



## Penerapan Algoritma C4.5 Untuk Memprediksi Besarnya Penggunaan Listrik Rumah Tangga di Kota Batam

Yulia<sup>a</sup>, Nurul Azwanti<sup>b</sup>

<sup>a</sup>Universitas Putera Batam, Fakultas Teknik dan Komputer, Teknik Informatika, yulia\_edwar@yahoo.com

<sup>b</sup>Universitas Putera Batam, Fakultas Teknik dan Komputer, Sistem Informasi, nurulazw@rocketmail.com

### Abstract

*Electricity has been proven from the activities carried out by everyday people. The use of electricity ranging from household appliances such as refrigerator, fan, iron, mixer, rice cooker to communication tools such as mobile phones, laptops and various other electronic devices desperately need electrical energy. The use of electricity today is the only thing that matters in people's lives. Batam is one of the major cities in Indonesia. The densely populated areas in Batam are the Batam Center area. Population density will also require electrical energy to be very high. Company as energy provider in Batam city is PT PLN Batam (b'right). If the population is not wise in using electricity, of course, will greatly affect the depletion of energy preparation listrik in Batam City itself. Therefore, techniques are needed to make predictions in the use of household appliances for the purpose of the government in list listing program. Techniques in making predictions can use data mining techniques by classifying data using C4.5 algorithm. Results from calculation Algorithm C4.5 Formation of decision tree where variable number of family members, Building area and length of time at home. Testing done with Weka has produced the same rule from the result tree that formed.*

*Keywords: algorithm C4.5, prediction use of electricity, decision tree*

### Abstrak

Listrik telah menjadi kebutuhan yang tidak terlepas dari kegiatan yang dilakukan oleh manusia sehari-hari. Penggunaan listrik mulai dari peralatan rumah tangga seperti kulkas, kipas angin, setrika, *mixer*, *rice cooker* hingga alat komunikasi seperti *handphone*, laptop dan berbagai macam alat eletronik lainnya sangat membutuhkan adanya energi listrik. Penggunaan tenaga listrik sekarang ini merupakan salah satu kebutuhan penting dalam kehidupan masyarakat. Batam merupakan salah satu kota besar yang ada di Indonesia. Daerah yang termasuk padat penduduk di Batam adalah daerah Batam Center. Kepadatan penduduk tersebut juga menyebabkan kebutuhan akan energi listrik menjadi sangat tinggi. Perusahaan sebagai penyedia sumber energi listrik di kota Batam adalah PT PLN Batam (b'right). Jika penduduk tidak bijak dalam menggunakan listrik, tentu saja akan sangat mempengaruhi menipisnya persediaan energi listrik di Kota Batam itu sendiri. Untuk itu diperlukan suatu teknik untuk melakukan prediksi dalam penggunaan listrik rumah tangga agar tujuan dari pemerintah dalam upaya menerapkan program penghematan listrik tercapai. Teknik dalam melakukan prediksi dapat menggunakan teknik *data mining* dengan melakukan klasifikasi data dengan menggunakan Algoritma C4.5. Hasil dari perhitungan Algoritma C4.5 membentuk pohon keputusan dimana variabel jumlah anggota keluarga, luas bangunan rumah dan lama waktu di rumah menjadi variabel penting dari prediksi besarnya penggunaan listrik. Pengujian dilakukan dengan Weka telah menghasilkan *rule* yang sama dari pohon keputusan yang terbentuk.

*Kata kunci: algoritma C4.5, prediksi, penggunaan listrik, pohon keputusan*

© 2018Jurnal RESTI

### 1. Pendahuluan

Listrik telah menjadi kebutuhan yang tidak terlepas dari kegiatan yang dilakukan oleh manusia sehari-hari. Peranan listrik tidak hanya menjadi kebutuhan sekunder, namun telah beralih menjadi kebutuhan primer dikarenakan tanpa adanya listrik segala kegiatan yang dilakukan akan sangat menghambat. Mulai dari peralatan rumah tangga seperti kulkas, kipas angin,

setrika, *mixer*, *rice cooker* hingga alat komunikasi seperti *handphone*, laptop dan berbagai macam alat eletronik lainnya sangat membutuhkan adanya energi listrik. Penggunaan tenaga listrik sekarang ini merupakan salah satu kebutuhan penting dalam kehidupan masyarakat dan seringkali dianggap sebagai salah satu tolak ukur taraf kesejahteraan masyarakat seiring dengan perkembangan teknologi [1]. Semakin

banyak penggunaan listrik, maka semakin banyak pula energi listrik yang dibutuhkan.

