150	102	137	90	57	L	5,4	11,3	6,9	Sedang	Sedang	Benar
80	170	111	207	168	80	L	7,9	31,7	2	Sedang	Sedang	Benar
81	140	200	122	80	62	L	7,7	25,5	1,3	Sedang	Sedang	Benar
82	150	128	237	154	50	P	4,5	12,4	0,7	Sedang	Sedang	Benar
83	140	92	161	102	50	P	4	12,8	1	Sedang	Sedang	Benar
91	200	167	263	158	78	L	8	15,8	1,4	Tinggi	Tinggi	Benar
92	180	421	277	165	56	L	4,7	14,9	1,5	Tinggi	Tinggi	Benar
93	200	136	294	195	60	P	4,8	16,2	0,8	Tinggi	Tinggi	Benar
94	170	150	190	130	65	L	6,2	20	1,5	Tinggi	Tinggi	Benar
95	190	127	133	165	65	P	3,8	18,2	0,5	Tinggi	Tinggi	Benar
96	180	136	177	136	65	L	7,3	13,5	1,2	Tinggi	Tinggi	Benar



**Tabel 3.26** Hasil prediksi data latih (lanjutan)

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA	SISTEM	HASIL
97	170	112	189	150	67	P	5,2	11,1	0,8	Tinggi	Tinggi	Benar
98	200	174	187	125	65	L	6,2	12,9	0,9	Tinggi	Tinggi	Benar
99	200	186	195	127	64	P	6,2	13,6	0,9	Tinggi	Tinggi	Benar
100	180	247	189	130	55	P	6,2	20,6	1,1	Tinggi	Tinggi	Benar
101	210	204	229	162	49	P	7,2	23,1	0,7	Tinggi	Tinggi	Benar
102	170	146	208	109	61	P	5	8,6	1	Tinggi	Tinggi	Benar
103	200	305	232	167	65	L	12,2	25,8	2,4	Tinggi	Tinggi	Benar
104	190	119	195	127	63	L	6,2	13,2	0,7	Tinggi	Tinggi	Benar
105	180	146	229	162	45	P	8	5,4	0,8	Tinggi	Tinggi	Benar
106	180	117	191	117	66	P	4	6,1	0,8	Tinggi	Tinggi	Benar
107	190	156	154	125	55	L	7,8	14,6	1,3	Tinggi	Tinggi	Benar
108	190	260	172	128	69	P	3,4	10	0,5	Tinggi	Tinggi	Benar
109	170	192	209	150	59	L	8,2	34,7	1,9	Tinggi	Tinggi	Benar
110	190	100	190	124	62	P	6	14,8	1,2	Tinggi	Tinggi	Benar
111	200	243	282	182	54	P	6,2	11,5	0,7	Tinggi	Tinggi	Benar
112	200	138	267	188	66	P	6,4	13	0,9	Tinggi	Tinggi	Benar
113	170	107	138	196	65	L	6,8	9,5	1,1	Tinggi	Tinggi	Benar
114	180	107	215	150	65	P	5,4	14,6	0,9	Tinggi	Tinggi	Benar
115	180	206	270	184	70	P	3,8	19,5	1,3	Tinggi	Tinggi	Benar
121	190	120	197	117	65	P	7,8	13,3	0,9	Tinggi	Tinggi	Benar
122	200	253	249	187	49	P	5,2	14,1	1	Tinggi	Tinggi	Benar
123	200	169	193	128	65	L	6,2	19,9	1,6	Tinggi	Tinggi	Benar
124	170	130	207	160	52	P	5,7	10,4	1,2	Tinggi	Tinggi	Benar
125	180	328	269	108	50	P	8,6	9,5	0,6	Tinggi	Tinggi	Benar

Dari hasil prediksi data latih sebanyak 90 data didapatkan hasil sebagai berikut yakni 90 data telah sesuai dengan target kelasnya. Pada **Tabel 3.26** dijelaskan bahwa klasifikasi yang dibuat berdasarkan perhitungan menghasilkan rule, sehingga rule tersebut dibuat untuk menentukan kelas status diagnosa. Secara lengkap akan ditampilkan pada **Tabel 3.27** dan evaluasi klasifikasi *Decision Tree C4.5* seperti *Precision*, *Recall* dan *F-measure* ditampilkan pada **Tabel 3.28**

**Tabel 3.27** Detai Prediksi Data Latih

Relevan	Terbaca		
	Rendah	Sedang	Tinggi
Rendah	30	0	0

<b>Sedang</b>	0	30	0
<b>Tinggi</b>	0	0	30

**Tabel 3.28** Evaluasi hasil prediksi data latih

Jenis Penilaian	Kelas			Akurasi Total	Laju Error
	Rendah	Sedang	Tinggi		
<b>Precision</b>	100%	100%	100%	100%	0%
<b>Recall</b>	100%	100%	100%		
<b>F-Measure</b>	100%	100%	100%		

Pada pengujian tabel data uji digunakan 38 data dimana terdapat 18 data dengan kelas asli diagnosa “Rendah”, 12 data dengan kelas asli diagnosa “Sedang”, dan 8 data dengan kelas asli diagnosa “Tinggi”. Penentuan diagnosa oleh sistem dilakukan dengan memperhatikan *rule* yang terbentuk dari hasil perhitungan klasifikasi *Decision Tree C4.5*. Hasil pengujian dapat dilihat pada **Tabel 3.29**.

