



## PENENTUAN DAYA LISTRIK MENGGUNAKAN METODE DECISION TREE DI DESA TEBARA

Aristo Tagu Bore <sup>1\*</sup>, Trisno <sup>2</sup>, Titus Kurra <sup>3</sup>

<sup>1\*,2,3</sup>. Program Studi Teknik Informatika, Fakultas Teknik, Universitas Stella Maris Sumba

\*email koresponden: [trisnomtf@gmail.com](mailto:trisnomtf@gmail.com)

### Abstrak

Listrik merupakan salah satu kebutuhan dasar masyarakat dan merupakan salah satu kebutuhan hajat hidup orang banyak, sehingga perlu diatur dan disediakan oleh negara sesuai amanah undang-undang 1945 pasal 33. Subsidi diberikan dengan tujuan agar ketersediaan listrik dapat terpenuhi, serta membantu pelanggan yang kurang mampu dan masyarakat yang belum terjangkau pelayanan PT. Berdasarkan hasil evaluasi BKF dengan German International Cooperation (GIZ) terhadap subsidi listrik yang diberikan kepada kelompok pelanggan R1-450 VA dan R1-900 VA yang berlaku saat ini menunjukkan subsidi listrik tidak tepat sasaran, karena 5,9 juta pelanggan R1-450 VA dan 14,4 juta pelanggan R1-900 VA adalah kelompok rumah tangga yang telah mampu karena termasuk dalam pengeluaran per kapita lebih dari Rp.1 juta per bulan. Guna mengurangi risiko salah sasaran tersebut, perlu dilakukan klasifikasi dalam pemberian subsidi listrik berdasarkan kriteria-kriteria tertentu yang telah menjadi standar di PT. PLN menggunakan data mining dengan metode *decision tree*. Berdasarkan evaluasi dan hasil pengujian bab sebelumnya dengan dataset 50 data, 70 data dan 100 data produksi yang telah dilakukan, maka hasil pengukuran *Confusion Matrix* dalam penerapan data mining untuk memprediksi target Desa Tebara dengan metode *Decision Tree (Algoritma C4.5)* pada pengujian 50 data menghasilkan akurasi 74%, *precision* 60% dan *recall* 83,33%. Pada pengujian 70 data menghasilkan akurasi 81,43%, *precision* 76,92% dan *recall* 74,07%. Pada pengujian 100 data menghasilkan akurasi 82%, *precision* 76,67% dan *recall* 67,65%. Jadi untuk data uji lebih banyak akan menghasilkan akurasi yang lebih tinggi.

**Kata Kunci:** Algoritma C45, Daya Listrik, Desa Tebara

### Abstract

Electricity is one of the basic needs of society and is one of the basic necessities of life for many people, so it needs to be regulated and provided by the state in accordance with the mandate of the 1945 law, article 33. Subsidies are given with the aim of ensuring that the availability of electricity can be met, as well as helping customers who are less fortunate and communities that have not been reached by PT services. Based on the results of BKF's evaluation with the German International Cooperation (GIZ) regarding the electricity subsidies given to the current R1-450 VA and R1-900 VA customer groups, it shows that the electricity subsidies are not on target, because 5.9 million R1-450 VA and R1-900 VA customers The 14.4 million R1-900 VA customers are a group of households that can afford it because they include per capita expenditure of more than IDR 1 million per month. In order to reduce the risk of wrong targets, it is necessary to classify the provision of electricity subsidies based on certain criteria that have become standards at PT. PLN uses data mining with the decision tree method. Based on the evaluation and test results of the previous chapter



with datasets of 50 data, 70 data and 100 production data that have been carried out, the results of the Confusion Matrix measurements in applying data mining to predict targets for Tebara Village using the Decision Tree method (Algorithm C4.5) on testing 50 data produces 74% accuracy, 60% precision and 83.33% recall. Testing 70 data resulted in 81.43% accuracy, 76.92% precision and 74.07% recall. Testing 100 data resulted in 82% accuracy, 76.67% precision and 67.65% recall. So more test data will produce higher accuracy.

**Keywords:** C45 algorithm, Electrical power, Tebara Village

## 1. PENDAHULUAN

Indonesia merupakan salah satu negara kepulauan terbesar di dunia. Secara administratif wilayah Indonesia terbagi ke dalam beberapa tingkat wilayah yaitu provinsi, kabupaten/kota, kecamatan, dan desa yang merupakan wilayah administratif terkecil (BPS, 2010). Selain itu, Indonesia juga terkenal akan sumber daya alamnya, namun hal ini belum dapat dimanfaatkan secara optimal guna mewujudkan kehidupan rakyat yang lebih sejahtera. Decision Tree atau pohon keputusan adalah suatu metode klasifikasi yang terkenal atau populer dan juga banyak digunakan karena kemudahannya untuk diinterpretasikan oleh manusia. Pohon keputusan memiliki model seperti namanya yakni struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah sejumlah atau suatu data menjadi suatu aturan keputusan. Pohon keputusan juga berguna dalam pengeksploasian data, menemukan suatu hubungan tersembunyi antara sejumlah variabel dengan variabel target. Dengan adanya pembangunan perdesaan sesuai dengan Undang-undang No. 6 Tahun 2014 tentang Desa, saat ini mulai memunculkan wilayah desa yang memiliki kehidupan yang lebih maju, baik dari segi ketahanan sosial, ekonomi dan lingkungan. Terlebih lagi seiring dengan perkembangan yang semakin pesat juga melahirkan beberapa wilayah desa yang masuk ke dalam kategori desa maju. Hal tersebut menyebabkan adanya pergeseran penentuan karakteristik status perdesaan dan perkotaan. Oleh karena itu, perlu diperlukan adanya keseragaman penggunaan konsep, definisi, dan kriteria wilayah perkotaan dan perdesaan di Indonesia

Penelitian menurut Ni Kadek Sumarwati, G. K. Gandhiadi, Tjokorda Bagus Oka, 2018, "Mengklasifikasikan Desa Perkotaan dan Desa Perdesaan di Kabupaten Klungkung Menggunakan Metode Mamdani". Penelitian ini bertujuan untuk mengklasifikasikan desa perkotaan dan desa perdesaan di Kabupaten Klungkung menggunakan metode mamdani. Penelitian ini menggunakan data sekunder yang bersumber dari Badan Pusat Statistik 2016. Hasil dari penelitian menunjukkan bahwa dengan metode Mamdani menghasilkan 52 desa terklasifikasi sebagai desa perkotaan dan 7 desa sebagai desa perdesaan dengan tingkat akurasi sebesar 93%. Selain itu, terdapat perbedaan pada total skor dan status desa antara hasil menggunakan metode Mamdani dengan data asli.

Penelitian yang dilakukan oleh Afianti Sonya Kurniasari, Diah Safitri, Sudarno, 2014, "Pemisahan Desa/Kelurahan di Kabupaten Semarang Menurut Status Daerah Menggunakan Analisis Diskriminan Kuadratik Klasik dan Diskriminan Kuadratik Robust" Penelitian ini bertujuan untuk mengklasifikasikan desa/kelurahan di Kabupaten Semarang ke dalam kategori desa perkotaan atau desa perdesaan menggunakan analisis diskriminan kuadratik klasik dan analisis diskriminan kuadratik robust. Serta mengetahui persentase ketepatan hasil pengklasifikasian desa/kelurahan menggunakan analisis diskriminan kuadratik klasik dan analisis diskriminan kuadratik robust. Data yang digunakan merupakan data Badan Pusat Statistik (BPS) hasil Pendataan Potensi Desa (PODES) Kabupaten Semarang Tahun 2011. Hasil penelitian menunjukkan bahwa dengan menggunakan metode analisis diskriminan kuadratik terdapat 183 desa/kelurahan berstatus perdesaan dan 52 desa/kelurahan berstatus perkotaan dengan tingkat akurasi sebesar 87,23%. Sedangkan dengan analisis diskriminan



kuadratik robust menghasilkan tingkat akurasi yang lebih tinggi yaitu sebesar 89,79% dengan hasil terdapat 167 desa/kelurahan berstatus perdesaan dan 68 desa/kelurahan berstatus perkotaan. Sehingga tujuan dari penelitian ini, diantaranya adalah untuk membandingkan model skoring yang dilakukan BPS dengan model decision tree, kemudian untuk mengetahui akurasi model decision tree dalam pengklasifikasian status desa/kalurahan, lalu untuk mengetahui standar minimum suatu indikator mengklasifikasikan menjadi desa perkotaan (kalurahan), serta untuk memberikan metode decision tree sebagai metode alternatif untuk pengklasifikasian status desa/kalurahan.

## 2. METODE PENELITIAN

Dalam penelitian ini, metode pengumpulan data terdiri dari wawancara kepada yang ada di tempat lokasi dan Observasi dilakukan turun lapangan secara langsung pada Desa Tebara.

## 3. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan data primer yang didapat dari departemen Casting dengan jumlah 100 data produksi. Untuk data yang digunakan dalam data perhitungan decision tree secara manual yaitu 100 data produksi ( 80 data training dan 20 data testing). Perhitungan entropy dan gain secara manual menggunakan data testing yaitu 20 data. Dimana pada data testing memiliki 7 data produksi target dan 13 data produksi tidak target. Berikut data testing yang digunakan.

Tabel Tabel data testing

### 4.1.1. Pengujian Data Menggunakan Algoritma C4.5

Berikut ini merupakan langkah-langkah perhitungan Entropy dan Gain pada setiap atribut yang memiliki label target dan tidak target :

#### 1. Perhitungan Nilai Entropy

Langkah awal algoritma C4.5 adalah mencari nilai Entropy. Pertama tentukan terlebih dahulu nilai entropy total dalam kasus. Rumusnya sebagai berikut :

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

Tabel Tabel sampel jumlah kasus tiap atribut

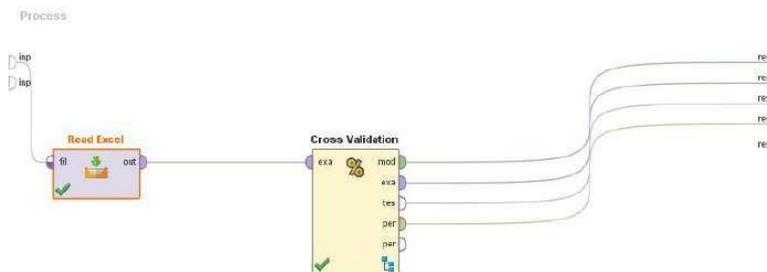
	ATRIBUT	KASUS	TARGET PROD	
			YA	TIDAK
TOTAL		20	7	13
WILAYAH MESIN				
	Area 1	5	1	4
	Area 2	8	5	3
	Area 3	5	1	4
	Area 4	2	0	2
DIES				
	CAVITY NG	2	1	1
	REPAIR DIES	3	0	3
	DIES (SOP)	10	5	5
	PIN PATAH	5	1	4
MESIN				
	MESIN (SOP)	15	7	8
	ROBOT TROUBLE	3	0	3
	LADLE TROUBLE	2	0	2
PARAMETER				

	SETTING (SOP)	18	7	11
	SETTING(DIES	2	0	2
	LAMA)			
MATERIAL				
	MATERIAL	16	6	10
	TEMPERATU RE TURUN	4	1	3
MAN				
	OPERATOR PINDAH	2	1	1
	OPERATOR	8	6	2

Nilai gain tertinggi dijadikan node akar pertama, untuk node akar selanjutnya, dilakukan perhitungan lagi dan seterusnya.

**Proses Validasi (Cross Validation)**

Melakukan validasi yaitu melakukan analisis berbagai model dan memilih model dengan kinerja prediksi yang baik. Pada gambar 4.2 merupakan proses validasi, setelah pembacaan berkas data, blok read excel dengan data tipe excel dihubungkan dengan blok cross validation. Proses data training dan testing berada didalam blok proses validation. Dari proses validasi tersebut dilakukan pengolahan pada tools RapidMiner dengan main process.

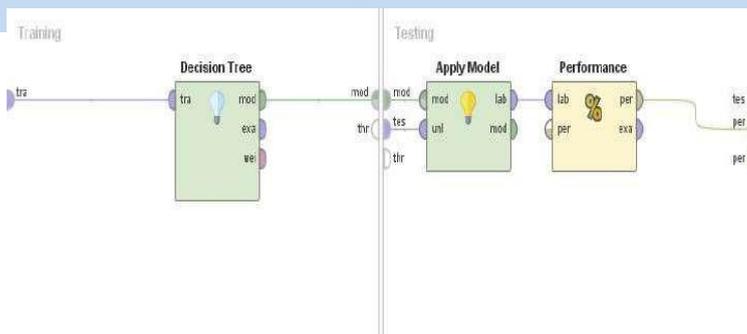


Gambar Proses Validasi

**Proses Training dan Testing**

Pada tahap proses training yaitu melakukan proses pelatihan data pada model C4.5 sedangkan proses testing yaitu melakukan pengujian data yang akan menghasilkan grafik dan pola. Pada gambar dijelaskan bahwa proses training digunakan untuk blok model Decision Tree, dihubungkan dengan garis penghubung pada blok Apply Model dan blok