Batam merupakan kawasan yang saat ini memiliki jumlah penduduk yang sangat padat. Berdasarkan data Badan Pusat Statistik Kepulauan Riau pada tahun 2016 jumlah penduduk di Kota Batam mencapai 1.236.399 jiwa yang tersebar di berbagai daerah di Kota Batam. Salah satu daerah yang termasuk padat penduduk ialah daerah Batam Center. Wilayah Batam Center termasuk wilayah pusat kota Batam karena wilayah tersebut dekat dengan Bandara Hang Nadim Batam dan Pelabuhan Internasional Batam Center. Kepadatan penduduk tersebut juga menyebabkan kebutuhan akan energi listrik menjadi sangat tinggi. Jika penduduk tidak bijak dalam menggunakan listrik, tentu saja akan sangat mempengaruhi menipisnya persediaan energi listrik di Kota Batam itu sendiri. Setiap penduduk wajib mengetahui besarnya penggunaan listrik terlebih lagi dalam penggunaan listrik rumah tangga. Hal ini didasari atas perlunya kesadaran oleh penduduk dalam melakukan penghematan energi listrik. Penghematan listrik dapat dilakukan secara efektif apabila setiap rumah tangga sudah memiliki perilaku penghematan listrik [2].

Salah satu perusahaan penyedia energi listrik adalah Perusahaan Listrik Negara (PLN). Berdasarkan keputusan Menteri Negara Penanaman Modal dan Pembinaan BUMN, selaku Pemegang saham PT PLN (Persero) dalam surat No S-23/M-PM-PBMUN/2000 tanggal 23 Agustus 2000, pada tanggal 3 Oktober 2000, status PT PLN (Persero) Wilayah Khusus Batam berubah menjadi PT Pelayanan Listrik Nasional Batam (PT PLN Batam) dengan status sebagai anak perusahaan PT PLN (Persero), sebagai unit mandiri yang mengelola kelistrikan dari hulu sampai hilir dan pada Juni 2008 PT PLN Batam melakukan *rebranding* menjadi b'right PLN Batam.

PT Pelayanan Listrik Nasional Batam atau PT PLN Batam (*b'right*) yang berkantor pusat di Batam Center adalah sebagai perusahaan utilitas penyedia layanan kebutuhan listrik kepada masyarakat. Untuk mengetahui penggunaan listrik rumah tangga diperlukan adanya prediksi. Prediksi penggunaan listrik rumah tangga dapat menggunakan teknik *data mining*. *Data mining* merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang statistik, pembelajaran mesin dan sistem manajemen basis data [3].

Teknik dalam melakukan prediksi dapat menggunakan teknik klasifikasi pada Algoritma C4.5. Algoritma C4.5 memiliki banyak kelebihan diantaranya dapat menghasilkan model berupa pohon atau aturan yang dapat dengan mudah diinterpretasikan. Pohon atau aturan yang terbentuk dapat membantu dalam membaca

prediksi yang akan diterapkan pada prediksi penggunaan listrik rumah tangga.

## 2. Tinjauan Pustaka

### 2.1. Knowledge Discovery in Databases (KDD)

*Data Mining* disebut juga *Knowledge Discovery in Database (KDD)* didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *Knowledge Discovery in Database* melibatkan hasil proses data mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami [4].

Istilah *data mining* dan *Knowledge Discovery in Databases (KDD)* sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* [5].

### 2.2. Data Mining

*Data Mining* mengacu pada proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data besar [6]. Definisi lain *Data Mining* adalah serangkaian proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [7]. Menurut [8] *Data mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*. *Data mining* adalah proses yang menggunakan teknik *statistic*, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar.

Jika dilihat dari definisi *mining* yang berarti penambangan dari tumpukan data. Tumpukan data disini adalah data yang telah dikumpulkan sebelumnya. Tujuan dari penambangan data ini untuk memperoleh sejumlah informasi atau pengetahuan yang tersimpan didalam penambangan data dan dapat dimanfaatkan untuk kepentingan masyarakat maupun organisasi itu sendiri.

