

Penentuan Daya Listrik Rumah Tangga Menggunakan Metode *Decision Tree*

Yudhi Pratama Tanjung, Steven Sentinuwo, Agustinus Jacobus
Teknik Informatika Universitas Sam Ratulangi
120216098@student.unsrat.ac.id, steven@unsrat.ac.id, a.jacobus@unsrat.ac.id

Abstrak –PT. Perusahaan Listrik Negara (PLN) Persero sebagai perusahaan penyedia energi listrik mengeluarkan inovasi yaitu listrik prabayar (listrik pintar) sebagai opsi terbaru dalam hal pemilihan tipe pelanggan selain listrik pascabayar. Semakin bervariasinya daya listrik maupun tipe pelanggan membuat pelanggan maupun calon pelanggan listrik PLN memiliki pilihan yang beragam saat menentukan daya listrik. Masalah penentuan daya listrik juga dialami di sisi PLN dimana harus menentukan pelanggan yang tepat untuk mendapatkan subsidi listrik yang disediakan pada daya 450 VA dan 900. Dengan pendekatan *data mining* menggunakan metode *decision tree*, penelitian ini mencari pola atau model *decision tree* yang dapat menentukan daya listrik dan tipe pelanggan yang sesuai. Algoritma ID3 dan C4.5 dipilih sebagai algoritma metode *decision tree* yang digunakan untuk kalkulasi pada saat proses *training* data. Keluaran dari penelitian ini adalah pola atau model *decision tree* yang dihasilkan kedua algoritma tersebut yang dapat digunakan untuk menentukan daya listrik. Dari hasil pengujian sendiri didapatkan hasil akurasi algoritma ID3 sebesar 62 % sedangkan hasil algoritma C4.5 sebesar 88 %.

Kata kunci : Listrik, Data, Mining, ID3, C4.5.

I. PENDAHULUAN

Listrik merupakan energi vital bagi keberlangsungan aktivitas manusia baik bagi individu, kelompok masyarakat maupun dunia industri. Dengan kata lain energi listrik dapat dimanfaatkan untuk melakukan aktivitas dengan manfaat yang sangat besar di mana berbagai peralatan untuk memenuhi kebutuhan hidup dioperasikan dengan menggunakan energi listrik. Kegiatan masyarakat cenderung meningkat dari waktu ke waktu. Peningkatan kegiatan mendorong peningkatan pengoperasian peralatan dengan tenaga listrik. PT. Perusahaan Listrik Negara (Persero) yang selanjutnya disebut sebagai PLN merupakan perusahaan penyedia energi listrik bagi pelanggan.

Berdasarkan segmentasi yang dilakukan PLN, pelanggan listrik dibagi menjadi 5 sektor yaitu sektor sosial, rumah tangga, bisnis/usaha, industri dan publik. Segmen perumahan (rumah tangga) merupakan jumlah pelanggan listrik terbesar mencapai 94 % dari jumlah seluruh pelanggan [1].

Dari segi ekonomi, berdasarkan data dari website Kementerian Keuangan, dalam APBN-P 2015 pemerintah telah mengalokasikan anggaran subsidi listrik sebesar Rp 73.1 Triliun dan sebagian besar (sekitar 85%) dinikmati oleh golongan R1-450 VA dan R1-900 VA. Sebagian pelanggan rumah tangga R1-450 VA dan R1-900 VA merupakan pelanggan yang tidak mampu, namun sebagian

yang lain pelanggan tersebut telah mampu secara ekonomi. Data Susenas dari BPS tahun 2014 menunjukkan bahwa 4.3 juta pelanggan R1-450 VA adalah kelompok rumah tangga yang telah mampu karena termasuk dalam kelompok pengeluaran per kapita di atas Rp 1.1 juta per bulan. Disamping itu, sekitar 7 juta pelanggan R1-900 VA merupakan kelompok rumah tangga yang telah mampu karena termasuk dalam kelompok pengeluaran per kapita di atas Rp 1.7 juta per bulan. Apabila dikaitkan dengan Undang-undang Nomor 30 Tahun 2007 tentang Energi pada pasal 7 dinyatakan bahwa subsidi energi diperuntukkan bagi masyarakat tidak mampu, maka data di atas membuktikan bahwa subsidi listrik bagi pelanggan R1-450 VA dan R1-900 VA dinilai belum tepat sasaran [2].

Selain itu dari segi inovasi, listrik pintar atau listrik prabayar yang ditawarkan PLN dalam hal pemakaian daya listrik memberikan opsi baru tipe pelanggan bagi pelanggan listrik PLN. Dengan listrik pintar kita bisa menentukan kebutuhan listrik sesuai kebutuhan dengan menggunakan pulsa listrik (voucher/token listrik isi ulang) yang terdiri dari 20 digit nomor dan dimasukkan pada MPB (Meter Prabayar).