Performance dibagian proses testing sebagai penampil informasi hasil dari pengujian data, hasil dari pengujian ini akan menghasilkan algoritma C4.5.



Gambar 4.3 Proses Training dan Testing

### Kurva ROC/AUC

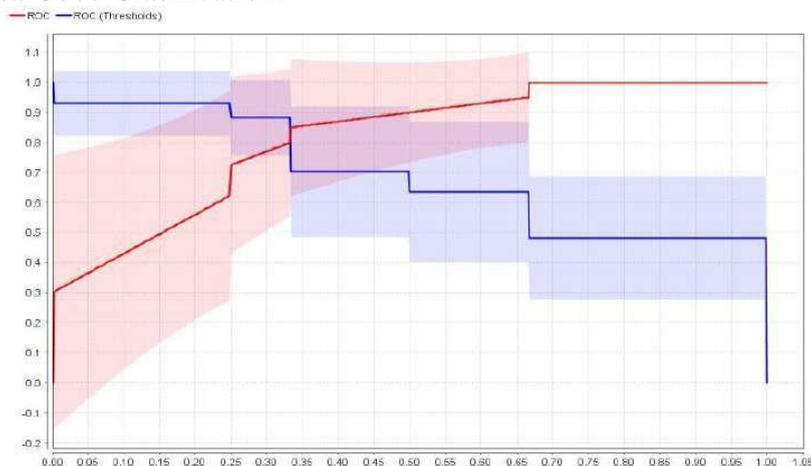
Kurva ROC/AUC adalah ilustrasi grafis dari kemampuan diskriminan dan biasanya diterapkan untuk masalah klasifikasi biner (Saumi Ramadhani, 2016). AUC mengukur kinerja diskriminatif dengan memperkirakan probabilitas output dari sampel yang dipilih secara acak dari populasi positif atau negatif, semakin besar AUC, maka semakin kuat klasifikasi yang digunakan. AUC memiliki tingkat nilai diagnosa yaitu :

- Akurasi bernilai 0,90 – 1,00 = Excellent Classification
- Akurasi bernilai 0,80 – 0,90 = Good Classification
- Akurasi bernilai 0,70 – 0,80 = Fair Classification
- Akurasi bernilai 0,60 – 0,70 = Poor Classification
- Akurasi bernilai 0,50 – 0,60 = Failure

Pada pengukuran kurva ROC/AUC menggunakan tool RapidMiner ada 3 data uji yaitu 50 data, 70 data dan 100 data.

### Kurva ROC/AUC 50 data

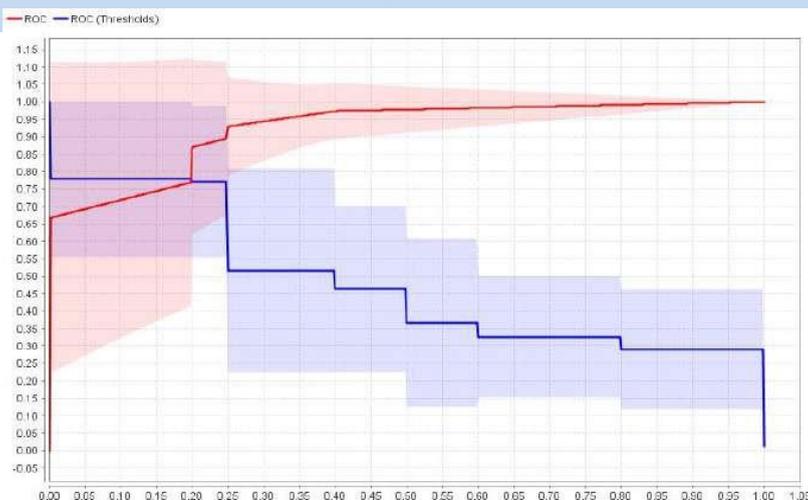
Hasil yang didapat dari pengolahan ROC/AUC 50 data untuk algoritma C4.5 dengan menggunakan tool RapidMiner sebesar 0,812 dapat dilihat pada gambar berikut ini dengan tingkat diagnosa Good Clasification.