### 2.3. Klasifikasi

Teknik klasifikasi adalah pendekatan sistematis untuk membangun model klasifikasi dari kumpulan data masukan. Misalnya, teknik pohon keputusan, *Bayesian (Naive Bayesian dan Bayesian Belief Networks)*, Jaringan Saraf Tiruan (*Backpropagation*), teknik yang berbasis konsep dari penambangan aturan-aturan asosiasi, dan teknik lain (*K-Nearest Neighbor*,

algoritma genetik, teknik dengan pendekatan himpunan *rough* dan *fuzzy*). Klasifikasi merupakan teknik mengklasifikasikan data. Perbedaannya dengan metode *clustering* terletak pada data, dimana pada *clustering* variabel dependen tidak ada, sedangkan pada *classification* diharuskan ada variabel dependen [9].

#### 2.4. Pohon Keputusan (*Decision Tree*)

Pohon keputusan menggunakan representasi struktur pohon (*tree*) di mana setiap *node* merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut dan daun merepresentasikan kelas. *Node* yang paling atas dari pohon keputusan disebut sebagai *root*. Pohon keputusan merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami [7].

Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter yang disebut sebagai kriteria dalam pembentukan pohon. Manfaat utama dari penggunaan pohon keputusan adalah kemampuannya untuk mem-*break down* proses pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan (Elmande, 2012). Misalkan untuk menentukan main tenis, kriteria yang diperhatikan adalah cuaca, angin, dan suhu. Salah satu atribut merupakan atribut yang menyatakan data solusi per *item* data yang disebut atribut hasil.

Pada pohon keputusan terdapat 3 jenis *node*, yaitu:

1. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
2. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
3. *Leaf node* atau *terminal node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output*.

#### 2.5. Algoritma C4.5

Algoritma C4.5 yaitu sebuah algoritma yang digunakan untuk membangun *decision tree* (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma induksi pohon keputusan yaitu ID3 (*Iterative Dichotomiser 3*). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel *training*, label *training* dan atribut. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continuu data*, dan *pruning* [10].

Sebuah obyek yang diklasifikasikan dalam pohon harus dites nilai *Entropy* -nya. *Entropy* adalah ukuran dari teori informasi yang dapat mengetahui karakteristik

dari *impurity* dan *homogeneity* dari kumpulan data. Dari nilai *Entropy* tersebut kemudian dihitung nilai *information gain* (IG) masing-masing atribut. *Entropy* ( $S$ ) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel  $S$ . *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai *Entropy* maka akan semakin *Entropy* digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidakkaslian  $S$ .

Untuk memilih atribut akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti yang tertera dalam persamaan berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (1)$$

Di mana :

- $S$  : himpunan kasus
- $A$  : atribut
- $N$  : jumlah partisi atribut  $A$
- $|S_i|$  : jumlah kasus pada partisi ke- $i$
- $|S|$  : jumlah kasus dalam  $S$

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 2 berikut.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2)$$

Di mana :

- $S$  : himpunan kasus
- $A$  : fitur
- $N$  : jumlah partisi  $S$
- $p_i$  : proporsi dari  $S_i$  terhadap  $S$

Algoritma C4.5 dimulai dari proses memilih atribut dengan *gain* tertinggi sebagai akar pohon, kemudian membuat cabang untuk tiap-tiap nilai, lalu membagi kasus dalam cabang, setelah itu mengulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama [11]. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap-tiap nilai.
- c. Bagi kasus dalam cabang.
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

#### 2.6. Tinjauan Umum Listrik Rumah Tangga

Energi listrik adalah energi akhir yang dibutuhkan bagi peralatan listrik untuk menggerakkan motor, lampu penerangan, memanaskan, mendinginkan ataupun untuk menggerakkan kembali suatu peralatan mekanik untuk menghasilkan bentuk energi yang lain. [12]. Dalam rumah tangga, contoh pemakaian energi listrik seperti mesin cuci, kipas angin, kulkas, setrika dan lain sebagainya. Pemakaian listrik ini juga

dilatarbelakangi dari kemajuan teknologi yang mempermudah manusia dalam melakukan pekerjaannya. Bisa kita bayangkan jika saat ini kita belum pernah sama sekali menyentuh fungsi dari listrik. Namun, dibalik penggunaannya yang sangat membantu, konsumsi energi listrik juga harus diseimbangkan. Pemerintah Indonesia saat ini sedang gencar-gencarnya mengkampanyekan program penghematan energi listrik, dimana pemakai listrik rumah tangga diharapkan mampu menghemat konsumsi listrik.

