



INNOVATIVE: Journal Of Social Science Research

Volume 5 Nomor 4 Tahun 2025 Page 8579-8592

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

Klasifikasi Penggunaan Listrik Rumah Tangga Menggunakan Metode Algoritma C4.5, Random Forest, Dan SVM pada PT. PLN ULP Benjeng

Bagas Diki Saputra^{1✉}, Alfi Zuhriya Khoirunnisaa²

Universitas Muhammadiyah Gresik

Email: bagasdiki46@gmail.com^{1✉}

Abstrak

Penggunaan listrik rumah tangga terus meningkat seiring dengan bertambahnya kebutuhan energi akibat perkembangan teknologi. Untuk mendukung efisiensi energi dan pengelolaan daya yang lebih baik, dibutuhkan sistem klasifikasi yang mampu memprediksi tingkat konsumsi listrik. Penelitian ini bertujuan untuk mengklasifikasikan penggunaan listrik rumah tangga menggunakan metode algoritma C4.5, SVM, dan random forest berbasis decision tree yang diimplementasikan melalui perangkat lunak Python. Data yang digunakan diperoleh dari PT. PLN (Persero) ULP Benjeng dan mencakup atribut seperti daya, tarif, pemakaian per kWh, serta komponen biaya lainnya. Proses meliputi tahap prapemrosesan data, pembentukan model klasifikasi, dan evaluasi akurasi dengan confusion matrix. Hasil penelitian diharapkan dapat membantu masyarakat dan penyedia layanan listrik dalam mengelola konsumsi energi secara efisien dan tepat guna.

Kata Kunci : *C4.5, Data Mining, Klasifikasi, Listrik Rumah Tangga, Pohon Keputusan*

Abstract

Household electricity consumption continues to rise in line with increasing energy demands driven by technological development. To support energy efficiency and better power management, a classification system capable of predicting electricity usage levels is needed. This study aims to classify household electricity consumption using the C4.5, SVM, and random forest decision tree algorithm implemented through Python software. The dataset used was obtained from PT. PLN (Persero) ULP Benjeng and includes attributes such as power capacity, tariff, usage per kWh, and other cost components. The process involves data preprocessing, classification model development, and accuracy evaluation using a confusion matrix. The results are expected to assist both consumers and electricity providers in managing energy consumption more efficiently and effectively.

Keywords: *C4.5, Data Mining, Classification, Household Electricity, Decision Tree*

PENDAHULUAN

Kemajuan teknologi di segala bidang meningkat dengan cepat, kemajuan ini akan menimbulkan peningkatan kebutuhan daya listrik. Listrik adalah salah satu bentuk energi yang paling dibutuhkan, karena mudah disalurkan dan dapat diubah menjadi energi lain. Listrik sudah menjadi kebutuhan pokok bagi manusia modern, sebuah elemen yang tak terpisahkan dari rutinitas harian kita. Peranan listrik bukan lagi sekedar kebutuhan sekunder, tetapi telah bertransformasi menjadi kebutuhan primer. Hal ini disebabkan karena tanpa listrik, semua aktivitas yang ingin kita lakukan akan terhambat.

Energi listrik memiliki peran yang sangat vital dalam menunjang berbagai aktivitas manusia, baik secara individu, kelompok, dalam kehidupan bermasyarakat, maupun di sektor industri. Listrik memberikan kontribusi besar terhadap pengoperasian berbagai peralatan yang membutuhkannya. PT Perusahaan Listrik Negara (Persero), atau yang dikenal dengan PLN, adalah perusahaan yang bertugas menyediakan pasokan listrik kepada masyarakat. Berdasarkan klasifikasi pelanggan yang diterapkan oleh PLN, terdapat lima kategori utama pengguna listrik, yaitu sektor sosial, rumah tangga, usaha atau bisnis, industri, dan sektor publik. Dari seluruh kategori tersebut, sektor rumah tangga tercatat sebagai pengguna listrik terbesar.

Dengan penggunaan listrik tersebut menyebabkan kepadatan penduduk, maka akan membutuhkan energi listrik. Dan jika energi listrik dikonsumsi berlebih akan menyebabkan menipisnya persediaan listrik. Oleh karena itu, perlu adanya prediksi diperlukan untuk mengetahui penggunaan listrik rumah tangga, salah satunya menggunakan data mining. Data mining adalah proses untuk mendapatkan data dari pengguna listrik dan diolah menggunakan algoritma kecerdasan buatan untuk memprediksi

penggunaan listrik rumah tangga. Telah banyak penelitian-penelitian yang mengangkat permasalahan ini diantara penelitian yang dikerjakan oleh Yulia dan Azwanti berjudul "Data Mining Prediksi Besarnya Penggunaan Listrik Rumah Tangga di Kota Batam dengan Menggunakan Algoritma C4.5" dapat menyimpulkan bahwa metode C4.5 dapat diterapkan sebagai cara untuk mengklasifikasikan dalam meramalkan seberapa besar konsumsi listrik di kota Batam dengan mempertimbangkan nilai penguatan maksimum.[9] Dengan hasil perhitungan menggunakan metode pohon keputusan mampu memberikan panduan aturan prediksi yang merefleksikan proses yang berkaitan. Dengan empat faktor seperti total barang elektronik, tingkat konsumsi, durasi tinggal dirumah, dan ukuran bangunan.

Selanjutnya penelitian yang dikerjakan oleh Alfa Saleh berjudul "Implementasi Metode Naïve Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga" Dengan metode naïve bayes menggunakan data pelatihan untuk memperoleh nilai kemungkinan dari setiap atribut, untuk kelas yang berbeda mengklasifikasikan 47 data dari 60 data yang diuji. Sehingga metode naïve bayes berhasil memprediksi besarnya penggunaan listrik rumah tangga dengan persentasi keakuratannya sebesar 78,33%. Penelitian selanjutnya dilakukan oleh M. Alfian Riyadi dan rekan-rekannya dengan judul "Data Mining Peramalan Konsumsi Listrik dengan Pendekatan Cluster Time Series sebagai Preprocessing." Dalam penelitian ini, dilakukan analisis terhadap tujuh klaster menggunakan metode ARIMA dan Artificial Neural Network (ANN). Hasilnya menunjukkan bahwa klaster keempat memiliki jumlah anggota terbanyak, yaitu 120 klien, sedangkan klaster pertama memiliki jumlah anggota paling sedikit, yaitu hanya 4 klien. Klaster ketujuh tercatat memiliki rata-rata konsumsi daya tertinggi, yakni sebesar 29.875,13 kWh, sementara konsumsi daya terendah terdapat pada klaster pertama dengan rata-rata sebesar 2.131,83 kWh. Berdasarkan pemodelan yang dilakukan, diketahui bahwa dari 348 klien yang dianalisis, sebanyak 259 klien menunjukkan bahwa metode ANN memberikan hasil yang lebih baik dibandingkan dengan metode ARIMA[5]. Penelitian berikutnya dilakukan oleh Yudhi Pratama Tanjung dan tim, dengan judul "Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree." Dalam penelitian ini, digunakan algoritma pohon keputusan, yaitu ID3 dan C4.5, untuk mengidentifikasi konsumsi listrik pada rumah tangga. Berdasarkan hasil analisis akurasi menggunakan confusion matrix terhadap total 300 data yang terdiri dari 150 data pelatihan dan 150 data pengujian diketahui bahwa algoritma C4.5 memiliki tingkat akurasi sebesar 88%, yang lebih tinggi dibandingkan dengan algoritma ID3 yang mencatat akurasi sebesar 62%.

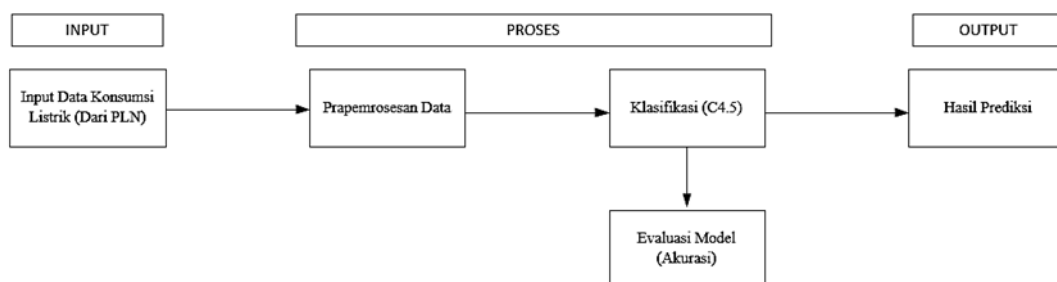
Pada penelitian ini, penulis akan melakukan penelitian yang sama dengan peneliti sebelumnya. Dengan menggunakan judul "Klasifikasi Penggunaan Listrik Rumah Tangga

Menggunakan Metode Algoritma C4.5, Random Forest, Dan SVM Pada PT. PLN ULP Benjeng". Penulis menggunakan metode itu dikarenakan untuk menentukan tingkat akurasi mana yang telah tinggi dan mendeploy dengan menggunakan streamlit. dengan dataset yang diambil dari dataset PT.PLN (PERSERO) ULP BENJENG.

METODE PENELITIAN

Perancangan Sistem

Tahap perancangan sistem dalam penelitian ini secara khusus diarahkan untuk membangun suatu sistem yang dapat memprediksi konsumsi listrik rumah tangga dengan menggunakan algoritma klasifikasi C4.5. sistem ini dirancang berbasis perangkat lunak (software-based), yang diimplementasikan menggunakan bahasa pemrograman python karena fleksibilitas dan kemampuannya dalam pengolahan data serta penerapan algoritma pembelajaran mesin (machine learning) [11]. Untuk memberikan gambaran umum mengenai alur kerja sistem secara keseluruhan, digunakan konsep blok sistem. Konsep ini menjelaskan tahapan sistem. Mulai dari pengumpulan data, prapemrosesan, proses klasifikasi menggunakan algoritma C4.5, hingga evaluasi hasil prediksi yang diperoleh [12]. Sistem ini bertujuan untuk mengidentifikasi dan memprediksi tingkat konsumsi listrik rumah tangga berdasarkan data historis, sehingga dapat memberikan informasi yang bermanfaat bagi pengguna maupun penyedia layanan listrik seperti PLN.



Gambar 1. Blog Diagram

Contoh Perhitungan:

$$\text{Entropy}(S) = \sum_{i=1}^n P_i \log_2 P_i$$

$$\text{Entropy (Total)} =$$

$$\left(-\frac{5}{15} * \log_2 \left(\frac{5}{15}\right)\right) + \left(-\frac{5}{15} * \log_2 \left(\frac{5}{15}\right)\right) + \left(-\frac{5}{15} * \log_2 \left(\frac{5}{15}\right)\right) = 1,584962501$$

$$\text{Entropy (RPBP)} > 287303,33 = \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2}\right)\right) + \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2}\right)\right) + \left(-\frac{2}{2} * \log_2 \left(\frac{2}{2}\right)\right) = 0$$

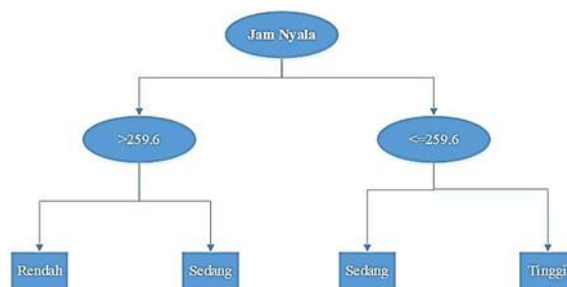
$$\text{Entropy (RPBP)} \leq 287303,33 = \left(-\frac{5}{13} * \log_2 \left(\frac{5}{13}\right)\right) + \left(-\frac{5}{13} * \log_2 \left(\frac{5}{13}\right)\right) + \left(-\frac{3}{13} * \log_2 \left(\frac{3}{13}\right)\right) = 1,5485806$$

$$\begin{aligned}
Gain (Total, RPBP) &= Entropy (Total) - \sum_{i=1}^n \frac{|RPBP|}{|Total|} * Entropy (RPBP) \\
&= 1,584962501 - \left(\left(\frac{2}{15} * 0 \right) + \left(\frac{13}{15} * 1,5485806 \right) \right) \\
&= 0,2428
\end{aligned}$$

Tabel 1. Contoh Perhitungan Algoritma C4.5.

Fitur		Jumlah	Rendah	Sedang	Tinggi	Entropy	Info Gain
Total		15	5	5	5	1.584962501	
RPBP	>287303.33	2	0	0	2	0	0.2428
	<=287303.33	13	5	5	3	1.5485806	
RPUJL	>111926.66	7	0	4	3	0	0.8923
	<=111926.66	8	5	1	2	1.2987949	
PemKwh	>224.87	7	1	2	4	1.3787835	0.1918
	<=224.87	8	4	3	1	1.4056391	
Jam Nyala	>259.6	7	5	2	0	0	1.5849
	<=259.6	8	0	3	5	0	

Dari hasil tersebut nilai gain tertinggi adalah Jam Nyala yaitu 1.5849, sehingga dapat menjadi node akar, dan jika nilai entropy 0 maka node akar akan selesai. Sehingga berhenti pada bagian jam nyala >259,6 dengan hasil 5 rendah dan 2 sedang, sedangkan <=259,6 dengan hasil 3 sedang dan 5 tinggi.