Gambar Kurva AUC 50 data

### Kurva ROC/AUC 70 data

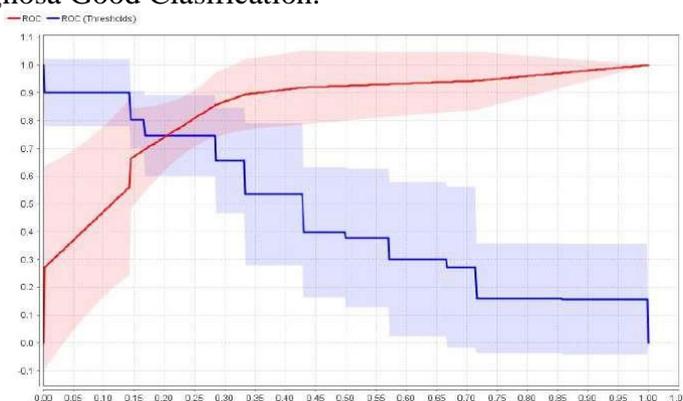
Hasil yang didapat dari pengolahan ROC/AUC 70 data untuk algoritma C4.5 dengan menggunakan tool RapidMiner sebesar 0,922 dapat dilihat pada gambar berikut ini dengan tingkat diagnose Excellent Clasification.



Gambar Kurva AUC 70 data

### Kurva ROC/AUC 100 data

Hasil yang didapat dari pengolahan ROC/AUC 100 data untuk algoritma C4.5 dengan menggunakan tool RapidMiner sebesar 0,839 dapat dilihat pada gambar berikut ini dengan tingkat diagnosa Good Clasification.



Gambar Kurva AUC 100 data

## 5. DAFTAR PUSTAKA

- Afianti, Sonya K, Diah Safitri Sudarno. 2014. Klasifikasi Wilayah De Perdesaan da Desa Perkotaan Wilayah Kabupaten Semarang Dengan Support Vector Machine. *Jurnal Gaussian Vol. 3 No. 4*, 751-760.
- Greenbarg. 2016. *Meta-Learning in Decision Tree Induction*. Switzerland: Springer International Publishing.
- Lewis, R.C., Booms, B.H. (1983). The marketing aspects of service quality. *Proceeding Marketing: American Marketing Association*, Chicago.
- Marsudi, D. 2006. *Operasi Sistem Tenaga Listrik*. Edisi Ke 2. Yogyakarta: Graha Ilmu
- Nisa, I. M. K., & Nooraeni, R. (2020). Penerapan Metode Random Forest Untuk Klasifikasi Wanita Usia Subur Di Perdesaan Dalam Menggunakan Internet (SDKI 2017). *Jurnal MSA (Matematika dan Statistika sertaA plikasinya)*, 8(1), 72-76.
- Shanefield, D. J. 2001. *Industrial Electronics for Engineers, Chemists, and Technicians*. New York: William Andre Publishing.
- Sari, M. S. (2014). Klasifikasi Wilayah Desa Perdesaan dan Desa Perkotaan Wilayah Kabupaten Semarang Dengan Support Vector Machine. *Jurnal Gaussian Vol. 3 No. 4*, 751-760.
- Sumarwati, N. K. (2018). Mengklasifikasikan Desa Perkotaan dan Desa Perdesaan Di Kabupaten Klungkung Menggunakan Metode Mamdani. *E-Journal Matematika Vol. 7(3)*, 203-2010.



- Yhudi. P. Tanjung, S. Sentinuwo, and A. Jacobus. 2016. Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree. [ejournal.unsrat.ac.id](http://ejournal.unsrat.ac.id).
- Badan Pusat Statistik., 2015, *Indonesia - Survei Sosial Ekonomi Nasional 2014 (Gabungan)*, Jakarta.
- Kementerian PUPR. (2014). *Bahan Informasi Rencana Program dan Kegiatan Di Pulau Jawa, Bali dan Nusa Tenggara 2015*. Lombok: Kementerian PUPR.