

## Sistem Rekomendasi Lokasi Optimal dan Potensi Penghematan Energi Pemasangan PLTS Atap Berbasis AI di Pulau Jawa

Supanto<sup>1\*,2</sup>, Hadi Saputra<sup>1,2</sup>, Amir Aminuddin<sup>1,2</sup>, Neng Ayu Herawati<sup>1,3</sup>, Ayu Purwarianti<sup>1,3</sup>, Nugraha Priya Utama<sup>1,3</sup>

<sup>1</sup> Sekolah Teknik Elektro dan Informatika (STEI), Institut Teknologi Bandung, Jl. Ganesa No.10, Coblong, Kota Bandung, Jawa Barat 40132

<sup>2</sup> PT PLN (Persero), Jl. Trunojoyo Blok M-I No. 135, Jakarta 12160, Indonesia

<sup>3</sup> Pusat Artificial Intelligence: Computer Vision, NLP, and Big Data Analytics, Institut Teknologi Bandung, Jl. Ganesa No.10, Coblong, Kota Bandung, Jawa Barat 40132

\*Corresponding Author: [23524025@std.stei.itb.ac.id](mailto:23524025@std.stei.itb.ac.id)

---

### Abstrak

Transisi menuju energi terbarukan di Indonesia menuntut pendekatan berbasis data dalam menentukan lokasi optimal pemasangan Pembangkit Listrik Tenaga Surya (PLTS) atap dan dalam memperkirakan dampak ekonomisnya. Penelitian ini mengembangkan sistem rekomendasi berbasis Artificial Intelligence (AI) yang mengintegrasikan data penyinaran matahari dari BMKG dan data konsumsi listrik dari PLN untuk mendukung perencanaan PLTS atap di Pulau Jawa. Pendekatan dilakukan melalui tiga metode pembelajaran mesin utama: klasifikasi untuk mengevaluasi kelayakan pelanggan, klusterisasi wilayah menggunakan algoritma clustering, dan regresi untuk memprediksi potensi penghematan energi. Lima algoritma klasifikasi dibandingkan, dengan LightGBM menunjukkan performa tertinggi (akurasi 87%). Segmentasi wilayah optimal diperoleh melalui KMeans (silhouette score 0,5566). Estimasi penghematan energi paling akurat dihasilkan oleh XGBoost Regressor dengan koefisien determinasi ( $R^2$ ) sebesar 0,9999. Hasil ini menunjukkan bahwa pendekatan integratif berbasis AI dapat menyediakan informasi prediktif yang akurat dan aplikatif bagi penyusunan strategi promosi dan investasi PLTS atap, sekaligus memberikan estimasi manfaat kuantitatif bagi pelanggan. Penelitian ini memberikan kontribusi signifikan dalam pengembangan sistem pendukung keputusan untuk energi terbarukan berbasis analitik spasial dan perilaku konsumsi.

**Keywords:** AI, Clustering, Klasifikasi, PLTS Atap, Regresi

### Abstract

The shift towards renewable energy in Indonesia necessitates a data-centric methodology for identifying ideal sites for Photovoltaic Rooftop (PV Rooftop) and assessing their economic repercussions. This research formulated an AI-driven recommendation system leveraging solar irradiance data from BMKG and electricity usage data from PLN to facilitate PV Rooftop planning on Java Island. The methodology utilizes three primary machine learning techniques: classification for evaluating customer eligibility, region clustering via clustering algorithms, and regression for forecasting potential energy savings. A comparison of five classification algorithms revealed LightGBM as the most effective, achieving 87% accuracy. Optimal regional segmentation was achieved through KMeans, yielding a silhouette score of 0.5566. The XGBoostRegressor produced the most precise energy saving estimates, with a coefficient of determination ( $R^2$ ) of 0.9999. These findings underscore the potential of AI-integrated approaches to deliver precise and practical predictive insights for the promotion and investment strategies of PV Rooftop, alongside providing quantitative benefits for consumers. This study significantly advances the development of decision support systems for renewable energy utilizing spatial analytics and consumption patterns.

