

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.

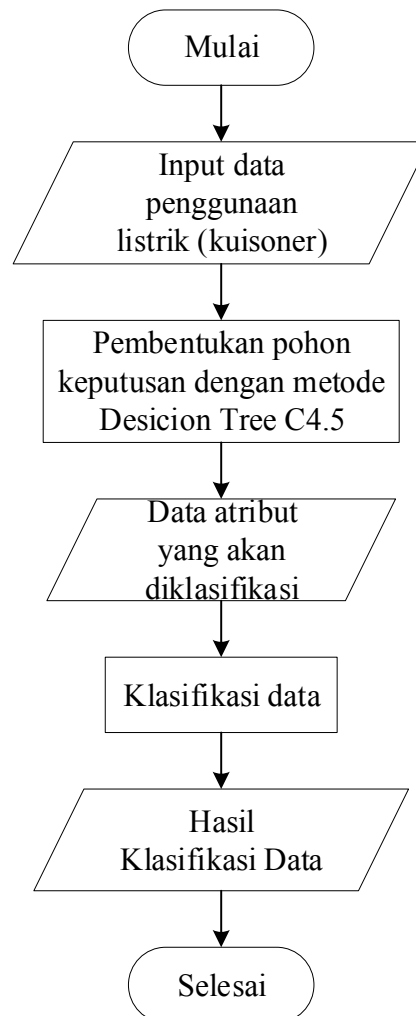
Adanya perubahan tarif dasar listrik membuat pelanggan harus mengeluarkan uang lebih untuk pembayaran listrik perbulan. Hal ini membuat pelanggan keberatan terutama pelanggan golongan R-1/900 VA-RTM tarif listrik tidak jauh berbeda dengan golongan 1.300 VA atau di atasnya. Untuk mengatasi hal tersebut masyarakat harus bijak dalam penggunaan listrik sehingga tarif listrik tidak menjadi tinggi atau membengkak. Masyarakat diharapkan dapat memprediksi besarnya tarif listrik per bulannya. Maka dari itu, permasalahan yang akan diteliti adalah menerapkan pada sebuah sistem prediksi agar dapat ditentukan apakah tarif listrik tersebut sedang, atau tinggi. Proses klasifikasi menggunakan metode *Decision Tree C4.5*.

3.2 Hasil Analisis

Hasil analisis yang dapat dilakukan dari aplikasi data mining untuk memprediksi besarnya tarif listrik rumah tangga yang dibangun dapat membantu untuk mengetahui besarnya tarif listrik rumah tangga dengan kategori sedang dan tinggi. Pembuatan aplikasi data mining untuk memprediksi besarnya tarif listrik rumah tangga diperlukan data pembelajaran, data tersebut diperoleh dari data wawancara (kuisoner) yang dibagikan kepada masyarakat Gresik dan sekitarnya, yang nantinya akan diolah dengan menggunakan metode *Decision Tree C4.5*. Hasil yang diperoleh dari perhitungan metode *Decision Tree C4.5* berupa hasil prediksi besarnya tarif listrik rumah tangga yang dapat membantu masyarakat dalam mengatur penggunaan listrik. Terdapat beberapa atribut yang dibutuhkan untuk

memprediksi besarnya penggunaan listrik rumah tangga diantaranya: luas rumah, daya listrik, jumlah perlengkapan yang dimiliki, jumlah anggota keluarga, dan pendapatan perbulan.

Gambar 3.1 menjelaskan alur sistem pada aplikasi data mining untuk memprediksi besarnya penggunaan listrik rumah tangga.



Gambar 3.1 *Flowchart System*

Penjelasan gambar 3.1:

1. Pertama memasukkan data training (data penggunaan listrik) yang 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 C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai simpul akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

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* digunakan untuk menentukan node yang akan menjadi pemecah data latih dan 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. *Flowchart* algoritma *Decision Tree* C4.5 dapat dilihat pada gambar 2.2.

3.2.1 Sumber Data

Tahapan awal yang dilakukan dalam penelitian ini adalah menyiapkan data, dimana data diperoleh dari pemberian kuisioner kepada masyarakat Gresik dan sekitarnya. Data yang diperoleh akan digunakan dalam penelitian ini berupa data yang berkaitan dengan prediksi besarnya tarif listrik rumah tangga yaitu luas rumah, daya listrik, jumlah perlengkapan yang dimiliki, jumlah anggota keluarga, dan pendapatan perbulan. Jumlah data yang digunakan sebanyak 80 *record* dengan kelas “Sedang” dan “Tinggi” masing-masing berjumlah 40 dan 40 yang akan dibagi menjadi data latih dan data uji.

Data yang didapatkan tersebut dibagi menjadi 7 atribut seperti pada tabel 3.1:

Tabel 3.1 Atribut Data Responden

No	Atribut	Keterangan
1	Nama	Nama Responden
2	Alamat	Alamat Responden
3	Luas Rumah	Luas Rumah Responden
4	Daya Listrik	Daya Listrik Rumah Responden
5	Jumlah Perlengkapan	Jumlah Perlengkapan Elektronik Responden
6	Jumlah Anggota Keluarga	Jumlah Anggota Keluarga Responden
7	Pendapatan Perbulan	Pendapatan Perbulan Responden

3.2.2 Persiapan Data

Data yang akan diproses untuk klasifikasi prediksi besarnya tarif listrik rumah tangga diperoleh dari kuisioner sebanyak 80. Dari data-data tersebut yang dipilih untuk dijadikan sebagai atribut adalah luas rumah, daya listrik, jumlah perlengkapan, jumlah anggota keluarga, dan pendapatan perbulan. Nilai atribut-atribut tersebut memiliki tipe kategorikal dan numerik serta kelas bertipe kategorikal, rinciannya adalah sebagai berikut:

Tabel 3.2 Data atribut

No	Atribut	Keterangan	Tipe
1	Luas Rumah	Luas Rumah Responden	Numerik
2	Daya Listrik	Daya Listrik Rumah Responden	Kategorikal
3	Jumlah Perlengkapan	Jumlah Perlengkapan Elektronik Responden	Numerik
4	Jumlah Anggota Keluarga	Jumlah Anggota Keluarga Responden	Numerik
5	Pendapatan Perbulan	Pendapatan Perbulan Responden	Kategorikal
6	Kelas	Sedang dan Tinggi	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.

Pada tabel 3.3 di tampilkan data keseluruhan yang akan di gunakan. Sedangkan data latih pada table 3.6 dan data uji pada table 3.8

Tabel 3.3 Data Keseluruhan

No.	Luas Rumah (m²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
6	51	900	6	4	Sedang	Sedang
7	84	900	9	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
13	26	900	10	3	Rendah	Sedang
14	120	1300	6	4	Tinggi	Sedang
15	70	450	10	4	Tinggi	Sedang
16	108	450	8	3	Tinggi	Sedang
17	112	1300	20	2	Tinggi	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
20	96	1300	13	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
22	126	900	10	3	Tinggi	Sedang
23	102	450	16	4	Rendah	Sedang
24	102	900	19	4	Sedang	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang

Tabel 3.3 Data Keseluruhan (lanjutan)

30	112	2200	9	5	Tinggi	Tinggi
31	77	1300	14	4	Tinggi	Tinggi
32	40	900	6	5	Tinggi	Tinggi
33	80	2200	20	4	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
35	72	2200	8	3	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
37	176	900	11	6	Sedang	Tinggi
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
44	140	2200	9	4	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
54	40	900	8	3	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
57	40	900	12	6	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi
59	72	900	9	3	Sedang	Sedang
60	61	1300	9	2	Tinggi	Sedang
61	92	1300	7	1	Sedang	Sedang
62	72	900	7	1	Tinggi	Sedang
63	77	900	7	4	Sedang	Sedang
64	133	900	6	4	Tinggi	Sedang
65	72	1300	9	0	Rendah	Sedang
66	100	900	11	3	Tinggi	Sedang
67	72	900	7	4	Sedang	Sedang
68	28	900	20	5	Tinggi	Sedang
69	40	450	19	4	Sedang	Sedang

Tabel 3.3 Data Keseluruhan (lanjutan)

70	133	1300	10	5	Sedang	Tinggi
71	120	900	8	4	Tinggi	Tinggi
72	72	900	9	2	Sedang	Tinggi
73	42	1300	10	2	Tinggi	Tinggi
74	72	1300	13	3	Tinggi	Tinggi
75	108	1300	12	5	Tinggi	Tinggi
76	72	1300	17	3	Sedang	Tinggi
77	72	900	7	3	Sedang	Tinggi
78	168	1300	15	5	Tinggi	Tinggi
79	67	1300	9	3	Tinggi	Tinggi
80	120	1300	12	4	Tinggi	Tinggi

Tabel 3.4 Data Latih

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
6	51	900	6	4	Sedang	Sedang
7	84	900	9	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
13	26	900	10	3	Rendah	Sedang
14	120	1300	6	4	Tinggi	Sedang
15	70	450	10	4	Tinggi	Sedang
16	108	450	8	3	Tinggi	Sedang
17	112	1300	20	2	Tinggi	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
20	96	1300	13	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
22	126	900	10	3	Tinggi	Sedang
23	102	450	16	4	Rendah	Sedang

Tabel 3.4 Data Latih (lanjutan)

24	102	900	19	4	Sedang	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
30	112	2200	9	5	Tinggi	Tinggi
31	77	1300	14	4	Tinggi	Tinggi
32	40	900	6	5	Tinggi	Tinggi
33	80	2200	20	4	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
35	72	2200	8	3	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
37	176	900	11	6	Sedang	Tinggi
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
44	140	2200	9	4	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
54	40	900	8	3	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
57	40	900	12	6	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi

Tabel 3.5 Data Uji

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	72	900	9	3	Sedang	Sedang
2	61	1300	9	2	Tinggi	Sedang
3	92	1300	7	1	Sedang	Sedang
4	72	900	7	1	Tinggi	Sedang
5	77	900	7	4	Sedang	Sedang
6	133	900	6	4	Tinggi	Sedang
7	72	1300	9	0	Rendah	Sedang
8	100	900	11	3	Tinggi	Sedang
9	72	900	7	4	Sedang	Sedang
10	28	900	20	5	Tinggi	Sedang
11	40	450	19	4	Sedang	Sedang
12	133	1300	10	5	Sedang	Tinggi
13	120	900	8	4	Tinggi	Tinggi
14	72	900	9	2	Sedang	Tinggi
15	42	1300	10	2	Tinggi	Tinggi
16	72	1300	13	3	Tinggi	Tinggi
17	108	1300	12	5	Tinggi	Tinggi
18	72	1300	17	3	Sedang	Tinggi
19	72	900	7	3	Sedang	Tinggi
20	168	1300	15	5	Tinggi	Tinggi
21	67	1300	9	3	Tinggi	Tinggi
22	120	1300	12	4	Tinggi	Tinggi

3.3 Representasi Model

Perhitungan *decision tree c4.5* ini akan menggunakan data pada tabel 3.6 (*data training*). 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 yakni \leq dan $>$
- Posisi v yang digunakan pada atribut Luas Rumah (LR) adalah nilai antara $\{41, 46, 51, 56, 61, 66, 71, 76, 81, 86, 91, 96, 101, 106, 111, 116, 121\}$.

- Posisi v yang digunakan pada atribut Jumlah Perlengkapan (JP) adalah nilai antara {4, 5, 6, 7, 8, 9, 10, 11, 12, 13}.
- Posisi v yang digunakan pada atribut Jumlah Anggota (JT) adalah nilai antara {1, 2, 3, 4, 5, 6}.

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.

$$\begin{aligned} Entropy(S) &= -\frac{29}{58} * \log_2\left(\frac{29}{58}\right) - \frac{29}{58} * \log_2\left(\frac{29}{58}\right) \\ &= 0,5000 + 0,5000 = 1,0000 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur, untuk fitur yang bertipe *numeric*, harus ditentukan posisi v yang terbaik untuk pemecahan. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 41$. Maka untuk fitur LR dilakukan diskretisasi pada $v = 41$ ketika menghitung *Entropy* dan *Gain* pada semua fitur.

Tabel 3.6 Posisi v untuk pemecahan fitur LR di *node* akar

	41		46		51		56		61	
	<=	>	<=	>	<=	>	<=	>	<=	>
Sedang	1	28	2	27	3	26	4	25	7	22
Tinggi	6	23	7	22	7	22	7	22	10	19
Jumlah	7	51	9	49	10	48	11	47	17	41
Entropy	0,5917	0,9931	0,7642	0,9925	0,8813	0,9950	0,9457	0,9971	0,9774	0,9961
Gain	0,0554		0,0429		0,0246		0,0127		0,0094	

Tabel 3.7 Posisi v untuk pemecahan fitur LR di *node* akar (lanjutan)

	66		71		76		81		86	
	<=	>	<=	>	<=	>	<=	>	<=	>
Sedang	7	22	10	19	11	18	12	17	15	14
Tinggi	10	19	13	16	16	13	18	11	18	11
Jumlah	17	41	23	35	27	31	30	28	33	25
Entropy	0,9774	0,9961	0,9877	0,9947	0,9751	0,9812	0,9710	0,9666	0,9940	0,9896
Gain	0,0094		0,0081		0,0217		0,0311		0,0079	

Tabel 3.8 Posisi v untuk pemecahan fitur LR di *node* akar (lanjutan)

	91		96		101		106		111	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
Sedang	16	13	19	10	21	8	23	6	24	5
Tinggi	19	10	21	8	23	6	23	6	23	6
Jumlah	35	23	40	18	44	14	46	12	47	11
Entropy	0,9947	0,9877	0,9982	0,9911	0,9985	0,9852	1,0000	1,0000	0,9997	0,9940
Gain	0,0081		0,0040		0,0047		0,0000		0,0014	

Tabel 3.9 Posisi v untuk pemecahan fitur LR di *node* akar (lanjutan)

	116		121	
	\leq	$>$	\leq	$>$
Sedang	25	4	27	2
Tinggi	24	5	24	5
Jumlah	49	9	51	7
Entropy	0,9997	0,9911	0,9975	0,8631
Gain	0,0016		0,0187	

Selanjutnya, untuk Jumlah Perlengkapan (JP) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 8$. Maka untuk fitur JP dilakukan diskretisasi pada $v = 8$ ketika menghitung *Entropy* dan *Gain* pada semua fitur.

Tabel 3.10 Posisi v untuk pemecahan fitur JP di *node* akar

	4		5		6		7		8	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
Sedang	0	29	2	27	8	21	13	16	19	10
Tinggi	0	29	0	29	3	26	5	24	8	21
Jumlah	0	58	2	56	11	47	18	40	27	31
Entropy	0,0000	1,0000	0,0000	0,9991	0,8454	0,9918	0,8524	0,9710	0,8767	0,9072
Gain	0,0000		0,0354		0,0360		0,0658		0,1070	

Tabel 3.11 Posisi v untuk pemecahan fitur JP di *node* akar (lanjutan)

	9		10		11		12		13	
	<=	>	<=	>	<=	>	<=	>	<=	>
Sedang	20	9	23	6	23	6	23	6	24	5
Tinggi	11	18	17	12	18	11	22	7	22	7
Jumlah	31	27	40	18	41	17	45	13	46	12
Entropy	0,9383	0,9183	0,9837	0,9183	0,9892	0,9367	0,9996	0,9957	0,9986	0,9799
Gain	0,0710		0,0366		0,0262		0,0012		0,0052	

Selanjutnya, untuk Jumlah Anggota Keluarga (JT) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 4$. Maka untuk fitur JT dilakukan diskretisasi pada $v = 4$ ketika menghitung *Entropy* dan *Gain* pada semua fitur.