Dari sisi pelanggan PLN, dengan adanya listrik pintar (prabayar), membuat pemilihan daya listrik dan tipe pelanggan makin bervariasi, sehingga menuntut untuk calon pelanggan atau pelanggan lama PLN untuk bisa menentukan daya listrik sesuai kebutuhan dan karakteristik rumah tangga pada saat pemasangan baru maupun penambahan daya.

Dari sisi PLN, perlu adanya sistem yang bisa menentukan daya yang sesuai karakteristik rumah tangga agar bisa memudahkan dalam mengenal pelanggan yang layak menerima subsidi listrik dari pemerintah yang dikriteriakan pada daya 450 VA dan 900 VA agar subsidi tersebut bisa tepat sasaran.

Maka dari itu, dalam membantu calon pelanggan dan pelanggan listrik rumah tangga dalam menentukan daya listrik dan tipe pelanggan yang sesuai dengan kebutuhan rumah tangga itu sendiri baik dalam pemasangan baru maupun penambahan daya serta untuk membantu PLN dalam menentukan pelanggan rumah tangga yang layak menerima pemasangan daya listrik bersubsidi, maka dengan menerapkan data mining, akan dicari pola pemakaian listrik rumah tangga yang dapat menentukan daya listrik yang efisien berdasarkan faktor faktor yang dianggap mempengaruhi kebutuhan daya listrik rumah tangga menggunakan data mining.

Saat ini kita tahu bahwa manusia setiap hari memproduksi banyak sekali data, mulai dari bidang ekonomi hingga pendidikan. Data sendiri adalah sesuatu yang belum mempunyai arti bagi penerimanya dan masih memerlukan adanya suatu pengolahan. Salah satu bidang ilmu yang mempelajari tentang pengolahan data adalah

data mining. Dimana dalam data mining, data akan diproses untuk menemukan korelasi baru, pola atau tren dengan memilah-milah sejumlah besar data yang disimpan dalam sebuah repository dengan teknik tertentu.

Salah satu metode data mining adalah decision tree, dimana decision treemerupakan metode yang akan dihasilkan pola atau model dalam bentuk pohon keputusan. Di dalam metode decision tree juga terdapat beberapa algoritma yang dapat digunakan. Pada penelitian ini akan digunakan algoritma ID3 dan C4.5. Dengan pendekatan data mining, akan dicari aturan/pola yang dapat digunakan untuk menentukan daya listrik yang efisien rumah tangga dan bagi pelanggan PLN. Dengan adanya penelitian ini diharapkan dapat memberikan saran kepada pelanggan listrik PLN dalam menentukan daya listrik dan tipe pelanggan yang efisien rumah tangga dan sekaligus bagi PLN, dapat membedakan atau memfilter pelanggan yang layak maupun tidak layak dalam menerima daya yang bersubsidi sehingga pendistribusian subsidi listrik lebih tepat sasaran.

II. LANDASAN TEORI

A. Data

Menurut Andoko, “Data are raw facts ; That is , facts that have not yet been processed to reveal their meaning to the end user”, yang artinya adalah data merupakan fakta fakta dasar; yaitu, fakta yang belum diproses dan belum berarti bagi penggunaanya. Jadi dapat disimpulkan bahwa data adalah sesuatu yang belum mempunyai arti bagi penerimanya dan masih memerlukan adanya suatu pengolahan [3].

B. Data Mining

Data mining adalah proses menemukan korelasi baru ,pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam sebuah repository, menggunakan teknologi pengenalan pola serta statistik dan teknik matematika [4]. Sedangkan menurut Han[5], data mining adalah proses menemukan pola yang menarik dan pengetahuan dari sejumlah besar data. Sumber data dapat mencakup database, data gudang, web, repositori informasi lainnya, atau data yang mengalir ke dalam sistem dinamis.

C. Metode Decision Tree

Menurut Yuliharyani [6], Metode Decision Tree (Pohon Keputusan) merupakan bagian dari metode klasifikasi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan rules.

Kelebihan dari metode pohon keputusan adalah:

- 1). Daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel dan spesifik
- 2). Eliminasi perhitungan-perhitungan yang tidak diperlukan, karena ketika menggunakan metode pohon keputusan maka sampel diuji hanya berdasarkan kriteria atau kelas tertentu
- 3). Fleksibel untuk memilih fitur dari node internal yang berbeda, fitur yang terpilih akan membedakan suatu kriteria dibandingkan kriteria yang lain dalam node yang sama. Kefleksibelan metode pohon keputusan ini meningkatkan kualitas keputusan yang dihasilkan jika dibandingkan ketika menggunakan metode penghitungan satu tahap yang lebih konvensional.

4). Dalam analisis multivarian, dengan kriteria dan kelas yang jumlahnya sangat banyak, seorang penguji biasanya perlu mengestimasi baik itu distribusi dimensi tinggi ataupun parameter tertentu dari distribusi kelas tersebut. Metode pohon keputusan dapat menghindari munculnya permasalahan ini dengan menggunakan kriteria yang jumlahnya lebih sedikit pada setiap node internal tanpa banyak mengurangi kualitas keputusan yang dihasilkan.