2.7. Penelitian Terdahulu

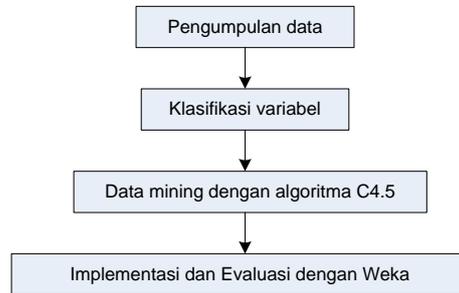
Adapun penelitian terdahulu yang mengacu pada pembahasan penelitian ini seperti terlihat pada tabel 1 berikut ini.

Tabel 1 Penelitian Terdahulu

Judul Penelitian	Hasil Penelitian
Implementasi Metode Klasifikasi <i>Naive Bayes</i> Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga [3]	Metode <i>Naive Bayes</i> memanfaatkan data <i>training</i> untuk menghasilkan probabilitas setiap kriteria untuk <i>class</i> yang berbeda, mengklasifikasikan 47 data dari 60 data yang diuji. Sehingga metode <i>Naive Bayes</i> berhasil memprediksi besarnya penggunaan listrik rumah tangga dengan persentase keakuratan sebesar 78,3333%.
<i>Data Mining</i> Peramalan Konsumsi Listrik dengan Pendekatan <i>Cluster Time Series</i> sebagai <i>Preprocessing</i> [13]	Hasil yang diperoleh adalah dihasilkan sebanyak tujuh <i>cluster</i> dengan anggota <i>cluster</i> terbanyak pada <i>cluster</i> ke empat yakni sebanyak 120 <i>client</i> . Selanjutnya model peramalan dengan menggunakan ANN lebih baik dibandingkan ARIMA. Diperoleh sebanyak 259 dari 348 <i>client</i> yang menyatakan bahwa permodelan dengan menggunakan ANN lebih baik dibandingkan ARIMA.
Penentuan Daya Listrik Rumah Tangga Menggunakan Metode <i>Decision Tree</i> [14]	Penggunaan algoritma dari metode <i>decision tree</i> baik algoritma ID3 maupun C4.5 dapat menentukan daya listrik rumah tangga. Hasil akurasi kedua algoritma menggunakan <i>confusion matrix</i> dengan data sejumlah 300 yang terdiri dari 150 data <i>training</i> dan 150 data <i>testing</i> menunjukkan algoritma C4.5 dengan akurasi 88 % lebih akurat daripada algoritma ID3 yang memiliki akurasi 62 %.

3. Metodologi Penelitian

Metodologi merupakan langkah-langkah yang akan dilakukan pada penelitian ini. Langkah – langkah tersebut digambarkan pada gambar 1.



Gambar 1 Metodologi Penelitian

Uraian dari gambar 1 adalah langkah awal dari penelitian ini yaitu peneliti terlebih dahulu melakukan pengumpulan data. Dari data yang dikumpul dilakukan pengklasifikasian pada variabel agar mudah diolah. Kemudian, dilakukan analisa dengan teknik algoritma C4.5 yang akan membentuk pohon keputusan. Dari hasil pohon keputusan tersebut akan dibandingkan dengan hasil pengolahan menggunakan Weka.

4. Hasil dan Pembahasan

4.1. Klasifikasi Data

Data yang telah dikumpulkan dalam kuisisioner harus diklasifikasikan untuk mempermudah dalam perhitungan. Adapun pengklasifikasian data pada variabel penelitian ini ada pada tabel 2 – 6.