**Tabel 3.29** Hasil Prediksi Data Uji

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA	SISTEM	HASIL
26	110	110	162	89	45	L	6	27	1,1	Rendah	Rendah	Benar
27	106	121	161	56	49	P	4,7	29	0,8	Rendah	Rendah	Benar
28	100	129	141	78	43	L	5,1	17,5	0,7	Rendah	Rendah	Benar
29	100	155	173	133	56	L	7,1	31,6	0,9	Rendah	Rendah	Benar
30	110	110	195	127	45	L	6,2	27,4	0,8	Rendah	Rendah	Benar
36	90	100	125	121	40	P	5,1	21,5	0,6	Rendah	Rendah	Benar
37	110	106	190	130	45	P	4,4	20,7	0,9	Rendah	Rendah	Benar
38	100	100	195	127	28	L	6,5	16,2	0,6	Rendah	Rendah	Benar
39	90	163	132	75	41	P	4,5	25	0,7	Rendah	Rendah	Benar
40	110	115	170	91	32	P	3,5	9,3	1	Rendah	Rendah	Benar
41	110	120	162	96	37	P	3	17,9	1	Rendah	Rendah	Benar
42	100	109	137	92	36	L	8	38	3,6	Rendah	Rendah	Benar
43	110	115	229	173	50	L	5,5	12,3	1,1	Rendah	Rendah	Benar
44	90	92	221	164	48	P	4,1	16,9	0,6	Rendah	Rendah	Benar
45	100	135	113	63	50	L	6,1	27,1	0,8	Rendah	Rendah	Benar
46	100	102	137	90	47	L	5,4	11,3	0,7	Rendah	Rendah	Benar
47	110	111	117	52	43	L	7,3	31,6	0,9	Rendah	Rendah	Benar
48	100	109	137	92	60	L	6,5	25,7	0,6	Rendah	Rendah	Benar
84	145	135	113	63	80	L	8,1	39,4	2,9	Sedang	Sedang	Benar

NO	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	DIAGNOSA	SISTEM	HASIL
85	145	115	229	173	71	L	5,5	12,3	1,1	Sedang	Sedang	Benar
86	150	117	198	113	56	L	9	9,8	1,1	Sedang	Sedang	Benar
87	150	98	147	95	60	L	2,8	21,5	1,5	Sedang	Sedang	Benar
88	140	108	222	146	40	P	3,3	11,6	0,5	Sedang	Sedang	Benar
89	140	78	243	182	59	P	4,9	11,7	0,6	Sedang	Sedang	Benar
90	150	91	162	125	55	P	4,8	20,2	1	Sedang	Sedang	Benar
75	140	98	176	106	52	L	10,2	14,4	1,1	Sedang	Sedang	Benar
76	150	220	168	106	47	P	3	7,4	0,5	Sedang	Sedang	Benar
77	145	104	192	112	52	L	8,2	15,1	1,1	Sedang	Sedang	Benar
78	150	116	159	92	55	P	4,2	12	0,8	Sedang	Sedang	Benar
79	140	128	155	102	42	L	8	10,5	0,9	Sedang	Sedang	Benar
116	180	257	195	140	56	P	5	14,7	0,7	Tinggi	Tinggi	Benar
117	180	140	194	134	69	L	8,6	28,7	1,7	Tinggi	Tinggi	Benar
118	190	173	147	147	70	L	6,6	21,5	0,9	Tinggi	Tinggi	Benar
119	170	86	214	106	80	L	7,8	9,6	0,8	Tinggi	Tinggi	Benar
120	180	82	130	185	72	P	4,7	12,7	0,5	Tinggi	Tinggi	Benar
126	170	152	195	127	55	L	6,2	14,9	1,2	Tinggi	Tinggi	Benar
127	170	223	375	205	60	P	6,2	79,2	3,8	Tinggi	Tinggi	Benar
128	190	231	183	134	76	L	9,8	17,9	1,2	Tinggi	Tinggi	Benar

Dari hasil prediksi data uji sebanyak 38 data didapatkan hasil, 38 data telah sesuai dengan target kelasnya, dengan perincian 18 data kelas Rendah, 12 data kelas Sedang dan 8 data kelas Tinggi. Secara lengkap akan ditampilkan pada **Tabel 3.30** dan untuk evaluasi klasifikasi *Decision Tree C4.5* yang dilakukan untyuk data uji seperti *Precision*, *Recall* dan *F-measure* ditampilkan pada **Tabel 3.31**.

**Tabel 3.30** Detail Prediksi Data Uji

Relevan	Terbaca		
	Rendah	Sedang	Tinggi
<b>Rendah</b>	18	0	0
<b>Sedang</b>	0	12	0
<b>Tinggi</b>	0	0	8

**Tabel 3.30** menjelaskan bahwa perbandingan kelas asli dan kelas yang dihasilkan berdasarakan rule menunjukkan tingkat kebenaran yang bagus, dimana semua kelas asli Rendah dikenali oleh sistem sebagai Rendah, begitu juga kelas Sedang maupun kelas Tinggi..

**Tabel 3.31** Evaluasi Hasil Prediksi Data Uji

Jenis Penilaian	Kelas			Akurasi Total	Laju Error
	Rendah	Sedang	Tinggi		
<b>Precision</b>	100%	100%	100%	100%	0%
<b>Recall</b>	100%	100%	100%		
<b>F-Measure</b>	100%	100%	100%		

### 3.4 Kebutuhan Pembuatan Sistem

#### 1. Kebutuhan Perangkat Keras

Perangkat keras adalah alat yang digunakan untuk menunjang dalam pembuatan sistem. Dalam pembuatan sistem ini perangkat keras yang digunakan yaitu laptop dengan spesifikasi :

- a. Processor Intel Core i3
- b. RAM 4 GB
- c. HDD 500 GB
- d. Monitor 12,5"
- e. Mouse

#### 2. Kebutuhan Perangkat Lunak

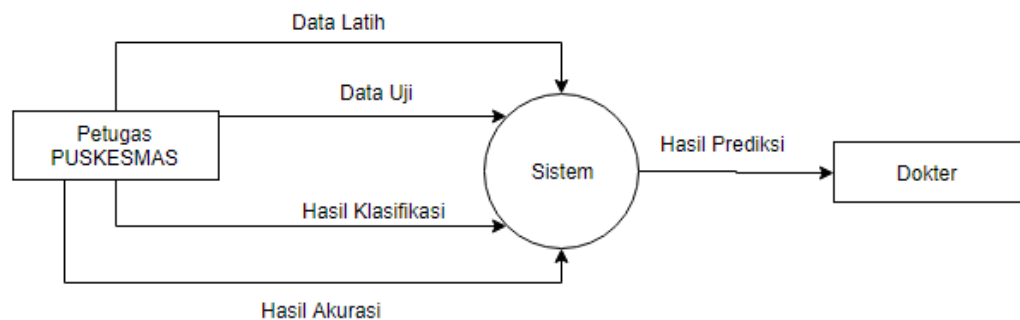
Perangkat lunak adalah program atau aplikasi yang digunakan untuk membangun sistem. Perangkat lunak yang dibutuhkan dalam pembuatan sistem ini adalah :

- a. Sistem Operasi : Windows 10
- b. *Web Server* : XAMPP v3.2.1
- c. *Database Server* : MySQL
- d. Bahasa Pemrograman : PHP
- e. *Software Development* : NetBeans IDE 8.0
- f. *Tool Basis Data* : Sqlyog
- g. *Web Browser* : Google Chrome

### 3.5 Perancangan Sistem

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

#### 3.5.1 Context Diagram

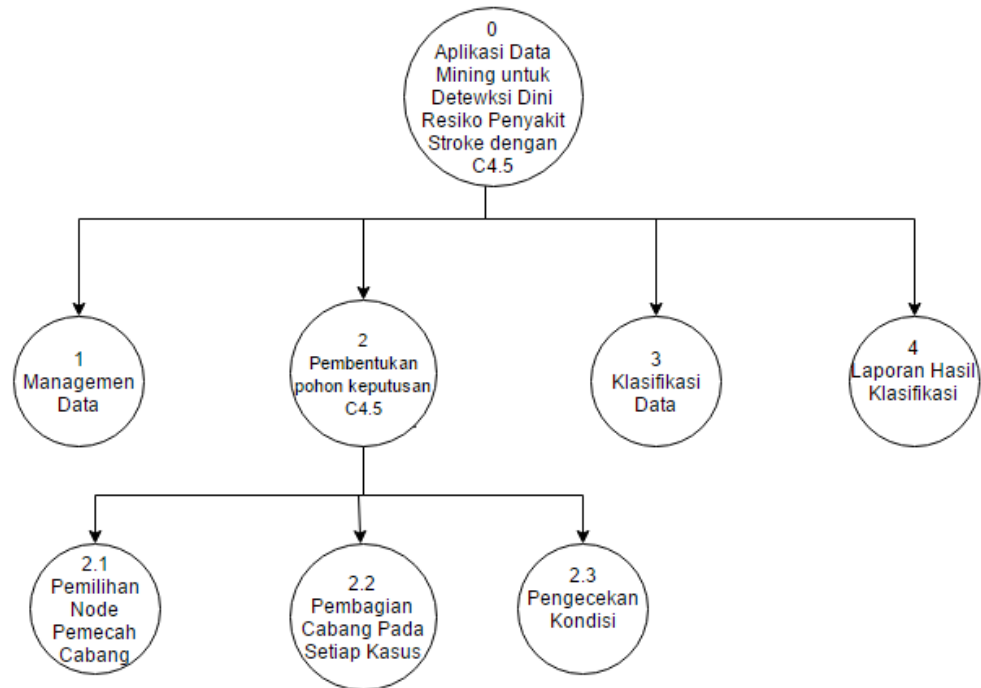


**Gambar 3.7** Context Diagram

Penjelasan dari **Gambar 3.7**, terlihat bahwa yang terlibat (*entity*) dalam sistem ini adalah Petugas PUSKESMAS dan Dokter. Petugas, memasukkan data rekam medis pasien berupa Tekanan Darah, Kadar Gula, Kolestrol Total, LDL, Usia, Jenis Kelamin, Asam Urat, BUN dan Kreatinin serta kelas diagnosa. Data tersebut digunakan sebagai *data training* atau data yang akan diproses untuk pembentukan pohon keputusan. Keluaran dari sistem untuk user adalah hasil klasifikasi diagnosa resiko penyakit stroke berupa kelas diagnosa resiko berdasarkan data yang telah dimasukkan serta hasil akurasi penghitungan. Sedangkan Dokter dapat melihat laporan atau daftar hasil klasifikasi diagnosa resiko penyakit stroke yang telah melalui proses klasifikasi berupa status resiko pasien.

#### 3.5.2 Diagram Berjenjang

Diagram berjenjang merupakan suatu gambaran struktur dari sistem berupa suatu bagan berjenjang yang menggambarkan semua semua proses yang ada disistem. Dipergunakan untuk mempersiapkan penggambaran DFD ke level lebih bawah lagi. Diagram berjenjang pada sistem ini dapat dilihat pada **Gambar 3.8**

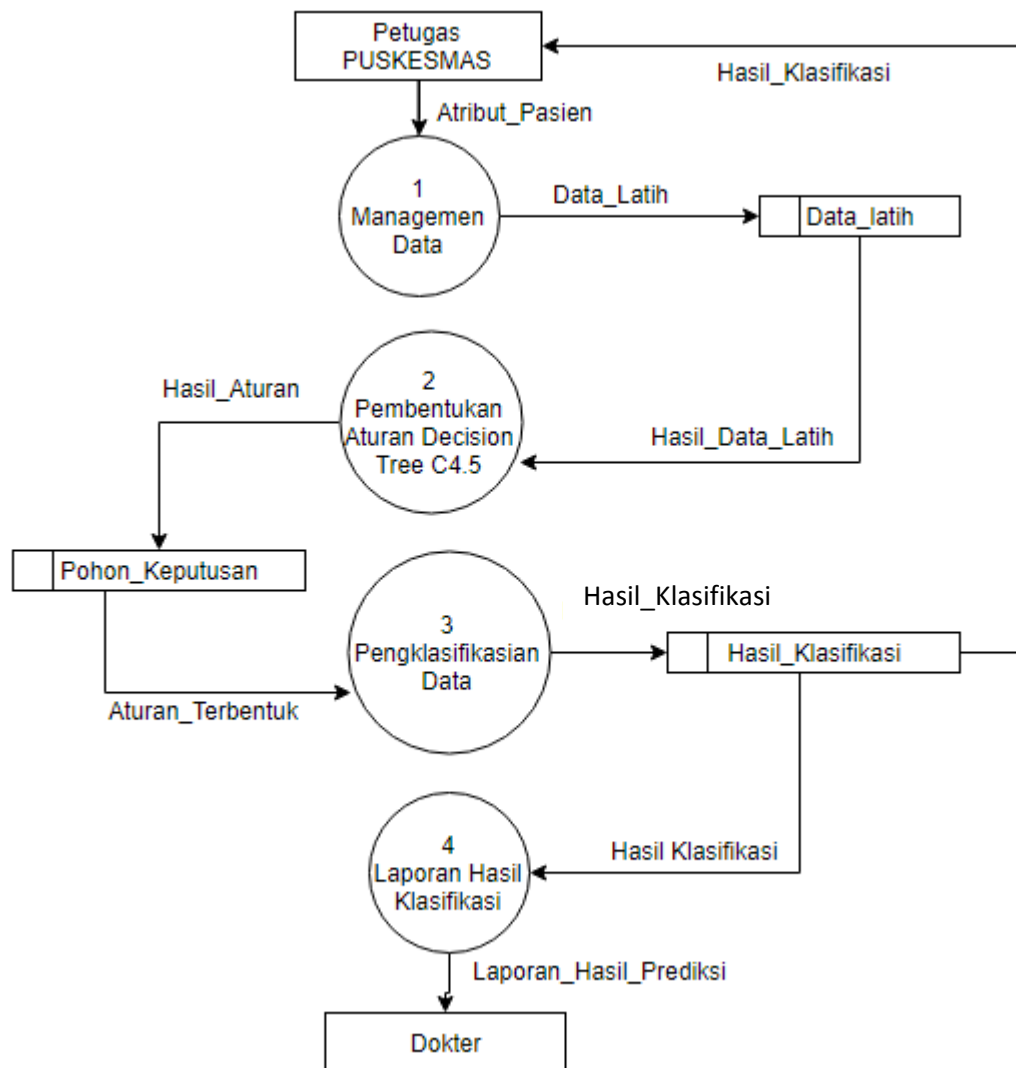


**Gambar 3.8** Diagram Berjenjang

Diagram berjenjang disajikan pada **Gambar 3.8.** berikut penjelasannya:

1. *Top level* : Aplikasi Data Mining untuk Deteksi Dini Resiko Penyakit Stroke menggunakan Metode *Decision Tree C4.5*
2. *Level 0* : 1 Manajemen data, merupakan proses pengolahan data 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.  
4 Pembuatan Laporan Hasil Klasifikasi.
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

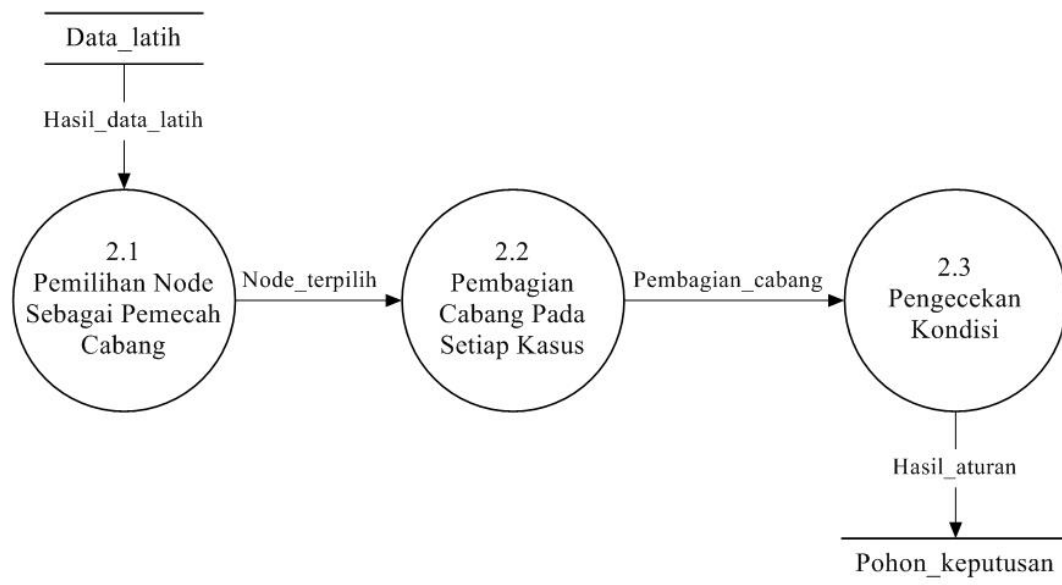


**Gambar 3.9** DFD *Level 0*

DFD *level 0* pada **Gambar 3.9** menjelaskan aliran data pada sistem. Terdapat empat proses didalam sistem tersebut. Proses satu adalah manajemen data yang diinputkan oleh Petugas PUSKESMAS. Data atribut rekam medis pasien akan menjadi data latih untuk proses pembentukan pohon keputusan. Proses dua adalah pembentukan aturan (pohon keputusan) yang akan digunakan pada proses pengklasifikasian data uji. Hasilnya akan diberikan kepada Petugas dan akan disimpan dalam hasil klasifikasi. Proses empat adalah pembuatan laporan hasil klasifikasi deteksi dini resiko penyakit

stroke yang akan diberikan kepada user dengan mengambil data dari tabel hasil klasifikasi.

### 3.5.4 Data Flow Diagram Level 1



**Gambar 3.10** 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 *data base*.

### 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 sumber data (PUSKESMAS). Struktur dari tabel ini dapat dilihat pada **Tabel 3.31**.

**Tabel 3.31** Struktur Tabel Data Latih

	Field_name	Type	Length	Key
	Id	Int	11	Primary key
	tekanan_darah	Int		
	kadar_gula	Int		
	kolestrol_total	Int		
	Ldl	int		
	Usia	int		
	jenis_kelamin	char	2	
	asam_urat	double		
	bun	double		
	kreatinin	deouble		
	diagnosa	varchar	20	

b. data\_uji

Tabel ini digunakan untuk menyimpan data pengujian, yaitu untuk menguji tingkat akurasi dari pohon keputusan yang terbentuk. Strukturnya

**Tabel 3.32** Struktur Tabel Data Uji

	Field_name	Type	Length	Key
	id	Int	11	Primary key
	tekanan_darah	Int		
	kadar_gula	int		
	kolestrol_total	int		
	ldl	int		
	usia	int		
	jenis_kelamin	char	2	
	asam_urat	double		
	bun	double		

**Tabel 3.32** Struktur tabel data uji (lanjutan)

	Field_name	Type	Length	Key
	kreatinin	double		
	diagnosa	varchar	20	
	diganosa_sistem	varchar	20	

## c. hasil\_klasifikasi

Tabel ini digunakan untuk menyimpan data hasil klasifikasi. Strukturnya sama dengan tabel data latih namun *field* diagnosa diganti dengan hasil.

**Tabel 3.33** Struktur Tabel Hasil Klasifikasi

	Field_name	Type	Length	Key
	id	Int	11	Primary key
	tekanan_darah	Int		
	kadar_gula	int		
	kolestrol_total	int		
	ldl	int		
	usia	int		
	jenis_kelamin	char	2	
	asam_urat	double		
	bun	double		
	kreatinin	double		
	hasil	varchar	20	

## d. gain

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

**Tabel 3.34** Struktur tabel *gain*

	Name_field	Type	Length	Key
	id	Int	11	Primary key
	node_id	Int	11	

	Name_field	Type	Length	Key
	atribut	Varchar	40	
	gain	Double		

e. t\_user

Tabel *user* ini dibuat untuk secara khusus agar bisa mengakses aplikasi ini. Data dari *user* tersebut tersimpan dalam tabel *user*. Struktur dari tabel *user* dapat dilihat pada tabel 3.29

**Tabel 3.35** Struktur t\_user

	Name_field	Type	Length	Key
	user_id	Int	25	Primary key
	nama	Varchar	50	
	username	Varchar	30	
	password	Varchar	30	
	type	Char	1	

f. t\_keputusan

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

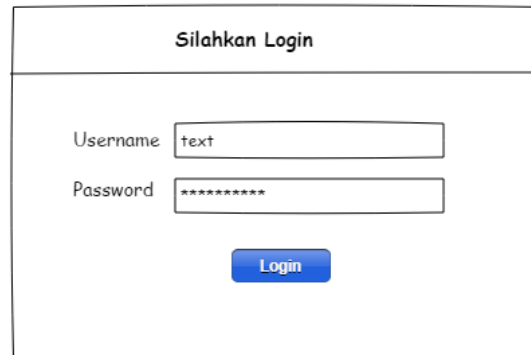
**Tabel 3.36** Struktur t\_keputusan

	Field_name	Type	Length	Key
	id	Int	11	Primary key
	parent	Text		
	akar	Text		
	keputusan	Varchar	10	

### 3.5.6 Desain Antar Muka

#### a. Halaman *LogIn* (*admin* maupun *user*)

Halaman *login* diperlukan untuk mengetahui hak akses pengguna yang masuk kedalam sistem yaitu pengguna sebagai *admin* atau sebagai *user*.

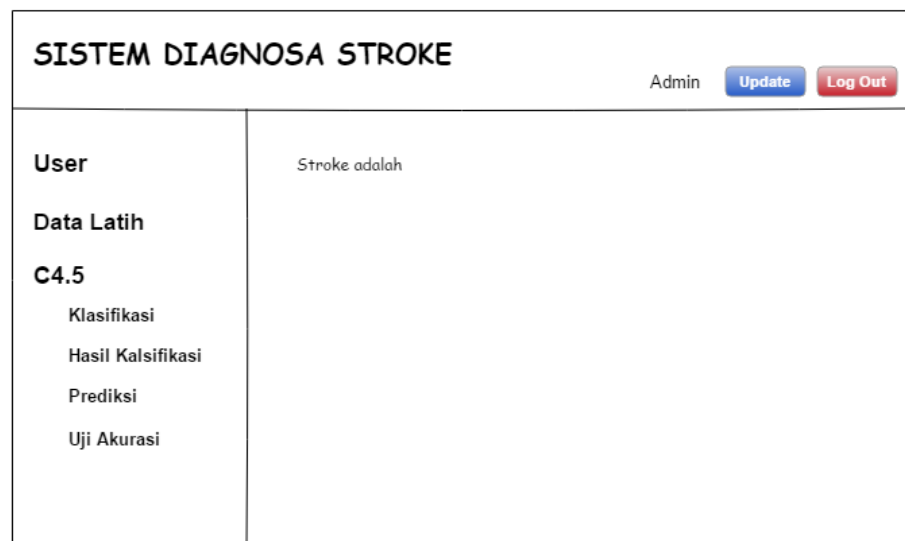


The image shows a login form titled "Silahkan Login". It contains two input fields: "Username" with a placeholder "text" and "Password" with a placeholder "\*\*\*\*\*". Below the fields is a blue button labeled "Login".

**Gambar 3.11** Rancangan halaman *login*

#### b. Halaman Utama

Menu yang ditampilkan untuk pengguna sebagai *admin* yaitu daftar *user*, Data Latih, dan Klasifikasi C4.5. Sedangkan pengguna sebagai *user* adalah laporan hasil klasifikasi. Pada halaman utama ini terdapat tombol *update* yang digunakan untuk merubah data pengguna serta tombol *log out* untuk keluar dari aplikasi. Pada tampilan halaman utama berisi pengetahuan umum tentang penyakit stroke.



The image shows the main page layout for "SISTEM DIAGNOSA STROKE". At the top right, there is a navigation bar with "Admin", a blue "Update" button, and a red "Log Out" button. On the left side, there is a vertical menu with the following items: "User", "Data Latih", "C4.5", "Klasifikasi", "Hasil Kalsifikasi", "Prediksi", and "Uji Akurasi". The main content area on the right displays the text "Stroke adalah".

**Gambar 3.12** Rancangan halaman utama

### c. Halaman Data Latih

Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan. Halaman ini hanya bisa diakses oleh admin saja. Tombol tambah data digunakan untuk menambahkan data latih.

**Data Latih**

Tambah Data

No	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	Diagnosa
1	120	125	195	100	34	L	6,2	20	0,7	Rendah

**Gambar 3.13** Rancangan halaman data latih

### d. Data Uji

Tampilan ini adalah halaman untuk menampilkan data uji yang dibuat untuk melakukan proses evaluasi dari hasil klasifikasi yang terbentuk sesuai dengan pohon keputusan. Tombol tambah data digunakan untuk menambahkan data uji.

**Data Uji**

Tambah Data

No	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	Diagnosa
1	120	125	195	100	34	L	6,2	20	0,7	Rendah

**Gambar 3.14** Halaman data uji

e. Halaman Klasifikasi

Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan. Tombol *mining* digunakan untuk melakukan perhitungan pada data latih untuk membentuk pohon keputusan.

**Hasil Klasifikasi**

Tabel Data Latih

No	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	Diagnosa
1	120	125	195	100	34	L	6,2	20	0,7	Rendah

Mining

Tabel Perhitungan Node 1

Total		jml	Sedang	Tinggi	Entropy	Gain
		60	30	30	1	
Tekanan Darah	<=165	29	29	0	0	0.8938
	>165	31	1	30	0.2056	

**Gambar 3.15** Rancangan Halaman Klasifikasi

f. Halaman Hasil Klasifikasi

Halaman ini menampilkan pohon keputusan atau aturan yang didapat dari proses *mining*. Tombol uji *rule* digunakan untuk melihat nilai akurasi dan laju *error* dari klasifikasi data uji.

**Hasil Klasifikasi**

Uji Rule

Aturan (rule) yang terbentuk

Id	Aturan yang Terbentuk
1	Jika tekanan darah > 120 maka dianosa SEDANG
2	Jika ..... maka ..... diagnosa .....
3	Jika ..... maka ..... diagnosa .....

**Gambar 3.16** Rancangan Halaman Pohon Keputusan

### g. Halaman Uji Akurasi

Halaman ini digunakan untuk menguji tingkat akurasi pohon keputusan yang terbentuk dari proses *mining*. Tombol hitung akurasi digunakan untuk melihat nilai akurasi dan laju *error*.

### Hasil Klasifikasi

Hitung Akurasi

Tabel Data Latih berjumlah .....

No	TD	KG	KT	LDL	U	JK	AU	BUN	KRE	Diagnosa	Kelas
1	110	125	195	100	35	L	6,3	20	0,7	Rendah	Rendah

Akurasi : 78 %

Laju Error : 2 %

**Gambar 3.17** Rancangan Halaman Uji Pohon Keputusan

### h. Halaman Prediksi

Halaman hasil ini akan menampilkan *form* yang digunakan untuk menambahkan data pasien baru yang akan dilakukan prediksi diagnosa oleh sistem.

### Form Data Rekam Medis

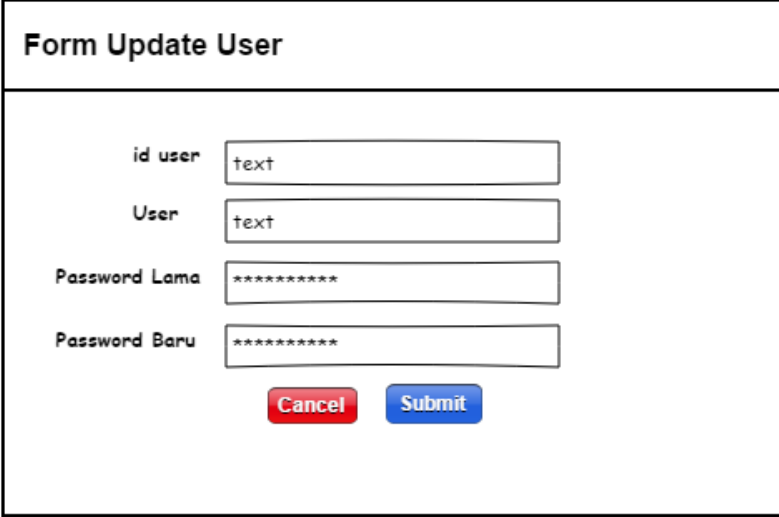
id	<input type="text"/>	Usia	<input type="text"/>
Tekanan Darah	<input type="text"/>	Jenis Kelamin	<input type="checkbox"/> Laki-laki <input checked="" type="checkbox"/> Perempuan
Kadar Gula	<input type="text"/>	Asam Urat	<input type="text"/>
Kolestrol Total	<input type="text"/>	BUN	<input type="text"/>
LDL	<input type="text"/>	Kreatinin	<input type="text"/>

Cancel
Submit

**Gambar 3.18** Rancangan Form Penambahan Prediksi

i. Halaman Ubah *Password*

Halaman ubah *password* ini adalah halaman untuk mengganti *password* user.