Kekurangan pada pohon keputusan adalah:

- 1). Terjadi overlapping terutama ketika kelas-kelas dan kriteria yang digunakan jumlahnya sangat banyak. Hal tersebut juga dapat menyebabkan meningkatnya waktu pengambilan keputusan dan jumlah memori yang diperlukan.
- 2). Pengakumulasi jumlah kesalahan dari setiap tingkat dalam sebuah pohon keputusan yang besar
- 3). Kesulitan dalam mendesain pohon keputusan yang optimal
- 4). Hasil kualitas keputusan yang didapatkan dari metode pohon keputusan sangat tergantung pada bagaimana pohon tersebut didesain.

D. Algoritma ID3

Algoritma pohon keputusan ID3 atau Iterative Dichotomiser 3 (ID3) merupakan sebuah metode decision tree yang telah dikembangkan oleh J. Ross Quinlan sejak tahun 1986. Algoritma pada metode ini menggunakan konsep dari entropi informasi. Algoritma ini melakukan pencarian secara rakus/menyeluruh (greedy) pada kemungkinan pohon keputusan. Gambar 1 Akan menjelaskan proses algoritma ID3 melalui diagram alur (flowchart).

$$Entropy(S) = \sum_{i=1}^c -p_i * \log_2(p_i) \quad (1)$$

Dimana definisi notasi algoritma ID3 adalah :

- S : Ruang (data) sample
- c : Banyaknya atribut
- Pi: Hasil bagi antara jumlah pada setiap value dengan c

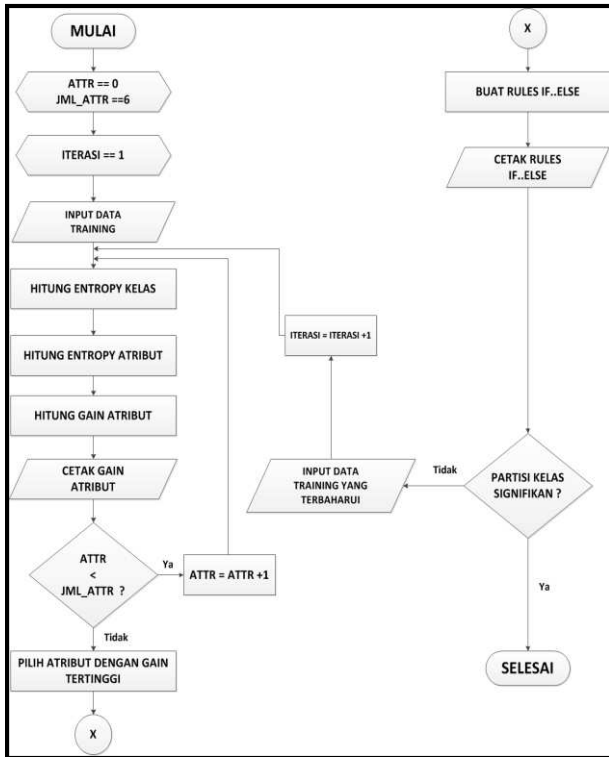
E. Algoritma C4.5

Algoritma C4.5 merupakan pengembangan dari algoritma ID3. Perbedaan algoritma ID3 dan C4.5 terletak pada pemilihan atribut berdasarkan gain ratio tertinggi dan pemangkasan pohon keputusan yang tidak dilakukan pada algoritma ID3. Gambar 2 akan menjelaskan secara terstruktur cara kerja dari algoritma C4.5. Formulasi algoritma C4.5 adalah :

$$GainRatio(S, A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (2)$$

Dimana :

- A : Atribut
- Gain(S,A) : Information gain pada atribut A
- SplitInfo(S,A) : Split information pada atribut A



Gambar 1. Diagram Alur Algoritma ID3

F. Confusion Matrix

Confusion Matrix adalah sebuah tabel untuk mengevaluasi informasi dari sistem dapat dilakukan dengan cara menghitung akurasi sistem berdasar inputan *data training* dan data uji. Untuk permasalahan dalam klasifikasi, pengukuran yang biasa digunakan adalah *precision*, *recall* dan *accuracy* [7]. *Precision* adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. *Accuracy* didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