Gambar 2. Pohon Keputusan Algoritma C4.5

Sistem prediksi konsumsi listrik rumah tangga menggunakan algoritma C4.5 dirancang untuk menganalisis pola konsumsi listrik berdasarkan data historis dari PT. PLN (PERSERO) ULP BENJENG. Tahapan pertama dalam sistem ini adalah pengambilan data konsumsi listrik yang mencakup variabel seperti daya listrik yang digunakan, Tarif, RPPTL, RPBPJU, Pemakaian per Kwh, Dan lain-lain. Data ini kemudian melalui tahap prapemrosesan untuk memastikan kualitas dan relevansinya. Prapemrosesan meliputi pembersihan data (data cleaning) untuk membersihkan nilai yang tidak terpakai atau tidak benar, kemudian menormalisasi supaya data ada di rentang yang semestinya, serta seleksi fitur untuk memilih

variabel yang paling berpengaruh terhadap konsumsi listrik.

Setelah data siap, proses selanjutnya adalah pengelompokan data dilakukan memakai algoritma C4.5, yang beroperasi melalui pembentukan struktur pohon keputusan berdasarkan pola yang ditemukan dalam data historis. Algoritma ini menentukan aturan klasifikasi dengan mempertimbangkan faktor-faktor seperti jumlah pemakaian dan lama penggunaan listrik untuk memprediksi pola konsumsi listrik rumah tangga. Keunggulan utama algoritma C4.5 adalah kemampuannya dalam menangani data kategorikal maupun numerik serta memberikan hasil yang dapat diinterpretasikan dengan mudah melalui struktur pohon keputusan. Agar model prediksi dapat diandalkan, dilakukan evaluasi model dengan mengukur akurasi prediksi menggunakan metrik seperti confusion matrix, akurasi dalam persentase, serta precision, recall dan f1 score.

		Nilai Aktual	
		Positive	Negative
Nilai Prediksi	Positive	TP	FP
	Negative	FN	TN

Gambar 3. Rumus Nilai Prediksi

Confusion Matrix :

$$\text{Ak} \quad \text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Preci:} \quad \text{Precision} = \frac{TP}{TP + FP}$$

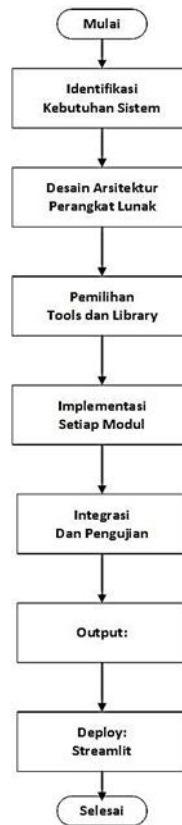
$$\text{Re} \quad \text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 !} \quad \text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Jika akurasi masih rendah, model dapat diperbaiki dengan menyesuaikan parameter pohon keputusan atau menggunakan metode optimasi lainnya. Setelah evaluasi selesai, sistem menghasilkan output berupa prediksi konsumsi listrik rumah tangga yang dapat disajikan dalam bentuk laporan tabel, grafik visualisasi, atau rekomendasi efisiensi energi. Output ini membantu pengguna dan PLN dalam mengatur penggunaan daya dengan lebih pintar serta membantu menyusun perkiraan keperluan energi di masa depan. Dengan pendekatan ini, sistem prediksi tidak hanya meningkatkan kesadaran pengguna terhadap penggunaan listrik, tetapi juga berkontribusi pada efisiensi energi secara lebih luas.

Perancangan Software

Perancangan perangkat lunak dilakukan untuk membangun sistem prediksi konsumsi listrik rumah tangga yang terstruktur, user-friendly, dan mampu menyajikan informasi yang akurat. Tahapan perancangan ini meliputi identifikasi kebutuhan sistem, perancangan arsitektur, pemilihan tools pengembangan, implementasi modul, integrasi, serta pengujian sistem. Alur kerja perangkat lunak secara umum digambarkan melalui flowchart pada Gambar 4.



Gambar 4. flowchart perancangan software

Keterangan Flowchart:

1. Mulai, merupakan titik awal dari proses perancangan dan pengembangan sistem prediksi konsumsi listrik rumah tangga.
2. Identifikasi Kebutuhan Sistem Pada tahap ini dilakukan analisis kebutuhan sistem, yang dibagi menjadi:
 - a) Kebutuhan Fungsional: Meliputi fitur input data konsumsi listrik, pemrosesan dengan algoritma C4.5, dan output prediksi.
 - b) Kebutuhan Non-Fungsional: Meliputi aspek dari kinerja sistem seperti kecepatan merespon, efisiensi pemrosesan data, skalabilitas sistem, dan tampilan antarmuka yang mudah digunakan.
3. Perancangan Arsitektur Perangkat Lunak Sistem dirancang secara modular agar mudah dikembangkan dan diuji. Modul utama meliputi:
 - a) Modul Input: Untuk memasukkan data daya listrik, jumlah perangkat, dan waktu

- penggunaan.
- b) Modul Proses: Untuk prapemrosesan data, penerapan algoritma, dan evaluasi.
 - c) Modul Output: Untuk menyajikan hasil prediksi dalam bentuk grafik dan laporan.
 - d) Antarmuka Pengguna (User Interface): Dirancang agar interaksi pengguna berjalan secara intuitif dan efisien.
4. Pemilihan Tools dan Library Beberapa tools dan library yang digunakan dalam pengembangan sistem antara lain:
- a) Python: Bahasa pemrograman utama.
 - b) Pandas: Untuk manipulasi dan analisis data.
 - c) Scikit-learn (sklearn): Untuk penerapan algoritma C4.5 dan evaluasi model.
 - d) Matplotlib: Untuk visualisasi grafik dan hasil analisis.
5. Implementasi Setiap Modul Tahap ini meliputi proses pengkodean berdasarkan desain sistem:
- a) Input Data Konsumsi: daya listrik yang digunakan, Tarif, RPPTL, RPBPJU, Pemakaian per Kwh, Dan lain-lain.
 - b) Prapemrosesan Data: Data dibersihkan, dinormalisasi, dan dilakukan seleksi fitur.
 - c) Penerapan Algoritma C4.5: Dimanfaatkan dalam merancang suatu model klasifikasi terkait pemakaian daya listrik.
 - d) Evaluasi dan Visualisasi: Performa model diukur dan ditampilkan dalam bentuk grafik serta laporan.
6. Integrasi dan Pengujian Setelah seluruh modul selesai diimplementasikan:
- a) Pengujian Modul: Menjamin setiap bagian sistem bekerja sesuai fungsinya.
 - b) Pengujian End-to-End: Memastikan seluruh sistem dari input hingga output berjalan secara konsisten dan akurat.
7. Streamlit adalah library Python yang digunakan untuk membuat aplikasi web interaktif secara cepat dan mudah, terutama untuk keperluan data science, machine learning, atau visualisasi data.
8. Output: Aplikasi Prediksi Siap Digunakan Sistem menghasilkan aplikasi yang siap digunakan oleh pengguna, dengan fitur:
- a) Prediksi konsumsi listrik berdasarkan data historis.
 - b) Tampilan hasil dalam bentuk grafik, tabel, dan rekomendasi efisiensi penggunaan listrik.
9. Selesai: Menandai akhir dari proses pengembangan perangkat lunak. Sistem telah siap untuk dioperasikan oleh pengguna akhir.

Perancangan Metode Klasifikasi dengan Algoritma C4.5

Perencanaan metode dilakukan untuk menyusun tahapan sistematis dalam membangun sistem prediksi konsumsi listrik rumah tangga menggunakan algoritma C4.5. Metode ini mencakup proses mulai dari input data, prapemrosesan, pembentukan model, klasifikasi, hingga evaluasi dan visualisasi hasil prediksi. Alur lengkap proses kerja sistem di sebutkan berikut ini. Keterangan:

1. Mulai, merupakan titik awal proses sistem prediksi konsumsi listrik.
2. Input Data Konsumsi, pengguna memasukkan data yang mencakup:
 - a) Idpel : Nomor Identitas Pelanggan.
 - b) Nama : Nama Pemilik listrik.
 - c) Tarif : Tarif sesuai kebutuhan.
 - d) Daya : Daya listrik yang digunakan pelanggan.
 - e) RPBP : Rupiah pemakaian listrik.
 - f) RPUJL : Pajak jalan umum.
 - g) PEMKWH : Pemakaian per Kwh.
 - h) Jam Nyala : Lama jam nyala pemakaian.
3. Prapemrosesan Data, Data yang dimasukkan dipersiapkan agar layak diproses, melalui beberapa tahap:
 - a) Data Cleaning: Menghapus data duplikat atau tidak valid.
 - b) Normalisasi: Menyelaraskan skala nilai agar sebanding.
 - c) Seleksi Fitur: Memilih fitur yang paling relevan untuk prediksi.
4. Pembentukan Model (Algoritma C4.5), Sistem ini memanfaatkan algoritma C4.5 untuk membangun model klasifikasi berbentuk pohon keputusan berdasarkan data pelatihan yang ada.
5. Proses Klasifikasi, Model yang telah dibentuk digunakan untuk mengklasifikasikan data baru dan membentuk struktur keputusan.
6. Evaluasi Akurasi, Sistem mengevaluasi apakah akurasi hasil prediksi sudah memenuhi target yang diharapkan. Jika:
 - a) Tidak, maka dilakukan proses optimasi model, misalnya dengan tuning parameter atau validasi silang (cross-validation).
 - b) Ya, maka sistem melanjutkan ke tahap berikutnya.
7. Output Hasil, Prediksi Hasil akhir dari sistem ditampilkan dalam beberapa bentuk:
 - a) Tabel: Menampilkan hasil prediksi dalam format data tabular.
 - b) Grafik: Menyajikan visualisasi konsumsi listrik.
 - c) Rekomendasi Efisiensi: Memberikan saran penghematan berdasarkan hasil prediksi.
8. Selesai, Proses metode prediksi selesai, dan aplikasi siap digunakan oleh pengguna.

Data Sheet dan Parameter Penelitian

Dalam penelitian ini, dataset yang digunakan merupakan data konsumsi listrik rumah tangga yang diperoleh dari PT. PLN (PERSERO) ULP BENJENG sumber simulasi berdasarkan pola penggunaan umum. Dataset ini terdiri dari beberapa atribut diantaranya:

Tabel 2. Atribut Input Data Konsumsi

Atribut	Keterangan
Idpel	Nomor Identitas Pelanggan
Nama	Nama Pemilik Listrik
Tarif	Tarif sesuai Kebutuhan
Daya	Daya Listrik yng digunakan Pelanggan
RPBP	Rupiah Biaya Penyambungan
RPUJL	Uang Jaminan Langgan
PEMKWH	Pemakaian per Kwh

Dataset terdiri dari total 150 data, yang dibagi menjadi dua bagian:

1. Data latih (training data): 70 data (70%) digunakan untuk membangun model klasifikasi.
2. Data uji (testing data): 30 data (30%) digunakan untuk menguji akurasi dan kinerja model.

Seluruh data telah melalui proses prapemrosesan yang meliputi normalisasi, seleksi fitur, dan penghapusan data outlier untuk meningkatkan kualitas input ke dalam model. Dataset akan dikategorikan berdasarkan 3 kategori sesuai dengan pemakaian daya:

- a. Daya 450VA dikategorikan sebagai daya Rendah.
- b. Daya 900VA dikategorikan sebagai daya Sedang.
- c. Daya 1300-2200VA dikategorikan sebagai daya Tinggi.



	A	B	C	D	E	F	G
3	202403	511220499389	NCAADAR02600	51	51GSK	51122	PJU *****
4	202403	511220499633	NCAADAR01300	51	51GSK	51122	PJU *****
5	202403	511220499292	NCAABGF38700	51	51GSK	51122	PJU *****
6	202403	511220499666	NCAADAR02400	51	51GSK	51122	PJU *****
7	202403	511220496191	NCAADAR03800	51	51GSK	51122	PJU *****
8	202403	511220499172	NCAADAR06400	51	51GSK	51122	PJU *****
9	202403	511220499180	NCAADAR05000	51	51GSK	51122	PJU *****
10	202403	511220499314	NCAADAR02700	51	51GSK	51122	PJU *****
11	202403	511220499475	NCAADAR00900	51	51GSK	51122	PJU *****
12	202403	511220499491	NCAADAR00500	51	51GSK	51122	PJU *****
13	202403	511220499505	NCAADAR04800	51	51GSK	51122	PJU *****
14	202403	511220499521	NCAADAR06600	51	51GSK	51122	PJU *****

Gambar 5. Contoh Dataset PT. PLN (PERSERO) ULP BENJENG

HASIL DAN PEMBAHASAN

Pengujian Program

Pengujian program dilakukan menggunakan tiga algoritma klasifikasi, yaitu C4.5, SVM, dan Random Forest, dengan mengukur performa masing-masing algoritma berdasarkan metrik Precision, Recall, F1-score, dan Akurasi. Evaluasi dilakukan terhadap tiga kelas, yaitu diambil dengan nilai algoritma tertinggi. Data tersebut dapat dilihat berikut ini.

Tabel 3. Algoritma C4.5

Algoritma C4.5	Precision	Recall	F1-score	Support
Rendah	0.95	1.00	0.97	19
Sedang	0.93	1.00	0.96	13
Tinggi	1.00	0.85	0.92	13
Akurasi	0.95			

Tabel 4. SVM

SVM	Precision	Recall	F1-score	Support
Rendah	1.00	1.00	1.00	19
Sedang	0.87	1.00	0.93	13
Tinggi	1.00	0.85	0.92	13
Akurasi	0.95			

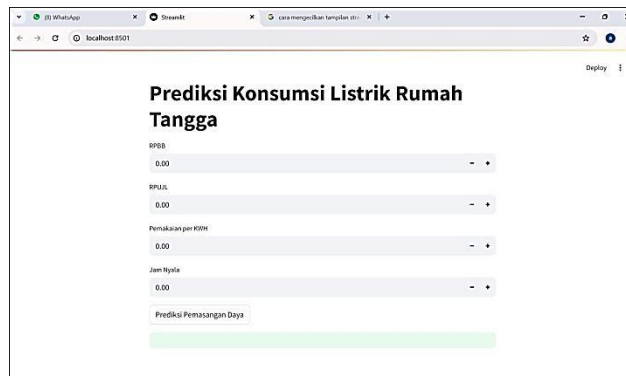
Tabel 5. Random Forest

Random Forest	Precision	Recall	F1-score	Support
Rendah	0.95	1.00	0.97	19
Sedang	1.00	1.00	1.00	13
Tinggi	1.00	0.92	0.96	13
Akurasi	0.97			

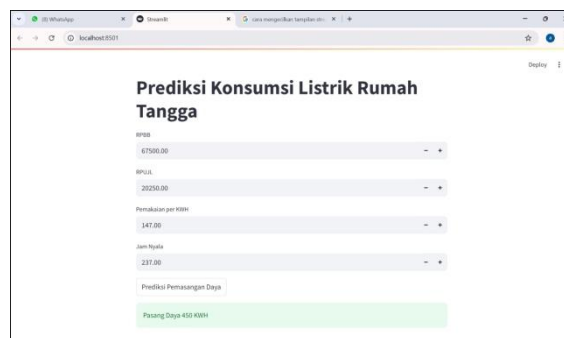
Berdasarkan hasil pengujian, algoritma dengan performa terbaik adalah Random Forest, karena menghasilkan akurasi tertinggi (0.97) dan nilai F1-score yang sangat baik di ketiga kelas. Oleh karena itu, algoritma Random Forest dipilih sebagai algoritma terbaik untuk implementasi program ini.

Hasil Pengujian Aplikasi

Aplikasi yang dikembangkan menggunakan antarmuka Streamlit memungkinkan pengguna untuk memasukkan beberapa parameter, dan berikut ini ialah tampilan dari hasil pengujian tersebut.



Gambar 6. Tampilan Data 1



Gambar 7. Tampilan Data 2

Setelah data dimasukkan, pengguna dapat mengklik tombol "Prediksi Pemasangan Daya", dan sistem akan memberikan rekomendasi pemasangan daya berdasarkan model klasifikasi yang digunakan (dalam hal ini, menggunakan algoritma Random Forest, yang memiliki akurasi tertinggi sebesar 97% dibandingkan algoritma lain seperti C4.5 dan SVM)

SIMPULAN

Berdasarkan hasil pengujian terhadap sistem prediksi konsumsi listrik rumah tangga yang telah dikembangkan, diperoleh beberapa poin kesimpulan. Tiga metode klasifikasi, yakni C4.5, SVM, dan Random Forest, telah diuji untuk mengidentifikasi algoritma dengan performa terbaik dalam memprediksi kebutuhan pemasangan daya listrik pada rumah tangga. Evaluasi dilakukan dengan menggunakan metrik Precision, Recall, F1-score, dan Akurasi terhadap tiga kategori konsumsi listrik, yaitu Rendah, Sedang, dan Tinggi. Dari hasil pengujian, algoritma C4.5 dan SVM masing-masing mencatat akurasi sebesar 0.95, dengan nilai F1-score yang cukup baik namun belum konsisten pada seluruh kelas. Di sisi lain, algoritma Random Forest menunjukkan hasil paling unggul dengan akurasi sebesar 0.97, serta F1-score yang tinggi dan stabil, yakni 0.97 untuk kelas Rendah, 1.00 untuk kelas Sedang, dan 0.96 untuk kelas Tinggi.

Oleh karena itu, algoritma Random Forest dipilih sebagai model terbaik dalam sistem

ini karena kemampuannya memberikan prediksi yang paling akurat dan andal. Aplikasi prediksi konsumsi listrik ini dibangun menggunakan antarmuka Streamlit. Melalui aplikasi ini, pengguna dapat memasukkan sejumlah data seperti RPBB (Rencana Penggunaan Beban Bulanan), RPUJL (Rencana Penggunaan Jumlah Listrik), pemakaian per KWH, dan jam nyala. Setelah input dimasukkan, sistem akan memproses data dan menampilkan rekomendasi pemasangan daya listrik yang sesuai. Sebagai ilustrasi, pada data uji tertentu, sistem memberikan hasil prediksi berupa rekomendasi "Pasang Daya 450 KWH", Sistem ini berpotensi membantu masyarakat maupun instansi terkait, seperti PLN atau lembaga energi lainnya, dalam menentukan kebutuhan pemasangan daya listrik secara lebih tepat, efisien, dan berdasarkan pola konsumsi listrik yang ada.

DAFTAR PUSTAKA

- A. A. Arifiyanti, R. M. Pradana, and I. F. Novian, "Klasifikasi Produk Retur dengan Algoritma Pohon Keputusan C4.5," *J. IPTEK*, vol. 22, no. 1, p. 79, 2018, doi: 10.31284/j.ipitek.2018.v22i1.243.
- A. Almira and A. Ikhwan, "Implementasi Data Mining Menggunakan Algoritma Fp-Growth pada Analisis Pola Pencurian Daya Listrik," *J. Inform. Univ. Pamulang*, vol. 6, no. 2, pp. 442–448, 2021, [Online]. Available: <http://openjournal.unpam.ac.id/index.php/informatika442>.
- A. Bahtiar and D. Firmansyah, "Prediksi Cacat Software Menggunakan Algoritma C4.5," vol. XX, no. Xx.
- D. A. C, "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes," vol. 2, no. 1, pp. 39–46.
- G. P. Kawani, "Implementasi Naive Bayes," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 73–81, 2019, doi: 10.20895/inista.v1i2.73.
- I. Junaedi, N. Nuswantari, and V. Yasin, "Perancangan dan Implementasi Algoritma C4.5 untuk Data Mining," *J. Inf. Syst. Informatics Comput.*, vol. 3, no. 1, pp. 29–44, 2019, [Online]. Available: <http://journal.stmikjayakarta.ac.id/index.php/jisicom/article/view/203>.
- M. A. A. Riyadi, K. Fithriasari, and Dwiatmono, "Data Mining Peramalan Konsumsi Listrik dengan Pendekatan Cluster Time Series sebagai Preprocessing," *J. Sains dan Seni ITS*, vol. 5, no. 1, pp. 2337–3520, 2016, [Online]. Available: <http://archive.ics.uci.edu/ml/datasets/ElectricityLoad>.
- M. Safii, "Implementasi Data Mining Dengan Metode Pohon Keputusan Algoritma ID3 Untuk Menentukan Status Mahasiswa," *J. Mantik Penusa*, vol. 2, no. 1, pp. 82–86, 2018,

[Online].

Available:

[http://download.garuda.kemdikbud.go.id/article.php?article=814312&val=11774&title=Implementasi Data Mining Dengan Metode Pohon Keputusan Algoritma ID3 Untuk Menentukan Status Mahasiswa](http://download.garuda.kemdikbud.go.id/article.php?article=814312&val=11774&title=Implementasi%20Data%20Mining%20Dengan%20Metode%20Pohon%20Keputusan%20Algoritma%20ID3%20Untuk%20Menentukan%20Status%20Mahasiswa).

- N. S. Syam et al., "Model Support Vector Machine untuk Prediksi pada Penggunaan Energi Listrik di Rumah Hemat Energi," *J. Inform.*, vol. 1, no. 2, pp. 56–59, 2022, [Online]. Available: <https://jurnal.uniraya.ac.id/index.php/JI>.
- Y. Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4.5," *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- Y. P. Tanjung, S. R. Sentinuwo, and A. Jacobus, "Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree," *J. Tek. Inform.*, vol. 9, no. 1, 2016, doi: 10.35793/jti.9.1.2016.14141.
- Yulia and N. Azwanti, "Data Mining Prediksi Besarnya Penggunaan Listrik Rumah Tangga di Kota Batam dengan Menggunakan Algoritma C4.5," *Semin. Nas. Ilmu Sos. dan Teknol.*, no. 1, pp. 175–180, 2018.