Tabel 2 Klasifikasi Pendidikan

Pendidikan	Klasifikasi
>SMA/K Sarjana	SMA S

Tabel 3 Klasifikasi Total Barang Elektronik

Total Barang Elektronik	Klasifikasi
<11	S
>11	B

Tabel 4 Klasifikasi Jumlah ANggota Keluarga

Jumlah Anggota Keluarga	Klasifikasi
<3	S
>3	B

Tabel 5 Klasifikasi Lama Waktu di Rumah

Lama Waktu di Rumah	Klasifikasi
<12	S
>12	L

Tabel 6 Klasifikasi Luas Bangunan Rumah

Luas Bangunan Rumah	Klasifikasi
Type 32	K
Type 34	S
Type 36	L

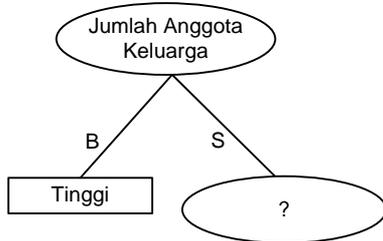
4.2. Hasil Penelitian

Dari data yang telah diklasifikasikan maka dilakukan perhitungan algoritma C4.5 untuk mencari *entropy* tertinggi dari setiap atribut. Hasil perhitungan node 1 dapat dilihat pada tabel 7.

Tabel 7 Perhitungan Node 1

Node	Jumlah Kasus(S)	Ya (S1)	Tidak (S2)	Entropy	Gain	
1 Total	50	35	15	0.881290899		
Pendidikan	SMA	23	17	6	0.828055725	0.004505515
	S	27	18	9	0.918295834	
Total Barang Elektronik	B	34	26	8	0.787126586	0.02966101
	S	16	9	7	0.988699408	
Jumlah Anggota Keluarga	B	27	25	2	0.380946586	0.221241189
	S	23	10	13	0.987692509	
Lama Waktu Di Rumah	L	46	34	12	0.828055725	0.054577382
	S	4	1	3	0.811278124	
Luas Bangunan Rumah	L	33	26	7	0.745517843	0.074425356
	S	13	8	5	0.961236605	
	K	4	1	3	0.811278124	

Pada tabel 7 didapat *entropy* tertinggi ada pada variabel Jumlah Anggota Keluarga dimana Jumlah Anggota Keluarga Banyak maka penggunaan listrik Tinggi, untuk Jumlah Anggota Keluarga Sedikit perlu dilakukan lagi perhitungan untuk mencari akarnya. Pohon keputusan yang terbentuk dari perhitungan Node 1 seperti pada gambar 2.



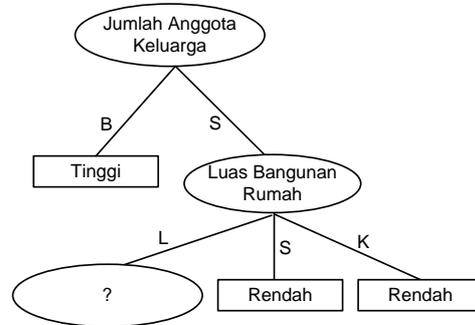
Gambar 2 Pohon Keputusan Node 1

Hasil perhitungan selanjutnya yaitu menyortir data Jumlah Anggota Sedikit dengan perhitungan node 1.1 dengan hasil perhitungan seperti pada Tabel 8.

Tabel 8 Perhitungan Node 1.1

Node	Jumlah Kasus(S)	Ya (S1)	Tidak (S2)	Entropy	Gain	
1.1 Jumlah Anggota Keluarga S	23	10	13	0.987692509		
Pendidikan	SMA	10	5	5	1	0.009602254
	S	13	5	8	0.961236605	
Total Barang Elcd	B	9	3	6	0.918295834	0.019663704
	S	14	7	7	1	
Lama Waktu Di F	L	19	9	10	0.998000884	0.022165148
	S	4	1	3	0.811278124	
Luas Bangunan R	L	17	10	7	0.977417818	0.265253252
	S	4	0	4	0	
	K	2	0	2	0	

Didapatkan lagi untuk *entropy* tertinggi ada pada variabel Luas Bangunan Rumah dengan kategori Luas Bangunan Rumah Sedang maka penggunaan listrik Rendah. Luas Bangunan Rumah Kecil juga penggunaan listrik Rendah. Namun, untuk Luas Bangunan Rumah Luas perlu dicari lagi atribut selanjutnya. Pohon keputusan yang terbentuk dapat dilihat pada gambar 3.