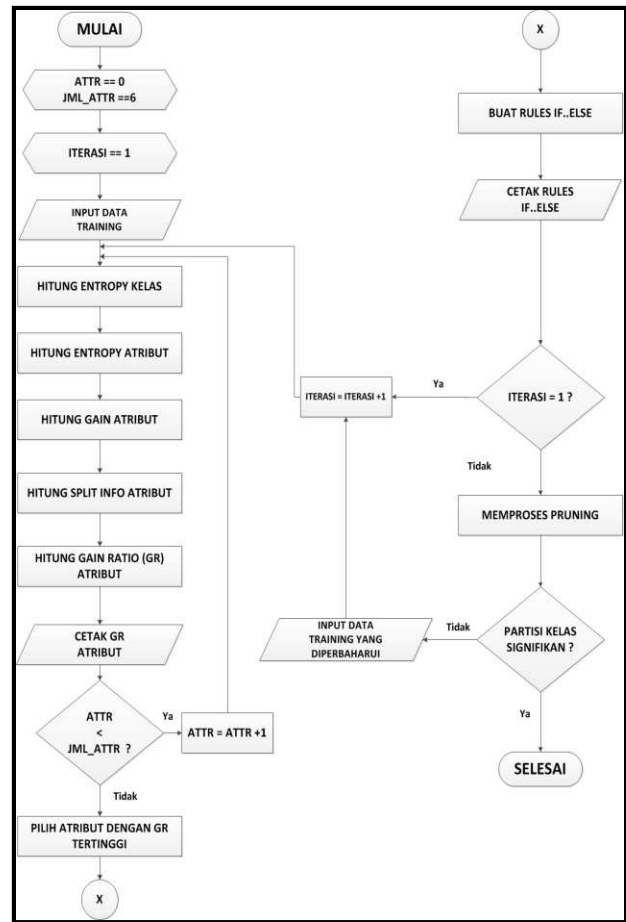
TABEL 1
CONFUSION MATRIX

		Valid Value	
		TRUE	FALSE
Prediction Value	TRUE	True Positive (TP) / Correct Value	False Positive (FP) / Unexpected Values
	FALSE	False Negative (FN) / Missing Value	True Negative (TN) (Correct of Opposite Value)

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (5)$$



Gambar 2. Diagram Alur Algoritma C4.5

III. METODOLOGI PENELITIAN

A. Data Penelitian

Data penelitian diambil dari 300 data kuisioner dengan menggunakan teknik *sampling* jenuh dan teknik *snowball sampling*. Data yang diambil antara lain : nama keluarga, alamat, jumlah tanggungan keluarga, pendapatan per bulan, pemakaian listrik per bulan, luas rumah, alat listrik rumah tangga, tipe pelanggan, dan daya.

B. Tahapan Penelitian

Banyak orang memperlakukan *data mining* sebagai sinonim untuk istilah populer lain, yaitu KDD (*Knowledge Discovery in Databases*) atau penemuan pengetahuan dalam database. Sementara yang lain melihat *data mining* hanya sebagai bagian dari langkah penting dalam proses penemuan pengetahuan [3]. Proses KDD ini terdiri dari :

1) *Data Cleaning & Integration* : Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* dan *integration* mencakup antara lain membuang duplikasi data, serta memeriksa data yang inkonsisten. Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal. Pada penelitian ini, data yang di *cleaning* adalah atribut/kolom “Alamat”, karena tidak akan diproses dalam metode *decision tree* pada proses *data mining*.

2) *Data Selection* : Pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap pengalihan informasi dalam KDD dimulai. Data akan di bagi menjadi 2 bagian yaitu data *training* dan data *testing*. Data *training* akan digunakan untuk pada saat proses *mining* untuk menemukan pola, sedangkan data *testing* digunakan untuk pengujian pola yang telah terbentuk pada proses *mining*. Pada penelitian ini data yang di gunakan yaitu 300 data, dibagi menjadi 150 untuk data *training*, dan 150 untuk data *testing*.

3) *Data Transformation* : *Data Transformation* adalah proses pengklasifikasikan data, adapun data yang akan diklasifikasikan adalah data jumlah tanggungan keluarga, pendapatan perbulan, biaya pemakaian listrik per bulan, luas rumah, perlengkapan listrik rumah tangga yang dipakai, perlengkapan listrik yang digunakan, serta atribut tipe pelanggan dan daya yang sebagai kelas. Kelasnya sendiri terdiri dari : Rendah_Reg (450 VA), Sedang_Reg (900 VA Pascabayar), Tinggi_Reg (1300 VA Pascabayar), Sedang_Pls (900 VA Prabayar), Tinggi_Pls (1300 VA Prabayar).