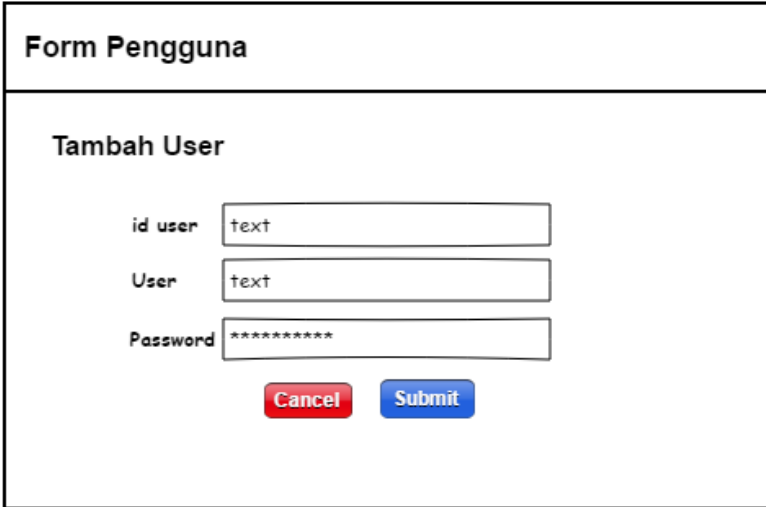


The image shows a web form titled "Form Update User". It contains four input fields: "id user" (text), "User" (text), "Password Lama" (password), and "Password Baru" (password). Below the fields are two buttons: "Cancel" (red) and "Submit" (blue).

**Gambar 3.19** Rancangan Halaman Ubah *Password*

j. Halaman Tambah User

Halaman ini digunakan untuk menambahkan *user*/pengguna untuk sistem ini.



The image shows a web form titled "Form Pengguna". It contains three input fields: "id user" (text), "User" (text), and "Password" (password). Below the fields are two buttons: "Cancel" (red) and "Submit" (blue).

**Gambar 3.20** Rancangan Halaman Penambahan Pengguna



#### k. Halaman Daftar *User*

Halaman ini menampilkan daftar user yang ada pada sistem ini. Pada halaman ini terdapat fitur tambah *user* untuk menambahkan pengguna baru, *update* status untuk mengubah status pengguna, *reset password* dan hapus pengguna.

SISTEM DIAGNOSA STROKE		Admin	Update	Log Out
<b>User</b>  <b>Data Latih</b>  <b>C4.5</b> Klasifikasi Hasil Kalsifikasi Prediksi Uji Akurasi	<b>Pengguna</b> <span style="float: right;">Tambah User</span>			
	<b>Id user</b>	<b>Nama</b>	<b>Username</b>	<b>Action</b>
	admin	Admin	Admin	Update Reset Password Hapus
	user	user	user	Update Reset Password Hapus

**Gambar 3.21** Rancangan Halaman Daftar Pengguna

### 3.6 Evaluasi Sistem

Evaluasi yang digunakan yaitu 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. Sedangkan untuk mengukur tingkat kesalahannya menggunakan rumus 2.6. Selain itu evaluasi yang digunakan juga dapat dilihat dengan menghitung nilai *Recall*, *Precision*, dan *F-Measure* sehingga dapat dilihat nilai akurasi total dan nilai laju *error*.

### 3.7 Skenario Pengujian Sistem

Sebelum membuat aplikasi klasifikasi untuk mendeteksi resiko dini penyakit stroke dengan metode *decision tree C4.5* ini, perlu dilakukan beberapa skenario pengujian sistem terlebih dahulu, agar sistem dapat berjalan sesuai dengan tujuan pembuatannya.

- a. Di sediakan 2 macam data yaitu data latih sebanyak 90 dan data uji sebanyak 38. 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, diklasifikasi berdasarkan pohon keputusan yang terbentuk. Kemudian dihitung akurasi yang menunjukkan baik atau tidaknya pohon keputusan yang sudah terbentuk.
- b. Pada uji hasil klasifikasi maka sistem dapat menentukan diagnosa resiko penyakit stroke dengan kelas resiko “Rendah”, “Sedang”, dan “Tinggi”.
- c. Akurasi sistem diperoleh dari hasil prosentase akurasi ketepatan sistem dan hasil prosentase akurasi kesalahan sistem.