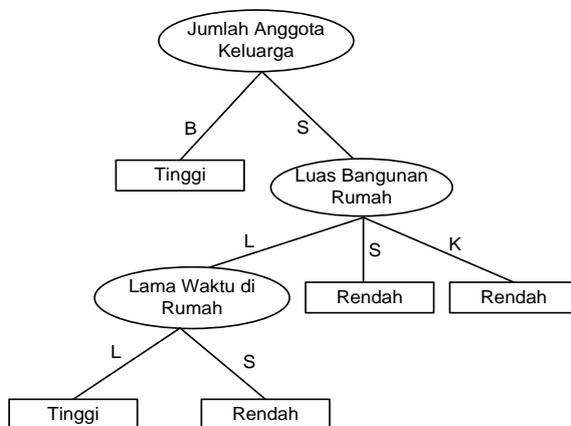
Gambar 3 Pohon Keputusan Node 1.1

Selanjutnya disortir data Luas Bangunan Rumah dengan kategori Luas. Adapun hasil perhitungan Node 1.2 dapat dilihat pada tabel 9.

Tabel 9 Perhitungan Node 1.2

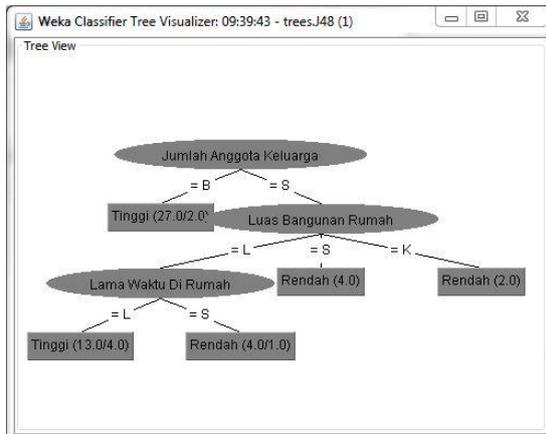
Node	Jumlah Kasus(S)	Ya (S1)	Tidak (S2)	Entropy	Gain	
1.2 Luas Bangun rumah L	17	10	7	0.977417818		
Pendidikan	SMA	6	5	1	0.650022422	0.1048021
	S	11	5	6	0.994030211	
Total Barang Elektronik	B	6	3	3	1	0.0125788
	S	11	7	4	0.945660305	
Lama Waktu Di Rumah	L	13	9	4	0.89049164	0.1055647
	S	4	1	3	0.811278124	

Dari tabel 9 dilihat *entropy* tertinggi ada pada variabel Lama Waktu di Rumah, dimana jika Lama maka penggunaan listrik Tinggi dan jika Sebentar maka penggunaan listrik Rendah. Pohon keputusan akhir yang terbentuk seperti pada gambar 4.



Gambar 4 Pohon Keputusan Node 1.2

Hasil pengujian yang dilakukan menggunakan *software* Weka dengan hasil pohon keputusan seperti gambar 5.



Gambar 5 Hasil Pengujian Weka

#### 4.2. Pembahasan

Pada pencarian *entropy* secara manual didapat *rule* dari hasil pohon keputusan sebagai berikut :

1. Jika jumlah anggota keluarga banyak maka penggunaan listrik tinggi.
2. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah sedang, maka penggunaan listrik rendah.
3. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah kecil, maka penggunaan listrik rendah.
4. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah luas dan lama waktu di rumah lama, maka penggunaan listrik tinggi.
5. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah luas dan lama waktu di rumah sebentar, maka penggunaan listrik rendah.

Kemudian kita bandingkan dengan hasil *rule* yang terbentuk dari pengujian Weka yaitu :

1. Jika jumlah anggota keluarga banyak maka penggunaan listrik tinggi.
2. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah sedang, maka penggunaan listrik rendah.
3. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah kecil, maka penggunaan listrik rendah.
4. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah luas dan lama waktu di rumah lama, maka penggunaan listrik tinggi.
5. Jika jumlah anggota keluarga sedikit dan luas bangunan rumah luas dan lama waktu di rumah sebentar, maka penggunaan listrik rendah.

Maka disimpulkan dari perhitungan manual dan pengujian dengan Weka *rule* yang dihasilkan sama dan pengujian ini sangat baik.