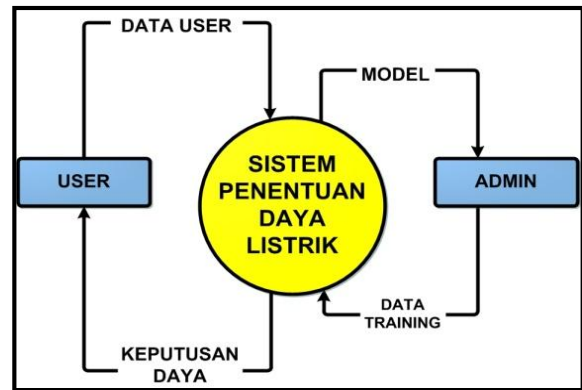
4) *Data Mining* : *Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Algoritma yang digunakan adalah ID3 dan C4.5

5) *Presentation & Evaluation* : Tahap ini mencakup ditampilkannya pola informasi dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses yang disebut dengan *presentation*. Pada tahap ini juga terdapat pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya yang disebut tahap *evaluation*. Pada tahap evaluasi akan digunakan metode *confusion matrix* sebagai metode yang akan menguji *precision*, *recall* maupun *accuracy* dari pola yang terbentuk. *Precision* adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. *Accuracy* didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

C. Pemodelan Proses Sistem

Pemodelan proses sistem adalah teknik mengelola dan mendokumentasikan struktur dan aliran data melalui proses sistem dan atau logika, kebijakan dan prosedur yang akan di implementasikan oleh proses sistem. Pemodelan proses pada penelitian ini dilakukan dengan menggunakan *Data Flow Diagram (DFD)*. *DFD* adalah diagram yang menggambarkan aliran data melalui sistem

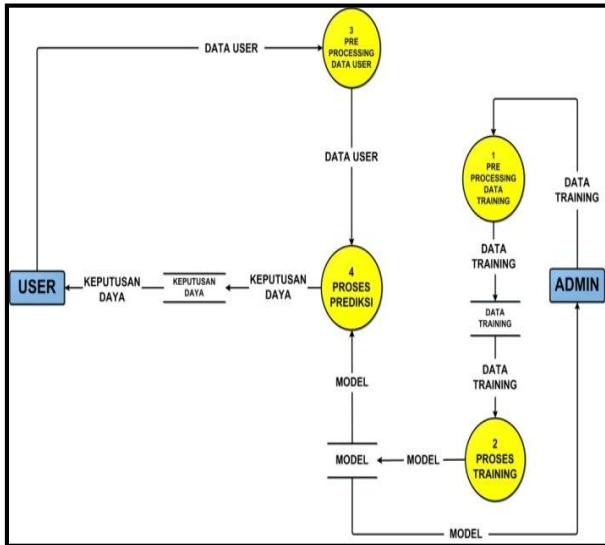
dan kerja atau pengolahan yang dilakukan oleh sistem tersebut. Diagram konteks adalah diagram yang menggambarkan aliran data dari sebuah sistem secara keseluruhan. Pada Gambar 1 dapat kita lihat bahwa sistem yang akan dibangun memiliki dua aktor yaitu *user* (Pelanggan PLN/Operator PLN) dan *admin* (Peneliti).



Gambar 3. Diagram Konteks

Pada Gambar 3 dijelaskan bahwa aktor *user* memasukkan inputan data *user* ke dalam sistem penentuan daya listrik dan sistem tersebut mengklasifikasikan data *user* yang di input kedalam kategori yang sesuai. Sedangkan pada aktor *admin* memasukkan inputan *data training* yang nanti akan menghasilkan model berupa *rules* yang akan menjadi acuan dalam proses pengklasifikasian.

Pertama, dilihat pada Gambar 4, dari sisi aktor *admin* dimana *admin* memasukkan data *training* untuk dilakukan proses *preprocessing* data yang selanjutnya disimpan pada *storage data training*. Yang kedua adalah proses inti yaitu proses *training* dimana data *training* hasil *preprocessing* diambil dan dikelola menggunakan algoritma dari metode *decision tree*. Hasil dari proses *training* adalah model berupa *rules* yang selanjutnya disimpan pada *storage rules* dan ditampilkan pada aktor *admin*. Yang ketiga dilihat dari sisi aktor *user* yaitu masuk dalam proses dimana *user* memasukan data *user* yang selanjutnya masuk dalam *preprocessing* data *user* agar menjadi data yang dapat dikelola oleh sistem. Hasil *preprocessing* data *user* kemudian akan dikelola dalam proses prediksi. Yang terakhir adalah proses prediksi dimana model atau *rules* yang disimpan pada *storage model* dan hasil *preprocessing* data *user* diambil untuk dilakukan proses prediksi/klasifikasi model yang telah terbentuk pada proses sebelumnya terhadap data *user* yang baru masuk.



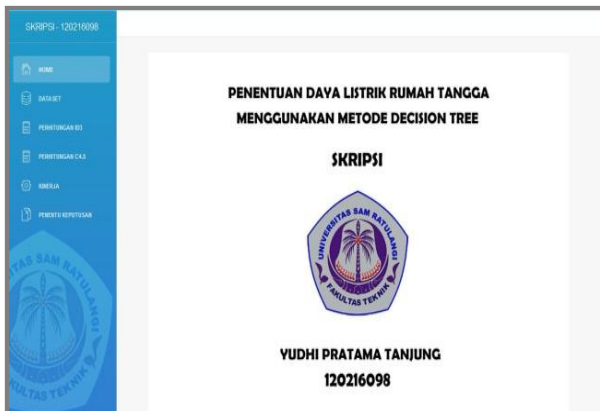
Gambar 4. Diagram Nol

Hasil dari proses prediksi atau klasifikasi ini adalah keputusan daya yang kemudian disimpan pada *storage* keputusan daya ditampilkan pada aktor *user* untuk menjadi *initial decision* sebelum mengambil keputusan dalam penentuan daya. Baik *data training*, dan *data user* terdiri dari atribut nama keluarga, jumlah tanggungan keluarga, pendapatan perbulan, biaya pemakaian listrik per bulan, luas rumah dan jumlah alat listrik rumah tangga dan daya sebagai atribut keputusan.

IV. HASIL DAN PEMBAHASAN

A. Interface Sistem

Gambar 5 merupakan *Interface Home* yang akan menunjukkan tampilan ketika sistem dijalankan. Tampilan ini akan pertama kali muncul ketikasistem ini diaks. Gambar 6 merupakan tampilan dari *Interface Dataset*. Tampilan ini akan memperlihatkan *dataset* yang akan digunakan pada sistem ini baik *data training* maupun *data testing*. Gambar 7 menunjukkan model yang telah dihasil oleh algoritma berupa pohon keputusan maupun *rules if else*. Dalam tampilan ini, terdapat fungsi untuk melakukan menginputan data baru serta upload multiple data via fungsi upload CSV.

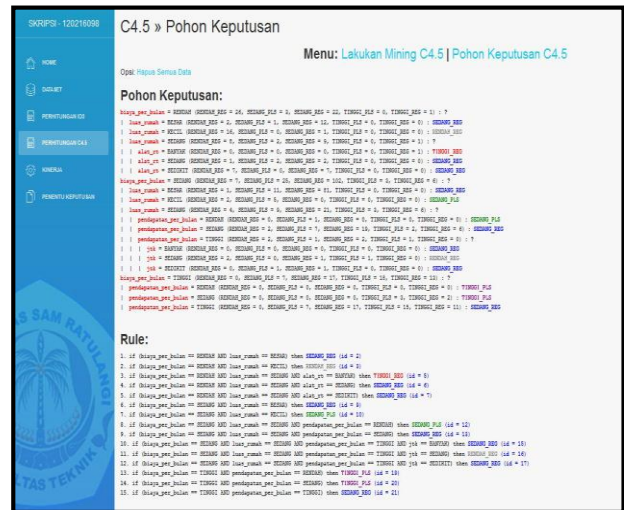


Gambar 5. Tampilan Depan

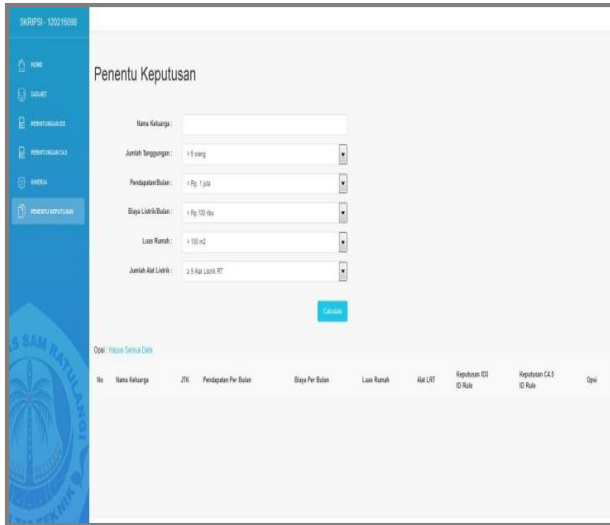
No	Nama Keluarga	JTK	Pendapatan Per Bulan	Biaya Per Bulan	Luas Rumah	Alat RT	Kelas (Daya)	Status Data	Opsi
1	SIDAKUNO	SEDANG	TINGGI	TINGGI	SEDANG	SEDANG	SEDANG_REG	Data Training	Edit Hapus
2	MASSE-RAWUNG	SEDIT	SEDANG	TINGGI	BESAR	SEDANG	TINGGI_PLS	Data Training	Edit Hapus
3	LASUTKARTYANI	SEDANG	SEDANG	SEDANG	SEDANG	SEDANG	SEDANG_PLS	Data Training	Edit Hapus
4	PONTOH-ROTTY	SEDANG	SEDANG	SEDANG	SEDANG	BANYAK	TINGGI_REG	Data Training	Edit Hapus
5	TULUNG-JANS	SEDANG	TINGGI	TINGGI	BESAR	SEDANG	TINGGI_REG	Data Training	Edit Hapus
6	RORIMPANDEY	SEDANG	SEDANG	SEDANG	SEDANG	SEDANG	SEDANG_REG	Data Training	Edit Hapus
7	KAPOH-TIMBLENG	SEDIT	SEDANG	SEDANG	SEDANG	SEDANG	TINGGI_REG	Data Training	Edit Hapus
8	MAWANTU-TAMPI	SEDANG	TINGGI	TINGGI	SEDANG	SEDANG	TINGGI_PLS	Data Training	Edit Hapus
9	KALUMATA	SEDANG	RENDAH	SEDANG	KEDUL	SEDANG	SEDANG_PLS	Data Training	Edit Hapus

Gambar 6. Interface Dataset

Gambar 8 menggambarkan *Interface Penentu Keputusan* yang berfungsi untuk memasukkan data baru untuk diklasifikasikan berdasarkan *rules* yang telah terbentuk sebelumnya. Data yang diinput berupa nama keluarga, jumlah tanggungan keluarga, pendapatan per bulan, biaya pemakaian per bulan, luas rumah, dan jumlah alat listrik. Hasil yang ditampilkan berupa keputusan daya yang menurut sistem, efisien untuk digunakan. Gambar 9 merupakan tampilan dari hasil saran daya yang dikeluarkan oleh sistem. Saran daya ini berupa tampilan *pop up* menggunakan *Javascript*, yang akan keluar jika selesai melakukan kalkulasi terhadap data baru yang baru saja diinput.



Gambar 7. Interface Model



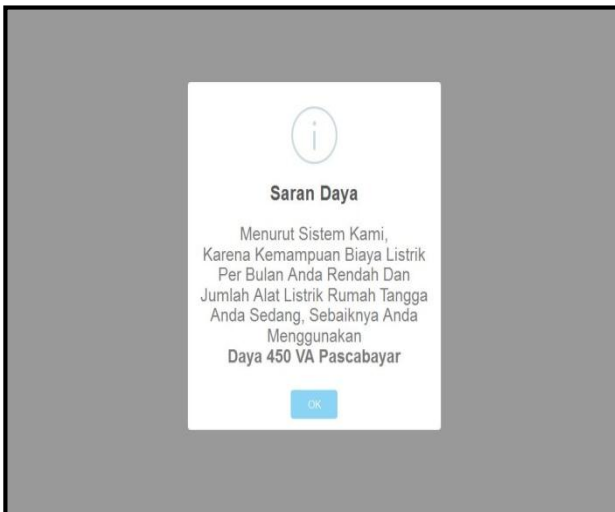
Gambar 8. Interface Penentu Keputusan

1	SANGKALIA-RUMAMPUK	SEDIKIT	TINGGI	SEDANG	KECIL	SEDANG	SEDANG_PLIS	SEDANG_PLIS	10
2	RAMBING-NGANGI	SEDIKIT	TINGGI	SEDANG	SEDANG	SEDANG	TINGGI_PLIS	SEDANG_REG	17
3	TIDOU-LUNTUNGAN	SEDIKIT	TINGGI	TINGGI	BESAR	SEDIKIT	TINGGI_REG	SEDANG_REG	21
4	AGUS ABDULLAH	SEDANG	SEDANG	SEDANG	KECIL	SEDANG	SEDANG_REG	SEDANG_PLIS	10
5	PONENGO-LAKASAN	SEDANG	TINGGI	TINGGI	BESAR	SEDANG	TINGGI_REG	SEDANG_REG	21
6	HAMEL SALAWADI	SEDIKIT	RENDAH	RENDAH	KECIL	SEDIKIT	RENDAH_REG	RENDAH_REG	3

Gambar 10. Tampilan Tabel Perbandingan

B. Pengujian Model

Sebelum melakukan pengujian model menggunakan *confusion matrix*, terlebih dahulu akan dikalkulasi tingkat kecocokan antara keputusan asli dan keputusan dari algoritma baik ID3 maupun C4.5 dengan Tabel Perbandingan untuk dimasukkan pada *confusion matrix*. Jika diperhatikan Gambar 8 maka dapat dilihat bahwa terdapat dua warna yaitu hijau dan merah. Hijau menunjukkan kecocokan antara nilai kecocokan dan keputusan dari algoritma. Sedangkan warna merah menunjukkan sebaliknya. Hasil kalkulasi dari tabel perbandingan akan dimasukkan pada *confusion matrix*, untuk dicari nilai *precision* dan *recall* tiap atribut dan nilai *accuracy* secara keseluruhan. Tinggi dan rendahnya *recall* menunjukkan seberapa besar jangkauan algoritma terhadap nilai yang sebenarnya. Tinggi dan rendahnya *precision* menunjukkan seberapa besar nilai ketepatan algoritma terhadap nilai yang telah ditemukan algoritma itu sendiri. Sedangkan *accuracy* menunjukkan tingkat ketelitian algoritma secara keseluruhan.



Gambar 9. Hasil Saran Daya

Berdasarkan Tabel 2 dapat dilihat bahwa algoritma C4.5 mengungguli algoritma ID3 baik per atribut (*recall* dan *precision*) maupun secara keseluruhan (*accuracy*). Dari segi *accuracy*, algoritma memiliki akurasi 88 % sedangkan algoritma ID3 memiliki akurasi 68 %.

TABEL II
HASIL UJI MODEL

Algoritma	ID3	C4.5
Atribut		
<i>Precision</i> Rendah_Reg	72 %	88 %
<i>Recall</i> Rendah_Reg	72 %	88 %
<i>Precision</i> Sedang_Reg	38.64 %	80.77 %
<i>Recall</i> Sedang_Reg	68 %	84 %
<i>Precision</i> Tinggi_Reg	74.19 %	92.31 %
<i>Recall</i> Tinggi_Reg	85.19 %	88.89 %
<i>Precision</i> Sedang_Pls	46.15 %	82.14 %
<i>Recall</i> Sedang_Pls	24 %	92 %
<i>Precision</i> Tinggi_Pls	91.67	100 %
<i>Recall</i> Tinggi_Pls	47.83 %	86.96 %
Accuracy	62 %	88 %

V. PENUTUP

A. Kesimpulan

Berdasarkan dari pembahasan-pembahasan pada bab-bab sebelumnya, maka dapat ditarik kesimpulan dari penelitian "Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree" adalah penggunaan algoritma dari metode decision tree baik algoritma ID3 maupun C4.5 dapat menentukan daya listrik rumah tangga. Hasil akurasi kedua algoritma menggunakan *confusion matrix* dengan data sejumlah 300 yang terdiri

dari 150 data *training* dan 150 data *testing* menunjukkan algoritma C4.5 dengan akurasi 88 % lebih akurat daripada algoritma ID3 yang memiliki akurasi 62 %.

B. Saran

Saran yang ingin disampaikan untuk pengembangan penelitian ini adalah sebagai berikut :

1). Sistem penentuan daya ini dapat disertakan bersamaan dengan sistem pemasangan / penambahan daya yang ada di situs PLN.

2). Penelitian ini bisa dianalisa menggunakan algoritma yang lain yang mungkin jauh lebih akurat agar bisa mendapatkan hasil yang akurat pula.

DAFTAR REFERENSI

- [1] Rohana, “*Analisis Perilaku Pelanggan PLN Dalam Pemakaian Daya Listrik*”, Thesis, Program Pasca Sarjana Teknik Industri Universitas Sumatera Utara. Medan, 2008.
- [2] Badan Kebijakan Fiskal Kemenkeu RI. *Mekanisme Kebijakan Subsidi Listrik Yang Lebih Tepat Sasaran*. <http://www.fiskal.kemenkeu.go.id/dw-konten-view.asp?id=20151230092250435820332>. 8 Maret 2016.
- [3] Andoko, Banni Satria, “*Transportasi Dan Transformasi Data Terjadwal Dengan Menggunakan SSIS*”. Jurnal Pendidikan Teknologi dan Kejuruan. Medan: Universitas Negeri Medan. 211-216, 2011.
- [4] Larose, Daniel (2005). “*Discovering Knowledge In Data*”. Canada: Willey Interscience.
- [5] Han Jiawei & Kamber Micheline (2012). “*Data mining Concepts and Techniques 3rd Edition*”. Waltham: Morgan Kaufmann.
- [6] Yuliharyani, Sisca, “*Algoritma Decision Tree C4.5 Untuk Klasifikasi Keluarga Peserta Jamkesmas Berdasarkan Kemiskinan*”. Skripsi Program S1 FMIPA Universitas Brawijaya. Malang, 2011.
- [7] Amin, Rafik Khairul dkk., “*Implementasi Klasifikasi Decision Tree Dengan Algoritma C4.5 Dalam Pengambilan Keputusan Permohonan Kredit Oleh Debitur (Studi Kasus: Bank Pasar Daerah Istimewa Yogyakarta)*”. Jurnal Tugas Akhir, Fakultas Informatika. Universitas Telkom, 2014.

SEKILAS TENTANG PENULIS



Yudhi Pratama Tanjung, lahir di Ternate, Maluku Utara pada tanggal 18 Juni 1995. Penulis merupakan seorang mahasiswa Program Studi Informatika, Jurusan Teknik Elektro, Fakultas Teknik, Universitas Sam Ratulangi pada tahun 2012 – 2016. Selama masa kuliah, penulistelah menjalani kerja praktek di

Kantor POS Indonesia, mengikuti kegiatan KuliahKerja Terpadu di Desa Talawaan, Kabupaten Minahasa Utara. Penulis mendapatkan sertifikasi dari Badan Nasional Sertifikasi Profesi (BNSP) untuk Pengelolaan Keamanan Informasi (*Information Security Management*) oleh Inixindo. Penulis memiliki ketertarikan dalam bidang *Machine Learning dan Artificial Intelegence*. Memiliki pengetahuan tentang Programming dan Networking.