Tabel 3.12 Posisi v untuk pemecahan fitur JT di *node* akar

	1		2		3		4		5	
	<=	>	<=	>	<=	>	<=	>	<=	>
Sedang	2	27	8	21	15	14	28	1	29	0
Tinggi	1	28	6	23	14	15	21	8	25	4
Jumlah	3	55	14	44	29	29	49	9	54	4
Entropy	0,9183	0,9998	0,9852	0,9985	0,9991	0,9991	0,9852	0,5033	0,9960	0,0000
Gain	0,0045		0,0047		0,0009		0,0896		0,0727	

Tabel 3.13 Posisi v untuk pemecahan fitur JT di *node* akar (lanjutan)

	6	
	<=	>
Sedang	29	0
Tinggi	28	1
Jumlah	57	1
Entropy	0,9998	0,0000
Gain	0,0175	

Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan.

$$Gain(DL) = 1.0000 - \left(\left(\frac{3}{58} * \left(-\frac{3}{3} * \log_2 \left(\frac{3}{3} \right) - \frac{0}{3} * \log_2 \left(\frac{0}{3} \right) \right) \right) \right)$$

$$\begin{aligned}
& + \frac{19}{58} * \left(-\frac{14}{19} * \log_2 \left(\frac{14}{19} \right) - \frac{5}{19} * \log_2 \left(\frac{5}{19} \right) \right) \\
& + \frac{32}{58} * \left(-\frac{12}{32} * \log_2 \left(\frac{12}{32} \right) - \frac{20}{32} * \log_2 \left(\frac{20}{32} \right) \right) \\
& + \frac{4}{58} * \left(-\frac{0}{4} * \log_2 \left(\frac{0}{4} \right) - \frac{4}{4} * \log_2 \left(\frac{4}{4} \right) \right) \\
& = 1,0000 - 0,7990 \\
& = 0,2010 \\
\\
Gain(PP) & = 1,0000 - \left(\left(\frac{5}{58} * \left(-\frac{5}{5} * \log_2 \left(\frac{5}{5} \right) - \frac{0}{5} * \log_2 \left(\frac{0}{5} \right) \right) \right. \right. \\
& \quad \left. \left. + \frac{12}{58} * \left(-\frac{8}{12} * \log_2 \left(\frac{8}{12} \right) - \frac{4}{12} * \log_2 \left(\frac{4}{12} \right) \right) \right. \right. \\
& \quad \left. \left. + \frac{41}{58} * \left(-\frac{16}{41} * \log_2 \left(\frac{16}{41} \right) - \frac{25}{41} * \log_2 \left(\frac{25}{41} \right) \right) \right) \right) \\
& = 1,0000 - 0,8721 \\
& = 0,1279
\end{aligned}$$

Tabel 3.14 Hasil perhitungan *entropy* dan *gain* untuk *node* akar

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		58	29	29	1,0000	
LR	<= 41	7	1	6	0,5917	0,0554
	> 41	51	28	23	0,9931	
DL	450	3	3	0	0,0000	0,2010
	900	19	14	5	0,8315	
	1.300	32	12	20	0,9544	
	2.200	4	0	4	0,0000	
JP	<= 8	27	19	8	0,8767	0,1070
	> 8	31	10	21	0,9072	
JT	<= 4	49	28	21	0,9852	0,0896
	> 4	9	1	8	0,5033	
PP	Rendah	5	5	0	0,0000	0,1279
	Sedang	12	8	4	0,9183	
	Tinggi	41	16	25	0,9650	

Hasil yang didapat di tabel 3.14 menunjukkan bahwa *Gain* tertinggi ada di fitur Daya Listrik, maka Daya Listrik dijadikan sebagai *node* akar. Selanjutnya, dihitung posisi split untuk fitur Daya Listrik dengan menghitung *Rasio Gain*, selengkapnya disajikan pada tabel 3.19.

Hasil perhitungan *rasio gain* posisi *split* untuk *opsi* satu sebagai berikut:

$$\begin{aligned} SplitInfo(Semua, DL) &= \left(-\frac{3}{58} * \log_2 \left(\frac{3}{58} \right) \right) + \left(-\frac{19}{58} * \log_2 \left(\frac{19}{58} \right) \right) \\ &\quad + \left(-\frac{32}{58} * \log_2 \left(\frac{32}{58} \right) \right) + \left(-\frac{4}{58} * \log_2 \left(\frac{4}{58} \right) \right) \\ &= 1,4879 \end{aligned}$$

$$\begin{aligned} RasioGain(Semua, DL) &= \frac{0,2010}{1,4879} \\ &= 0,1351 \end{aligned}$$

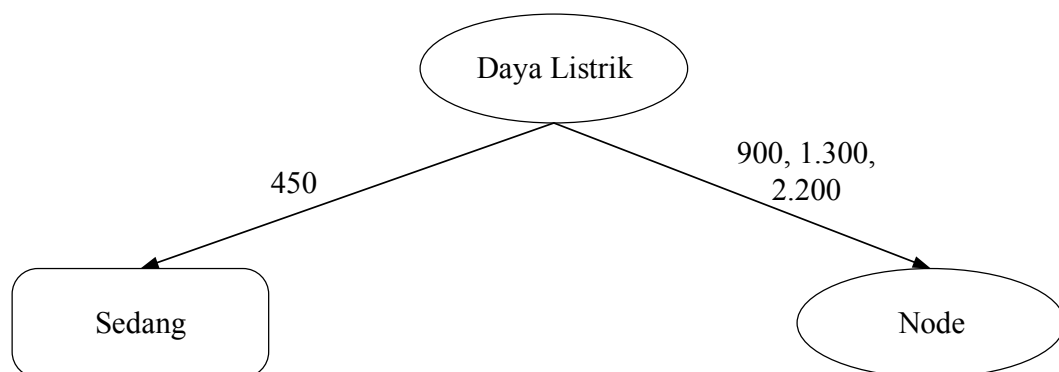
Dengan cara yang sama, akan didapatkan nilai *rasio gain* untuk *opsi* yang lain. Hasil di tabel 3.15 menunjukkan bahwa *rasio gain* tertinggi yaitu *split* {900, 1.300, 2.200} dengan {450}. Itu artinya, cabang untuk akar ada 2, yaitu: {900, 1.300, 2.200} dengan {450}, seperti ditunjukkan pada gambar 3.2.

Tabel 3.15 Perhitungan *Rasio Gain* untuk fitur DL

Total		Jumlah	Entropy	Gain	Rasio Gain
		58		0,2010	
DL	450	3	1,4879		0,1351
	900	19			
	1.300	32			
	2.200	4			
DL	450 dan 900	22	0,9576		0,2099
	1.300 dan 2.200	36			

Tabel 3.15 Perhitungan *Rasio Gain* untuk fitur DL (lanjutan)

DL	450 dan 1.300	35	0,9689		0,2075
	900 dan 2.200	23			
DL	450 dan 2.200	7	0,5313		0,3784
	900 dan 1.300	51			
DL	450, 900, 1.300	54	0,3621		0,5552
	2.200	4			
DL	450, 900, 2.200	26	0,9923		0,2026
	1.300	32			
DL	900, 1.300, 2.200	55	0,2937		0,6845
	450	3			
DL	450, 1.300, 2.200	39	0,9124		0,2203
	900	19			

Gambar 3.2 Hasil pembentukan cabang pada *node* akar

Tabel 3.16 Data pada kasus daya listrik 900, 1.300, dan 2.200

No.	Luas Rumah (m²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
6	51	900	6	4	Sedang	Sedang
7	84	900	9	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
13	26	900	10	3	Rendah	Sedang
14	120	1300	6	4	Tinggi	Sedang
17	112	1300	20	2	Tinggi	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
20	96	1300	13	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
22	126	900	10	3	Tinggi	Sedang
24	102	900	19	4	Sedang	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
30	112	2200	9	5	Tinggi	Tinggi
31	77	1300	14	4	Tinggi	Tinggi
32	40	900	6	5	Tinggi	Tinggi
33	80	2200	20	4	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
35	72	2200	8	3	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
37	176	900	11	6	Sedang	Tinggi
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi

Tabel 3.16 Data pada kasus daya listrik 900, 1.300, dan 2.200 (lanjutan)

40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
44	140	2200	9	4	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
54	40	900	8	3	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
57	40	900	12	6	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Daya Listrik dengan nilai data 450, 900, dan 1.300. Hasil pemecahan tersebut menyisakan 54 data yang terdiri dari 29 data dengan kelas Sedang dan 25 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{26}{55} * \log_2 \left(\frac{26}{55} \right) - \frac{29}{55} * \log_2 \left(\frac{29}{55} \right) \\ &= 0,9979 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur, untuk fitur yang bertipe *numeric*, harus ditentukan posisi v yang terbaik untuk pemecahan. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 41$. Untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 8$. Sedangkan untuk Jumlah

Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 4$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan.

$$\begin{aligned}
 \text{Gain(DL)} &= 0,9979 - \left(\left(\frac{19}{55} * \left(-\frac{14}{19} * \log_2 \left(\frac{14}{19} \right) - \frac{5}{19} * \log_2 \left(\frac{5}{19} \right) \right) \right) \right. \\
 &\quad \left. + \frac{32}{55} * \left(-\frac{12}{32} * \log_2 \left(\frac{12}{32} \right) - \frac{20}{32} * \log_2 \left(\frac{20}{32} \right) \right) \right) \\
 &\quad \left. + \frac{4}{55} * \left(-\frac{0}{4} * \log_2 \left(\frac{0}{4} \right) - \frac{4}{4} * \log_2 \left(\frac{4}{4} \right) \right) \right) \\
 &= 0,1553
 \end{aligned}$$

$$\begin{aligned}
 \text{Gain(PP)} &= 0,9979 - \left(\left(\frac{4}{55} * \left(-\frac{4}{4} * \log_2 \left(\frac{4}{4} \right) - \frac{0}{4} * \log_2 \left(\frac{0}{4} \right) \right) \right) \right. \\
 &\quad \left. + \frac{12}{55} * \left(-\frac{8}{12} * \log_2 \left(\frac{8}{12} \right) - \frac{4}{12} * \log_2 \left(\frac{4}{12} \right) \right) \right) \\
 &\quad \left. + \frac{39}{55} * \left(-\frac{14}{39} * \log_2 \left(\frac{14}{39} \right) - \frac{25}{39} * \log_2 \left(\frac{25}{39} \right) \right) \right) \\
 &= 0,1279
 \end{aligned}$$

Tabel 3.17 Hasil perhitungan *entropy* dan *gain* untuk *node* 3

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		55	26	29	0,9979	
LR	≤ 41	7	1	6	0,5917	0,0509
	> 41	48	25	23	0,9987	
DL	900	19	14	5	0,8315	0,1553
	1.300	32	12	20	0,9544	
	2.200	4	0	4	0,0000	
JP	≤ 8	26	18	8	0,8905	0,1288
	> 8	29	8	21	0,8498	
JT	≤ 4	46	25	21	0,9945	0,0837
	> 4	9	1	8	0,5033	
PP	Rendah	4	4	0	0,0000	0,1297
	Sedang	12	8	4	0,9183	
	Tinggi	39	14	25	0,9418	

Hasil yang didapat di tabel 3.17 menunjukkan bahwa *Gain* tertinggi ada di fitur Daya Listrik, maka Daya Listrik dijadikan sebagai *node* akar. Selanjutnya, dihitung posisi split untuk fitur Daya Listrik dengan menghitung *Rasio Gain*, selengkapnya disajikan pada tabel 3.32. Hasil perhitungan *rasio gain* posisi *split* untuk *opsi* satu sebagai berikut:

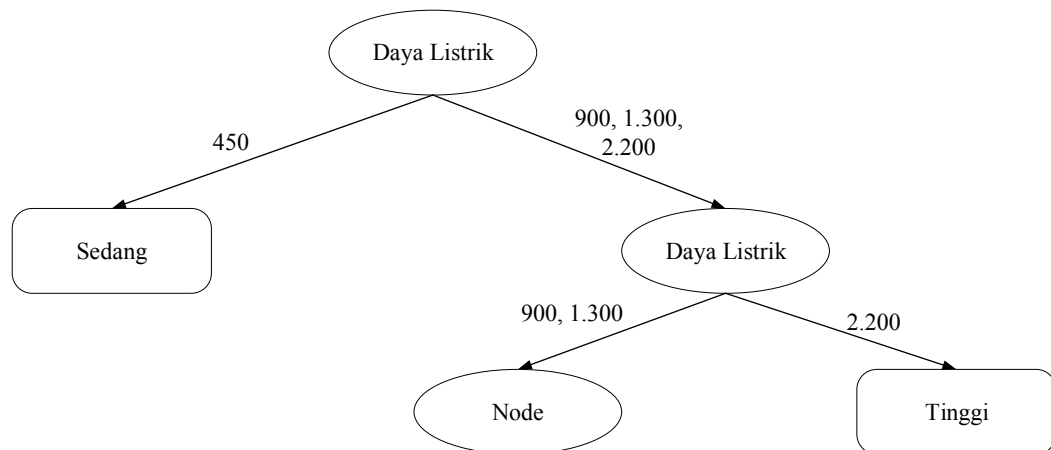
$$\begin{aligned} SplitInfo(Semua, DL) &= \left(-\frac{19}{55} * \log_2 \left(\frac{19}{55} \right) \right) + \left(-\frac{32}{55} * \log_2 \left(\frac{32}{55} \right) \right) \\ &\quad + \left(-\frac{4}{55} * \log_2 \left(\frac{4}{55} \right) \right) \\ &= 1,2593 \end{aligned}$$

$$RasioGain(Semua, DL) = \frac{0,1553}{1,2593} = 0,1233$$

Dengan cara yang sama, akan didapatkan nilai *rasio gain* untuk *opsi* yang lain. Hasil ditabel 3.18 menunjukkan bahwa *rasio gain* tertinggi ada di *opsi* 6 yaitu *split* {2.200} dengan {900, 1.300}. Itu artinya, cabang untuk akar ada 2, yaitu: {2.200} dengan {900, 1.300}, seperti ditunjukkan pada gambar 3.3.

Tabel 3.18 Perhitungan *Rasio Gain* untuk fitur DL

Total		Jumlah	Entropy	Gain	Rasio Gain
		55		0,1553	
DL	900	19	1,2593		0,1233
	1.300	32			
	2.200	4			
DL	900	19	0,9299		0,1670
	1.300 dan 2.200	36			
DL	1.300	32	0,9806		0,1584
	900 dan 2.200	23			
DL	2.200	4	0,376		0,4131
	900 dan 1.300	51			



Gambar 3.3 Pohon keputusan node 2

Tabel 3.19 Data pada kasus daya listrik 900, dan 1.300

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
6	51	900	6	4	Sedang	Sedang
7	84	900	9	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
13	26	900	10	3	Rendah	Sedang
14	120	1300	6	4	Tinggi	Sedang
17	112	1300	20	2	Tinggi	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
20	96	1300	13	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
22	126	900	10	3	Tinggi	Sedang

Tabel 3.19 Data pada kasus daya listrik 900, dan 1.300 (lanjutan)

24	102	900	19	4	Sedang	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
31	77	1300	14	4	Tinggi	Tinggi
32	40	900	6	5	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
37	176	900	11	6	Sedang	Tinggi
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
54	40	900	8	3	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
57	40	900	12	6	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Daya Listrik dengan nilai data 900, dan 1.300. Hasil pemecahan tersebut menyisakan 51 data yang terdiri dari 26 data dengan kelas Sedang dan 25 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

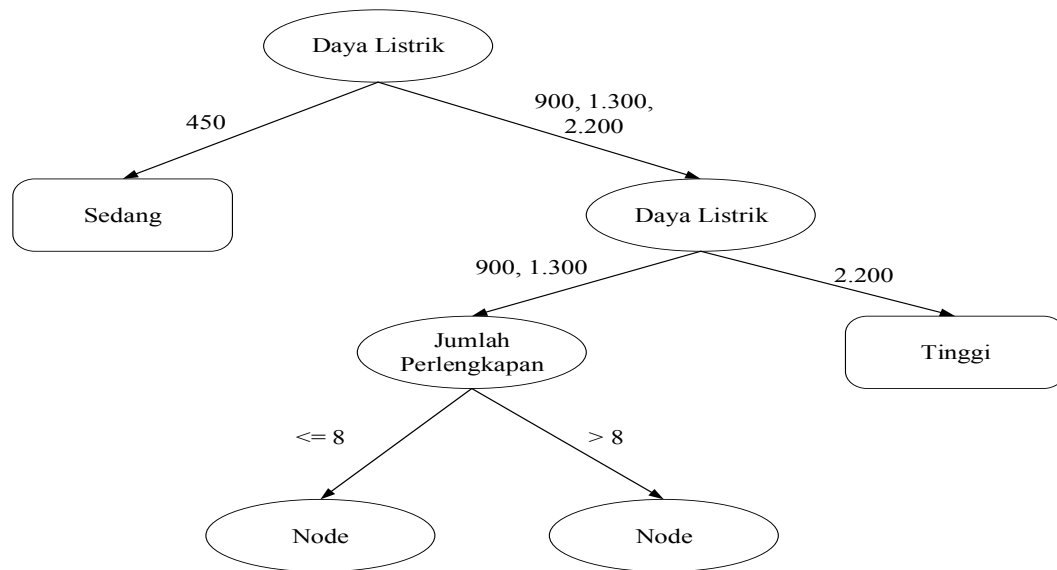
$$\begin{aligned} \text{Entropy}(S) &= -\frac{26}{51} * \log_2\left(\frac{26}{51}\right) - \frac{25}{51} * \log_2\left(\frac{25}{51}\right) \\ &= 0,9997 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 41$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 8$. Sedangkan, untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 4$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti dalam tabel 3.20 dibawah ini.

Tabel 3.20 Hasil perhitungan *entropy* dan *gain* untuk *node 3*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		51	26	25	0,9997	
LR	≤ 41	7	1	6	0,5917	0,0674
	> 41	44	25	19	0,9865	
DL	900	19	14	5	0,8315	0,0911
	1.300	32	12	20	0,9544	
JP	≤ 8	25	18	7	0,8555	0,1264
	> 8	26	8	18	0,8905	
JT	≤ 5	47	26	21	0,9918	0,0857
	> 5	4	0	4	0,0000	
PP	Rendah	4	4	0	0,0000	0,1173
	Sedang	12	8	4	0,9183	
	Tinggi	35	14	21	0,9710	

Dari perhitungan tersebut diperoleh fitur jumlah perlengkapan (JP) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.4 Pohon keputusan node 3

Tabel 3.21 Data pada kasus JP ≤ 8

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
6	51	900	6	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
14	120	1300	6	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
32	40	900	6	5	Tinggi	Tinggi

Tabel 3.21 Data pada kasus JP ≤ 8 (lanjutan)

38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi
54	40	900	8	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Jumlah Perlengkapan. Hasil pemecahan tersebut menyisakan 25 data yang terdiri dari 18 data dengan kelas Sedang dan 7 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{18}{25} * \log_2\left(\frac{18}{25}\right) - \frac{7}{25} * \log_2\left(\frac{7}{25}\right) \\ &= 0,8555 \end{aligned}$$

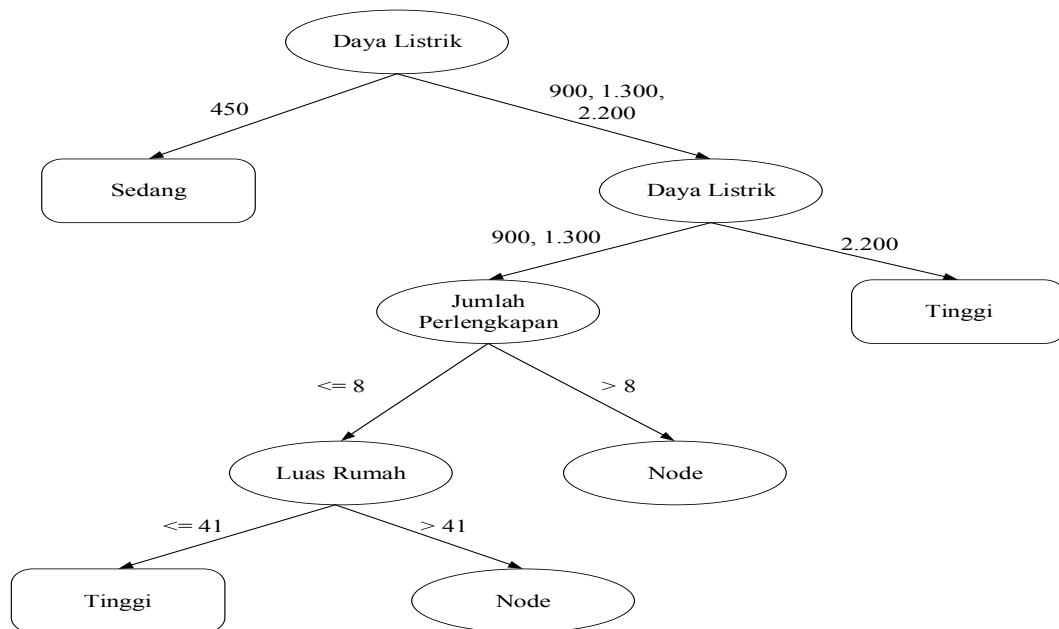
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 41$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 3$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti pada tabel 3.22 dibawah ini.

Tabel 3.22 Hasil perhitungan *entropy* dan *gain* untuk *node* 5

		Jumlah	Sedang	Tinggi	Entropy	Gain
Total		25	18	7	0,8555	
LR	≤ 41	2	0	2	0,0000	0,1605
	> 41	23	18	5	0,7554	
DL	900	11	8	3	0,8454	0,0001
	1.300	14	10	4	0,8631	

JP	<= 5	2	2	0	0,0000	0,0398
	> 5	23	16	7	0,8865	
JT	<= 3	15	9	6	0,9710	0,0853
	> 3	10	9	1	0,4690	
PP	Rendah	3	3	0	0,0000	0,0616
	Sedang	6	4	2	0,9183	
	Tinggi	16	11	5	0,8960	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.5 Pohon keputusan node 5

Tabel 3.23 Data pada kasus LR > 41

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
1	67,5	900	6	4	Rendah	Sedang
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang

6	51	900	6	4	Sedang	Sedang
8	120	1300	6	4	Rendah	Sedang
9	84	900	7	4	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
12	84	1300	7	4	Tinggi	Sedang
14	120	1300	6	4	Tinggi	Sedang
21	98	1300	7	4	Tinggi	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
27	61	900	6	5	Tinggi	Sedang
28	61	1300	6	4	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 23 data yang terdiri dari 18 data dengan kelas Sedang dan 5 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

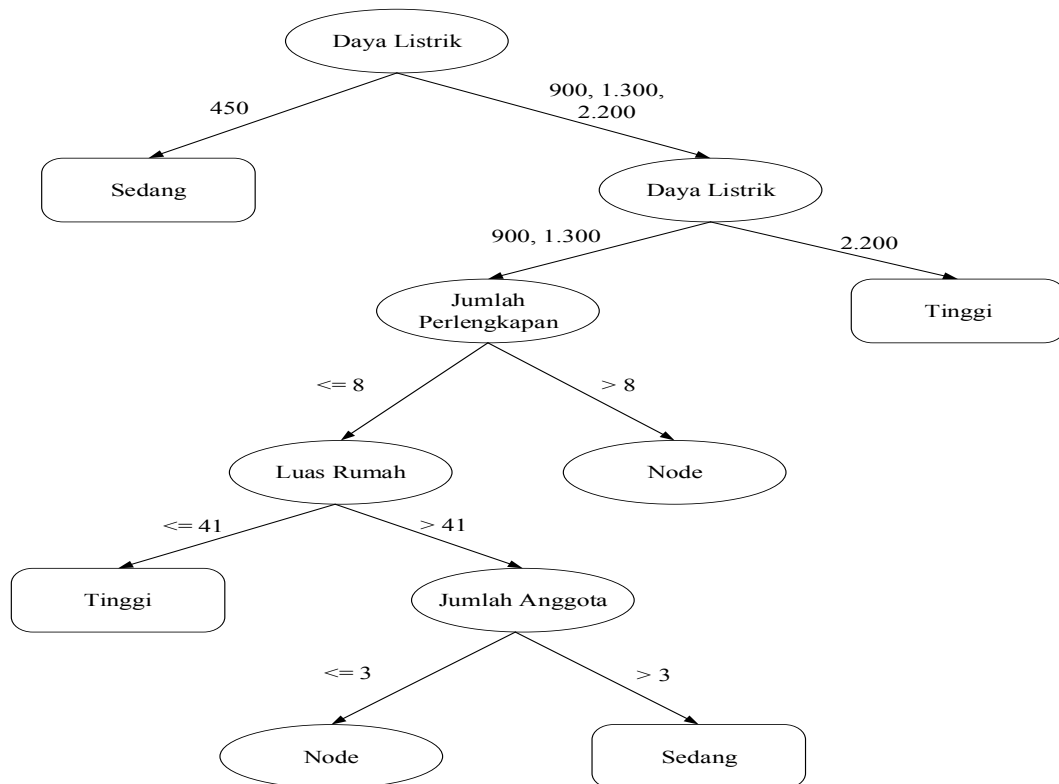
$$\begin{aligned} Entropy(S) &= -\frac{18}{23} * \log_2\left(\frac{18}{23}\right) - \frac{5}{23} * \log_2\left(\frac{5}{23}\right) \\ &= 0,7554 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 71$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 3$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil dalam tabel 3.24 dibawah ini.

Tabel 3.24 Hasil perhitungan *entropy* dan *gain* untuk *node 8*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		23	18	5	0,7554	
LR	≤ 71	6	6	0	0,0000	0,1094
	> 71	17	12	5	0,8740	
DL	900	9	8	1	0,5033	0,0331
	1.300	14	10	4	0,8631	
JP	≤ 5	2	2	0	0,0000	0,0324
	> 5	21	16	5	0,7919	
JT	≤ 3	14	9	5	0,9403	0,1830
	> 3	9	9	0	0,0000	
	Rendah	3	3	0	0,0000	
	Sedang	6	4	2	0,9183	
	Tinggi	14	11	3	0,7496	

Dari perhitungan tersebut diperoleh fitur Jumlah Anggota Keluarga (JT) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.6 Pohon keputusan node 8

Tabel 3.25 Data pada kasus JT ≤ 3

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
11	92	1300	5	3	Sedang	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
38	100	900	7	2	Tinggi	Tinggi
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 14 data yang terdiri dari 9 data dengan kelas Sedang dan 5 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

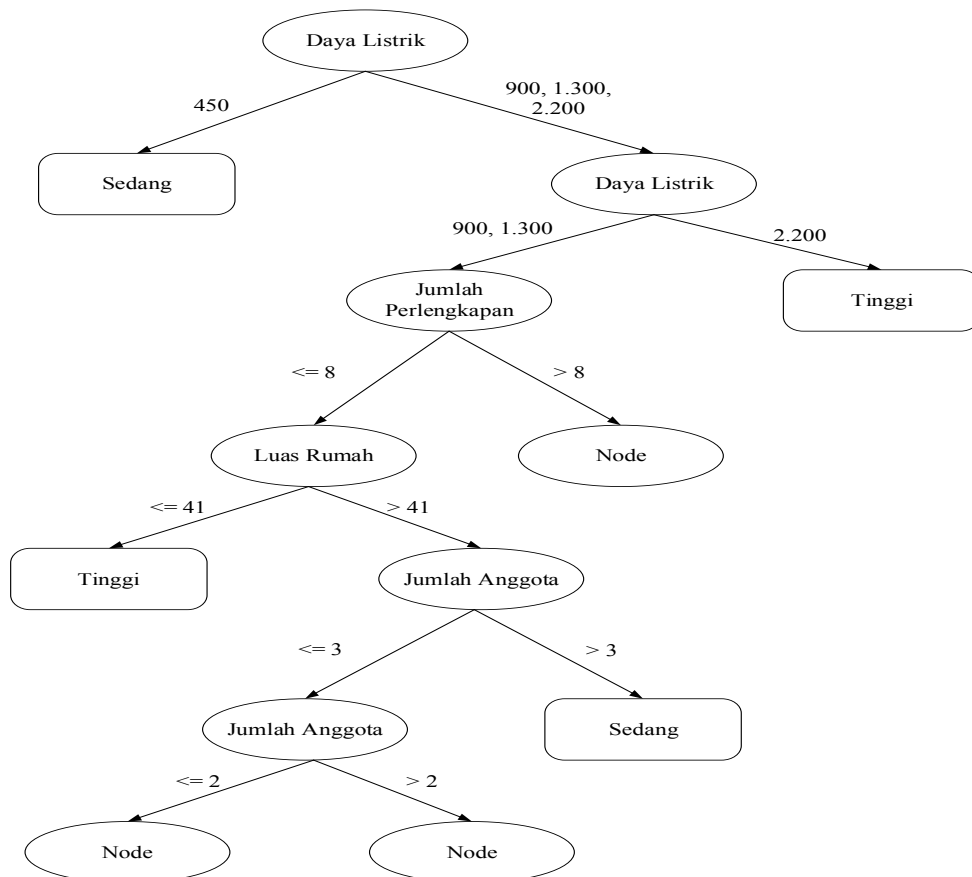
$$\begin{aligned} Entropy(S) &= -\frac{9}{14} * \log_2\left(\frac{9}{14}\right) - \frac{5}{14} * \log_2\left(\frac{5}{14}\right) \\ &= 0,9403 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 71$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 2$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti pada tabel dibawah ini.

Tabel 3.26 Hasil perhitungan *entropy* dan *gain* untuk *node 9*

		Jumlah	Sedang	Tinggi	Entropy	Gain
Total		14	9	5	0,9403	
LR	<= 71	2	2	0	0,0000	0,1004
	> 71	12	7	5	0,9799	
DL	900	5	4	1	0,7219	0,0453
	1.300	9	5	4	0,9911	
	> 5	12	7	5	0,9799	
JT	<= 2	7	6	1	0,5917	0,1518
	> 2	7	3	4	0,9852	
PP	Rendah	1	1	0	0,0000	0,0481
	Sedang	5	3	2	0,9710	
	Tinggi	8	5	3	0,9544	

Dari perhitungan tersebut diperoleh fitur Jumlah Anggota Keluarga (JT) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.7 Pohon keputusan node 9

Tabel 3.27 Data pada kasus JT ≤ 2

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
4	92	900	8	1	Tinggi	Sedang
5	81	900	8	1	Tinggi	Sedang
10	100	900	5	2	Rendah	Sedang
25	61	1300	7	2	Sedang	Sedang
26	133	1300	8	2	Tinggi	Sedang
29	73	900	7	2	Tinggi	Sedang
38	100	900	7	2	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 7 data yang terdiri dari 6 data dengan kelas Sedang dan 1 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{6}{7} * \log_2\left(\frac{6}{7}\right) - \frac{1}{7} * \log_2\left(\frac{1}{7}\right) \\ &= 0,5917 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 96$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 7$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 1$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti pada tabel dibawah ini.

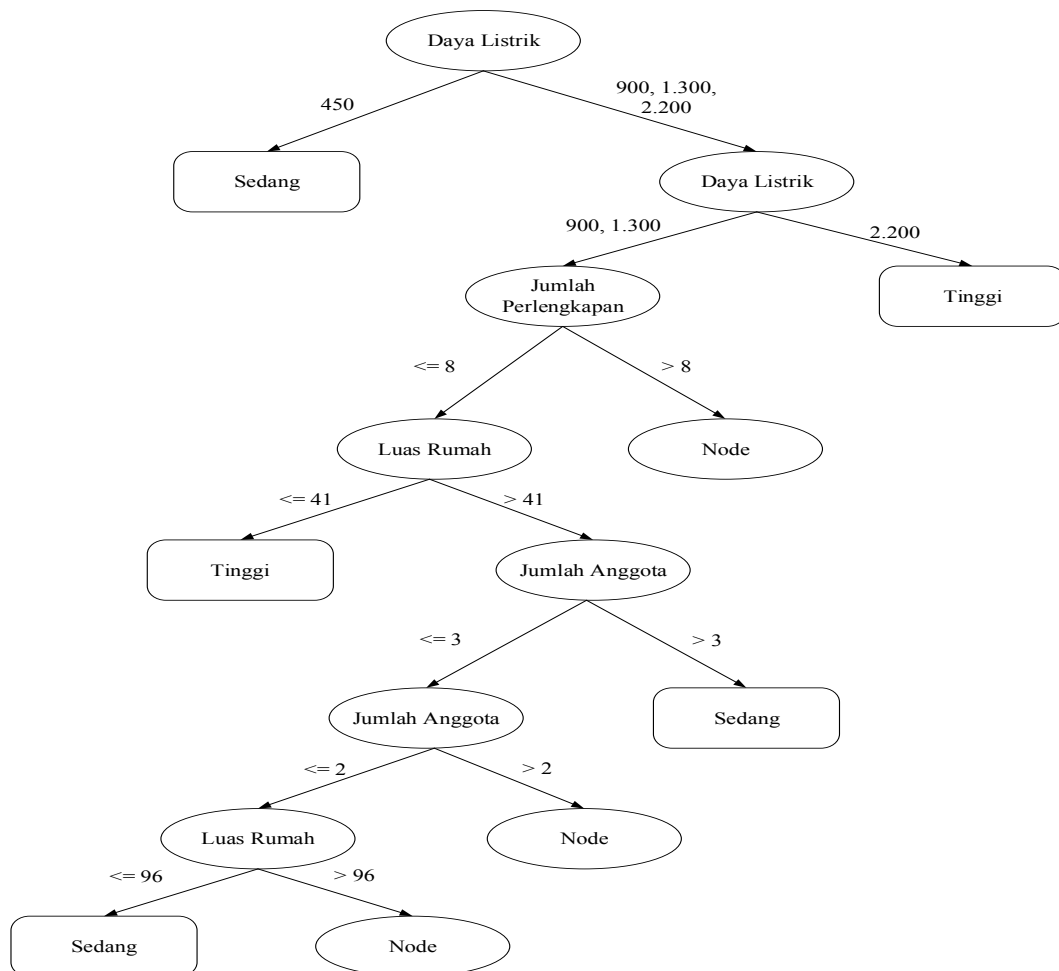
Tabel 3.28 Hasil perhitungan *entropy* dan *gain* untuk *node* 11

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		7	6	1	0,5917	
LR	≤ 96	4	4	0	0,0000	0,1981
	> 96	3	2	1	0,9183	

Tabel 3.28 Hasil perhitungan *entropy* dan *gain* untuk *node* 11 (lanjutan)

DL	900	5	4	1	0,7219	0,0760
	1.300	2	2	0	0,0000	
JP	<= 7	4	3	1	0,8113	0,1281
	> 7	3	3	0	0,0000	
JT	<= 1	2	2	0	0,0000	0,0760
	> 1	5	4	1	0,7219	
PP	Rendah	1	1	0	0,0000	0,0760
	Sedang	1	1	0	0,0000	
	Tinggi	5	4	1	0,7219	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.8 Pohon keputusan node 13

Tabel 3.29 Data pada kasus LR > 96

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
10	100	900	5	2	Rendah	Sedang
26	133	1300	8	2	Tinggi	Sedang
38	100	900	7	2	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 3 data yang terdiri dari 2 data dengan kelas Sedang dan 1 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{2}{3} * \log_2\left(\frac{2}{3}\right) - \frac{1}{3} * \log_2\left(\frac{1}{3}\right) \\ &= 0,9183 \end{aligned}$$