## 5. Kesimpulan

### 5.1. Simpulan

Simpulan dari penelitian ini adalah sebagai berikut :

1. Hasil perhitungan dengan menggunakan Algoritma C4.5 dapat melakukan prediksi besarnya penggunaan listrik di Kota Batam dengan memperhatikan nilai *gain* (penguatan) tertinggi dari lima atribut yang digunakan yaitu pendidikan, jumlah anggota keluarga, total barang elektronik, jumlah pemakai, lama waktu di rumah dan luas bangunan rumah.
2. Hasil perhitungan dengan Algoritma C4.5 dengan metode pohon keputusan dapat memberikan informasi *rule* prediksi untuk menggambarkan proses yang terkait dengan prediksi besarnya penggunaan listrik rumah tangga. Dimana, variabel jumlah anggota keluarga, luas bangunan rumah, dan lama waktu di rumah menjadi variabel penting dalam membentuk pohon keputusan.
3. Hasil pengujian yang dilakukan menggunakan Weka didapatkan *rule* yang dihasilkan dengan menggunakan perhitungan manual adalah sama. Sehingga pengujian dengan Weka sangat baik.

### 5.2. Saran

1. Perlu dilakukan penambahan variabel karena akan sangat mempengaruhi dari hasil prediksi untuk penelitian besarnya penggunaan listrik rumah tangga.
2. Perlu dilakukan penelitian untuk membandingkan hasil menggunakan algoritma lain guna mendapatkan hasil prediksi yang lebih baik lagi.

## Ucapan Terima Kasih

Terima kasih kepada DIKTI atas pendanaan penelitian ini sebagai Penelitian Dosen Pemula (PDP) Lampiran Surat Nomor : 0045/E3/LL/2018

## Daftar Rujukan

- [1] A. Rahman and K. Nanggalo, "Prakiraan Dan Analisa Kebutuhan Energi Listrik Provinsi Sumatera Barat Hingga Tahun 2024 Dengan Metode Analisis Regresi Linear Berganda," *J. Tek. Elektro ITP*, vol. 4, no. 2, 2015.
- [2] L. N. Y. Sari, Moh. Djemdjem Djamiludin, and Anggi Mayang, "Analisis Sikap dan Perilaku Penghematan Listrik Pada Sektor Rumah Tangga," vol. 4, no. 1, pp. 82–90, 2011.
- [3] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Citec J.*, vol. 2, pp. 207–217, 2015.
- [4] B. R. Siburian, "Aplikasi Data Mining Untuk Menampilkan Tingkat Kelulusan Mahasiswa Dengan Algoritma Apriori," *Pelita Inform. Budi Dharma*, vol. VII, pp. 56–61, 2014.
- [5] M. Le Cam, A. Daoud, and R. Zmeureanu, "Forecasting electric demand of supply fan using data mining techniques," *Energy*, vol. 101, pp. 541–557, 2016.
- [6] I. H. Selvia Lorena Br Ginting, Wendi Zarman, "Analisis Dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik," no. November, 2014.

- [7] A. Sijabat, "Penerapan Data Mining Untuk Pengolahan Data Siswa Dengan Menggunakan Metode Decision Tree ( Studi Kasus : Yayasan Perguruan," vol. V, pp. 7–12, 2015.
- [8] Kusri and Luthfi, *Algoritma Data Mining*. Yogyakarta: Penerbit Andi, 2009.
- [9] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," vol. VI, no. 1, pp. 15–20, 2014.
- [10] S. Faradillah, "Implementasi Data Mining Untuk Pengenalan Karakteristik Transaksi Customer," pp. 63–70, 2013.
- [11] Y. A. Fiandra, S. Defit, and Yuhandri, "Penerapan Algoritma C4.5 untuk Klasifikasi Data Rekam Medis berdasarkan International Classification Diseases (ICD-10)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 2, pp. 82–89, 2017.
- [12] W. Widjayanti, "Profil Konsumsi Energi Listrik Pada Hunian Rumah Tinggal Studi Kasus Rumah Desain Minimalis Ditinjau Dari Aspek Pencahayaan Buatan," *Enclosure*, vol. 6, no. 2, pp. 97–106, 2007.
- [13] M. A. A. Riyadi and K. Fithriasari, "Data Mining Peramalan Konsumsi Listrik dengan Pendekatan Cluster Time Series sebagai Preprocessing," vol. 5, no. 1, 2016.
- [14] Y. P. Tanjung, S. Sentinuwo, and A. Jacobus, "Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree."