

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Pada penelitian ini, tahapan pengumpulan data dilakukan di Dinas Perumahan Rakyat dan Kawasan Permukiman (DISPERKIM) Kabupaten Jepara terkhusus di Bagian Permukiman. Tahapan ini melalui proses observasi dan wawancara terkait dengan penerimaan program Hibah Air Minum Perdesaan (HAMP). Dalam kegiatan observasi dilakukan dengan mendatangi lokasi secara langsung yaitu kantor Dinas Perumahan Rakyat dan Kawasan Permukiman (DISPERKIM) Kabupaten Jepara yang beralamat di Jalan Ki Mangunsarkoro, Kelurahan Panggang, Kecamatan Jepara, Kabupaten Jepara, Provinsi Jawa Tengah. Sedangkan untuk kegiatan wawancara dilakukan dengan mengajukan beberapa pertanyaan yang berkaitan dengan penerimaan program Hibah Air Minum Perdesaan (HAMP) dan dijawab oleh tim pelaksana dari kegiatan tersebut.

Setelah dilakukan pengumpulan data melalui dua kegiatan diatas, maka diperoleh data calon penerima program Hibah Air Minum Perdesaan (HAMP) Tahun 2020 dalam bentuk Microsoft Excel dengan jumlah 1050 *record*. Data tersebut diperoleh dari *baseline survey* yang dilakukan oleh tim pelaksana lapangan. Adapun data yang diperoleh dapat dilihat pada Lampiran V.

4.2 Pengolahan Data Awal

Dari proses pengumpulan data telah diperoleh data calon penerima program Hibah Air Minum Perdesaan (HAMP) Tahun 2020 dengan jumlah 1050 *record* yang akan digunakan sebagai data *training* dan data *testing*. Atribut yang dimiliki oleh data tersebut diantaranya id calon penerima, kecamatan, desa/kelurahan, nama, alamat (RT/RW), nama responden, kesesuaian alamat, hubungan keluarga, jenis kelamin, daya listrik, status sambungan PDAM, kesediaan menyambung, hasil *survey*, alasan *ineligible*, dan lainnya. Selanjutnya, dilakukan pengolahan

data awal atau *pre-processing* untuk meningkatkan kualitas data menjadi lebih baik, sebagai berikut:

4.2.1 Menyusun dan Mengkategorikan Data

Dalam tahapan awal *pre-processing* data, dilakukan dengan mengkonversi atribut berdasarkan dengan kriteria variabel yang sudah ditentukan menggunakan Microsoft Excel. Adapun atribut yang dikonversi diantaranya merupakan kesesuaian alamat, hubungan keluarga, daya listrik, kesesuaian MBR, status sambungan PDAM, kesediaan menyambung SR, dan menerima bantuan lain. Untuk lebih jelas, kriteria variabel dapat dilihat pada tabel 3.1. Sedangkan untuk tipe data dari masing-masing atribut telah ditentukan dan dapat dilihat pada tabel 4.1, sehingga dari tahapan ini diperoleh hasil sebagai berikut:

Tabel 4.1 Tipe Data Atribut

No	Atribut	Tipe Data	Role
1	Nama	Polynomial	ID
2	Kesesuaian Alamat	Polynomial	
3	Hubungan Pemilik	Polynomial	
4	Daya Listrik	Polynomial	
5	Kesesuaian MBR	Polynomial	
6	Status Sambungan PDAM	Polynomial	
7	Kesediaan Menyambung SR	Polynomial	
8	Menerima Bantuan Lain	Polynomial	
9	Hasil	Polynomial	Label

Selanjutnya, pada atribut lain yang bukan merupakan atribut id dan atribut label akan disusun dan diubah datanya ke dalam bentuk kategori. Adapun tujuh atribut yang disusun dan diurutkan yaitu:

1. Atribut Kesesuaian Alamat dinotasikan sebagai A1, dengan kategori sebagai berikut:
 - a. Sesuai, dinotasikan sebagai S
 - b. Tidak Sesuai, dinotasikan sebagai TS

2. Atribut Hubungan Pemilik dinotasikan sebagai A2, dengan kategori sebagai berikut:
 - a. Pemilik, dinotasikan sebagai PM
 - b. Keluarga, dinotasikan sebagai K
 - c. Penyewa, dinotasikan sebagai PN
 - d. Lain-Lain, dinotasikan sebagai L
3. Atribut Daya Listrik dinotasikan sebagai A3, dengan kategori sebagai berikut:
 - a. 0 VA, dinotasikan sebagai = 0 VA
 - b. 450 VA, dinotasikan sebagai = 450 VA
 - c. 900 VA, dinotasikan sebagai = 900 VA
 - d. 1300 VA, dinotasikan sebagai = 1300 VA
 - e. 1300 VA, dinotasikan sebagai > 1300 VA
4. Atribut Kesesuaian MBR dinotasikan sebagai A4, dengan kategori sebagai berikut:
 - a. MBR, dinotasikan sebagai MBR
 - b. NON MBR, dinotasikan sebagai N-MBR
5. Atribut Status Sambungan PDAM dinotasikan sebagai A5, dengan kategori sebagai berikut:
 - a. Ada, dinotasikan sebagai AD
 - b. Tidak Ada, dinotasikan sebagai TA
6. Atribut Kesiediaan Menyambung SR dinotasikan sebagai A6, dengan kategori sebagai berikut:
 - a. Bersedia, dinotasikan sebagai BS
 - b. Tidak Bersedia, dinotasikan sebagai TB
7. Atribut Menerima Bantuan Lain dinotasikan sebagai A7, dengan kategori sebagai berikut:
 - a. Ya, dinotasikan sebagai Y
 - b. Tidak, dinotasikan sebagai T

Berikut merupakan tampilan dari data pada *Microsoft Excel* yang telah disusun dan dikategorikan sebagai tahapan awal dalam *pre-processing* data :

ID CALON PENERIMA	KECAMATAN	DESA/KELURAHAN	NAMA	ALAMAT	NAMA RESPONDER	AS1	AS2	AS3	AS4	AS5	AS6	AS7	HASIL
2332004001100723	MAYONG	BUARAN 1	NUR KHALIMAH	RT 008 004	NUR KHALIMAH	PM	= 450 VA	MBR	TA	BS			TIDAK MENVERDIA
2332004001100724	MAYONG	BUARAN 2	RUDARIYAH	RT 008 004	RUDARIYAH	PM	= 450 VA	MBR	TA	BS			TIDAK MENVERDIA
2332004001100725	MAYONG	BUARAN 3	TABMIN	RT 008 004	TABMIN	PM	= 450 VA	MBR	TA	BS			TIDAK MENVERDIA
2332004001100726	MAYONG	BUARAN 4	KEMENAH	RT 008 004	KEMENAH	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100727	MAYONG	BUARAN 5	MUBAROKAH	RT 008 004	MUBAROKAH	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100728	MAYONG	BUARAN 6	AHSUN ASLORI	RT 008 004	AHSUN ASLORI	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100729	MAYONG	BUARAN 7	MASUD	RT 008 004	MASUD	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100730	MAYONG	BUARAN 8	FREYONO	RT 008 004	FREYONO	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100731	MAYONG	BUARAN 9	MARRACHUMAH	RT 008 004	MARRACHUMAH	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100732	MAYONG	BUARAN 10	SUPARNO	RT 008 004	SUPARNO	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100733	MAYONG	BUARAN 11	SITULAMBI	RT 008 004	SITULAMBI	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100734	MAYONG	BUARAN 12	MABAROH	RT 008 004	MABAROH	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100735	MAYONG	BUARAN 13	DELUKHAN	RT 008 004	DELUKHAN	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100736	MAYONG	BUARAN 14	URMANTO	RT 008 004	URMANTO	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100737	MAYONG	BUARAN 15	AMIN GUPFRON	RT 008 004	AMIN GUPFRON	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100738	MAYONG	BUARAN 16	SYUTADAK	RT 008 004	SYUTADAK	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100739	MAYONG	BUARAN 17	SUYONO SOLIKHAN	RT 008 004	SUYONO SOLIKHAN	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100740	MAYONG	BUARAN 18	KASBI	RT 008 004	KASBI	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100741	MAYONG	BUARAN 19	BEO AHMAD SUKRON	RT 009 004	BEO AHMAD SUKRON	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100742	MAYONG	BUARAN 20	DIANGINGIH	RT 009 004	DIANGINGIH	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100743	MAYONG	BUARAN 21	ANDI KRIBWYANTO	RT 009 004	ANDI KRIBWYANTO	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100744	MAYONG	BUARAN 22	MOH ALI	RT 009 004	MOH ALI	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100745	MAYONG	BUARAN 23	HERY PURWANTO	RT 009 004	HERY PURWANTO	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100746	MAYONG	BUARAN 24	RUMMAN	RT 009 004	RUMMAN	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100747	MAYONG	BUARAN 25	KHORUSSOLEH	RT 009 004	KHORUSSOLEH	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100748	MAYONG	BUARAN 26	ANJARWADI	RT 009 004	ANJARWADI	PM	= 0 VA	MBR	TA	BS			MENVERDIA
2332004001100749	MAYONG	BUARAN 27	NAROWI	RT 009 004	NAROWI	PM	= 450 VA	MBR	TA	BS			TIDAK MENVERDIA
2332004001100750	MAYONG	BUARAN 28	RUMAWI	RT 009 004	RUMAWI	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100751	MAYONG	BUARAN 29	ABDULLAH JAMIN	RT 009 004	ABDULLAH JAMIN	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100752	MAYONG	BUARAN 30	DRIS ULIN NEDIA	RT 009 004	DRIS ULIN NEDIA	PM	= 900 VA	MBR	TA	BS			MENVERDIA
2332004001100753	MAYONG	BUARAN 31	YUDI PRYANTO	RT 009 004	YUDI PRYANTO	PM	= 450 VA	MBR	TA	BS			MENVERDIA
2332004001100754	MAYONG	BUARAN 32	KAMARI	RT 009 004	KAMARI	PM	= 450 VA	MBR	TA	BS			MENVERDIA

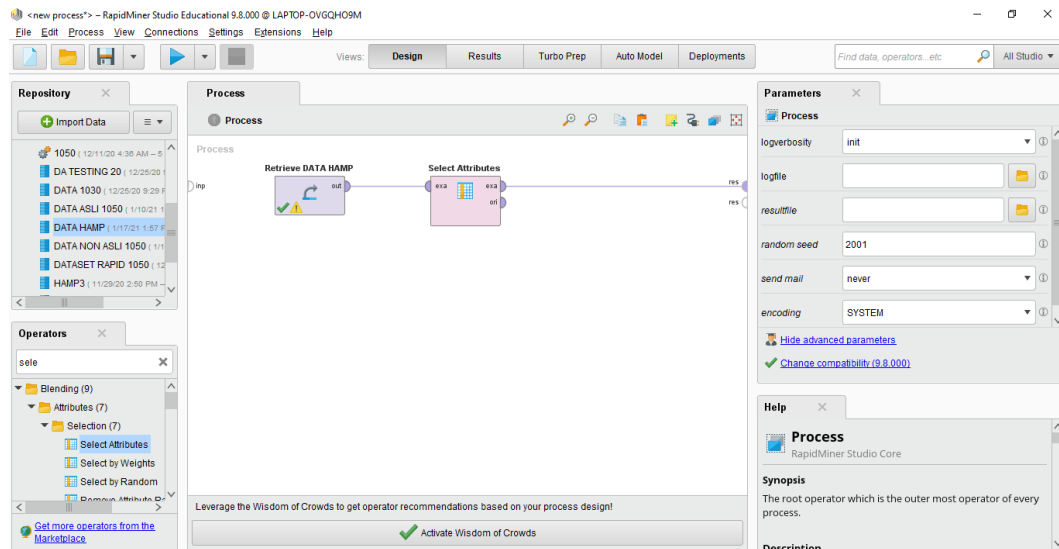
Gambar 4.1 Hasil Penyusunan dan Pengkategorian Data

4.2.2 Seleksi Atribut

Setelah menyusun dan mengkategorikan data, selanjutnya dilakukan seleksi atribut untuk mengambil variabel tertentu dari semua atribut yang ada pada data awal. Pada penelitian ini, variabel yang digunakan yaitu nama, kesesuaian alamat, hubungan keluarga, daya listrik, kesesuaian MBR, status sambungan PDAM, kesediaan menyambungkan SR, menerima bantuan lain, dan hasil yang digunakan sebagai target dengan *output* menerima atau tidak menerima.

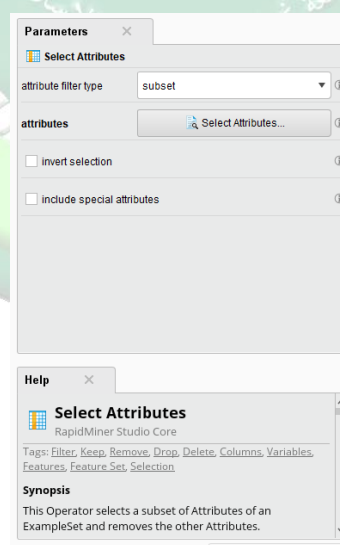
Atribut kesesuaian MBR dan menerima bantuan lain merupakan atribut yang ditambahkan peneliti dari atribut sebelumnya yang bernama alasan *ineligible*. Atribut tersebut berisi alasan mengapa seorang calon penerima tidak mendapatkan bantuan, diantaranya merupakan karena calon penerima NON MBR atau sudah menerima bentuk bantuan lain. Sehingga, bagi peneliti akan jauh lebih baik jika atribut tersebut dimodifikasi dan dijadikan sebagai variabel. Sedangkan beberapa atribut lainnya tidak digunakan dalam perhitungan. Berikut merupakan cara penyeleksian atribut menggunakan *software RapidMiner* versi 9.6:

1. *Import* data yang digunakan dan tambahkan operator *Select Attribute*. Kemudian hubungkan *Retrieve* data dengan operator *Select Attribute*, hingga terhubung dengan *result*, seperti gambar berikut ini:



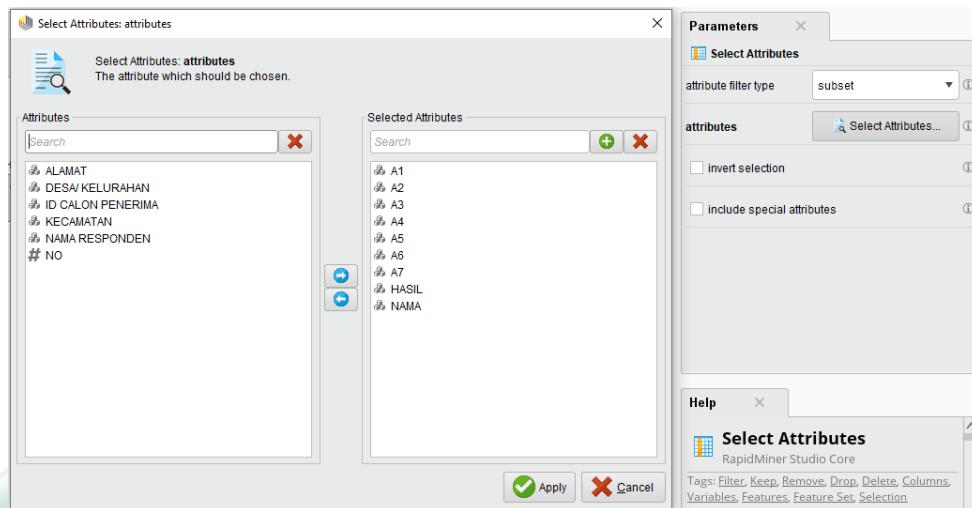
Gambar 4.2 Permodelan Seleksi Atribut

2. Setelah membuat permodelan seleksi atribut pada *software RapidMiner* versi 9.6, langkah selanjutnya adalah melakukan pengaturan pada *parameters select attributes*. Pada *attribute filter type*, masukkan *subset* sebagai tipe penyeleksian. *Subset* dipilih karena mampu memilih dan menyeleksi atribut yang nantinya akan diolah, seperti gambar berikut ini:



Gambar 4.3 Parameters Select Attributes

3. Kemudian, dilakukan penyeleksian atribut dari data yang telah dimasukkan dalam proses awal. Atribut yang digunakan akan dimasukkan ke dalam kolom kanan yaitu *Selected Attributes* yang selanjutnya dapat diolah, lebih jelas dapat dilihat pada gambar berikut ini:



Gambar 4.4 Seleksi Atribut Gambar 4. 5

4. Dari proses seleksi data yang telah dilakukan, menghasilkan tampilan data seperti gambar berikut ini:

The screenshot shows the 'Results' view in RapidMiner Studio. The table displays the output of the attribute selection process. The columns are: Row No., NAMA, HASIL SURV..., KESESUAI..., HUBUNGAN..., DAYA LISTRI..., KESESUAI..., STATUS SA..., KESEDIAAN..., and MENERII. The table contains 14 rows of data.

Row No.	NAMA	HASIL SURV...	KESESUAI...	HUBUNGAN ...	DAYA LISTRI...	KESESUAI...	STATUS SA...	KESEDIAAN ...	MENERII
1	NUR KHALIM...	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
2	SUDARIYAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
3	TASMIN	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
4	KEMINAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
5	MUJAROAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
6	AHSUN ASL...	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
7	MASUD	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
8	TRİYONO	MENERIMA	SESUAI	PEMILIK	= 900 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
9	MARKHUMAH	MENERIMA	SESUAI	PEMILIK	= 900 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
10	SUPARNO	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
11	SITULAMI	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
12	MAESAROH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
13	DZUKHAN	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
14	USMANTO	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK

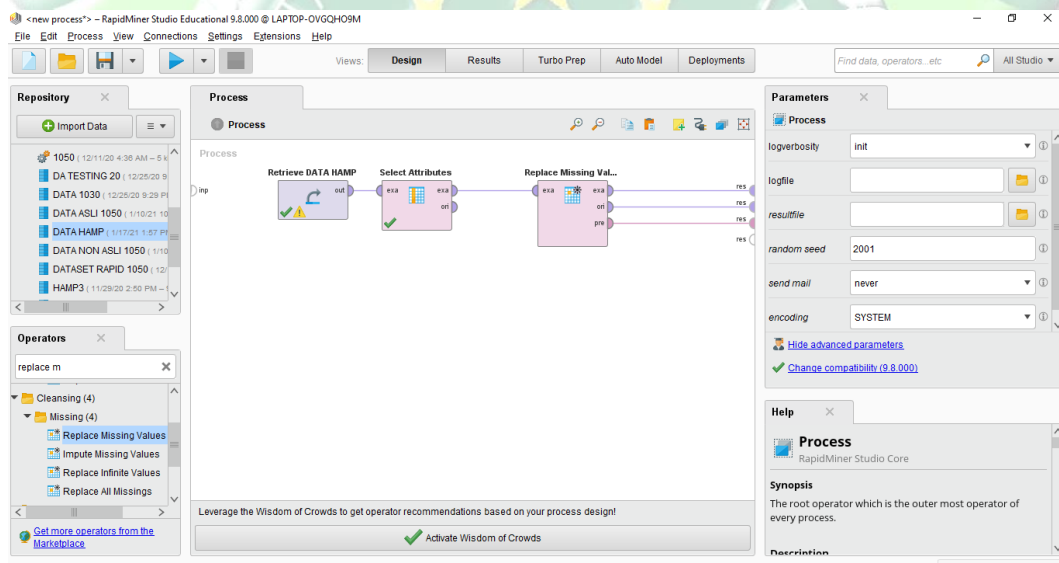
ExampleSet (1,050 examples, 2 special attributes, 7 regular attributes)

Gambar 4.6 Hasil Seleksi Atribut

4.2.3 Pembersihan Data

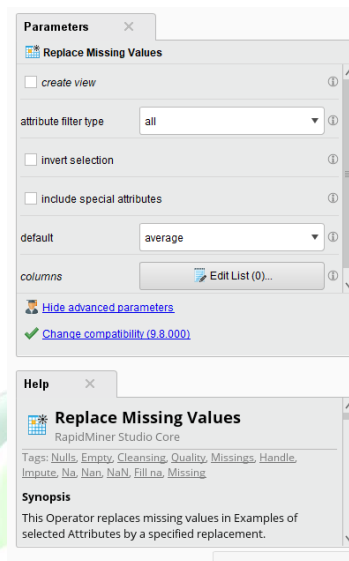
Tahapan ketiga dalam *pre-processing* adalah pembersihan data yang dilakukan untuk mengidentifikasi dan menghilangkan data yang kosong (*missing value*) menggunakan operator *replace missing values*. Missing value diartikan sebagai sebuah nilai yang tidak lengkap atau hilang dalam suatu data [25]. Atribut yang melalui proses pembersihan data diantaranya merupakan nama, kesesuaian alamat, hubungan keluarga, daya listrik, kesesuaian MBR, status sambungan PDAM, kesedian menyambung SR, menerima bantuan lain, dan hasil. Berikut merupakan cara pembersihan data menggunakan *software RapidMiner* versi 9.6:

1. Dari tahapan seleksi atribut, selanjutnya ditambahkan operator *replace missing values*. Kemudian, hubungkan operator *select attribute* dan *replace missing values* hingga terhubung dengan *result*, seperti gambar berikut ini:



Gambar 4.7 Permodelan *Replace Missing Values*

2. Setelah membuat permodelan pembersihan data pada *software RapidMiner* versi 9.6, langkah selanjutnya adalah melakukan pengaturan pada *parameters replace missing values*. Karena dilakukan pembersihan pada seluruh atribut, maka langkah selanjutnya yang perlu dilakukan adalah memasukkan *all* pada *attribute filter type*, seperti gambar berikut ini:



Gambar 4.8 *Parameters Replace Missing Values*

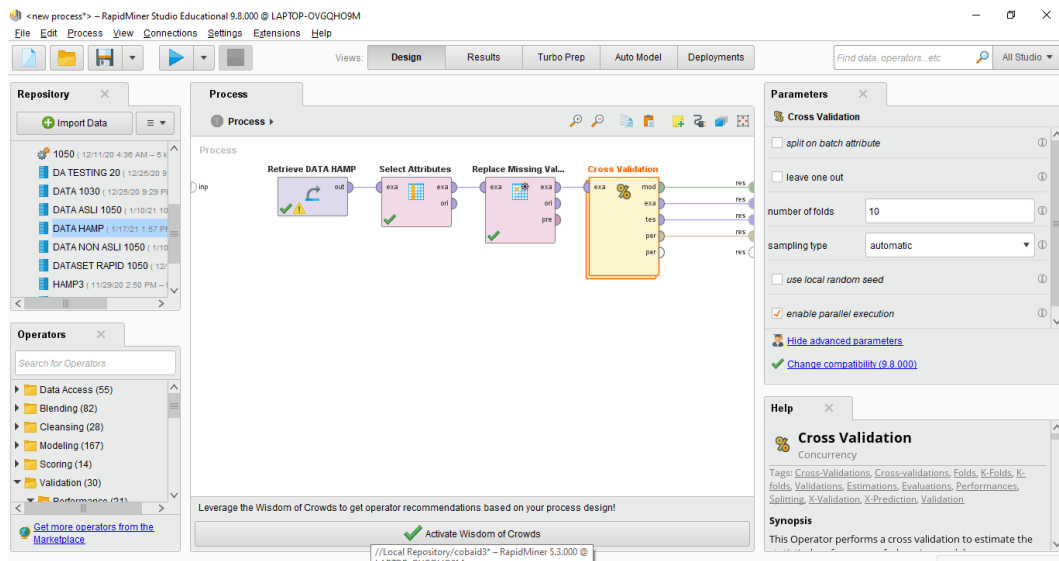
3. Dari proses pembersihan data yang telah dilakukan, menghasilkan tampilan data seperti gambar berikut ini:

Row No.	NAMA	HASIL SURV...	KESESUJIA...	HUBUNGAN ...	DAYA LISTRI...	KESESUJIA...	STATUS SA...	KESEDIAAN ...	MENERI...
1	NUR KHALIM...	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
2	SUDARIYAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
3	TASMIN	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
4	KEMINAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
5	MUJAROAH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
6	AHSUN ASL...	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
7	MASUD	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
8	TRIYONO	MENERIMA	SESUAI	PEMILIK	= 900 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
9	MARKHUMAH	MENERIMA	SESUAI	PEMILIK	= 900 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
10	SUPARNO	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
11	SITULAMI	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
12	MAESAROH	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK
13	DZUKHAN	MENERIMA	SESUAI	PEMILIK	= 450 VA	MBR	TIDAK ADA	BERSEDIA	TIDAK

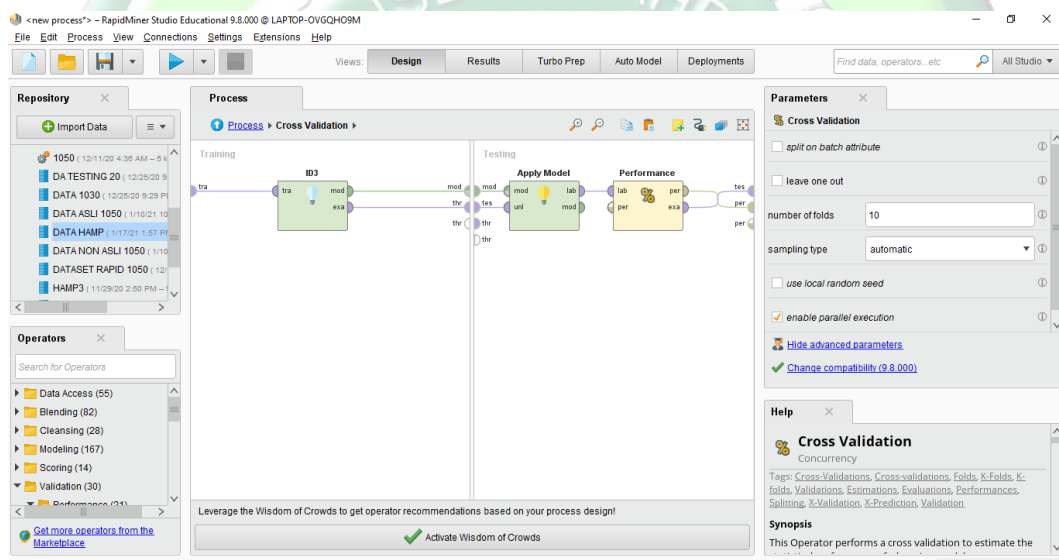
Gambar 4.9 Hasil Pembersihan Data

4.3 Pengujian Model *K-Fold Cross Validation*

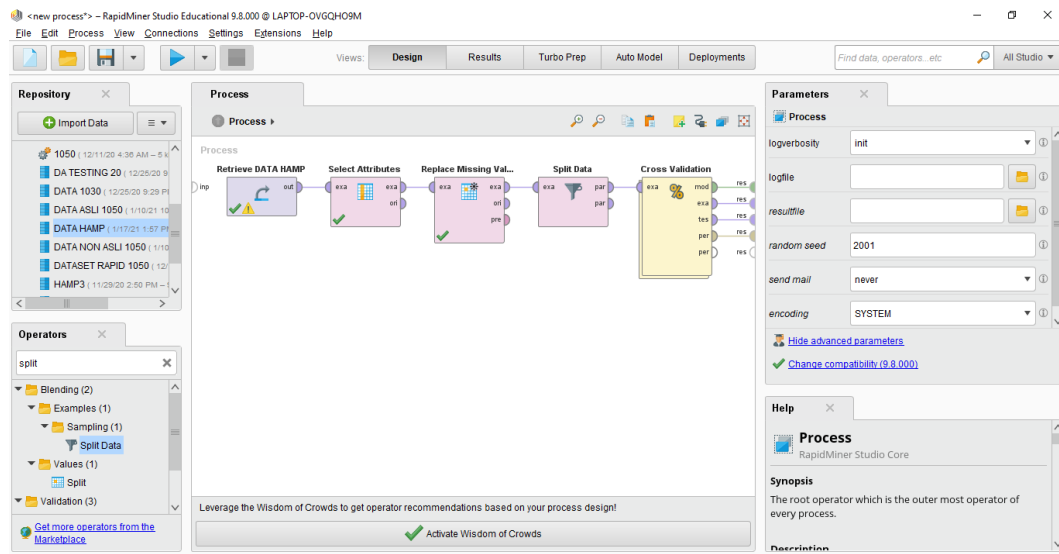
Tahapan pengujian model dilakukan dengan membagi dataset secara otomatis menggunakan *software RapidMiner* versi 9.6 menjadi 10 bagian dan selanjutnya dilakukan percobaan sebanyak 10 kali atau disebut dengan *10-fold cross validation*. Penerapan permodelan *10-fold cross validation* dapat dilihat pada gambar berikut ini:



Gambar 4.10 Permodelan K-Fold Cross Validation

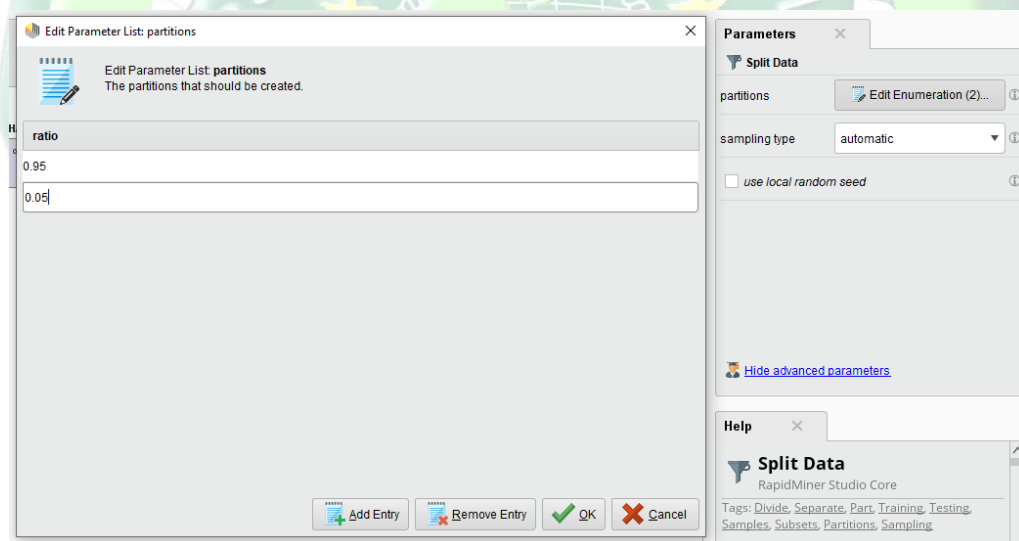
Gambar 4.11 Permodelan Algoritma *Iterative Dichotomiser Three* (ID3)

Pada penelitian ini, dilakukan pengujian model menggunakan *10-fold cross validation* sebanyak 10 kali. Selanjutnya, untuk mengetahui pembagian yang lebih baik antara banyaknya data *training* dan data *testing*, maka dilakukan pengujian untuk melihat perbandingan tingkat akurasi yang dihasilkan. Untuk melakukan hal tersebut, perlu adanya penambahan operator *Split Data* untuk membagi *ratio* antara data *training* dan data *testing*, seperti gambar di bawah ini:



Gambar 4.12 Permodelan Split Data

Pada parameters *split data*, klik *Edit Enumerations* pada *partitions* untuk selanjutnya memasukkan *ratio* antara data *training* dan data *testing*. Dimisalkan dengan data *training* 95% dan data *testing* 5%, seperti gambar di bawah ini:



Gambar 4.13 Parameters Split Data

Setelah dilakukan permodelan *split data* pada *RapidMiner* versi 9.6, berikut merupakan hasil pengujian model yang dilakukan:

1. Pengujian pertama dilakukan menggunakan ratio data *training* 55% dan data *testing* 45% sehingga menghasilkan nilai akurasi sebesar 98,61% seperti gambar berikut ini:

Table View

accuracy: 98.61% +/- 1.09% (micro average: 98.61%)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	50	2	96.15%
pred. MENERIMA	6	519	98.86%
class recall	89.29%	99.62%	

Gambar 4.14 Hasil Pengujian Model Pertama

2. Pengujian kedua dilakukan menggunakan ratio data *training* 65% dan data *testing* 35% sehingga menghasilkan nilai akurasi sebesar 98,53% seperti gambar berikut ini:

Table View

accuracy: 98.53% +/- 1.55% (micro average: 98.53%)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	60	4	93.75%
pred. MENERIMA	6	612	99.03%
class recall	90.91%	99.35%	

Gambar 4.15 Hasil Pengujian Model Kedua

3. Pengujian ketiga dilakukan menggunakan ratio data *training* 75% dan data *testing* 25% sehingga menghasilkan nilai akurasi sebesar 98,73% seperti gambar berikut ini:

Local Repository/2* - RapidMiner Studio Educational 9.8.000 @ LAPTOP-OVGGH09M

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

ExampleSet (Split Data) ExampleSet (Split Data) Tree (D3) ExampleSet (Local Repository/DATA HAMP) ExampleSet (Cross Validation)

Result History

Criterion
accuracy
precision
recall
AUC (optimistic)
AUC
AUC (pessimistic)

Performance

Description

Annotations

Table View Plot View

accuracy: 98.73% +/- 0.60% (micro average: 98.73%)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	71	4	94.67%
pred. MENERIMA	6	707	99.16%
class recall	92.21%	99.44%	

Gambar 4.16 Hasil Pengujian Model Ketiga

4. Pengujian keempat dilakukan menggunakan ratio data *training* 85% dan data *testing* 15% sehingga menghasilkan nilai akurasi sebesar 98,55% seperti gambar berikut ini:

Local Repository/2* - RapidMiner Studio Educational 9.8.000 @ LAPTOP-OVGGH09M

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

ExampleSet (Split Data) ExampleSet (Split Data) Tree (D3) ExampleSet (Local Repository/DATA HAMP) ExampleSet (Cross Validation)

Result History

Criterion
accuracy
precision
recall
AUC (optimistic)
AUC
AUC (pessimistic)

Performance

Description

Annotations

Table View Plot View

accuracy: 98.55% +/- 0.75% (micro average: 98.54%)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	78	4	95.12%
pred. MENERIMA	9	802	98.89%
class recall	89.66%	99.50%	

Gambar 4.17 Hasil Pengujian Model Keempat

5. Pengujian keempat dilakukan menggunakan ratio data *training* 95% dan data *testing* 5% sehingga menghasilkan nilai akurasi sebesar 98,60% seperti gambar berikut ini:

Table View

accuracy: 98.60% +/- 0.70% (micro average: 98.60%)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	87	4	95.60%
pred. MENERIMA	10	897	98.90%
class recall	89.69%	99.56%	

Gambar 4.18 Hasil Pengujian Model Kelima

Setelah dilakukan percobaan dalam pengujian dengan membagi *ratio* antara data *training* dan data *testing*, maka diperoleh hasil seperti tabel berikut ini:

Tabel 4.2 Hasil Pengujian Data *Training* dan Data *Testing*

	Pengujian Data <i>Training</i> : Data <i>Testing</i>				
	55% : 45%	65% : 35%	75% : 25%	85% : 15%	95% : 5%
Accuracy	98,61%	98,53%	98,73%	98,55%	98,60%
Precision	98,87%	99,05%	99,17%	98,90%	98,90%
Recall	99,62%	99,35%	99,44%	99,50%	99,56%
AUC	0,928	0,949	0,981	0,982	0,981

4.4 Algoritma *Iterative Dichotomiser Three* (ID3)

Dari proses pengujian yang telah dilakukan sebelumnya, diperoleh nilai akurasi tertinggi sebesar 98,73% dengan *ratio* data *training* dan data *testing* sebesar 75% dan 25%. Maka dari itu, penerapan algoritma *Iterative Dichotomiser Three* (ID3) akan dilakukan menggunakan data *training* sebanyak 788 *record*. Adapun data *training* penelitian yang akan diimplementasikan pada algoritma *Iterative Dichotomiser Three* (ID3) adalah sebagai berikut:

Tabel 4.3 Data *Training* Penelitian

NO	NAMA	A1	A2	A3	A4	A5	A6	A7	HASIL
1	NUR KHALIMAH	S	PM	= 450 VA	MBR	TA	BS	T	TIDAK MENERIMA
2	TASMIN	S	PM	= 450 VA	MBR	TA	BS	T	TIDAK MENERIMA
3	KEMINAH	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
4	MUJAROAH	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
5	AHSUN ASLORI	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
6	TRİYONO	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
7	MARKHUMAH	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
8	SUPARNO	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
9	SITULAMI	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
10	MAESAROH	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
11	DZUKHAN	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
12	USMANTO	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
13	AMIN GUFRON	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
14	SYUHADAK	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
15	SUYONO S.	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
16	BEJO AHMAD S.	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
17	DIANINGSIH	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
18	ANDI K.	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
19	MOH ALI	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
20	HERI PURWANTO	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
:	:	:	:	:	:	:	:	:	:
784	RUSMINI	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
785	TRIANAH	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
786	EKO SARTANTO	S	PM	= 900 VA	MBR	TA	BS	T	MENERIMA
787	SARWAN	S	PM	= 450 VA	MBR	TA	BS	T	MENERIMA
788	AHMAD JUNAIDI	S	PM	= 0 VA	MBR	TA	BS	T	MENERIMA

4.4.1 Menghitung Jumlah Sampel

Tahap awal dalam penerapan algoritma *Iterative Dichotomiser Three* (ID3) adalah dengan menentukan jumlah sampel dan kejadian untuk atribut Hasil yang merupakan atribut label dalam data training, sehingga diperoleh:

- Jumlah sampel, dinotasikan dengan $n = 788$
- Jumlah kejadian Hasil Menerima, dinotasikan dengan $n(M) = 711$
- Jumlah kejadian Hasil Tidak Menerima, dinotasikan dengan $n(TM) = 77$

4.4.2 Menghitung Entropy (Total) untuk Root Node

Langkah selanjutnya dilakukan dengan menghitung *entropy* total dari data *training* penelitian menggunakan persamaan (2.1), sehingga diperoleh:

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{711}{788} \right) * \frac{\log(711/788)}{\log(2)} + \left(\frac{77}{788} \right) * \frac{\log(77/788)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,461712)$$

$$Entropy(Total) = 0,461712$$

4.4.3 Menghitung Entropy Ratio Tiap Atribut untuk Root Node

Untuk menghitung nilai *entropy ratio* dari data *training* pada tabel 4.3, atribut yang akan melalui proses perhitungan adalah A1, A2, A3, A4, A5, A6, dan A7. Adapun perhitungan yang dilakukan adalah sebagai berikut:

1. A1 (Kesesuaian Alamat)

a. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 787$, dengan $n(M) = 711$ dan $n(TM) = 76$.

$$Entropy(S) = - \left(\left(\frac{711}{787} \right) * \frac{\log(711/787)}{\log(2)} + \left(\frac{76}{787} \right) * \frac{\log(76/787)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,458025)$$

$$Entropy(S) = 0,458025$$

b. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

2. A2 (Hubungan Pemilik)

a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 786$, dengan $n(M) = 710$ dan $n(TM) = 76$.

$$Entropy(S) = - \left(\left(\frac{710}{786} \right) * \frac{\log(710/786)}{\log(2)} + \left(\frac{76}{786} \right) * \frac{\log(76/786)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,458421)$$

$$Entropy(S) = 0,458421$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 1$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} + \left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. A3 (Daya Listrik)

a. Kategori = 0 VA

Jumlah kejadian A3 kategori = 0 VA, dinotasikan dengan $n(S) = 97$, dengan $n(M) = 92$ dan $n(TM) = 5$.

$$Entropy(S) = - \left(\left(\frac{92}{97} \right) * \frac{\log(92/97)}{\log(2)} + \left(\frac{5}{97} \right) * \frac{\log(5/97)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,292929)$$

$$Entropy(S) = 0,292929$$

b. Kategori = 450 VA

Jumlah kejadian A3 kategori = 450 VA, dinotasikan dengan $n(S) = 520$, dengan $n(M) = 468$ dan $n(TM) = 52$.

$$Entropy(S) = - \left(\left(\frac{468}{520} \right) * \frac{\log(468/520)}{\log(2)} + \left(\frac{52}{520} \right) * \frac{\log(52/520)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,468995)$$

$$Entropy(S) = 0,468995$$

c. Kategori = 900 VA

Jumlah kejadian A3 kategori = 900 VA, dinotasikan dengan $n(S) = 159$, dengan $n(M) = 145$ dan $n(TM) = 14$.

$$Entropy(S) = - \left(\left(\frac{145}{159} \right) * \frac{\log(145/159)}{\log(2)} + \left(\frac{14}{159} \right) * \frac{\log(14/159)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,429928)$$

$$Entropy(S) = 0,429928$$

d. Kategori = 1300 VA

Jumlah kejadian A3 kategori = 1300 VA, dinotasikan dengan $n(S) = 11$, dengan $n(M) = 6$ dan $n(TM) = 5$.

$$Entropy(S) = -\left(\left(\frac{6}{11}\right) * \frac{\log(6/11)}{\log(2)} + \left(\frac{5}{11}\right) * \frac{\log(5/11)}{\log(2)}\right)$$

$$Entropy(S) = -(-0,994030)$$

$$Entropy(S) = 0,994030$$

e. Kategori > 1300 VA

Jumlah kejadian A3 kategori > 1300 VA, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = -\left(\left(\frac{0}{1}\right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1}\right) * \frac{\log(1/1)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. A4 (Kesesuaian MBR)

a. Kategori MBR

Jumlah kejadian A4 kategori MBR, dinotasikan dengan $n(S) = 733$, dengan $n(M) = 711$ dan $n(TM) = 22$.

$$Entropy(S) = -\left(\left(\frac{711}{733}\right) * \frac{\log(711/733)}{\log(2)} + \left(\frac{22}{733}\right) * \frac{\log(22/733)}{\log(2)}\right)$$

$$Entropy(S) = -(-0,194460)$$

$$Entropy(S) = 0,194460$$

b. Kategori N-MBR

Jumlah kejadian A4 kategori N-MBR, dinotasikan dengan $n(S) = 55$, dengan $n(M) = 55$ dan $n(TM) = 0$.

$$Entropy(S) = -\left(\left(\frac{55}{55}\right) * \frac{\log(55/55)}{\log(2)} + \left(\frac{0}{55}\right) * \frac{\log(0/55)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

5. A5 (Status Sambungan PDAM)

a. Kategori AD

Jumlah kejadian A5 kategori AD, dinotasikan dengan $n(S) = 18$, dengan $n(M) = 0$ dan $n(TM) = 18$.

$$Entropy(S) = -\left(\left(\frac{0}{18}\right) * \frac{\log(0/18)}{\log(2)} + \left(\frac{18}{18}\right) * \frac{\log(18/18)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

b. Kategori TA

Jumlah kejadian A5 kategori TA, dinotasikan dengan $n(S) = 770$, dengan $n(M) = 711$ dan $n(TM) = 59$.

$$Entropy(S) = -\left(\left(\frac{711}{770}\right) * \frac{\log(711/770)}{\log(2)} + \left(\frac{59}{770}\right) * \frac{\log(59/770)}{\log(2)}\right)$$

$$Entropy(S) = -(-0,390168)$$

$$Entropy(S) = 0,390168$$

6. A6 (Kesediaan Menyambung SR)

a. Kategori BS

Jumlah kejadian A6 kategori BS, dinotasikan dengan $n(S) = 783$, dengan $n(M) = 711$ dan $n(TM) = 72$.

$$Entropy(S) = -\left(\left(\frac{711}{783}\right) * \frac{\log(711/783)}{\log(2)} + \left(\frac{72}{783}\right) * \frac{\log(72/783)}{\log(2)}\right)$$

$$Entropy(S) = -(-0,442958)$$

$$Entropy(S) = 0,442958$$

b. Kategori TB

Jumlah kejadian A6 kategori TB, dinotasikan dengan $n(S) = 5$, dengan $n(M) = 0$ dan $n(TM) = 5$.

$$Entropy(S) = -\left(\left(\frac{0}{5}\right) * \frac{\log(0/5)}{\log(2)} + \left(\frac{5}{5}\right) * \frac{\log(5/5)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

7. A7 (Menerima Bantuan Lain)

a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 3$, dengan $n(M) = 2$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{2}{3} \right) * \frac{\log(2/3)}{\log(2)} + \left(\frac{1}{3} \right) * \frac{\log(1/3)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,918295)$$

$$Entropy(S) = 0,918295$$

b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 785$, dengan $n(M) = 709$ dan $n(TM) = 76$.

$$Entropy(S) = - \left(\left(\frac{709}{785} \right) * \frac{\log(709/785)}{\log(2)} + \left(\frac{76}{785} \right) * \frac{\log(76/785)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,458818)$$

$$Entropy(S) = 0,458818$$

4.4.4 Menghitung *Information Gain* Tiap Atribut dan Menentukan Nilai Tertinggi untuk *Root Node*

Untuk mengetahui atribut dengan nilai tertinggi, maka perlu dilakukan perhitungan untuk mengetahui *information gain* yang diperlihatkan pada perhitungan dan tabel sebagai berikut:

1. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{787}{788} * 0,458020 \right) + \left(\frac{1}{788} * 0 \right) \right)$$

$$Gain = 0,004268$$

2. A2 (Hubungan Pemilik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{786}{788} * 0,458421 \right) + \left(\frac{1}{788} * 0 \right) + \left(\frac{0}{788} * 0 \right) + \left(\frac{1}{788} * 0 \right) \right)$$

$$Gain = 0,004454$$

3. A3 (Daya Listrik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(= 0 VA)}{n} * Entropy n(= 0 VA) \right) + \left(\frac{n(= 450 VA)}{n} * Entropy n(= 450 VA) \right) + \left(\frac{n(= 990 VA)}{n} * Entropy n(= 990 VA) \right) + \left(\frac{n(= 1300 VA)}{n} * Entropy n(= 1300 VA) \right) + \left(\frac{n(> 1300 VA)}{n} * Entropy n(> 1300 VA) \right) \right)$$

$$Gain = 0,461712 -$$

$$\left(\left(\frac{97}{788} * 0,292929 \right) + \left(\frac{520}{788} * 0,468995 \right) + \left(\frac{159}{788} * 0,429928 \right) + \left(\frac{11}{788} * 0,994030 \right) + \left(\frac{1}{788} * 0 \right) \right)$$

$$Gain = 0,015538$$

4. A4 (Kesesuaian MBR)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(MBR)}{n} * Entropy n(MBR) \right) + \left(\frac{n(N - MBR)}{n} * Entropy n(N - MBR) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{733}{788} * 0,194460 \right) + \left(\frac{55}{788} * 0 \right) \right)$$

$$Gain = 0,280825$$

5. A5 (Status Sambungan PDAM)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(AD)}{n} * Entropy n(AD) \right) + \left(\frac{n(TA)}{n} * Entropy n(TA) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{18}{788} * 0 \right) + \left(\frac{770}{788} * 0,390168 \right) \right)$$

$$Gain = 0,080456$$

6. A6 (Kesediaan Menyambung SR)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(BS)}{n} * Entropy n(BS) \right) + \left(\frac{n(TB)}{n} * Entropy n(TB) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{783}{788} * 0,442958 \right) + \left(\frac{5}{788} * 0 \right) \right)$$

$$Gain = 0,021564$$

7. A7 (Menerima Bantuan Lain)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,461712 - \left(\left(\frac{3}{788} * 0,918295 \right) + \left(\frac{785}{788} * 0,458818 \right) \right)$$

$$Gain = 0,001145$$

Tabel 4.4 Hasil Perhitungan *Information Gain* Tertinggi pada Node 1

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1	A1	Total	788	711	77	0,4617125	0,0042687
		S	787	711	76	0,4580251	
		TS	1	0	1	0	
	A2	PM	786	710	76	0,4584213	0,0044548
		K	1	1	0	0	
		PN	0	0	0	0	
	A3	L	1	0	1	0	0,0155389
		= 0 VA	97	92	5	0,29293	
		= 450 VA	520	468	52	0,4689956	
		= 900 VA	159	145	14	0,4299283	
		= 1300 VA	11	6	5	0,9940302	
	A4	> 1300 VA	1	0	1	0	0,280825
		MBR	733	711	22	0,1944603	
	A5	N-MBR	55	0	55	0	0,0804568
		AD	18	0	18	0	
		TA	770	711	59	0,3901682	

A6	BS	783	711	72	0,4429587	0,0215645
	TB	5	0	5	0	
A7	Y	3	2	1	0,9182958	0,001145
	T	785	709	76	0,4588182	

Karena atribut A4 memiliki nilai *information gain* terbesar, maka atribut A4 menjadi *root node*. Pada atribut A4, kategori N-MBR memiliki 55 kasus dengan semua jawabannya adalah TM (Tidak Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A4 = N-MBR menjadi daun atau *leaf*.

4.4.5 Menentukan Cabang dari Tiap-Tiap Atribut

4.4.5.1 Cabang dari Atribut A4 = MBR

Setelah diperoleh *root node* 1. A4, maka langkah selanjutnya adalah mencari atribut sebagai *node* 1.1. Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A4 = MBR untuk membuat tabel *Node* 1.1. Langkah ini dilakukan dengan membuat cabang untuk tiap-tiap nilai, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A4 = MBR

Jumlah kejadian A4 = MBR, dinotasikan dengan $n(S) = 733$, dengan $n(M) = 713$ dan $n(TM) = 20$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{713}{733} \right) * \frac{\log(713/733)}{\log(2)} + \left(\frac{20}{733} \right) * \frac{\log(20/733)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,180588)$$

$$Entropy(Total) = 0,180588$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

- a. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 732$, dengan $n(M) = 711$ dan $n(TM) = 21$.

$$Entropy(S) = - \left(\left(\frac{711}{732} \right) * \frac{\log(711/732)}{\log(2)} + \left(\frac{21}{732} \right) * \frac{\log(21/732)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,187771)$$

$$Entropy(S) = 0,187771$$

b. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = -\left(\left(\frac{0}{1}\right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1}\right) * \frac{\log(1/1)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Kesesuaian Alamat)

a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 732$, dengan $n(M) = 710$ dan $n(TM) = 22$.

$$Entropy(S) = -\left(\left(\frac{710}{732}\right) * \frac{\log(710/732)}{\log(2)} + \left(\frac{22}{732}\right) * \frac{\log(22/732)}{\log(2)}\right)$$

$$Entropy(S) = -(-0,194665)$$

$$Entropy(S) = 0,194665$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 1$ dan $n(TM) = 0$.

$$Entropy(S) = -\left(\left(\frac{1}{1}\right) * \frac{\log(1/1)}{\log(2)} + \left(\frac{0}{1}\right) * \frac{\log(0/1)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = -\left(\left(\frac{0}{0}\right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0}\right) * \frac{\log(0/0)}{\log(2)}\right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A3 (Daya Listrik)

a. Kategori = 0 VA

Jumlah kejadian A3 kategori = 0 VA, dinotasikan dengan $n(S) = 94$, dengan $n(M) = 92$ dan $n(TM) = 2$.

$$Entropy(S) = - \left(\left(\frac{92}{94} \right) * \frac{\log(92/94)}{\log(2)} + \left(\frac{2}{94} \right) * \frac{\log(2/94)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,148549)$$

$$Entropy(S) = 0,148549$$

b. Kategori = 450 VA

Jumlah kejadian A3 kategori = 450 VA, dinotasikan dengan $n(S) = 483$, dengan $n(M) = 468$ dan $n(TM) = 15$.

$$Entropy(S) = - \left(\left(\frac{468}{483} \right) * \frac{\log(468/483)}{\log(2)} + \left(\frac{15}{483} \right) * \frac{\log(15/483)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,199659)$$

$$Entropy(S) = 0,199659$$

c. Kategori = 900 VA

Jumlah kejadian A3 kategori = 900 VA, dinotasikan dengan $n(S) = 145$, dengan $n(M) = 145$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{145}{145} \right) * \frac{\log(145/145)}{\log(2)} + \left(\frac{0}{145} \right) * \frac{\log(0/145)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori = 1300 VA

Jumlah kejadian A3 kategori = 1300 VA, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(S) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,970950)$$

$$Entropy(S) = 0,970950$$

e. Kategori > 1300 VA

Jumlah kejadian A3 kategori > 1300 VA, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

5. Menghitung *entropy ration* dari atribut A5 (Status Sambungan PDAM)

a. Kategori AD

Jumlah kejadian A5 kategori AD, dinotasikan dengan $n(S) = 12$, dengan $n(M) = 0$ dan $n(TM) = 12$.

$$Entropy(S) = - \left(\left(\frac{0}{12} \right) * \frac{\log(0/12)}{\log(2)} + \left(\frac{12}{12} \right) * \frac{\log(12/12)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

b. Kategori TA

Jumlah kejadian A5 kategori TA, dinotasikan dengan $n(S) = 721$, dengan $n(M) = 711$ dan $n(TM) = 10$.

$$Entropy(S) = - \left(\left(\frac{711}{721} \right) * \frac{\log(711/721)}{\log(2)} + \left(\frac{10}{721} \right) * \frac{\log(10/721)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,105472)$$

$$Entropy(S) = 0,105472$$

6. Menghitung *entropy ration* dari atribut A6 (Kesediaan Menyambung SR)

a. Kategori BS

Jumlah kejadian A6 kategori BS, dinotasikan dengan $n(S) = 729$, dengan $n(M) = 711$ dan $n(TM) = 18$.

$$Entropy(S) = - \left(\left(\frac{711}{729} \right) * \frac{\log(711/729)}{\log(2)} + \left(\frac{18}{729} \right) * \frac{\log(18/729)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,167026)$$

$$Entropy(S) = 0,167026$$

b. Kategori TB

Jumlah kejadian A6 kategori TB, dinotasikan dengan $n(S) = 4$, dengan $n(M) = 0$ dan $n(TM) = 4$.

$$Entropy(S) = - \left(\left(\frac{0}{4} \right) * \frac{\log(0/4)}{\log(2)} + \left(\frac{4}{4} \right) * \frac{\log(4/4)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

7. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 3$, dengan $n(M) = 2$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{2}{3} \right) * \frac{\log(2/3)}{\log(2)} + \left(\frac{1}{3} \right) * \frac{\log(1/3)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,918295)$$

$$Entropy(S) = 0,918295$$

b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 730$, dengan $n(M) = 709$ dan $n(TM) = 21$.

$$Entropy(S) = - \left(\left(\frac{709}{730} \right) * \frac{\log(709/730)}{\log(2)} + \left(\frac{21}{730} \right) * \frac{\log(21/730)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,188170)$$

$$Entropy(S) = 0,188170$$

8. Menghitung Information Gain pada atribut A1, A2, A3, A5, A6, dan A7 untuk mengetahui atribut dengan nilai Gain tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

- a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,180588 - \left(\left(\frac{732}{733} * 0,187771 \right) + \left(\frac{1}{733} * 0 \right) \right)$$

$$Gain = -0,006926$$

- b. A2 (Hubungan Pemilik)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$$Gain = 0,180588 -$$

$$\left(\left(\frac{732}{733} * 0,194665 \right) + \left(\frac{1}{733} * 0 \right) + \left(\frac{0}{733} * 0 \right) + \left(\frac{0}{733} * 0 \right) \right)$$

$$Gain = -0,013811$$

- c. A3 (Daya Listrik)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(= 0 VA)}{n} * Entropy n(= 0 VA) \right) + \left(\frac{n(= 450 VA)}{n} * Entropy n(= 450 VA) \right) + \left(\frac{n(= 990 VA)}{n} * Entropy n(= 990 VA) \right) + \left(\frac{n(= 1300 VA)}{n} * Entropy n(= 1300 VA) \right) + \left(\frac{n(> 1300 VA)}{n} * Entropy n(> 1300 VA) \right) \right)$$

$$Gain = 0,180588 -$$

$$\left(\left(\frac{94}{733} * 0,148549 \right) + \left(\frac{483}{733} * 0,199659 \right) + \left(\frac{145}{733} * 0 \right) + \left(\frac{10}{733} * 0,994030 \right) + \left(\frac{1}{733} * 0 \right) \right)$$

$$Gain = 0,016729$$

d. A5 (Status Sambungan PDAM)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(AD)}{n} * Entropy n(AD) \right) + \left(\frac{n(TA)}{n} * Entropy n(TA) \right) \right)$$

$$Gain = 0,180588 - \left(\left(\frac{12}{733} * 0 \right) + \left(\frac{721}{733} * 0,105472 \right) \right)$$

$$Gain = 0,076842$$

e. A6 (Kesediaan Menyambung SR)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(BS)}{n} * Entropy n(BS) \right) + \left(\frac{n(TB)}{n} * Entropy n(TB) \right) \right)$$

$$Gain = 0,180588 - \left(\left(\frac{729}{733} * 0,167026 \right) + \left(\frac{4}{733} * 0 \right) \right)$$

$$Gain = 0,014473$$

f. A7 (Menerima Bantuan Lain)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,180588 - \left(\left(\frac{3}{733} * 0,918295 \right) + \left(\frac{730}{733} * 0,188170 \right) \right)$$

$$Gain = -0,010570$$

Tabel 4.5 Hasil Perhitungan Information Gain Tertinggi pada Node 1.1

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1.1	A4 = MBR		733	713	20	0,1805886	-
	A1	S	732	711	21	0,1877716	-0,006927
		TS	1	0	1	0	
	A2	PM	732	710	22	0,1946658	-0,013812
		K	1	1	0	0	
		PN	0	0	0	0	

	L	0	0	0	0	
A3	= 0 VA	94	92	2	0,1485495	0,0167293
	= 450 VA	483	468	15	0,1996598	
	= 900 VA	145	145	0	0	
	= 1300 VA	10	6	4	0,9709506	
	> 1300 VA	1	0	1	0	
A5	AD	12	0	12	0	0,0768428
	TA	721	711	10	0,1054726	
A6	BS	729	711	18	0,1670268	0,0144733
	TB	4	0	4	0	
A7	Y	3	2	1	0,9182958	-0,01057
	T	730	709	21	0,1881709	

Karena atribut A5 memiliki nilai information gain terbesar, maka atribut A5 menjadi cabang atau *internal node* dari A4 = MBR. Pada atribut A5, kategori AD memiliki 12 kasus dengan semua jawabannya adalah TM (Tidak Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A5 = AD menjadi daun atau *leaf*. Selanjutnya akan dilakukan perhitungan untuk menentukan cabang dari atribut A5 = TA.

4.4.5.2 Cabang dari Atribut A5 = TA

Setelah diperoleh *root node* 1.1 A5, maka langkah selanjutnya adalah mencari atribut sebagai *node* 1.1.2 Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A5 = TA untuk membuat tabel *Node* 1.1.2. Langkah ini dilakukan dengan membuat cabang untuk tiap-tiap nilai, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A5 = TA

Jumlah kejadian A5 = TA, dinotasikan dengan $n(S) = 721$, dengan $n(M) = 711$ dan $n(TM) = 10$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{711}{721} \right) * \frac{\log(711/721)}{\log(2)} + \left(\frac{10}{721} \right) * 20 \right)$$

$$\text{Entropy (Total)} = -(-0,105472)$$

$$\text{Entropy (Total)} = 0,105472$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

c. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 720$, dengan $n(M) = 711$ dan $n(TM) = 9$.

$$\text{Entropy}(S) = -\left(\left(\frac{711}{720}\right) * \frac{\log(711/720)}{\log(2)} + \left(\frac{9}{720}\right) * \frac{\log(9/720)}{\log(2)}\right)$$

$$\text{Entropy}(S) = -(-0,096944)$$

$$\text{Entropy}(S) = 0,096944$$

d. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$\text{Entropy}(S) = -\left(\left(\frac{0}{1}\right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1}\right) * \frac{\log(1/1)}{\log(2)}\right)$$

$$\text{Entropy}(S) = -(-0)$$

$$\text{Entropy}(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Hubungan Pemilik)

a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 720$, dengan $n(M) = 710$ dan $n(TM) = 10$.

$$\text{Entropy}(S) = -\left(\left(\frac{710}{720}\right) * \frac{\log(710/720)}{\log(2)} + \left(\frac{10}{720}\right) * \frac{\log(10/720)}{\log(2)}\right)$$

$$\text{Entropy}(S) = -(-0,105591)$$

$$\text{Entropy}(S) = 0,105591$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 1$ dan $n(TM) = 0$.

$$\text{Entropy}(S) = -\left(\left(\frac{1}{1}\right) * \frac{\log(1/1)}{\log(2)} + \left(\frac{0}{1}\right) * \frac{\log(0/1)}{\log(2)}\right)$$

$$\text{Entropy}(S) = -(-0)$$

$$\text{Entropy}(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A3 (Daya Listrik)

a. Kategori = 0 VA

Jumlah kejadian A3 kategori = 0 VA, dinotasikan dengan $n(S) = 93$, dengan $n(M) = 92$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{92}{93} \right) * \frac{\log(92/93)}{\log(2)} + \left(\frac{1}{93} \right) * \frac{\log(1/93)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,085742)$$

$$Entropy(S) = 0,085742$$

b. Kategori = 450 VA

Jumlah kejadian A3 kategori = 450 VA, dinotasikan dengan $n(S) = 473$, dengan $n(M) = 468$ dan $n(TM) = 5$.

$$Entropy(S) = - \left(\left(\frac{468}{473} \right) * \frac{\log(468/473)}{\log(2)} + \left(\frac{5}{473} \right) * \frac{\log(5/473)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,084554)$$

$$Entropy(S) = 0,084554$$

c. Kategori = 900 VA

Jumlah kejadian A3 kategori = 900 VA, dinotasikan dengan $n(S) = 145$, dengan $n(M) = 145$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{145}{145} \right) * \frac{\log(145/145)}{\log(2)} + \left(\frac{0}{145} \right) * \frac{\log(0/145)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori = 1300 VA

Jumlah kejadian A3 kategori = 1300 VA, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(S) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,970950)$$

$$Entropy(S) = 0,970950$$

e. Kategori > 1300 VA

Jumlah kejadian A3 kategori > 1300 VA, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

5. Menghitung *entropy ration* dari atribut A6 (Kesediaan Menyambung SR)

a. Kategori BS

Jumlah kejadian A6 kategori BS, dinotasikan dengan $n(S) = 718$, dengan $n(M) = 711$ dan $n(TM) = 7$.

$$Entropy(S) = - \left(\left(\frac{711}{718} \right) * \frac{\log(711/718)}{\log(2)} + \left(\frac{7}{718} \right) * \frac{\log(7/718)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,079126)$$

$$Entropy(S) = 0,079126$$

b. Kategori TB

Jumlah kejadian A6 kategori TB, dinotasikan dengan $n(S) = 3$, dengan $n(M) = 0$ dan $n(TM) = 3$.

$$Entropy(S) = - \left(\left(\frac{0}{3} \right) * \frac{\log(0/3)}{\log(2)} + \left(\frac{3}{3} \right) * \frac{\log(3/3)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

c. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 3$, dengan $n(M) = 2$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{2}{3} \right) * \frac{\log(2/3)}{\log(2)} + \left(\frac{1}{3} \right) * \frac{\log(1/3)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,918295)$$

$$Entropy(S) = 0,918295$$

d. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 718$, dengan $n(M) = 709$ dan $n(TM) = 9$.

$$Entropy(S) = - \left(\left(\frac{709}{718} \right) * \frac{\log(709/718)}{\log(2)} + \left(\frac{9}{718} \right) * \frac{\log(9/718)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,097164)$$

$$Entropy(S) = 0,097164$$

6. Menghitung *Information Gain* pada atribut A1, A2, A3, dan A7 untuk mengetahui atribut dengan nilai *Gain* tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,105472 - \left(\left(\frac{720}{721} * 0,096944 \right) + \left(\frac{1}{721} * 0 \right) \right)$$

$$Gain = 0,008662$$

b. A2 (Hubungan Pemilik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$Gain = 0,105472 -$

$$\left(\left(\frac{720}{721} * 0,105591 \right) + \left(\frac{1}{721} * 0 \right) + \left(\frac{0}{721} * 0 \right) + \left(\frac{0}{721} * 0 \right) \right)$$

$Gain = 2,79664E - 05$

c. A3 (Daya Listrik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(= 0 VA)}{n} * Entropy n(= 0 VA) \right) + \left(\frac{n(= 450 VA)}{n} * Entropy n(= 450 VA) \right) + \left(\frac{n(= 990 VA)}{n} * Entropy n(= 990 VA) \right) + \left(\frac{n(= 1300 VA)}{n} * Entropy n(= 1300 VA) \right) + \left(\frac{n(> 1300 VA)}{n} * Entropy n(> 1300 VA) \right) \right)$$

$Gain = 0,105472 -$

$$\left(\left(\frac{93}{721} * 0,085742 \right) + \left(\frac{473}{721} * 0,084554 \right) + \left(\frac{145}{721} * 0 \right) + \left(\frac{10}{721} * 0,970950 \right) + \left(\frac{1}{721} * 0 \right) \right)$$

$Gain = 0,025475$

d. A6 (Kesediaan Menyambung SR)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(BS)}{n} * Entropy n(BS) \right) + \left(\frac{n(TB)}{n} * Entropy n(TB) \right) \right)$$

$Gain = 0,105472 - \left(\left(\frac{718}{721} * 0,079126 \right) + \left(\frac{3}{721} * 0 \right) \right)$

$Gain = 0,026675$

e. A7 (Menerima Bantuan Lain)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,105472 - \left(\left(\frac{3}{721} * 0,918295 \right) + \left(\frac{718}{721} * 0,097164 \right) \right)$$

$$Gain = 0,004891$$

Tabel 4. 6 Hasil Perhitungan Information Gain Tertinggi pada Node 1.1.2

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1.1.2	A5 = AD		721	711	10	0,10547	-
	A1	S	720	711	9	0,09694	0,0086624
		TS	1	0	1	0	
	A2	PM	720	710	10	0,10559	2,7966E-05
		K	1	1	0	0	
		PN	0	0	0	0	
	A3	L	0	0	0	0	0,02547584
		= 0 VA	93	92	1	0,08574	
		= 450 VA	473	468	5	0,08455	
		= 900 VA	145	145	0	0	
		= 1300 VA	10	6	4	0,97095	
	A6	> 1300 VA	0	0	0	0	0,02667523
		BS	718	711	7	0,07913	
	A7	TB	3	0	3	0	0,00489189
		Y	3	2	1	0,9183	
	T	718	709	9	0,09716		

Karena atribut A6 memiliki nilai information gain terbesar, maka atribut A6 menjadi cabang atau *internal node* dari A5 = AD. Pada atribut A6, kategori TB memiliki 3 kasus dengan semua jawabannya adalah TM (Tidak Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A6 = TB menjadi daun atau *leaf*. Selanjutnya akan dilakukan perhitungan untuk menentukan cabang dari atribut A6 = BS.

4.4.5.3 Cabang dari Atribut A6 = BS

Setelah diperoleh *node* 1.1.2. A6, maka langkah selanjutnya adalah mencari atribut sebagai *node* 1.1.2.1 Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A6 = BS untuk membuat tabel *Node* 1.1.2.1. Langkah ini dilakukan dengan membuat cabang untuk tiap-tiap nilai, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A6 = BS

Jumlah kejadian A6 = BS, dinotasikan dengan $n(S) = 718$, dengan $n(M) = 710$ dan $n(TM) = 8$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{710}{718} \right) * \frac{\log(710/718)}{\log(2)} + \left(\frac{8}{718} \right) * \frac{\log(8/718)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,088272)$$

$$Entropy(Total) = 0,088272$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

a. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 717$, dengan $n(M) = 711$ dan $n(TM) = 6$.

$$Entropy(S) = - \left(\left(\frac{711}{717} \right) * \frac{\log(711/717)}{\log(2)} + \left(\frac{6}{717} \right) * \frac{\log(6/717)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,069769)$$

$$Entropy(S) = 0,069769$$

b. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Hubungan Pemilik)

a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 717$, dengan $n(M) = 710$ dan $n(TM) = 7$.

$$Entropy(S) = - \left(\left(\frac{710}{717} \right) * \frac{\log(710/717)}{\log(2)} + \left(\frac{7}{717} \right) * \frac{\log(7/717)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,079217)$$

$$Entropy(S) = 0,079217$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A3 (Daya Listrik)

a. Kategori = 0 VA

Jumlah kejadian A3 kategori = 0 VA, dinotasikan dengan $n(S) = 93$, dengan $n(M) = 92$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{92}{93} \right) * \frac{\log(92/93)}{\log(2)} + \left(\frac{1}{93} \right) * \frac{\log(1/93)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,085742)$$

$$Entropy(S) = 0,085742$$

b. Kategori = 450 VA

Jumlah kejadian A3 kategori = 450 VA, dinotasikan dengan $n(S) = 470$, dengan $n(M) = 468$ dan $n(TM) = 2$.

$$Entropy(S) = - \left(\left(\frac{468}{470} \right) * \frac{\log(468/470)}{\log(2)} + \left(\frac{2}{470} \right) * \frac{\log(2/470)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,039643)$$

$$Entropy(S) = 0,039643$$

c. Kategori = 900 VA

Jumlah kejadian A3 kategori = 900 VA, dinotasikan dengan $n(S) = 145$, dengan $n(M) = 145$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{145}{145} \right) * \frac{\log(145/145)}{\log(2)} + \left(\frac{0}{145} \right) * \frac{\log(0/145)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori = 1300 VA

Jumlah kejadian A3 kategori = 1300 VA, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(S) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,970950)$$

$$Entropy(S) = 0,970950$$

- e. Kategori > 1300 VA

Jumlah kejadian A3 kategori > 1300 VA, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

5. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

- a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 2$, dengan $n(M) = 2$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{2}{2} \right) * \frac{\log(2/2)}{\log(2)} + \left(\frac{0}{2} \right) * \frac{\log(0/2)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

- b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 716$, dengan $n(M) = 709$ dan $n(TM) = 7$.

$$Entropy(S) = - \left(\left(\frac{709}{716} \right) * \frac{\log(709/716)}{\log(2)} + \left(\frac{7}{716} \right) * \frac{\log(7/716)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,097164)$$

$$Entropy(S) = 0,188170$$

6. Menghitung Information Gain pada atribut A1, A2, A3, A6, dan A7 untuk mengetahui atribut dengan nilai Gain tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

- a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,088272 - \left(\left(\frac{717}{718} * 0,069769 \right) + \left(\frac{1}{718} * 0 \right) \right)$$

$$Gain = 0,018699$$

b. A2 (Hubungan Pemilik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$Gain = 0,088272 -$

$$\left(\left(\frac{717}{718} * 0,079217 \right) + \left(\frac{1}{718} * 0 \right) + \left(\frac{0}{718} * 0 \right) + \left(\frac{0}{718} * 0 \right) \right)$$

$Gain = 0,00916$

c. A3 (Daya Listrik)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(= 0 VA)}{n} * Entropy n(= 0 VA) \right) + \left(\frac{n(= 450 VA)}{n} * Entropy n(= 450 VA) \right) + \left(\frac{n(= 990 VA)}{n} * Entropy n(= 990 VA) \right) + \left(\frac{n(= 1300 VA)}{n} * Entropy n(= 1300 VA) \right) + \left(\frac{n(> 1300 VA)}{n} * Entropy n(> 1300 VA) \right) \right)$$

$Gain = 0,088272 -$

$$\left(\left(\frac{93}{721} * 0,085742 \right) + \left(\frac{470}{721} * 0,039643 \right) + \left(\frac{145}{721} * 0 \right) + \left(\frac{10}{721} * 0,970950 \right) + \left(\frac{0}{721} * 0 \right) \right)$$

$Gain = 0,037693$

d. A7 (Menerima Bantuan Lain)

$Gain = Entropy (Total) -$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,088272 - \left(\left(\frac{2}{721} * 0 \right) + \left(\frac{716}{721} * 0,079308 \right) \right)$$

$Gain = 0,009185$

Tabel 4.7 Hasil Perhitungan *Information Gain* Tertinggi pada *Node* 1.1.2.1

<i>Node</i>	<i>Atribut</i>	<i>Kategori</i>	<i>Jumlah Sampel (n)</i>	<i>M</i>	<i>TM</i>	<i>Entropy</i>	<i>Gain</i>
1.1.2.1	A6 = BS		718	710	8	0,08827	-
	A1	S	717	711	6	0,06977	0,01859985
		TS	1	0	1	0	
	A2	PM	717	710	7	0,07922	0,00916576
		K	1	1	0	0	
		PN	0	0	0	0	
		L	0	0	0	0	
	A3	= 0 VA	93	92	1	0,08574	0,03769344
		= 450 VA	470	468	2	0,03964	
		= 900 VA	145	145	0	0	
		= 1300 VA	10	6	4	0,97095	
		> 1300 VA	0	0	0	0	
	A7	Y	2	2	0	0	0,00918549
		T	716	709	7	0,07931	

Karena atribut A3 memiliki nilai *information gain* terbesar, maka atribut A3 menjadi cabang atau *internal node* dari A6 = BS. Pada atribut A3, kategori = 900 VA memiliki 145 kasus dengan semua jawabannya adalah M (Menerima) dan pada perhitungan sebelumnya kategori >1300 VA hanya memiliki 1 kasus dengan jawaban adalah TM (Tidak Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A3 = 900 VA dan A3 > 1300 VA menjadi daun atau *leaf*. Selanjutnya akan dilakukan perhitungan untuk menentukan cabang dari atribut A3 = 0 VA.

4.4.5.4 Cabang dari A3 = 0 VA

Setelah diperoleh *node* 1.1.2. A3, maka langkah selanjutnya adalah mencari atribut sebagai *node* 1.1.2.1.1. Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A3 = 0 VA untuk membuat tabel *Node* 1.1.2.1.1. Langkah ini dilakukan dengan membuat cabang untuk tiap-tiap nilai, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A3 = 0 VA

Jumlah kejadian A3 = 0 VA, dinotasikan dengan $n(S) = 93$, dengan $n(M) = 92$ dan $n(TM) = 1$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{92}{93} \right) * \frac{\log(92/93)}{\log(2)} + \left(\frac{1}{93} \right) * \frac{\log(1/93)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,085742)$$

$$Entropy(Total) = 0,085742$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

- a. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 92$, dengan $n(M) = 92$ dan $n(TM) = 0$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{92}{92} \right) * \frac{\log(92/92)}{\log(2)} + \left(\frac{0}{92} \right) * \frac{\log(0/92)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0)$$

$$Entropy(Total) = 0$$

- b. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 1$, dengan $n(M) = 0$ dan $n(TM) = 1$.

$$Entropy(S) = - \left(\left(\frac{0}{1} \right) * \frac{\log(0/1)}{\log(2)} + \left(\frac{1}{1} \right) * \frac{\log(1/1)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Hubungan Pemilik)

- a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 93$, dengan $n(M) = 92$ dan $n(TM) = 1$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{92}{93} \right) * \frac{\log(92/93)}{\log(2)} + \left(\frac{1}{93} \right) * \frac{\log(1/93)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,085742)$$

$$Entropy(Total) = 0,085742$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 93$, dengan $n(M) = 92$ dan $n(TM) = 1$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{92}{93} \right) * \frac{\log(92/93)}{\log(2)} + \left(\frac{1}{93} \right) * \frac{\log(1/93)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,085742)$$

$$Entropy(Total) = 0,085742$$

5. Menghitung *Information Gain* pada atribut A1, A2, dan A7 untuk mengetahui atribut dengan nilai *Gain* tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,085742 - \left(\left(\frac{92}{93} * 0 \right) + \left(\frac{1}{93} * 0 \right) \right)$$

$$Gain = 0,085742$$

b. A2 (Hubungan Pemilik)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$$Gain = 0,085742 -$$

$$\left(\left(\frac{93}{93} * 0,085742 \right) + \left(\frac{0}{93} * 0 \right) + \left(\frac{0}{93} * 0 \right) + \left(\frac{0}{93} * 0 \right) \right)$$

$$Gain = 0$$

c. A7 (Menerima Bantuan Lain)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,085742 - \left(\left(\frac{0}{93} * 0 \right) + \left(\frac{93}{93} * 0,085742 \right) \right)$$

$$Gain = 0$$

Tabel 4.8 Hasil Perhitungan Information Gain Tertinggi pada Node 1.1.2.1.1

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1.1.2.1.1	A3 = 0 VA		93	92	1	0,08574	
	A1	S	92	92	0	0	0,08574
		TS	1	0	1	0	
	A2	PM	93	92	1	0,08574	0
		K	0	0	0	0	
		PN	0	0	0	0	
		L	0	0	0	0	
	A7	Y	0	0	0	0	0
		T	93	92	1	0,08574	

Karena atribut A1 memiliki nilai information gain terbesar, maka atribut A1 menjadi cabang atau *internal node* dari A3 = 0 VA. Pada atribut A1, kategori S memiliki 92 kasus dengan semua jawabannya adalah M (Menerima) dan pada kategori TS hanya memiliki 1 kasus dengan jawaban adalah TM (Tidak Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A1 = S dan A1 = TS menjadi daun atau *leaf*.

4.4.5.5 Cabang dari A3 = 450 VA

Setelah diperoleh cabang dari atribut A3 = 0 VA, maka langkah selanjutnya adalah mencari cabang dari atribut A3 = 450 VA. Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A3 = 450 VA untuk membuat tabel *Node* 1.1.2.1.2. Langkah ini dilakukan dengan membuat cabang dari atribut A3 = 450 VA, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A3 = 450 VA

Jumlah kejadian A3 = 450 VA, dinotasikan dengan $n(S) = 470$, dengan $n(M) = 469$ dan $n(TM) = 1$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{469}{470} \right) * \frac{\log(469/470)}{\log(2)} + \left(\frac{1}{470} \right) * \frac{\log(1/470)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,021952)$$

$$Entropy(Total) = 0,021952$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

- a. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 470$, dengan $n(M) = 470$ dan $n(TM) = 0$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{470}{470} \right) * \frac{\log(470/470)}{\log(2)} + \left(\frac{0}{470} \right) * \frac{\log(0/470)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0)$$

$$Entropy(Total) = 0$$

- b. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Hubungan Pemilik)

- a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 470$, dengan $n(M) = 468$ dan $n(TM) = 2$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{468}{470} \right) * \frac{\log(468/470)}{\log(2)} + \left(\frac{2}{470} \right) * \frac{\log(2/470)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,039643)$$

$$Entropy(Total) = 0,039643$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 2$, dengan $n(M) = 2$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{2}{2} \right) * \frac{\log(2/2)}{\log(2)} + \left(\frac{0}{2} \right) * \frac{\log(0/2)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 468$, dengan $n(M) = 466$ dan $n(TM) = 2$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{466}{468} \right) * \frac{\log(466/468)}{\log(2)} + \left(\frac{2}{468} \right) * \frac{\log(2/468)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,039786)$$

$$Entropy(Total) = 0,039786$$

5. Menghitung *Information Gain* pada atribut A1, A2, dan A7 untuk mengetahui atribut dengan nilai *Gain* tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,021952 - \left(\left(\frac{470}{470} * 0 \right) + \left(\frac{0}{470} * 0 \right) \right)$$

$$Gain = 0,021952$$

b. A2 (Hubungan Pemilik)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$$Gain = 0,021952 -$$

$$\left(\left(\frac{470}{470} * 0,039786 \right) + \left(\frac{0}{470} * 0 \right) + \left(\frac{0}{470} * 0 \right) + \left(\frac{0}{470} * 0 \right) \right)$$

$$Gain = -0,017690$$

c. A7 (Menerima Bantuan Lain)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,021952 - \left(\left(\frac{2}{470} * 0 \right) + \left(\frac{468}{470} * 0,039786 \right) \right)$$

$$Gain = -0,017690$$

Tabel 4.9 Hasil Perhitungan Information Gain Tertinggi pada Node 1.1.2.1.2

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1.1.2.1.2	A3 = 450 VA		470	469	1	0,02195	
	A1	S	470	470	0	0	0,02195
		TS	0	0	0	0	
	A2	PM	470	468	2	0,03964	-0,0177
		K	0	0	0	0	
		PN	0	0	0	0	
		L	0	0	0	0	
	A7	Y	2	2	0	0	-0,0177
		T	468	466	2	0,03979	

Karena atribut A1 memiliki nilai information gain terbesar, maka atribut A1 menjadi cabang atau *internal node* dari A3 = 450 VA. Pada atribut A1, kategori S memiliki 470 kasus dengan semua jawabannya adalah M (Menerima). Sehingga tidak diperlukan perhitungan lebih lanjut dan A1 = S dan menjadi daun atau *leaf*.

4.4.5.6 Cabang dari A3 = 1300 VA

Setelah diperoleh cabang dari atribut A3 = 450 VA, maka langkah selanjutnya adalah mencari cabang dari atribut A3 = 1300 VA. Untuk memudahkan, dataset difilter dengan mengambil data yang dimiliki A3 = 1300 VA untuk membuat tabel *Node* 1.1.2.1.4. Langkah ini dilakukan dengan membuat cabang dari atribut A3 = 1300 VA, diantaranya sebagai berikut:

1. Menghitung *entropy* total dari atribut A3 = 1300 VA

Jumlah kejadian A3 = 1300 VA, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,970950)$$

$$Entropy(Total) = 0,970950$$

2. Menghitung *entropy ration* dari atribut A1 (Kesesuaian Alamat)

c. Kategori S

Jumlah kejadian A1 kategori S, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,970950)$$

$$Entropy(Total) = 0,970950$$

d. Kategori TS

Jumlah kejadian A1 kategori TS, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

3. Menghitung *entropy ration* dari atribut A2 (Hubungan Pemilik)

a. Kategori PM

Jumlah kejadian A2 kategori PM, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(Total) = - \left(\left(\frac{n(M)}{n} \right) * \frac{\log(n(M)/n)}{\log(2)} + \left(\frac{n(TM)}{n} \right) * \frac{\log(n(TM)/n)}{\log(2)} \right)$$

$$Entropy(Total) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(Total) = -(-0,970950)$$

$$Entropy(Total) = 0,970950$$

b. Kategori K

Jumlah kejadian A2 kategori K, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

c. Kategori PN

Jumlah kejadian A2 kategori PN, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

d. Kategori L

Jumlah kejadian A2 kategori L, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

4. Menghitung *entropy ration* dari atribut A7 (Menerima Bantuan Lain)

a. Kategori Y

Jumlah kejadian A7 kategori Y, dinotasikan dengan $n(S) = 0$, dengan $n(M) = 0$ dan $n(TM) = 0$.

$$Entropy(S) = - \left(\left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} + \left(\frac{0}{0} \right) * \frac{\log(0/0)}{\log(2)} \right)$$

$$Entropy(S) = -(-0)$$

$$Entropy(S) = 0$$

b. Kategori T

Jumlah kejadian A7 kategori T, dinotasikan dengan $n(S) = 10$, dengan $n(M) = 6$ dan $n(TM) = 4$.

$$Entropy(S) = - \left(\left(\frac{6}{10} \right) * \frac{\log(6/10)}{\log(2)} + \left(\frac{4}{10} \right) * \frac{\log(4/10)}{\log(2)} \right)$$

$$Entropy(S) = -(-0,970950)$$

$$Entropy(S) = 0,970950$$

5. Menghitung *Information Gain* pada atribut A1, A2, dan A7 untuk mengetahui atribut dengan nilai *Gain* tertinggi menggunakan persamaan (2.2), sehingga diperoleh:

a. A1 (Kesesuaian Alamat)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(S)}{n} * Entropy n(S) \right) + \left(\frac{n(TS)}{n} * Entropy n(TS) \right) \right)$$

$$Gain = 0,970950 - \left(\left(\frac{10}{10} * 0,970950 \right) + \left(\frac{0}{0} * 0 \right) \right)$$

$$Gain = 0$$

b. A2 (Hubungan Pemilik)

$$Gain = Entropy(Total) -$$

$$\left(\left(\frac{n(PM)}{n} * Entropy n(PM) \right) + \left(\frac{n(K)}{n} * Entropy n(K) \right) + \left(\frac{n(PN)}{n} * Entropy n(PN) \right) + \left(\frac{n(L)}{n} * Entropy n(L) \right) \right)$$

$$Gain = 0,970950 -$$

$$\left(\left(\frac{10}{10} * 0,970950 \right) + \left(\frac{0}{10} * 0 \right) + \left(\frac{0}{10} * 0 \right) + \left(\frac{0}{10} * 0 \right) \right)$$

$$Gain = 0$$

c. A7 (Menerima Bantuan Lain)

$$Gain = Entropy (Total) -$$

$$\left(\left(\frac{n(Y)}{n} * Entropy n(Y) \right) + \left(\frac{n(T)}{n} * Entropy n(T) \right) \right)$$

$$Gain = 0,970950 - \left(\left(\frac{0}{10} * 0 \right) + \left(\frac{10}{10} * 0,970950 \right) \right)$$

$$Gain = 0$$

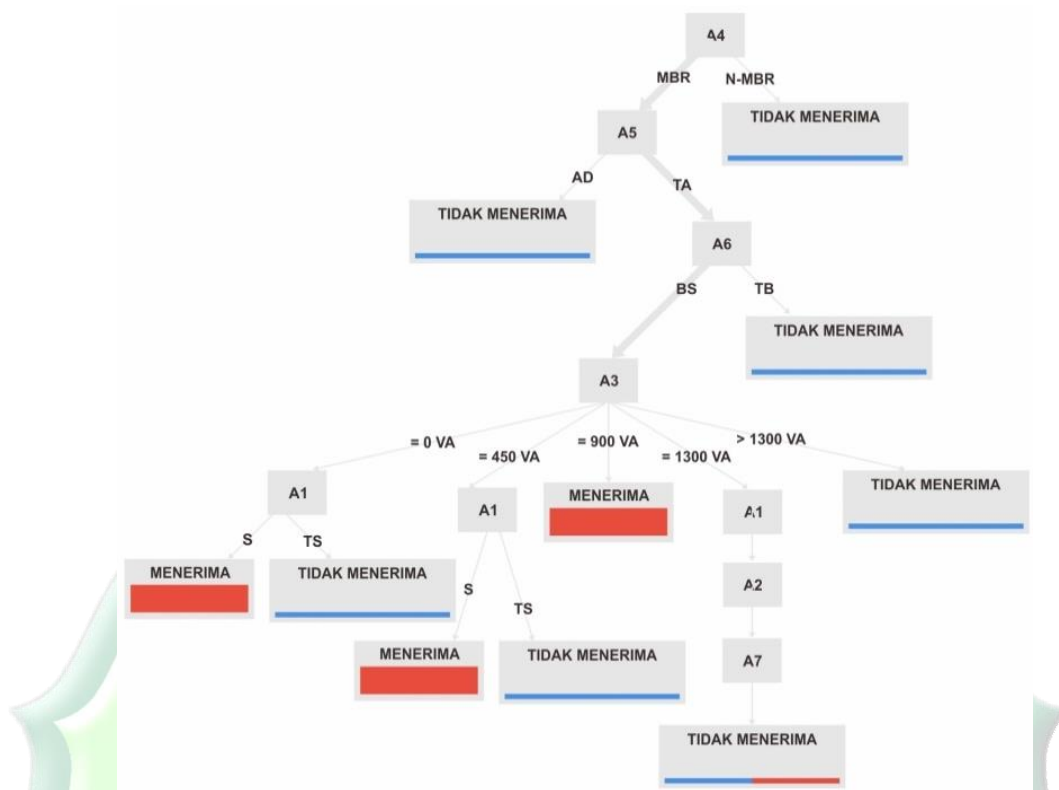
Tabel 4.10 Hasil Perhitungan *Information Gain* Tertinggi pada Node 1.1.2.1.4

Node	Atribut	Kategori	Jumlah Sampel (n)	M	TM	Entropy	Gain
1.1.2.1.4	A3 = 1300 VA		10	6	4	0,97095	0
	A1	S	10	6	4	0,97095	
		TS	0	0	0	0	
	A2	PM	10	6	4	0,97095	0
		K	0	0	0	0	
		PN	0	0	0	0	
		L	0	0	0	0	
	A7	Y	0	0	0	0	0
		T	10	6	4	0,97095	

Karena semua atribut memiliki nilai *information gain* sebesar 0, maka tidak diperlukan perhitungan lebih lanjut. *Leaf* dari perhitungan ini menghasilkan perbandingan antara M (Menerima) dan TM (Tidak Menerima). yaitu 6 : 4.

4.5 Pohon Keputusan dan Aturan

Berdasarkan perhitungan dengan menerapkan algoritma *Iterative Dichotomiser Three* (ID3) pada *dataset* berjumlah 1050 *record* yang telah dilakukan sebelumnya, dihasilkan pohon keputusan sebagai berikut:



Gambar 4.19 Pohon Keputusan

Adapun aturan *rules* dari perhitungan pada *dataset* berjumlah 1050 *record* adalah sebagai berikut:

1. Jika $A4 = N\text{-MBR}$, maka $HASIL = TIDAK\ MENERIMA\ (TM)$.
2. Jika $A4 = MBR$ dan $A5 = AD$, maka $HASIL = TIDAK\ MENERIMA\ (TM)$.
3. Jika $A4 = MBR$ dan $A5 = TA$ dan $A6 = TB$, maka $HASIL = TIDAK\ MENERIMA\ (TM)$.
4. Jika $A4 = MBR$ dan $A5 = TA$ dan $A6 = BS$ dan $A3 = 900\ VA$, maka $HASIL = MENERIMA\ (M)$.
5. Jika $A4 = MBR$ dan $A5 = TA$ dan $A6 = BS$ dan $A3 = > 1300\ VA$, maka $HASIL = TIDAK\ MENERIMA\ (TM)$.
6. Jika $A4 = MBR$ dan $A5 = TA$ dan $A6 = BS$ dan $A3 = 0\ VA$ dan $A1 = S$, maka $HASIL = MENERIMA\ (M)$.
7. Jika $A4 = MBR$ dan $A5 = TA$ dan $A6 = BS$ dan $A3 = 0\ VA$ dan $A1 = TS$, maka $HASIL = TIDAK\ MENERIMA\ (TM)$.

8. Jika A4 = MBR dan A5 = TA dan A6 = BS dan A3 = 450 VA dan A1 = S, maka HASIL = MENERIMA (M).
9. Jika A4 = MBR dan A5 = TA dan A6 = BS dan A3 = 450 VA dan A1 = TS, maka HASIL = TIDAK MENERIMA (TM).
10. Jika A4 = MBR dan A5 = TA dan A6 = BS dan A3 = 1300 VA dan A1 dan A2 dan A7, maka HASIL = TIDAK MENERIMA.

4.6 Evaluasi dan Validasi Hasil

Pengujian dalam penelitian ini dilakukan menggunakan *confusion matrix* sebagai performance yang menghasilkan nilai *accuracy*, *precision* dan *recall*. Selain itu, dilakukan pengujian menggunakan kurva ROC (*Receiver Operating Characteristic*) yang menghasilkan nilai AUC (*Area Under Curve*).

4.6.1 Confusion Matrix

Tabel 4.11 *Performance Vector*

	<i>True Menerima</i>	<i>True Tidak Menerima</i>
<i>Pred. Menerima</i>	707	6
<i>Pred. Tidak Menerima</i>	4	71

Berdasarkan hasil *performance vector* pada tabel 4.11, dapat disimpulkan *data training* yang diprediksi dengan benar menggunakan algoritma *Iterative Dichotomiser Three* (ID3) sejumlah 788 *record* dengan 707 pendaftar menerima bantuan program Hibah Air Minum Perdesaan (PPA) dan 71 pendaftar tidak menerima. Kemudian sebanyak 4 pendaftar menerima, tetapi diklasifikasikan tidak menerima. Serta terdapat sebanyak 6 pendaftar yang tidak menerima, tetapi diklasifikasikan menerima, sehingga dihasilkan:

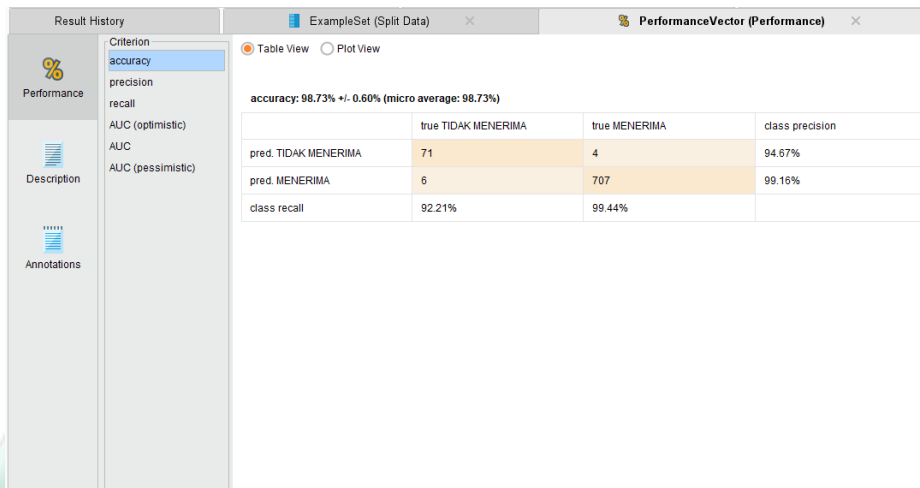
- a. *Accuracy*

$$Accuracy = \frac{(TP + TN)}{(TP + FN + FP + TN)} * 100\%$$

$$Accuracy = \frac{(707 + 71)}{(707 + 6 + 4 + 71)} * 100\%$$

$$Accuracy = 98,73\%$$

Berdasarkan uji validitas yang telah dilakukan dapat disimpulkan bahwa persentase *accuracy* terhadap algoritma *Iterative Dichotomiser Three* (ID3) adalah sebesar 98,73%. Berikut merupakan hasil *performance* pada *RapidMiner* versi 9.6:



The screenshot shows the PerformanceVector (Performance) window in RapidMiner. The left sidebar lists various performance criteria: accuracy, precision, recall, AUC (optimistic), AUC, and AUC (pessimistic). The main area displays the 'Table View' of the performance metrics. The overall accuracy is 98.73% with a +/- 0.60% (micro average) margin. A confusion matrix table is provided below.

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	71	4	94.67%
pred. MENERIMA	6	707	99.16%
class recall	92.21%	99.44%	

Gambar 4.20 Performance Accuracy

b. *Precision*

$$Precision = \frac{TP}{(TP + FP)} * 100\%$$

$$Precision = \frac{707}{(707 + 6)} * 100\%$$

$$Precision = 99,17\%$$

Berdasarkan uji validitas yang telah dilakukan dapat disimpulkan bahwa persentase *precision* terhadap algoritma *Iterative Dichotomiser Three* (ID3) adalah sebesar 99,17%. Berikut merupakan hasil *performance* pada *RapidMiner* versi 9.6:

Result History | ExampleSet (Split Data) | PerformanceVector (Performance)

Criterion: accuracy, precision, recall, AUC (optimistic), AUC, AUC (pessimistic)

Performance: precision

Description: AUC (optimistic), AUC, AUC (pessimistic)

Annotations:

Table View | Plot View

precision: 99.17% +/- 0.72% (micro average: 99.16%) (positive class: MENERIMA)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	71	4	94.67%
pred. MENERIMA	6	707	99.16%
class recall	92.21%	99.44%	

Gambar 4.21 Performance Precision

c. Recall

$$\text{Recall} = \frac{TP}{(TP + FN)} * 100\%$$

$$\text{Recall} = \frac{707}{(707 + 4)} * 100\%$$

$$\text{Recall} = 99,44\%$$

Berdasarkan uji validitas yang telah dilakukan dapat disimpulkan bahwa persentase *recall* terhadap algoritma *Iterative Dichotomiser Three* (ID3) adalah sebesar 99,44%. Berikut merupakan hasil *performance* pada *RapidMiner* versi 9.6:

Result History | ExampleSet (Split Data) | PerformanceVector (Performance)

Criterion: accuracy, precision, recall, AUC (optimistic), AUC, AUC (pessimistic)

Performance: recall

Description: AUC (optimistic), AUC, AUC (pessimistic)

Annotations:

Table View | Plot View

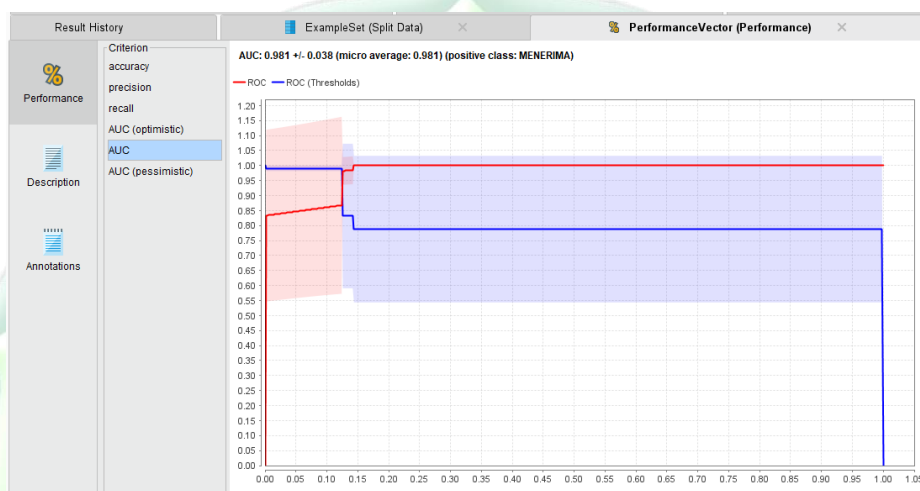
recall: 99.44% +/- 0.98% (micro average: 99.44%) (positive class: MENERIMA)

	true TIDAK MENERIMA	true MENERIMA	class precision
pred. TIDAK MENERIMA	71	4	94.67%
pred. MENERIMA	6	707	99.16%
class recall	92.21%	99.44%	

Gambar 4.22 Performance Recall

4.6.2 Kurva ROC (*Receiver Operating Characteristic*)

Berdasarkan uji validitas yang telah dilakukan dapat disimpulkan bahwa nilai AUC (*Area Under Curva*) dari Kurva ROC (*Receiver Operating Characteristic*) terhadap algoritma *Iterative Dichotomiser Three* (ID3) adalah sebesar 0,981 sehingga dikategorikan sebagai klasifikasi yang sangat baik. Berikut merupakan gambaran kurva ROC (*Receiver Operating Characteristic*) pada *RapidMiner* versi 9.6:



Gambar 4.23 Kurva ROC (*Receiver Operating Characteristic*)