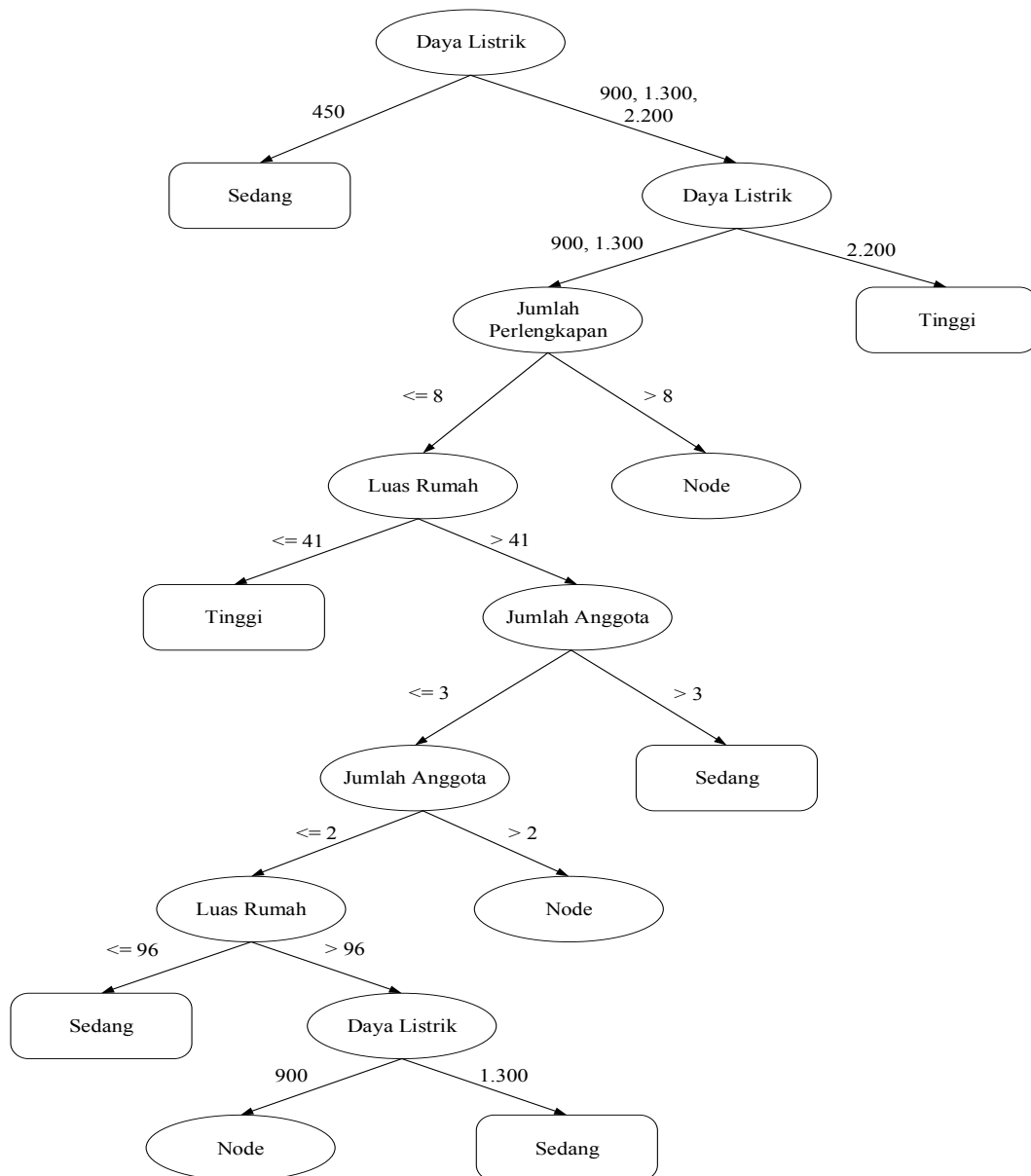
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 101$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 1$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti dibawah ini.

Tabel 3.30 Hasil perhitungan *entropy* dan *gain* untuk node 14

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		3	2	1	0,9183	
LR	<= 101	2	1	1	1,0000	0,2516
	> 101	1	1	0	0,0000	
DL	900	2	1	1	1,0000	0,2516
	1.300	1	1	0	0,0000	
JP	<= 5	1	1	0	1,0000	0,2516
	> 5	2	1	1	0,0000	

JT	≤ 1	0	0	0	0,0000	0,0000
	> 1	3	2	1	0,9183	
PP	Rendah	1	1	0	0,0000	0,2516
	Sedang	0	0	0	0,0000	
	Tinggi	2	1	1	1,0000	

Dari perhitungan tersebut diperoleh fitur Daya Listrik (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.9 Pohon keputusan node 14

Tabel 3.31 Data pada kasus DL = 900

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
10	100	900	5	2	Rendah	Sedang
38	100	900	7	2	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Daya Listrik. Hasil pemecahan tersebut menyisakan 2 data yang terdiri dari 1 data dengan kelas Sedang dan 1 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{1}{2} * \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} * \log_2 \left(\frac{1}{2}\right) \\ &= 1,0000 \end{aligned}$$

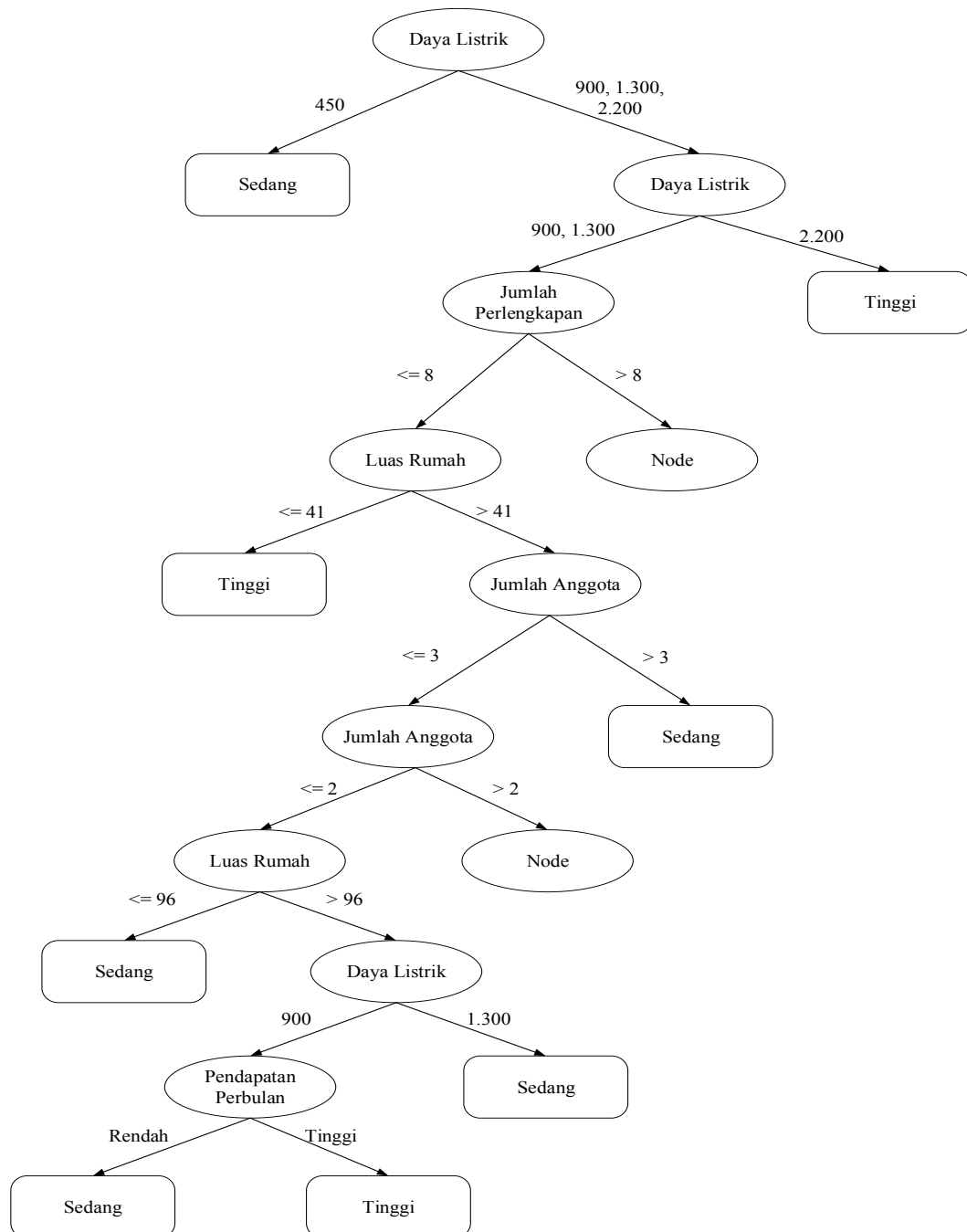
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 116$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 1$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti tabel dibawah ini.

Tabel 3.32 Hasil perhitungan *entropy* dan *gain* untuk node 15

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		2	1	1	1,0000	
LR	≤ 116	2	1	1	1,0000	0,0000
	> 116	0	0	0	0,0000	
DL	900	2	1	1	1,0000	0,0000
	1.300	0	0	0	0,0000	
JP	≤ 5	1	1	0	0,0000	1,0000
	> 5	1	0	1	0,0000	
JT	≤ 1	0	0	0	0,0000	0,0000
	> 1	2	1	1	1,0000	

PP	Rendah	1	1	0	0,0000	1,0000
	Sedang	0	0	0	0,0000	
	Tinggi	1	0	1	0,0000	

Dari perhitungan tersebut diperoleh fitur Jumlah Perlengkapan memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.10 Pohon keputusan node 15

Tabel 3.33 Data pada kasus JT > 2

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
2	90	1300	8	3	Tinggi	Sedang
3	68	1300	8	3	Sedang	Sedang
11	92	1300	5	3	Sedang	Sedang
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 7 data yang terdiri dari 3 data dengan kelas Sedang dan 4 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{3}{7} * \log_2 \left(\frac{3}{7} \right) - \frac{4}{7} * \log_2 \left(\frac{4}{7} \right) \\ &= 0,9852 \end{aligned}$$

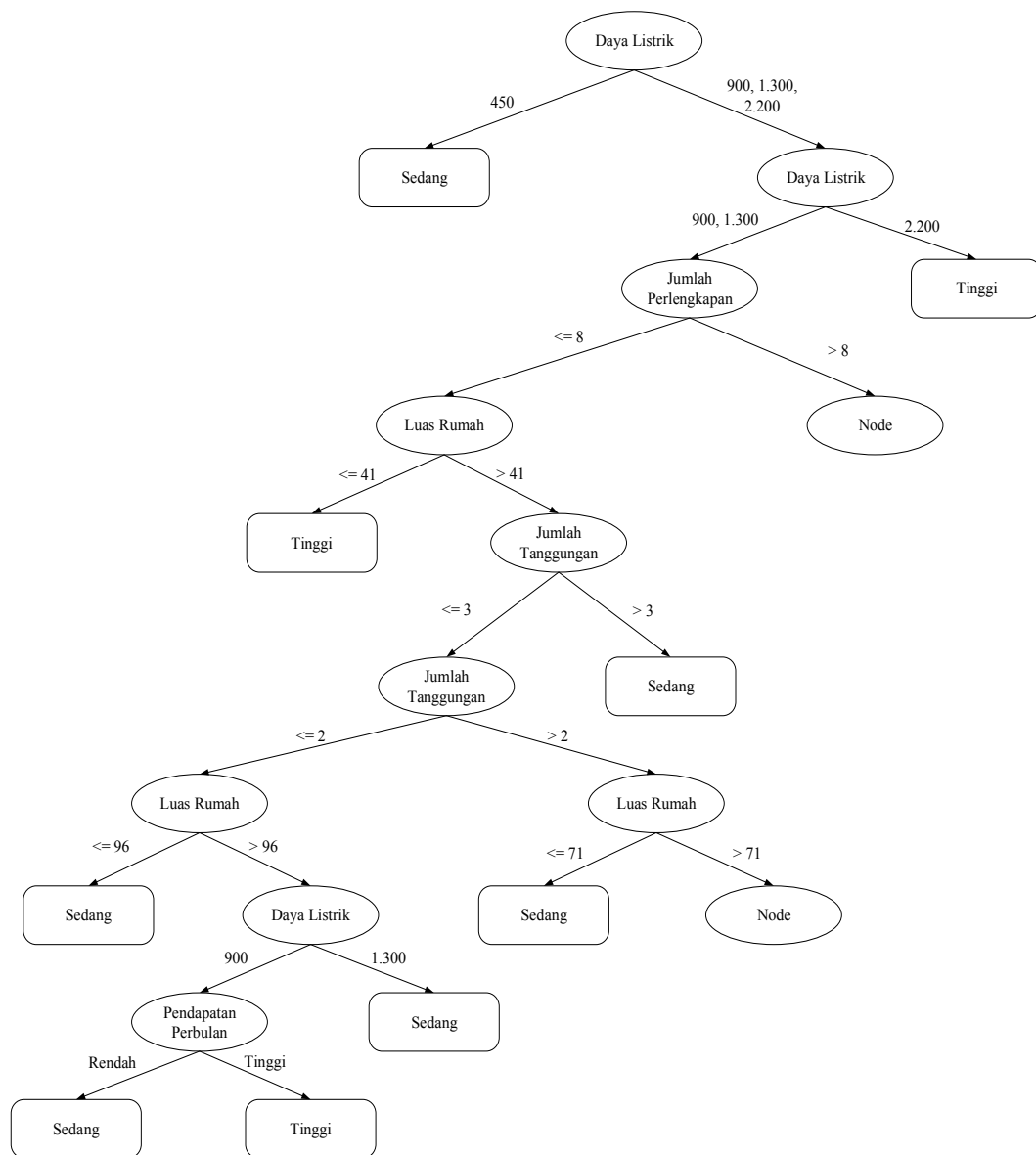
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 71$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti dibawah ini.

Tabel 3.34 Hasil perhitungan *entropy* dan *gain* untuk *node* 12

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		7	3	4	0,9852	
LR	≤ 71	1	1	0	0,0000	0,1981
	> 71	6	2	4	0,9183	
DL	900	0	0	0	0,0000	0,0000
	1.300	7	3	4	0,9852	

JP	≤ 5	1	1	0	0,0000	0,1981
	> 5	6	2	4	0,9183	
PP	Rendah	0	0	0	0,0000	0,0202
	Sedang	4	2	2	1,0000	
	Tinggi	3	1	2	0,9183	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.11 Pohon keputusan node 12

Tabel 3.35 Data pada kasus LR > 71

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
2	90	1300	8	3	Tinggi	Sedang
11	92	1300	5	3	Sedang	Sedang
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 6 data yang terdiri dari 2 data dengan kelas Sedang dan 4 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{2}{6} * \log_2\left(\frac{2}{6}\right) - \frac{4}{6} * \log_2\left(\frac{4}{6}\right) \\ &= 0,9183 \end{aligned}$$

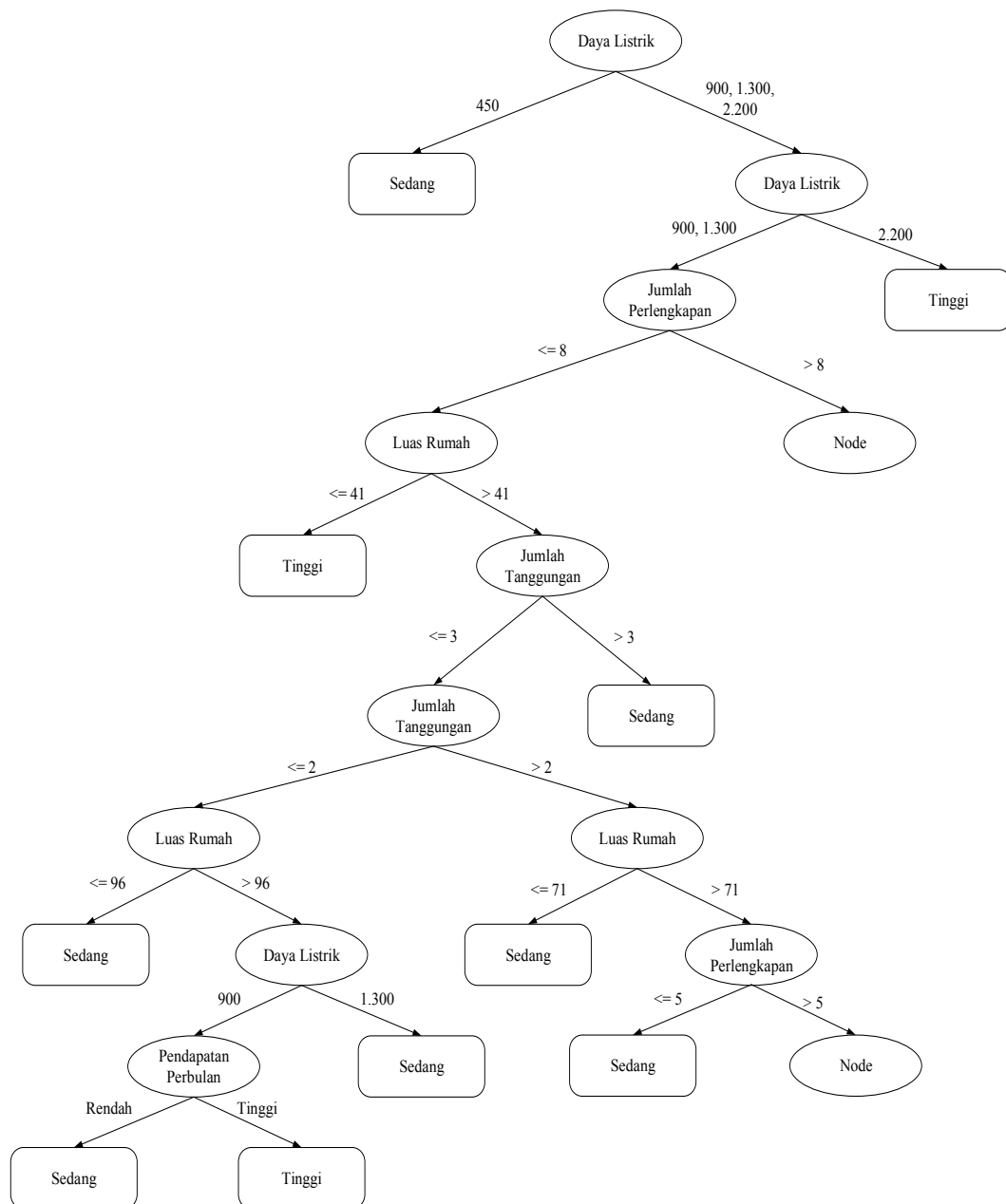
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 86$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 5$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan dengan hasil seperti pada tabel dibawah ini.

Tabel 3.36 Hasil perhitungan *entropy* dan *gain* untuk *node* 20

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		6	2	4	0,9183	
LR	<= 86	2	0	2	0,0000	0,2516
	> 86	4	2	2	1,0000	
DL	900	0	0	0	0,0000	0,0000
	1.300	6	2	4	0,9183	

JP	≤ 5	1	1	0	0,0000	0,3167
	> 5	5	1	4	0,7219	
PP	Rendah	0	0	0	0,0000	0,0000
	Sedang	3	1	2	0,9183	
	Tinggi	3	1	2	0,9183	

Dari perhitungan tersebut diperoleh fitur Jumlah Perlengkapan (JP) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.12 Pohon keputusan node 20

Tabel 3.37 Data pada kasus $JP > 5$

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
2	90	1300	8	3	Tinggi	Sedang
39	92	1300	8	3	Tinggi	Tinggi
40	72	1300	6	3	Tinggi	Tinggi
41	72	1300	7	3	Sedang	Tinggi
42	92	1300	6	3	Sedang	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 5 data yang terdiri dari 1 data dengan kelas Sedang dan 4 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{1}{5} * \log_2 \left(\frac{1}{5} \right) - \frac{4}{6} * \log_2 \left(\frac{4}{6} \right) \\ &= 0,7219 \end{aligned}$$

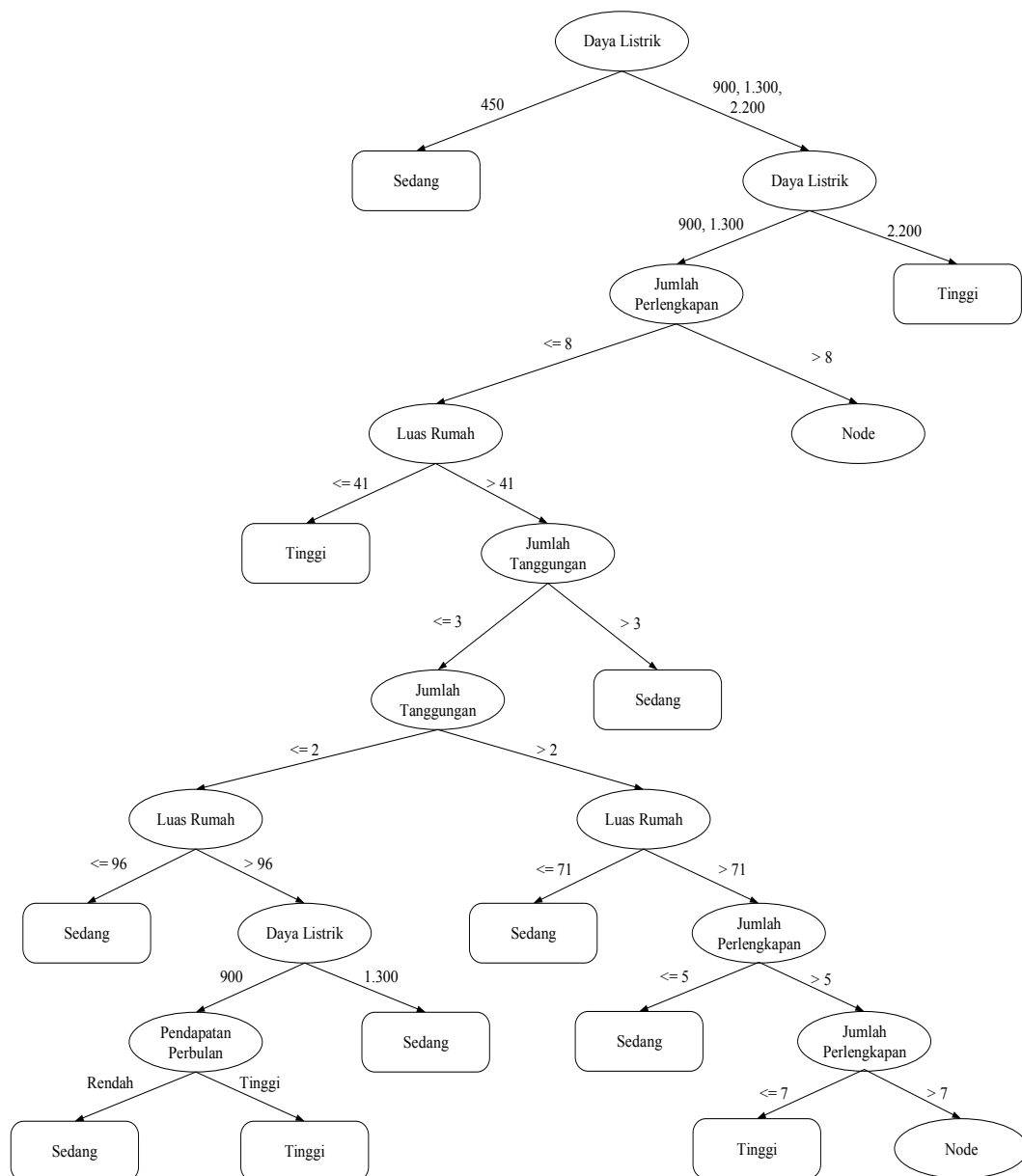
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 76$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 7$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan (PP) seperti pada tabel dibawah ini.

Tabel 3.38 Hasil perhitungan *entropy* dan *gain* untuk *node 22*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		5	1	4	0,7219	
LR	≤ 76	2	0	2	0,0000	0,1710
	> 76	3	1	2	0,9183	
DL	900	0	0	0	0,0000	0,0000
	1.300	5	1	4	0,7219	

JP	≤ 7	3	0	3	0,0000	0,3219
	> 7	2	1	1	1,0000	
PP	Rendah	0	0	0	0,0000	0,1710
	Sedang	2	0	2	0,0000	
	Tinggi	3	1	2	0,9183	

Dari perhitungan tersebut diperoleh fitur Jumlah Perlengkapan (JP) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.13 Pohon keputusan node 22

Tabel 3.39 Data pada kasus $JP > 7$

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
2	90	1300	8	3	Tinggi	Sedang
39	92	1300	8	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 2 data yang terdiri dari 1 data dengan kelas Sedang dan 1 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

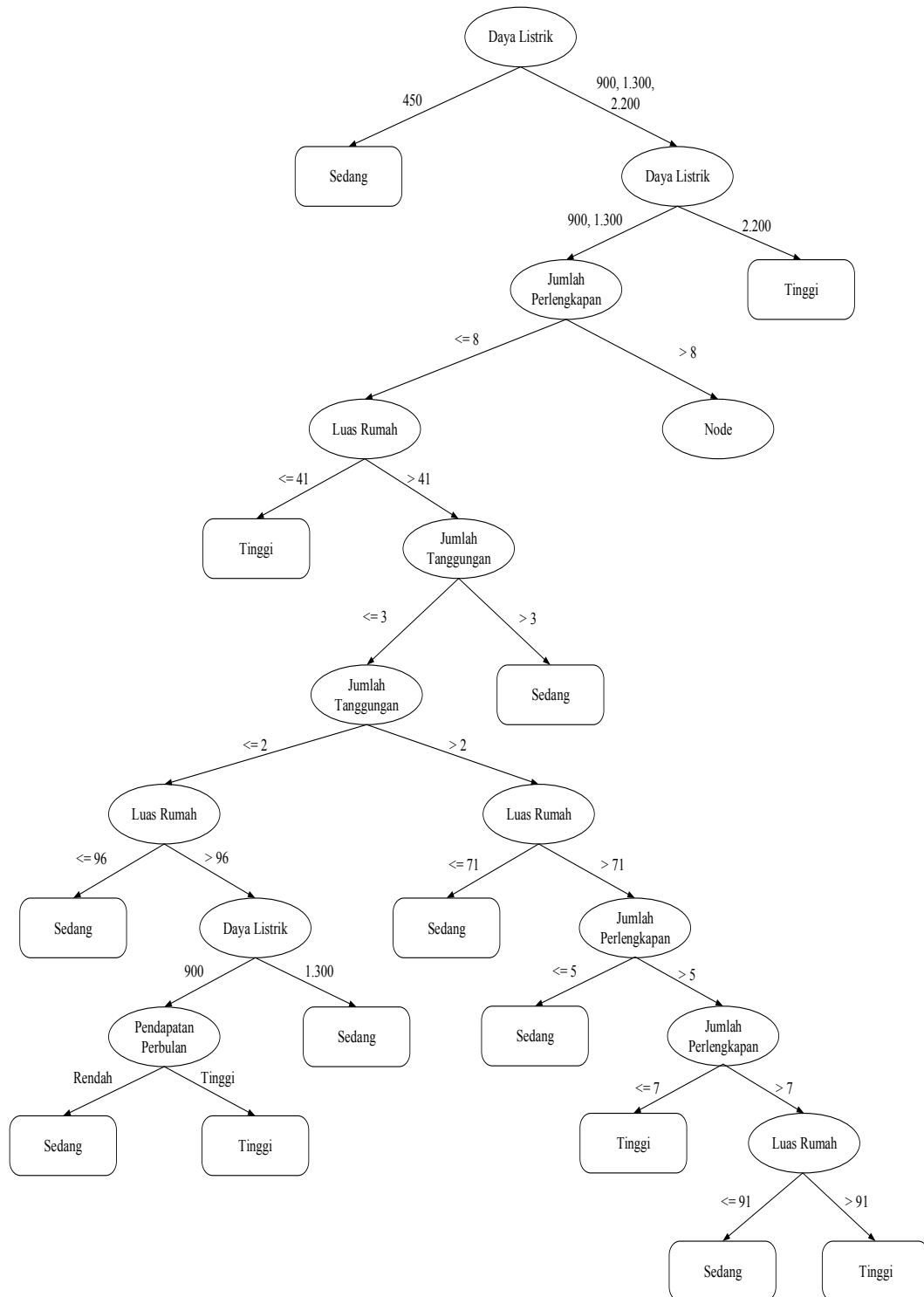
$$\begin{aligned} Entropy(S) &= -\frac{1}{2} * \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} * \log_2 \left(\frac{1}{2} \right) \\ &= 1,0000 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 91$. Kemudian, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan (PP) seperti pada tabel dibawah ini.

Tabel 3.40 Hasil perhitungan *entropy* dan *gain* untuk *node 24*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		2	1	1	1,0000	
LR	≤ 91	1	1	0	0,0000	1,0000
	> 91	1	0	1	0,0000	
DL	900	0	0	0	0,0000	0,0000
	1.300	2	1	1	1,0000	
PP	Rendah	0	0	0	0,0000	0,0000
	Sedang	0	0	0	0,0000	
	Tinggi	2	1	1	1,0000	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.14 Pohon keputusan node 24

Tabel 3.41 Data pada kasus $JP > 8$

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
7	84	900	9	4	Sedang	Sedang
13	26	900	10	3	Rendah	Sedang
17	112	1300	20	2	Tinggi	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
20	96	1300	13	4	Tinggi	Sedang
22	126	900	10	3	Tinggi	Sedang
24	102	900	19	4	Sedang	Sedang
31	77	1300	14	4	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
37	176	900	11	6	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
57	40	900	12	6	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 26 data yang terdiri dari 8 data dengan kelas Sedang dan 18 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

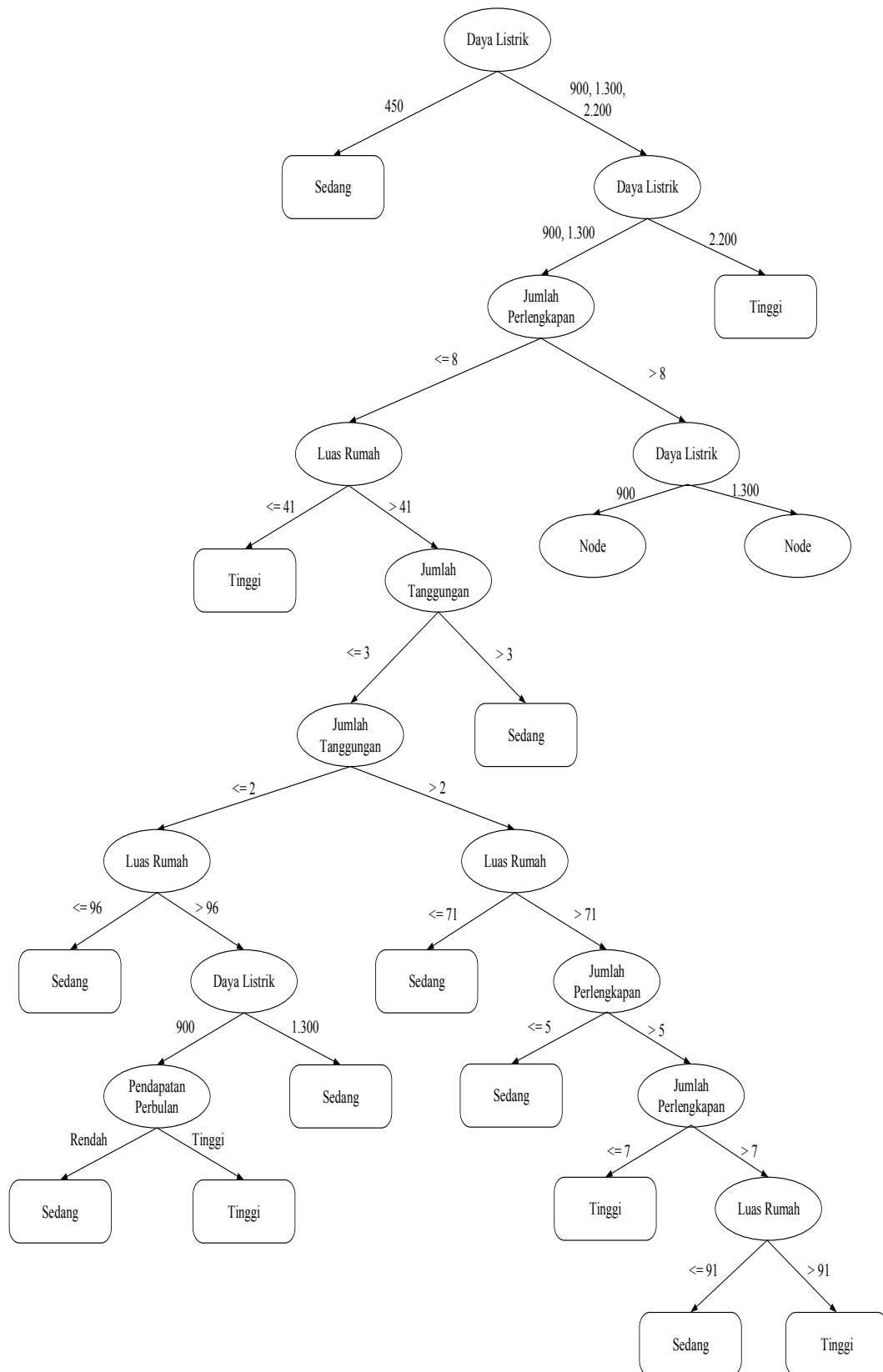
$$\begin{aligned} Entropy(S) &= -\frac{8}{26} * \log_2\left(\frac{8}{26}\right) - \frac{18}{26} * \log_2\left(\frac{18}{26}\right) \\ &= 0,8905 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 81$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai gain tertinggi didapatkan pada posisi $v = 12$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai gain tertinggi didapatkan pada posisi $v = 4$. Selanjutnya, hitung nilai *gain* untuk fitur Daya Listrik (DL) dan Pendapatan Perbulan seperti pada tabel dibawah ini.

Tabel 3.42 Hasil perhitungan *entropy* dan *gain* untuk *node 6*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		26	8	18	0,8905	
LR	≤ 81	15	3	12	0,7219	0,0534
	> 81	11	5	6	0,9940	
DL	900	8	6	2	0,8113	0,2925
	1.300	18	2	16	0,5033	
JP	≤ 12	15	3	12	0,7219	0,0534
	> 12	11	5	6	0,9940	
JT	≤ 4	20	8	12	0,9710	0,1436
	> 4	6	0	6	0,0000	
PP	Rendah	1	1	0	0,0000	0,2187
	Sedang	6	4	2	0,9183	
	Tinggi	19	3	16	0,6292	

Dari perhitungan tersebut diperoleh fitur Daya Listrik (DL) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.15 Pohon keputusan node 6

Tabel 3.43 Data pada kasus DL = 900

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
7	84	900	9	4	Sedang	Sedang
13	26	900	10	3	Rendah	Sedang
18	52	900	21	3	Sedang	Sedang
19	42	900	16	2	Sedang	Sedang
22	126	900	10	3	Tinggi	Sedang
24	102	900	19	4	Sedang	Sedang
37	176	900	11	6	Sedang	Tinggi
57	40	900	12	6	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 8 data yang terdiri dari 6 data dengan kelas Sedang dan 2 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{6}{8} * \log_2 \left(\frac{6}{8} \right) - \frac{2}{8} * \log_2 \left(\frac{2}{8} \right) \\ &= 0,8113 \end{aligned}$$

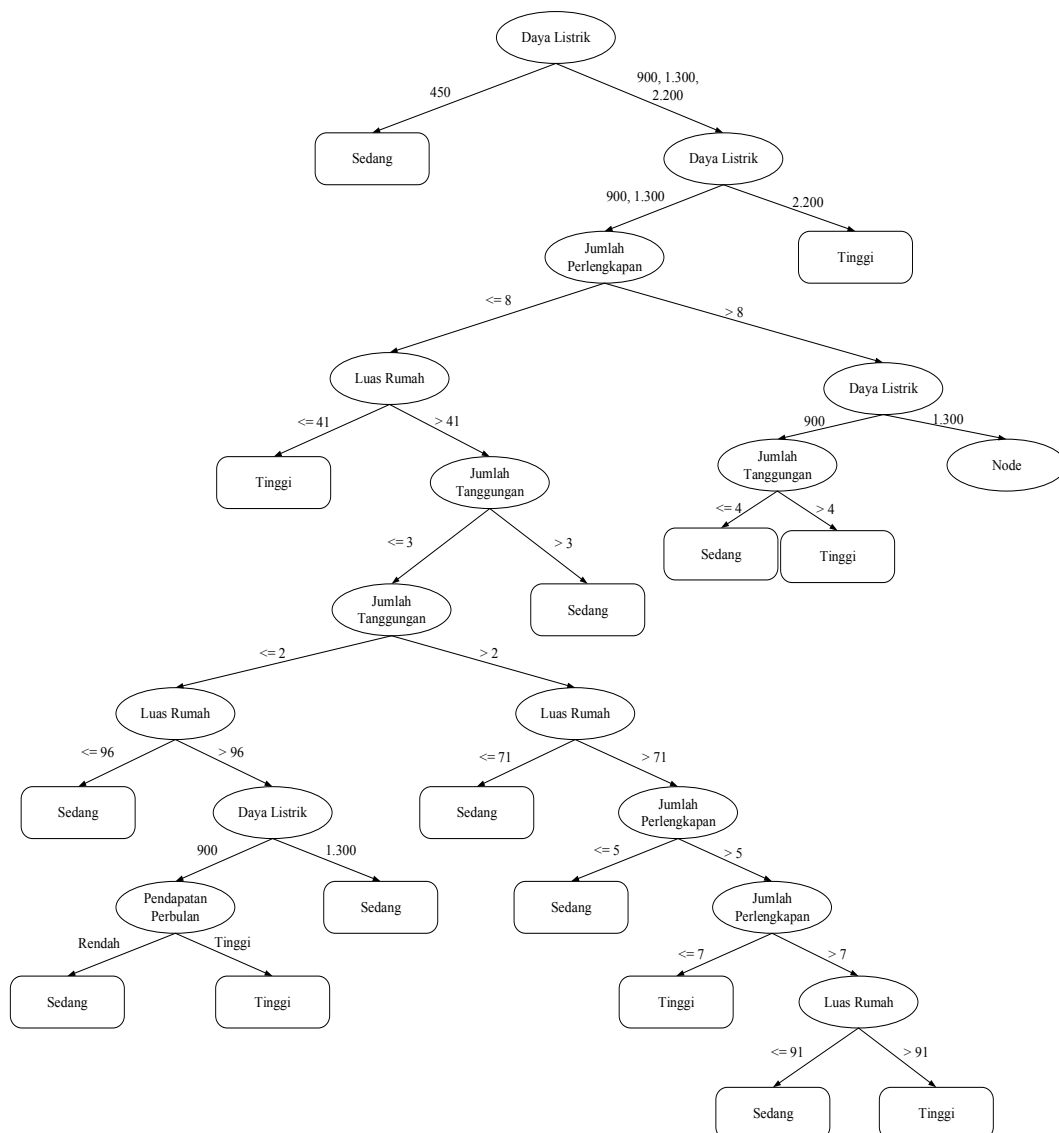
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai gainnya. Nilai gain tertinggi didapatkan pada posisi $v = 41$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai gain tertinggi didapatkan pada posisi $v = 12$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai gain tertinggi didapatkan pada posisi $v = 4$. Selanjutnya, hitung nilai *gain* untuk fitur Pendapatan Perbulan dengan hasil seperti pada tabel 3.44 dibawah ini.

Tabel 3.44 Hasil perhitungan *entropy* dan *gain* untuk *node 27*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		8	6	2	0,8113	
LR	<= 41	2	1	1	1,0000	0,0738
	> 41	6	5	1	0,6500	

JP	≤ 12	5	3	2	0,9710	0,2044
	> 12	3	3	0	0,0000	
JT	≤ 4	6	6	0	0,0000	0,8113
	> 4	2	0	2	0,0000	
PP	Rendah	1	1	0	0,0000	0,1101
	Sedang	5	4	1	0,7219	
	Tinggi	2	1	1	1,0000	

Dari perhitungan tersebut diperoleh fitur Jumlah Anggota Keluarga (JT) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.16 Pohon keputusan node 27

Tabel 3.45 Data pada kasus DL = 1.300

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
17	112	1300	20	2	Tinggi	Sedang
20	96	1300	13	4	Tinggi	Sedang
31	77	1300	14	4	Tinggi	Tinggi
34	144	1300	18	9	Tinggi	Tinggi
36	67	1300	9	2	Sedang	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
45	90	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
47	60	1300	10	5	Tinggi	Tinggi
48	45	1300	12	3	Tinggi	Tinggi
49	70	1300	10	4	Tinggi	Tinggi
50	60	1300	15	4	Tinggi	Tinggi
51	40	1300	12	6	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi
53	40	1300	10	5	Tinggi	Tinggi
55	70	1300	10	4	Tinggi	Tinggi
56	60	1300	15	4	Tinggi	Tinggi
58	20	1300	12	3	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 18 data yang terdiri dari 2 data dengan kelas Sedang dan 16 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

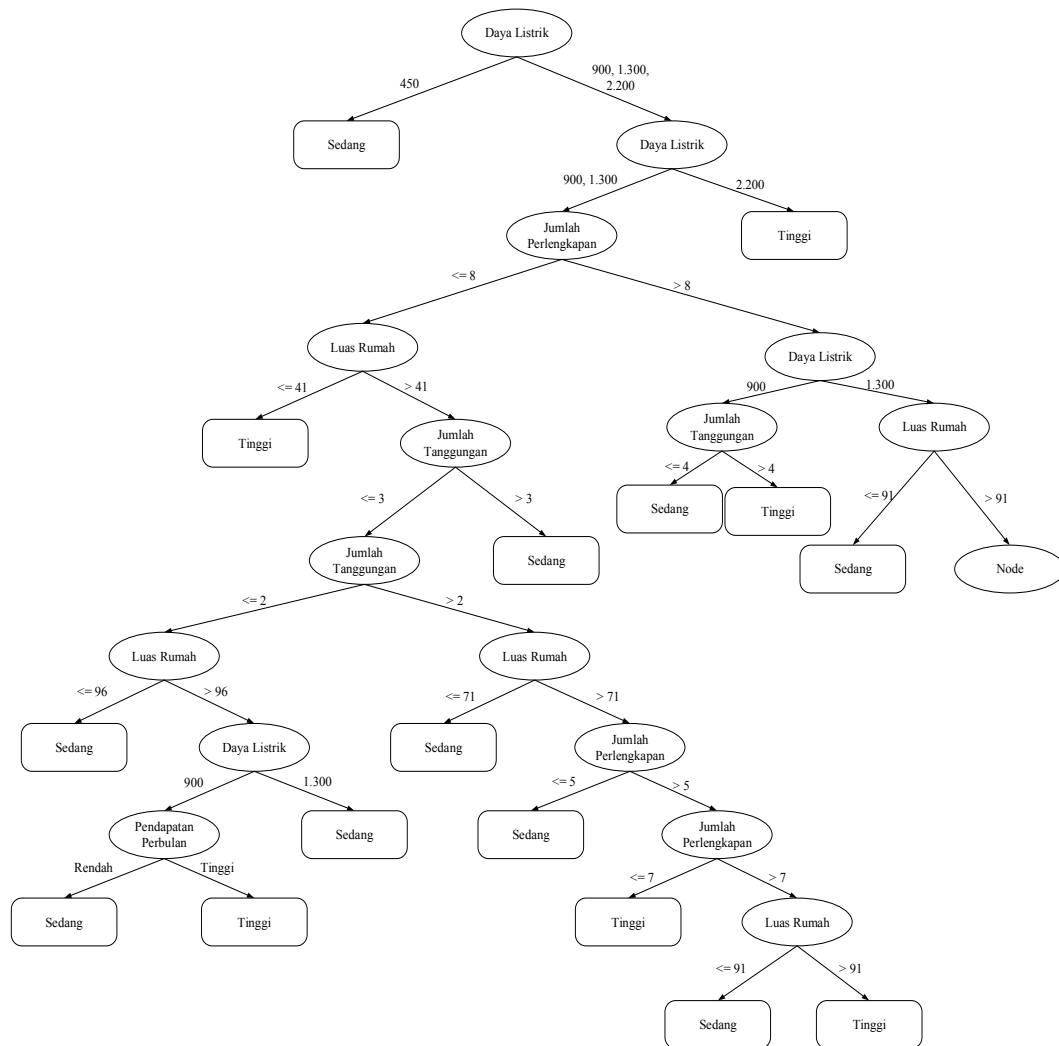
$$\begin{aligned}
 Entropy(S) &= -\frac{2}{18} * \log_2\left(\frac{2}{18}\right) - \frac{16}{18} * \log_2\left(\frac{16}{18}\right) \\
 &= 0,5033
 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai gainnya. Nilai gain tertinggi didapatkan pada posisi $v = 91$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai gain tertinggi didapatkan pada posisi $v = 12$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai gain tertinggi didapatkan pada posisi $v = 4$. Selanjutnya, hitung nilai *gain* untuk fitur Pendapatan Perbulan dengan hasil seperti pada tabel 3.46 dibawah ini.

Tabel 3.46 Hasil perhitungan *entropy* dan *gain* untuk *node 28*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		18	2	16	0,5033	
LR	<= 91	12	0	12	0,0000	0,1972
	> 91	6	2	4	0,9183	
JP	<= 12	10	0	10	0,0000	0,1427
	> 12	8	2	6	0,8113	
JT	<= 4	14	2	12	0,5917	0,0431
	> 4	4	0	4	0,0000	
PP	Rendah	0	0	0	0,0000	0,0097
	Sedang	1	0	1	0,0000	
	Tinggi	17	2	15	0,5226	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti gambar 3.17 dibawah ini.



Gambar 3.17 Pohon keputusan node 28

Tabel 3.47 Data pada kasus LR > 91

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
17	112	1300	20	2	Tinggi	Sedang
20	96	1300	13	4	Tinggi	Sedang
34	144	1300	18	9	Tinggi	Tinggi
43	98	1300	10	2	Tinggi	Tinggi
46	182	1300	15	1	Tinggi	Tinggi
52	134	1300	15	2	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 6 data yang terdiri dari 2 data dengan kelas Sedang dan 4 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{2}{6} * \log_2\left(\frac{2}{6}\right) - \frac{4}{6} * \log_2\left(\frac{4}{6}\right) \\ &= 0,9183 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai gain tertinggi didapatkan pada posisi $v = 116$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai gain tertinggi didapatkan pada posisi $v = 10$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai gain tertinggi didapatkan pada posisi $v = 4$. Kemudian, hitung nilai *gain* untuk fitur Pendapatan Perbulan.

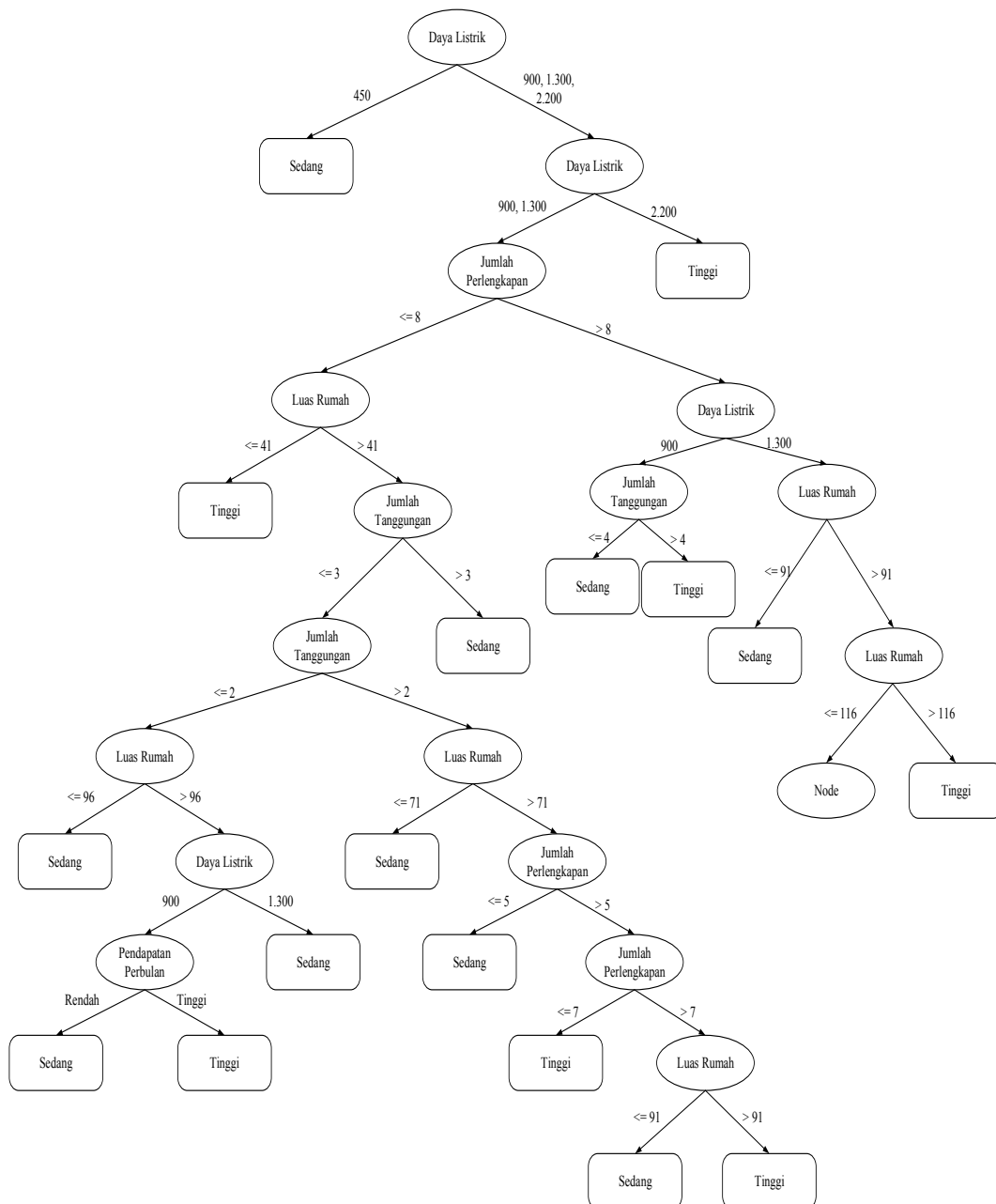
$$\begin{aligned} Gain(PP) &= 0,5033 - \left(\left(\frac{0}{6} * \left(-\frac{0}{6} * \log_2\left(\frac{0}{6}\right) - \frac{0}{6} * \log_2\left(\frac{0}{6}\right) \right) \right. \right. \\ &\quad \left. \left. + \frac{0}{6} * \left(-\frac{0}{6} * \log_2\left(\frac{0}{6}\right) - \frac{0}{6} * \log_2\left(\frac{0}{6}\right) \right) \right) \right. \\ &\quad \left. + \frac{6}{6} * \left(-\frac{2}{6} * \log_2\left(\frac{2}{6}\right) - \frac{4}{6} * \log_2\left(\frac{4}{6}\right) \right) \right) \\ &= 0,9183 - 0,9183 \\ &= 0,0000 \end{aligned}$$

Tabel 3.48 Hasil perhitungan *entropy* dan *gain* untuk *node 32*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		6	2	4	0,9183	
LR	≤ 116	3	2	1	0,9183	0,4591
	> 116	3	0	3	0,0000	
JP	≤ 10	1	0	1	0,0000	0,1092
	> 10	5	2	3	0,9710	
JT	≤ 4	5	2	3	0,9710	0,1092
	> 4	1	0	1	0,0000	

PP	Rendah	0	0	0	0,0000	0,0000
	Sedang	0	0	0	0,0000	
	Tinggi	6	2	4	0,9183	

Dari perhitungan tersebut diperoleh fitur Luas Rumah (LR) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.18 Pohon keputusan node 32

Tabel 3.49 Data pada kasus LR ≤ 116

No.	Luas Rumah (m ²)	Daya Listrik	Jumlah Perlengkapan	Jumlah Anggota Keluarga	Pendapatan Perbulan	Penggunaan Listrik
17	112	1300	20	2	Tinggi	Sedang
20	96	1300	13	4	Tinggi	Sedang
43	98	1300	10	2	Tinggi	Tinggi

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk fitur Luas Rumah. Hasil pemecahan tersebut menyisakan 3 data yang terdiri dari 2 data dengan kelas Sedang dan 1 data dengan kelas Tinggi. Perhitungan yang dilakukan sama dengan perhitungan sebelumnya yaitu terlebih dahulu menentukan nilai untuk entropy keseluruhan.

$$\begin{aligned} Entropy(S) &= -\frac{2}{3} * \log_2 \left(\frac{2}{3} \right) - \frac{1}{3} * \log_2 \left(\frac{1}{3} \right) \\ &= 0,9183 \end{aligned}$$

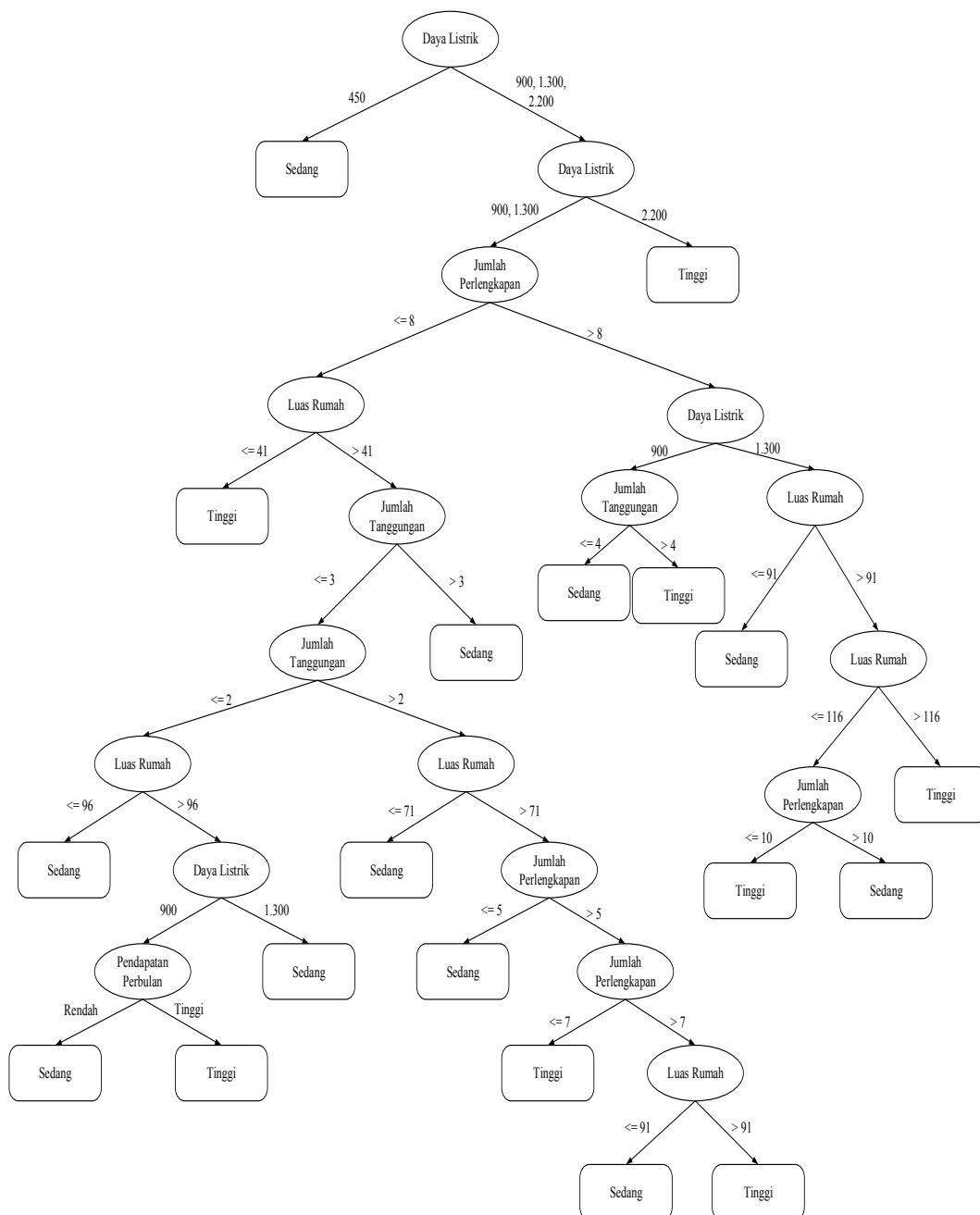
Selanjutnya menghitung nilai *gain* untuk setiap fitur. Untuk fitur Luas Rumah (LR) dengan menghitung nilai Gainnya. Nilai Gain tertinggi didapatkan pada posisi $v = 96$. Selanjutnya, untuk Jumlah Perlengkapan (JP) nilai Gain tertinggi didapatkan pada posisi $v = 10$. Sedangkan untuk Jumlah Anggota Keluarga (JT) nilai Gain tertinggi didapatkan pada posisi $v = 2$. Kemudian hitung nilai *gain* untuk fitur Pendapatan Perbulan dengan hasil seperti pada tabel 3.50 dibawah ini.

Tabel 3.50 Hasil perhitungan *entropy* dan *gain* untuk *node 33*

Total		Jumlah	Sedang	Tinggi	Entropy	Gain
		3	2	1	0,9183	
LR	≤ 96	1	1	0	0,0000	0,2516
	> 96	2	1	1	1,0000	
JP	≤ 10	1	0	1	0,0000	0,9183
	> 10	2	2	0	0,0000	
JT	≤ 2	2	1	1	1,0000	0,2516
	> 2	1	1	0	0,0000	

PP	Rendah	0	0	0	0,0000	0,0000
	Sedang	0	0	0	0,0000	
	Tinggi	3	2	1	0,9183	

Dari perhitungan tersebut diperoleh fitur Jumlah Perlengkapan (JP) memiliki nilai gain lebih tinggi dibanding dengan fitur lainnya sehingga diperoleh pohon keputusan seperti berikut.



Gambar 3.19 Pohon keputusan node 33

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

- RULE 1 : IF daya listrik = 450 THEN prediksi SEDANG
- RULE 2 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 2.200 THEN prediksi TINGGI
- RULE 3 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah \leq 41 THEN prediksi TINGGI
- RULE 4 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan \leq 2 AND luas rumah \leq 96 THEN prediksi SEDANG
- RULE 5 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan \leq 2 AND luas rumah $>$ 96 AND daya listrik = 900 AND pendapatan perbulan = Rendah THEN prediksi SEDANG
- RULE 6 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan \leq 2 AND luas rumah $>$ 96 AND daya listrik = 900 AND pendapatan perbulan = Tinggi THEN prediksi TINGGI
- RULE 7 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan \leq 2 AND luas rumah $>$ 96 AND daya listrik = 1300 THEN prediksi SEDANG

- RULE 8 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan $>$ 2 AND luas rumah \leq 71 THEN prediksi SEDANG
- RULE 9 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan $>$ 2 AND luas rumah $>$ 71 AND jumlah perlengkapan \leq 5 THEN prediksi SEDANG
- RULE 10 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan $>$ 2 AND luas rumah $>$ 71 AND jumlah perlengkapan $>$ 5 AND jumlah perlengkapan \leq 7 THEN prediksi TINGGI
- RULE 11 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan $>$ 2 AND luas rumah $>$ 71 AND jumlah perlengkapan $>$ 5 AND jumlah perlengkapan $>$ 7 AND luas rumah \leq 91 THEN prediksi SEDANG
- RULE 12 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan \leq 3 AND jumlah tanggungan $>$ 2 AND luas rumah $>$ 71 AND jumlah perlengkapan $>$ 5 AND jumlah perlengkapan $>$ 7 AND luas rumah $>$ 91 THEN prediksi TINGGI
- RULE 13 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300 AND jumlah perlengkapan \leq 8 AND luas rumah $>$ 41 AND jumlah tanggungan $>$ 3 THEN prediksi SEDANG

RULE 14 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 900 AND jumlah
tanggungan <= 4 THEN prediksi SEDANG

RULE 15 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 900 AND jumlah
tanggungan > 4 THEN prediksi TINGGI

RULE 16 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 1300 AND luas
rumah <= 91 THEN prediksi TINGGI

RULE 17 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 1300 AND luas
rumah > 91 AND luas rumah <= 116 AND jumlah perlengkapan <= 10
THEN prediksi TINGGI

RULE 18 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 1300 AND luas
rumah > 91 AND luas rumah <= 116 AND jumlah perlengkapan > 10
THEN prediksi SEDANG

RULE 19 : IF daya listrik = 900, 1.300, 2.200 AND daya listrik = 900, 1.300
AND jumlah perlengkapan > 8 AND daya listrik = 1300 AND luas
rumah > 91 AND luas rumah > 116 THEN prediksi TINGGI

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 :

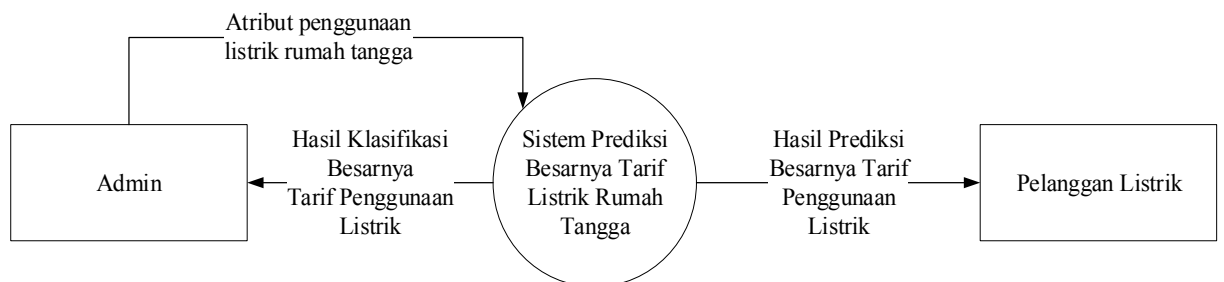
- a. Processor Intel Core i5
- b. RAM 4 GB

- c. HDD 500 GB
 - d. Monitor 14”
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. Windows 8.1 pro
 - b. Web Server : Apache
 - c. Database Server : MySQL
 - d. Bahasa Pemrograman : PHP
 - e. Sublime Text
 - f. SQLyog

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

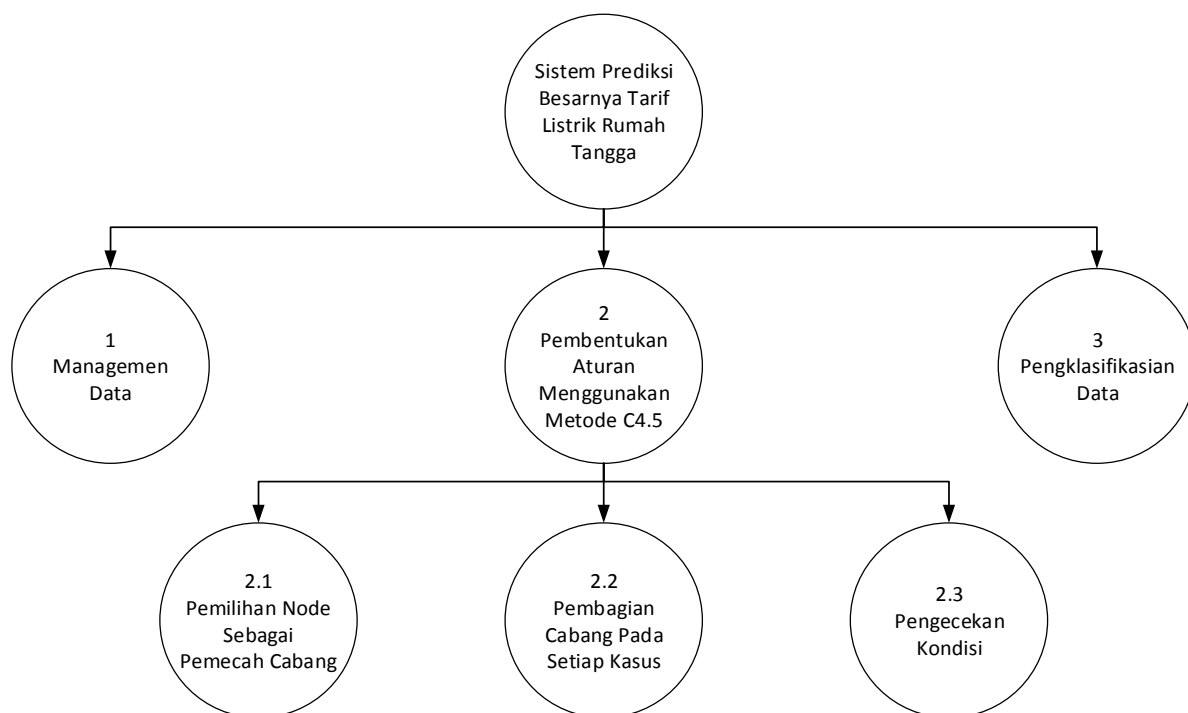


Gambar 3.20 *Context Diagram*

Penjelasan dari gambar 3.20, terlihat bahwa yang terlibat (*entity*) dalam sistem ini adalah petugas pelayanan pembayaran listrik dan Pelanggan Listrik. Petugas pelayanan pembayaran listrik, memasukkan data penggunaan listrik rumah tangga sebagai data latih. Data tersebut digunakan sebagai *data training* atau data yang akan diproses untuk pembentukan pohon keputusan. Keluaran

dari sistem untuk petugas pelayanan pembayaran listrik adalah hasil klasifikasi besarnya penggunaan listrik rumah tangga berdasarkan data yang telah dimasukkan. Sedangkan pelanggan listrik juga dapat melihat hasil klasifikasi dari penggunaan listrik rumah tangganya yang telah melalui proses klasifikasi.

3.5.2 Diagram Berjenjang



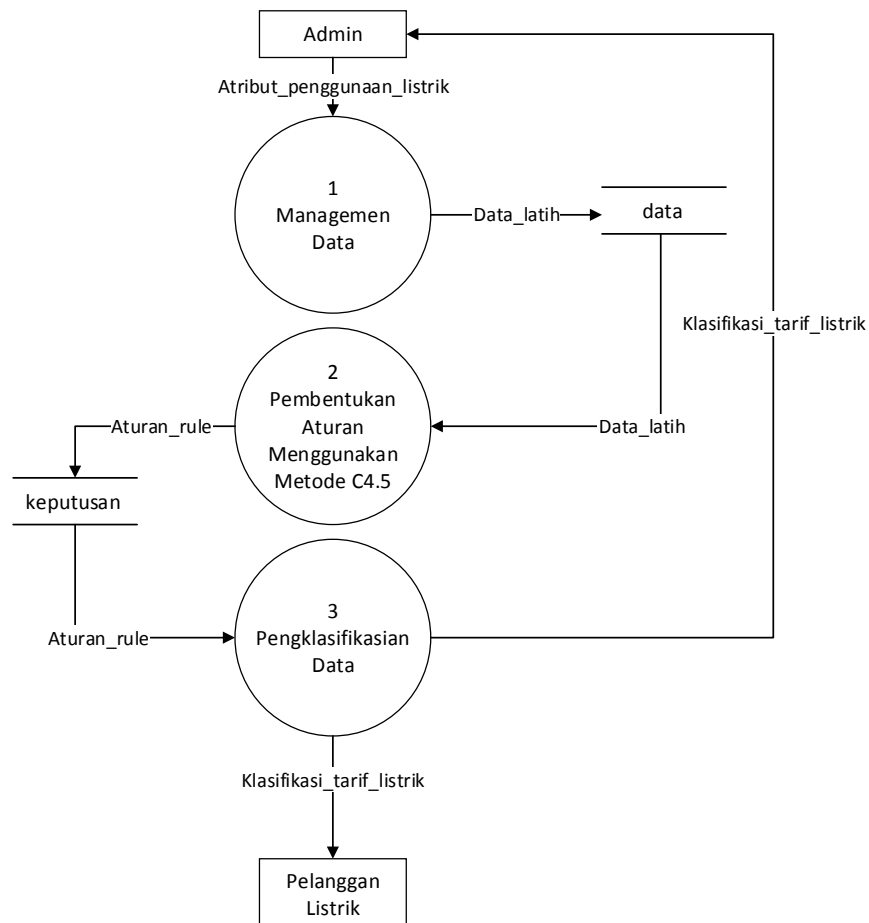
Gambar 3.21 Diagram Berjenjang

Diagram berjenjang disajikan pada gambar 3.21, berikut penjelasannya:

1. *Top level* : Sistem prediksi besarnya penggunaan listrik rumah tangga.
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.
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.

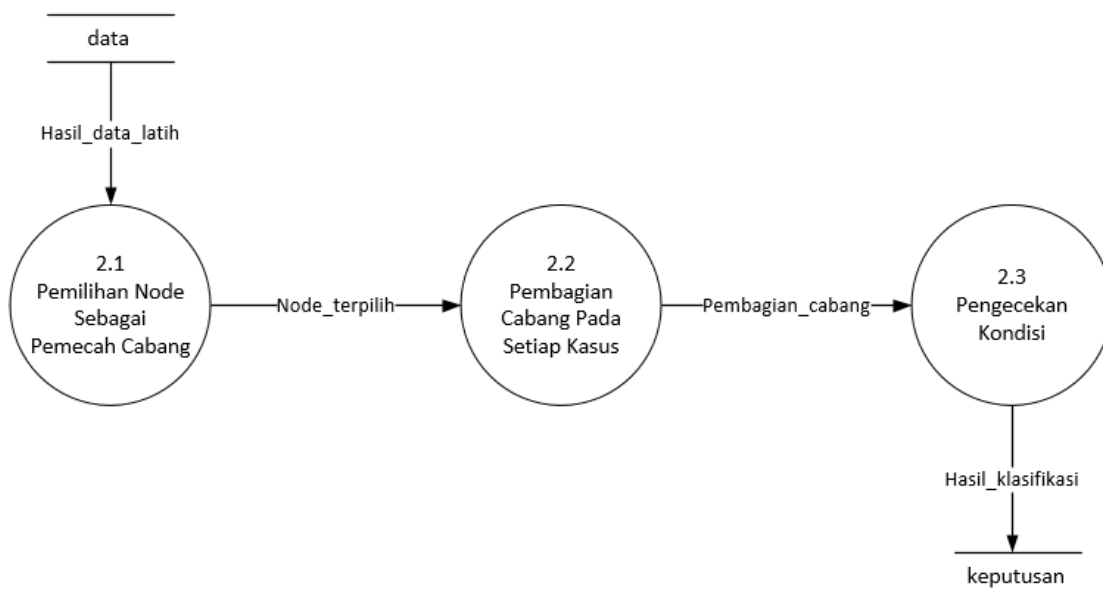
3.5.3 Data Flow Diagram Level 0



Gambar 3.22 DFD *Level 0*

DFD *level 0* pada gambar 3.22 menjelaskan aliran data pada sistem. Terdapat empat proses didalam sistem tersebut. Proses satu adalah manajemen data yang diinputkan oleh admin. Data atribut penggunaan listrik akan menjadi data latih untuk proses pembentukan pohon keputusan. Proses dua adalah pembentukan aturan (pohon keputusan) yang akan digunakan pada proses pengklasifikasian data uji. Proses ketiga adalah pegklasifikasian data dengan hasilnya akan menjadi output berupa klasifikasi tarif listrik untuk admin dan pelanggan listrik.

3.5.4 Data Flow Diagram Level 1



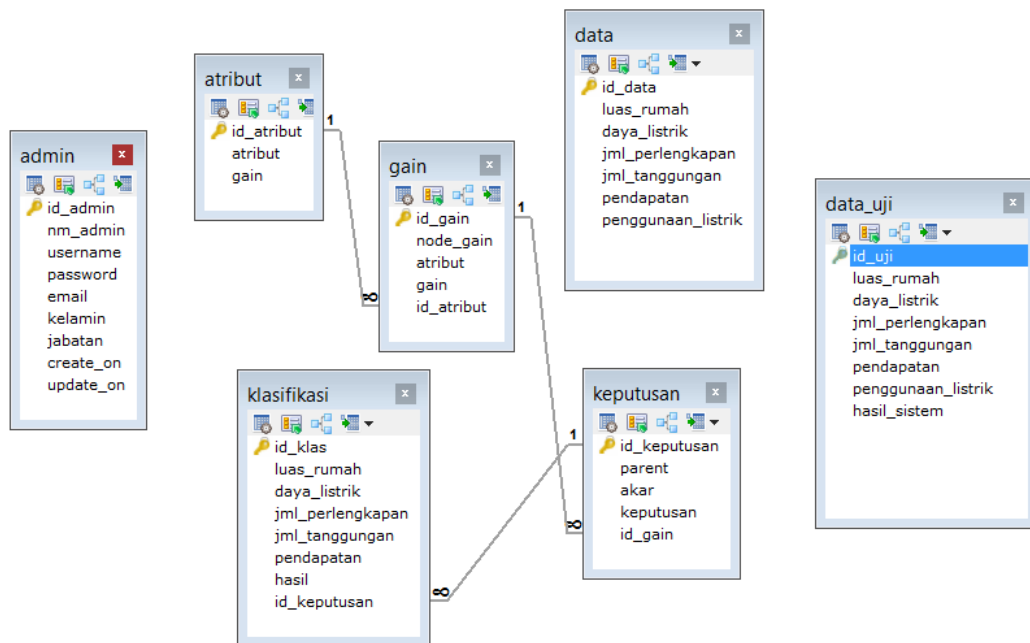
Gambar 3.23 DFD *Level 1* proses pembentukan aturan (pohon keputusan)

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 Entity Relationship Diagram (ERD)

Relasi tabel yang digunakan untuk membuat program aplikasi ini seperti gambar 3.24 dibawah. Pada gambar tersebut dijelaskan hubungan antar tabel pada database *db_listrik*.

Tabel *gain* memiliki relasi dengan tabel *atribut* dengan kolom *id_atribut* menjadi *foreign key* pada tabel *gain*. Sedangkan *id_gain* menjadi *foreign key* untuk tabel keputusan. Tabel keputusan juga mempunyai relasi dengan tabel klasifikasi dengan menjadikan *primary key* tabel keputusan menjadi *foreign key* pada tabel klasifikasi yaitu kolom *id_keputusan*.



Gambar 3.24 Entity Relationship Diagram

3.5.6 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

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

Tabel 3.51 Struktur tabel data

No	Field_name	Type	Length	Key
1	id_data	int	11	Primary key
2	luas_rumah	int	11	
3	daya_listrik	int	11	
4	jml_perlengkapan	int	11	
5	jml_tanggungan	int	11	
6	pendapatan	varchar	10	
7	penggunaan_listrik	varchar	10	

b. data_uji

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

Tabel 3.52 Struktur tabel data_uji

No	Field_name	Type	Length	Key
1	id_data	int	11	Primary key
2	luas_rumah	int	11	
3	daya_listrik	int	11	
4	jml_perlengkapan	int	11	
5	jml_tanggungan	int	11	
6	pendapatan	varchar	10	
7	penggunaan_listrik	varchar	10	
8	hasil_sistem	varchar	10	

c. klasifikasi

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

Tabel 3. 53 Struktur tabel hasil klasifikasi

No	Field_name	Type	Length	Key
1	id_data	int	11	Primary key
2	luas_rumah	int	11	
3	daya_listrik	int	11	
4	jml_perlengkapan	int	11	
5	jml_tanggungan	int	11	
6	pendapatan	varchar	10	
7	Hasil	varchar	10	
8	id_keputusan	int	11	Foreign Key

d. gain

Tabel ini merupakan *temporary* digunakan untuk menyimpan hasil perhitungan gain.

Tabel 3.54 Struktur tabel gain

No	Name_field	Type	Length	Key
1	id_gain	Int	11	Primary key
2	node_gain	Int	11	
3	atribut	Varchar	100	
4	gain	Double		
5	id_atribut	int	11	Foreign Key

e. atribut

Tabel ini merupakan *temporary* untuk menyimpan variable yang digunakan dalam data.

Tabel 3.55 Struktur tabel atribut

No	Name_field	Type	Length	Key
1	id_atribut	Int	11	Primary key
2	atribut	Varchar	20	
3	gain	double		

f. keputusan

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

Tabel 3.56 Struktur tabel keputusan

No	Field_name	Type	Length	Key
1	id_keputusan	Int	11	Primary key
2	parent	Text		
3	akar	Text		
4	keputusan	Varchar	100	
5	id_gain	int	11	Foreign Key

g. admin

Tabel *admin* ini dibuat untuk secara khusus agar bisa mengakses aplikasi ini. Data dari *admin* tersebut tersimpan dalam tabel *admin*.

Tabel 3.57 Struktur tabel admin

No	Name_field	Type	Length	Key
1	id_admin	int	11	Primary key
2	nm_admin	varchar	50	
3	username	varchar	50	
4	password	varchar	225	
5	email	varchar	50	
6	kelamin	varchar	20	
7	jabatan	varchar	10	
8	create_on	varchar	20	
9	update_on	varchar	20	

3.5.7 Desain Antar Muka

Tampilan antar muka pengguna sebagai admin yaitu halaman *sign in*, halaman utama, data latih, klasifikasi, hasil klasifikasi, uji akurasi, prediksi, kelola admin, profil dan *logout*.

a. Halaman *Sign In*

Halaman *sign in* diperlukan untuk mengetahui hak akses pengguna yang masuk kedalam sistem yaitu petugas pelayanan pembayaran listrik.



Aplikasi Data Mining Prediksi Besarnya
Penggunaan Listrik Rumah Tangga
Menggunakan Metode Decision Tree C4.5

Username

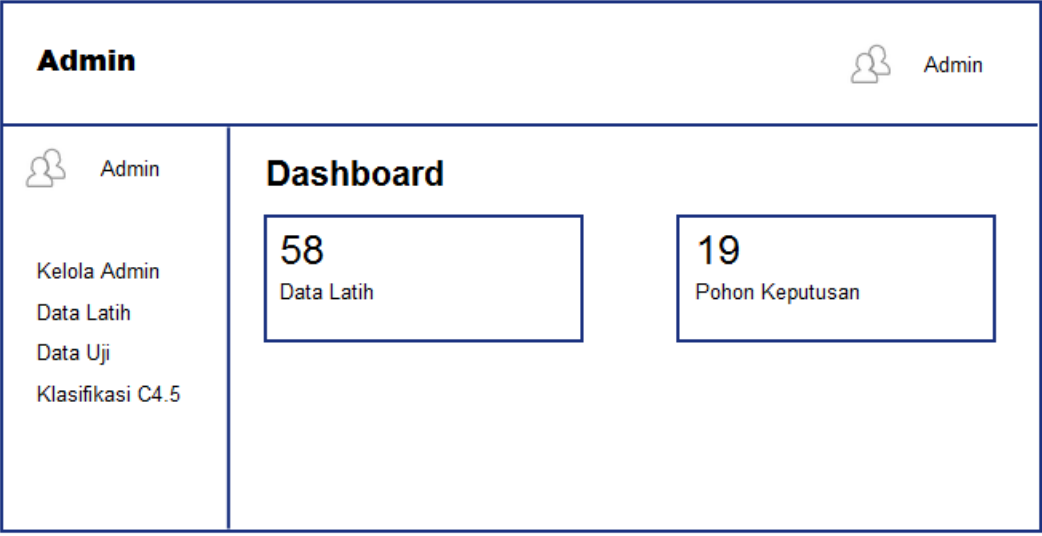
Password

Sign In

Gambar 3.25 Rancangan halaman *sign in*

b. Halaman Utama

Menu yang ditampilkan untuk pengguna sebagai admin yaitu kelola admin, Data Latih, dan Klasifikasi C4.5.



Admin Admin

Admin

Kelola Admin
Data Latih
Data Uji
Klasifikasi C4.5

Dashboard

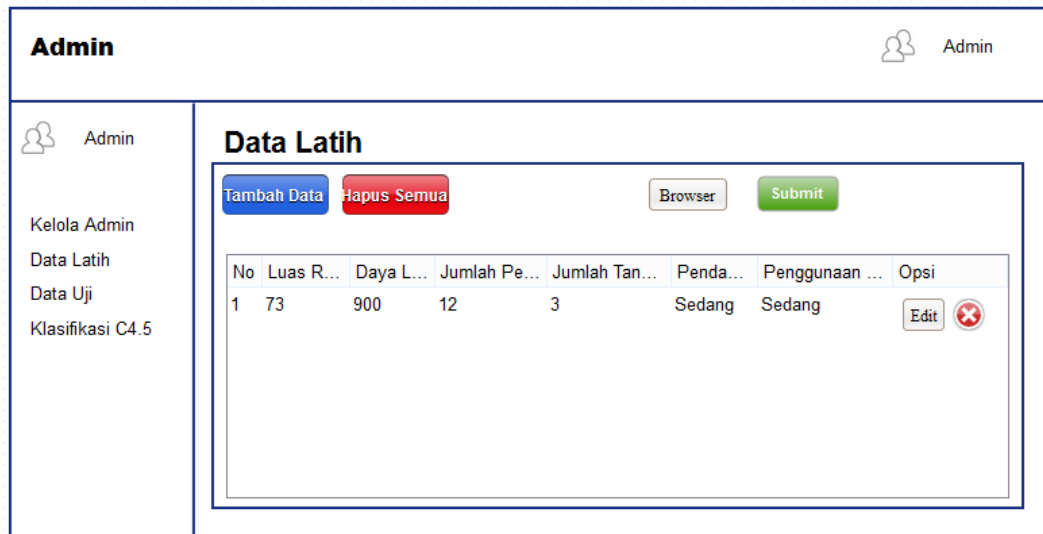
58
Data Latih

19
Pohon Keputusan

Gambar 3.26 Rancangan halaman utama

c. Halaman Data Latih

Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan.



Gambar 3.27 Rancangan halaman data latih

d. Halaman Klasifikasi

Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan dan perhitungan mining.



Gambar 3.28 Rancangan halaman klasifikasi

e. Halaman Hasil Klasifikasi

Halaman ini menampilkan pohon keputusan atau aturan yang didapat dari proses *mining*. Pohon keputusan ini juga dapat diuji keakurasiannya.



Gambar 3.29 Rancangan halaman hasil klasifikasi

f. Halaman Uji Akurasi

Halaman ini digunakan untuk menguji tingkat akurasi pohon keputusan yang terbentuk dari proses *mining*.



Gambar 3.30 Rancangan halaman uji akurasi

g. Halaman Prediksi

Halaman hasil ini akan menampilkan form yang digunakan untuk menambahkan data penggunaan listrik baru yang akan dilakukan prediksi oleh sistem.

The screenshot shows a web interface for an Admin user. The page title is 'Admin' and the user is logged in as 'Admin'. The main content area is titled 'Klasifikasi'. On the left, there is a sidebar menu with options: 'Kelola Admin', 'Data Latih', 'Data Uji', and 'Klasifikasi C4.5'. The 'Klasifikasi' form contains the following fields:

- 'Luas Rumah': a text input field.
- 'Daya Listrik': a dropdown menu with the option 'Pilih Daya Listrik'.
- 'Jumlah Perlengkapan': a text input field.
- 'Jumlah Tanggungan': a text input field.
- 'Pendapatan Perbulan': a dropdown menu with the option 'Pilih Pendapatan Perbulan'.

 At the bottom of the form are 'Reset' and 'Submit' buttons.

Gambar 3.31 Rancangan form penambahan prediksi

h. Halaman Tambah Administrator

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

The screenshot shows a web interface for an Admin user. The page title is 'Admin' and the user is logged in as 'Admin'. The main content area is titled 'Tambah Administrator'. On the left, there is a sidebar menu with options: 'Kelola Admin', 'Data Latih', and 'Klasifikasi C4.5'. The 'Tambah Administrator' form contains the following fields:

- 'Nama Lengkap': a text input field.
- 'Email': a text input field.
- 'Jenis Kelamin': radio buttons for 'Laki-laki' and 'Perempuan'.
- 'Username': a text input field.
- 'Password': a text input field.

 At the bottom of the form are 'Reset' and 'Submit' buttons.

Gambar 3.32 Rancangan form tambah administrator

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 3.1. Sedangkan untuk mengukur tingkat kesalahannya menggunakan rumus 3.2.

$$Akurasi = \frac{Jumlah\ h\ data\ yang\ diklasifikasi\ secara\ benar}{Jumlah\ h\ klasifikasi\ yang\ dilakukan} \times 100\% \dots\dots\dots(3.1)$$

$$Laju\ Error = \frac{Jumlah\ h\ data\ yang\ diklasifikasi\ secara\ salah}{Jumlah\ h\ klasifikasi\ yang\ dilakukan} \times 100\% \dots\dots\dots(3.2)$$

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 prediksi penggunaan listrik rumah tangga dengan metode decision tree C4.5 ini, perlu dilakukan beberapa skenario pengujian sistem terlebih dahulu, agar sistem dapat berjalan sesuai dengan tujuan pembuatannya.

- a. Disediakan 2 macam data yaitu data latih sebanyak 58 dan data uji sebanyak 22. 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 prediksi penggunaan listrik rumah tangga dengan kelas penggunaan “Sedang”, dan “Tinggi”.
- c. Akurasi sistem diperoleh dari hasil prosentase akurasi ketepatan sistem dan hasil prosentase akurasi kesalahan sistem.