**Keywords:** AI, Clustering, Classification, PV Rooftop, Regression

---

## PENDAHULUAN

Pemanasan global dan krisis iklim telah mendorong transisi menuju energi terbarukan di berbagai negara, termasuk Indonesia (Climate Transparency, 2024; Kementerian Energi dan Sumber Daya Mineral, n.d.; McKinsey & Company, 2023). Sebagai negara tropis dengan intensitas radiasi matahari mencapai 4,8 kWh/m<sup>2</sup> per hari, Indonesia memiliki potensi besar dalam pengembangan Pembangkit Listrik Tenaga Surya (PLTS) atap (Asian Development Bank, 2021; Kompas, 2023; World Bank Group, n.d.). Pemerintah Indonesia merespon peluang ini dengan menerbitkan Peraturan Menteri ESDM No. 2 Tahun 2024 untuk mendorong pemanfaatan energi surya melalui instalasi PLTS atap pada jaringan listrik nasional (Kementerian Energi dan Sumber Daya Mineral Republik Indonesia [Kementerian ESDM], 2024).

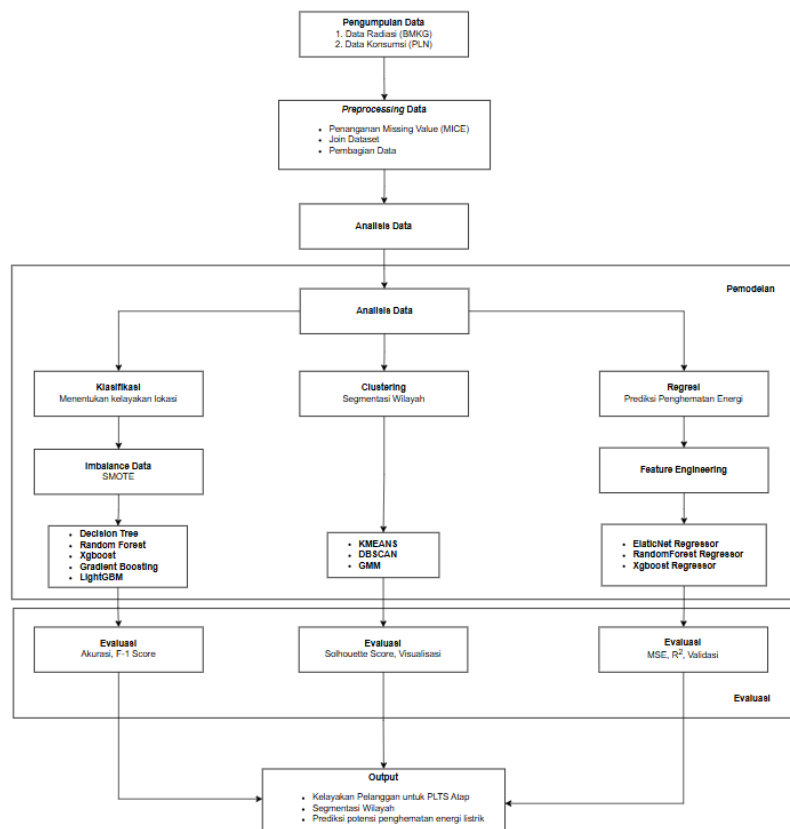
Sebagai penyedia utama listrik nasional, Perusahaan Listrik Negara (PLN) berperan penting untuk mendukung kesuksesan program pemerintah terkait pemanfaatan energi surya ini (Kementerian Energi dan Sumber Daya Mineral Republik Indonesia [Kementerian ESDM], 2021). Namun dalam pelaksanaannya, PLN menghadapi tantangan diantaranya dalam menentukan wilayah optimal untuk pemasangan PLTS atap, memastikan kelayakan pelanggan dan memperkirakan potensi penghematan energi bagi pelanggan. Di sisi lain, hasil survei dari Institute for Essential Services Reform (IESR) menunjukkan minat tinggi masyarakat terhadap PLTS atap, dengan 66% responden di Surabaya menyatakan ketertarikan (Citraningrum & Tumiwa, 2019). Kondisi ini menciptakan kebutuhan mendesak akan pendekatan berbasis data yang komprehensif dan adaptif untuk perencanaan PLTS atap oleh PLN.

Berbagai penelitian sebelumnya dalam analisis potensi PLTS atap umumnya hanya menggunakan satu jenis data secara terpisah, seperti data radiasi matahari atau data konsumsi listrik pelanggan (de Luis-Ruiz et al., 2024; Fakhraian et al., 2021; Rishitha et al., 2024). Selain itu, pendekatan berbasis *Artificial Intelligence* (AI) dan *Machine Learning* (ML) dalam studi-studi tersebut umumnya terbatas pada satu metode tunggal, seperti prediksi produksi atau segmentasi wilayah, tanpa adanya evaluasi komparatif terhadap berbagai metode yang berbeda. Selain itu, pendekatan yang digunakan juga cenderung terfokus pada satu jenis analisis, tanpa menggabungkan beberapa metode secara terpadu (Levent et al., 2025). Penelitian yang secara simultan mengintegrasikan kedua jenis data tersebut yaitu radiasi matahari dan konsumsi listrik dengan pendekatan klasifikasi, clustering, dan regresi untuk menghasilkan rekomendasi lokasi sekaligus estimasi potensi penghematan energi masih sangat terbatas. Oleh karena itu, penelitian ini bertujuan untuk menjawab kesenjangan tersebut dengan mengembangkan sistem rekomendasi berbasis AI yang mampu menentukan kelayakan pelanggan untuk PLTS atap (klasifikasi), mengelompokkan wilayah potensial pemasangan berdasarkan karakteristik teknis (clustering), dan memprediksi potensi penghematan energi listrik (regresi).

Kontribusi utama penelitian ini adalah pada integrasi simultan data radiasi matahari dan konsumsi listrik dalam satu sistem AI, yang mampu memberikan rekomendasi lokasi pemasangan PLTS atap sekaligus estimasi manfaat energi secara kuantitatif. Pendekatan ini diharapkan dapat memberikan dasar teknis yang kuat dalam pengambilan keputusan PLN untuk promosi, investasi, dan ekspansi PLTS atap di wilayah strategis di Pulau Jawa.

## METODOLOGI PENELITIAN

Bab ini menjelaskan secara sistematis tahapan yang dilakukan dalam penelitian untuk membangun sistem rekomendasi lokasi optimal pemasangan PLTS atap dan estimasi penghematan energi berbasis AI. Metodologi yang digunakan dirancang untuk menjawab tiga tujuan utama: (1) mengklasifikasikan kelayakan pelanggan dalam pemasangan PLTS atap, (2) mengelompokkan wilayah potensial berdasarkan kesamaan karakteristik teknis, dan (3) memprediksi besarnya penghematan energi listrik yang diperoleh pelanggan. Setiap tahap dalam proses penelitian dijelaskan secara rinci, mulai dari pengumpulan dan praproses data, pembentukan label klasifikasi, pemodelan menggunakan algoritma machine learning, evaluasi kinerja model, hingga visualisasi dan analisis hasil. Pendekatan ini dirancang untuk menghasilkan sistem berbasis data yang tidak hanya akurat, tetapi juga dapat direplikasi dan digunakan dalam pengambilan keputusan strategis di sektor energi terbarukan. Langkah-langkah penelitian digambarkan dalam bentuk alur diagram pada Gambar 1.



Gambar 1. Metodologi Penelitian

Gambar 1 menjelaskan bahwa metodologi yang digunakan dirancang agar sesuai dengan tujuan penelitian yaitu memberikan gambaran dan analisa terkait pemilihan rekomendasi lokasi yang optimal untuk pemasangan PLTS atap dan rekomendasi penghematan. Untuk memudahkan pemahaman dan alur lengkap dari awal hingga akhir, berikut penjelasannya.

## 1. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari dua sumber utama.

### 1.1. Data Online BMKG

Data *Online* Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) yang bersifat publik dan dapat diakses bebas. Dataset ini berisi informasi tentang lama intensitas penyinaran matahari dalam satuan jam per hari untuk beberapa kota besar di Pulau Jawa seperti DKI Jakarta, Banten, Jawa Barat, Jawa Tengah, Jawa Timur, dan Yogyakarta. Data yang diambil menggunakan sampel data pada periode Juli hingga Desember 2024, dengan total sebanyak 5.519 baris data dan 11 kolom atribut. Data yang digunakan di BMKG merupakan data “lama penyinaran matahari” karena ingin melihat bagaimana paparan sinar matahari yang akan digunakan untuk penelitian lokasi optimal dalam PLTS atap di pulau Jawa. Bentuk struktur data awal yang akan diolah dapat dilihat Gambar 2.

| TANGGAL   | SS          | ID WMO | NAMA STASIUN               | LINTANG  | BUJUR     | ELEVASI  | Kota              | Provinsi | TAHUN | BULAN_NAMA |
|-----------|-------------|--------|----------------------------|----------|-----------|----------|-------------------|----------|-------|------------|
| 7/1/2024  | 9.3         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/2/2024  | 5.1         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/3/2024  | 2.3         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/4/2024  | 5.8         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/5/2024  | 3           | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/6/2024  | 3.8         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/7/2024  | 0           | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/8/2024  | 5.680350083 | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/9/2024  | 7.8         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/10/2024 | 3.9         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |
| 7/11/2024 | 3.3         | 96733  | Stasiun Klimatologi Banten | -6.26151 | 106.75084 | 27 Meter | Tangerang Selatan | BANTEN   | 2024  | July       |

Gambar 2. Sampel Dataset BMKG yang digunakan

### 1.2. Data Konsumsi Listrik

Data konsumsi listrik dalam penelitian ini bersumber dari dataset milik PLN, yang berasal dari catatan penggunaan energi listrik yang dihimpun secara rutin. Untuk keperluan penelitian ini, data yang digunakan telah difilter sehingga hanya mencakup pelanggan dengan daya listrik di atas 1300 VA. Pemilihan ini didasarkan pada hasil wawancara dengan A. Soerahmat, praktisi PLTS atap di PLN [19 Februari 2025], yang menyatakan bahwa mayoritas pengguna PLTS atap saat ini berasal dari kalangan masyarakat menengah ke atas dengan daya listrik lebih dari 1300 VA, sejalan dengan kebutuhan investasi yang cukup besar untuk membangun sistem PLTS atap. Dataset hasil penyaringan ini terdiri dari 1.995.119 baris data dan 12 kolom, mencakup informasi seperti ID pelanggan, lokasi administratif (kelurahan, kecamatan, kabupaten/kota), daya listrik, serta konsumsi energi bulanan. Bentuk struktur data yang diolah untuk data PLN dapat dilihat pada Gambar 3.

| idpel        | nama_kelurahan | nama_kec | nama_kab    | nama_prov  | daya | bth_tagih | pemkwh | tipe_pelanggan | TAHUN | BULAN | NAMA_BULAN |
|--------------|----------------|----------|-------------|------------|------|-----------|--------|----------------|-------|-------|------------|
| 511211861729 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 248.1  | prabayar       | 2024  | 7     | July       |
| 511211730979 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 4.4  | 2024-07   | 0      | prabayar       | 2024  | 7     | July       |
| 511210080909 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 239.3  | prabayar       | 2024  | 7     | July       |
| 511210492947 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 3.5  | 2024-07   | 744.3  | prabayar       | 2024  | 7     | July       |
| 511211700014 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 0      | prabayar       | 2024  | 7     | July       |
| 511210616686 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 0      | prabayar       | 2024  | 7     | July       |
| 511210448796 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 198.8  | prabayar       | 2024  | 7     | July       |
| 511210413859 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 5.5  | 2024-07   | 192.8  | prabayar       | 2024  | 7     | July       |
| 511210546309 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 189    | prabayar       | 2024  | 7     | July       |
| 511210645114 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 2.2  | 2024-07   | 188.9  | prabayar       | 2024  | 7     | July       |
| 511210631618 | ABAR- ABIR     | BUNGAH   | KAB. GRESIK | JAWA TIMUR | 3.5  | 2024-07   | 165.3  | prabayar       | 2024  | 7     | July       |

Gambar 3. Sampel Dataset Konsumsi Listrik (PLN)

## 2. Preprocessing Data

Tahap ini dilakukan untuk menyiapkan data sebelum masuk ke proses analisis dan pemodelan. Data yang digunakan berasal dari dua sumber utama, yaitu data radiasi matahari dari BMKG dan data konsumsi listrik dari PLN. Langkah pertama adalah menangani nilai hilang menggunakan metode *Multiple Imputation by Chained Equations* (MICE) agar struktur data tetap konsisten dan akurat (Mora-Gaona et al., 2021). Setelah itu, dilakukan proses penggabungan data berdasarkan parameter lokasi dan waktu, sehingga masing-masing entri data memuat informasi lengkap dari kedua sumber. Terakhir, data dibagi menjadi data *training* dan data *testing* dengan tujuan agar model yang dikembangkan dapat dievaluasi menggunakan data yang belum pernah dilatih sebelumnya.

## 3. Pembagian Data

Dalam penelitian ini data dibagi menjadi dua bagian utama yaitu data *training* dan data *testing*. Pembagian data secara proposional dilakukan dengan rasio 80:20, di mana 80% dari total data digunakan sebagai data *training* dan 20% sisanya digunakan sebagai data *testing* (Seidu et al., 2022). Data *training* digunakan untuk melatih model prediksi dan data *testing* digunakan untuk mengevaluasi performa dan validasi model.

## 4. Analisa Data

Dalam penelitian ini, proses pemodelan dibagi ke dalam tiga tugas utama, yaitu menentukan kelayakan lokasi pelanggan, mengelompokkan wilayah berdasarkan karakteristik yang sama, dan memprediksi potensi penghematan

energi. Ketiga tugas ini mewakili pendekatan yang berbeda dalam *machine learning*, yaitu klasifikasi, clustering, dan regresi.

#### 4.1. Klasifikasi

Untuk tugas pertama yaitu penentuan kelayakan lokasi pelanggan menggunakan klasifikasi. Pemilihan model didasarkan pada efektivitasnya dalam studi sebelumnya di bidang energi terbarukan. Penelitian membandingkan 5 algoritma yaitu Decision Tree dipilih karena kemampuannya menyederhanakan proses pengambilan keputusan dan mengungkap pola dari data *non-linier* (Berlanger et al., 2023). Random Forest sebagai pengembangan dari Decision Tree memiliki performa lebih stabil dalam data spasial dan heterogen (Saqib et al., 2023). Gradient Boosting dan XGBoost digunakan karena keduanya memiliki akurasi tinggi dalam memproses data lingkungan pada sistem *photovoltaik* (Abdelsattar et al., 2025; Santos et al., 2025). Sementara itu, LightGBM dipertimbangkan karena efisiensi waktu pelatihan dan keandalannya dalam memodelkan radiasi surya secara bertahap (Zhang et al., 2020). Seluruh model tersebut telah terbukti relevan dan kompetitif dalam konteks analisis data energi surya sehingga layak digunakan dalam studi ini.

#### 4.2. Clustering

Tugas kedua melibatkan segmentasi lokasi melalui penerapan metodologi pengelompokan untuk mengkategorikan wilayah berdasarkan tingkat radiasi matahari dan atribut geografis. Pendekatan analitis ini memerlukan perbandingan algoritma Kmeans, *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN), dan *Gaussian Mixture Model* (GMM). Algoritma Kmeans dibedakan oleh kemampuan pemrosesan yang cepat dan interpretabilitas intuitifnya (Abdullah et al., 2024). Berbeda dengan Kmeans, yang memerlukan pra-definisi jumlah *cluster*, DBSCAN menawarkan keuntungan signifikan dari data pengelompokan tanpa pengetahuan sebelumnya tentang kuantitas *cluster*. Karakteristik ini sangat menguntungkan dalam situasi di mana jumlah *cluster* tetap tidak pasti, sehingga memfasilitasi proses pengelompokan yang lebih mudah beradaptasi (Cheng et al., 2021). Algoritma GMM menyajikan manfaat pengelompokan berdasarkan penilaian probabilistik, sehingga meningkatkan fleksibilitas dalam klasifikasi titik data dalam *cluster* (Wang & Jiang, 2021).

#### 4.3. Regresi

Tugas ketiga yaitu melakukan prediksi penghematan energi listrik dengan menggunakan regresi. Metode ini membandingkan algoritma ElasticNet, Random Forest Regressor, XGBoost Regressor dalam memodelkan prediksi penghematan energi (Mohanasundaram & Rangaswamy, 2025; Prashant, 2023; Gökçe & Duman, 2022). Ketiga algoritma tersebut dipilih karena telah mewakili pendekatan regresi yang saling melengkapi yaitu *linear* dengan regularisasi, *ensemble* berbasis pohon, dan *boosting*. Selain itu dari ketiga algoritma telah terbukti efektif dalam konteks prediksi energi di berbagai literatur. Model-model ini menawarkan keseimbangan antara akurasi, efisiensi komputasi, dan interpretabilitas, menjadikannya ideal untuk implementasi praktis dalam perencanaan strategi pemasaran PLTS atap berbasis data.

### 5. Evaluasi Model

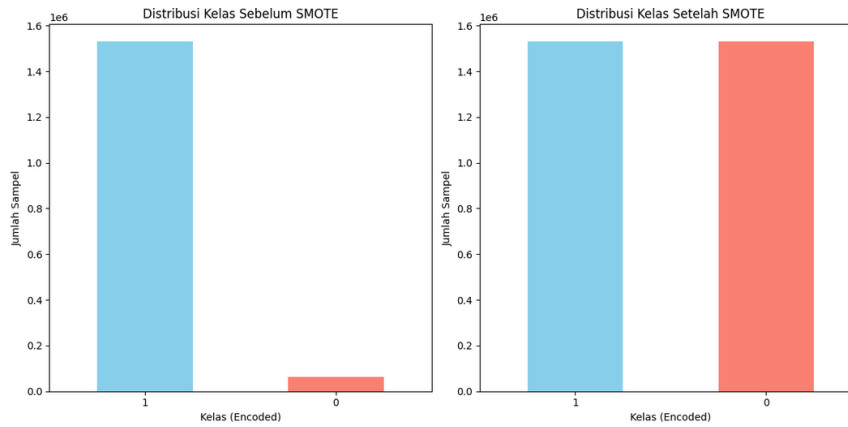
Evaluasi kinerja model untuk metode klasifikasi dengan akurasi, presisi, recall, dan f1-score. Untuk clustering, evaluasi model menggunakan matrik evaluasi silhouette score dan visualisasi untuk tiap algoritma yang digunakan. Sedangkan untuk metode regresi evaluasi yang dilakukan menggunakan MSE dan R2 dengan cross validation.

## HASIL DAN PEMBAHASAN

Penelitian ini mengevaluasi tiga pendekatan *machine learning* secara *independent* yaitu klasifikasi, clustering, dan regresi untuk memberikan rekomendasi lokasi dan estimasi penghematan energi dari PLTS atap. Masing-masing pendekatan diarahkan untuk menjawab pertanyaan berbeda namun saling melengkapi : klasifikasi untuk kelayakan pelanggan, clustering untuk segmentasi wilayah, dan regresi untuk estimasi kuantitatif dampak energi.

### 1. Klasifikasi

Pada tahap ini dilakukan penanganan terhadap ketidakseimbangan distribusi kelas dalam dataset, yaitu antara label layak dan tidak layak untuk pemasangan PLTS atap. Masalah ketidakseimbangan ini dapat menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas. Untuk mengatasi hal tersebut, digunakan metode *Synthetic Minority Over-sampling Technique* (SMOTE), yaitu teknik *oversampling* sintesis yang bekerja dengan membuat data baru pada kelas minoritas berdasarkan interpolasi antara sampel-sampel yang berdekatan. SMOTE membantu menyeimbangkan distribusi label dan meningkatkan kinerja klasifikasi terhadap kelas minoritas. Teknik ini pertama kali diperkenalkan oleh Chawla et al., dan telah digunakan secara luas dalam berbagai studi klasifikasi dengan data tidak seimbang (Chawla et al., 2002). Untuk melihat visualisasi sebelum dan sesudah dilakukan SMOTE dapat dilihat pada Gambar 4.



Gambar 4. Distribusi data sebelum dan sesudah dilakukan SMOTE

### 1.1. Pembentukan Label Kelayakan PLTS

Label target ditentukan berdasarkan tiga kriteria utama: intensitas penyinaran (SS), konsumsi listrik bulanan (pemkWh), dan daya tersambung pelanggan (daya). *Threshold* ditentukan secara dinamis menggunakan pendekatan kuantil, yaitu  $SS > Q3$ ,  $pemkWh > Q3$ , dan  $daya \geq Q2$ . Penggunaan pendekatan kuantil dalam pembentukan *threshold* mengacu pada metode *quantile-based classification* yang dinilai efektif dalam menangani distribusi data yang tidak normal dan heterogen, serta memberikan pemisahan kelas yang stabil dan akurat pada data berdimensi tinggi (Hennig & Viroli, 2013). Pelanggan yang memenuhi ketiga kriteria tersebut diklasifikasikan sebagai Layak, sedangkan sisanya dikategorikan sebagai Tidak Layak.

### 1.2. Pemodelan dengan Algoritma Klasifikasi

Model dilatih menggunakan data hasil *oversampling* dengan SMOTE dan diuji menggunakan data uji yang tidak mengalami *oversampling*. Berikut adalah ringkasan hasil evaluasi dari kelima model menggunakan metrik akurasi, precision, recall, dan f1-score:

Tabel 1. Evaluasi *Modeling* Klasifikasi

| Model             | Akurasi | Precision | Recall | F1-Score |
|-------------------|---------|-----------|--------|----------|
| Decision Tree     | 84.7%   | 0.95      | 0.85   | 0.89     |
| Random Forest     | 82.4%   | 0.96      | 0.76   | 0.85     |
| Gradient Boosting | 84.1%   | 0.95      | 0.84   | 0.88     |
| XGBoost           | 82.3%   | 0.96      | 0.76   | 0.85     |
| LightGBM          | 87.0%   | 0.98      | 0.78   | 0.86     |

Tabel 1 menunjukkan bahwa LightGBM memiliki akurasi tertinggi (87%), disusul oleh Decision Tree dan Gradient Boosting. Decision Tree mencatat nilai recall tertinggi (85%), yang menunjukkan kemampuannya dalam mengenali lebih banyak pelanggan potensial. Seluruh model dikonfigurasi dengan  $max\_depth=10$  untuk mencegah *overfitting*. Teknik SMOTE terbukti efektif dalam meningkatkan deteksi pada kelas minoritas ("Layak"), terutama pada metrik recall. LightGBM dipilih sebagai model terbaik, tidak hanya karena akurasinya tertinggi, tetapi juga karena kemampuannya menangani data berdimensi tinggi, relasi *non-linear* antar fitur, serta toleransinya terhadap data tidak seimbang saat dikombinasikan dengan SMOTE (Ke et al., 2017).

Untuk memastikan kinerja setiap model diperoleh secara optimal dan adil, dilakukan penyetelan parameter (*hyperparameter tuning*) dengan mempertimbangkan kombinasi nilai parameter yang umum digunakan dalam praktik pembelajaran mesin dan studi-studi sebelumnya. Adapun konfigurasi parameter yang digunakan adalah sebagai berikut:

- Decision Tree:  $max\_depth=10$ ,  $min\_samples\_split=2$ ,  $criterion='gini'$
- Random Forest:  $n\_estimators=100$ ,  $max\_depth=10$ ,  $random\_state=42$ ,  $criterion='gini'$
- Gradient Boosting:  $learning\_rate=0.1$ ,  $n\_estimators=100$ ,  $max\_depth=10$
- XGBoost:  $learning\_rate=0.1$ ,  $n\_estimators=100$ ,  $max\_depth=10$ ,  $subsample=0.8$
- LightGBM:  $learning\_rate=0.1$ ,  $n\_estimators=100$ ,  $max\_depth=10$ ,  $num\_leaves=31$ ,  $objective='binary'$

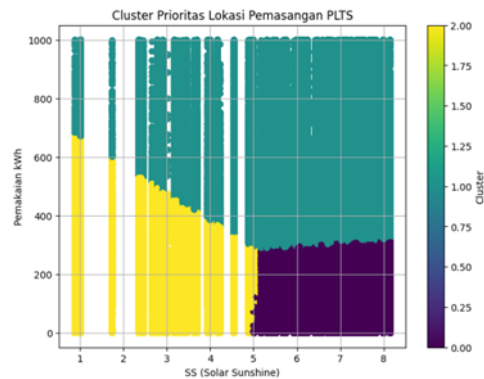
Pemilihan kombinasi parameter tersebut mengacu pada praktik umum di bidang pembelajaran mesin, seperti yang digunakan dalam studi oleh Pedregosa et al. (2011) untuk Scikit-learn dan oleh Ke et al. (2017) dalam pengembangan LightGBM.

## 2. Clustering

Proses sebelum pemodelan, dataset yang digunakan sudah dilakukan *preprocessing* terlebih dahulu sehingga siap untuk dilakukan pemodelan. Clustering sendiri merupakan metode *unsupervised learning* yang bertujuan untuk mengelompokkan data berdasarkan kemiripan atau kedekatan karakteristik antar data. Dalam konteks penelitian ini, clustering membantu mengidentifikasi area-area yang memiliki pola serupa dalam hal paparan sinar matahari dan konsumsi energi listrik, sehingga dapat digunakan sebagai dasar dalam menentukan lokasi yang optimal untuk pemasangan PLTS atap.

### 2.1. Clustering dengan K-Means

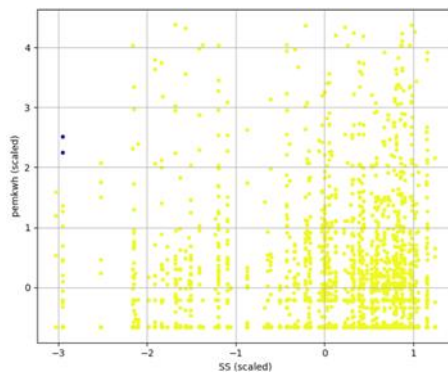
Sebelum dilakukan pemodelan dengan menggunakan algoritma K-Means, perlu menentukan jumlah *cluster* yang akan digunakan. Disini untuk menentukan jumlah *cluster* menggunakan *elbow method*. *Elbow method* digunakan untuk membantu menentukan jumlah *cluster* ( $k$ ) yang optimal dalam pengelompokan K-mean (Abdullah et al., 2024). Dari hasil yang dilakukan didapat jumlah *cluster* sebanyak 3. Dari jumlah *cluster* dilakukan *modeling* dan didapat visualisasinya seperti Gambar 5.



Gambar 5. Visualisasi Clustering dengan K-Means menggunakan K=3

## 2.2. Clustering dengan DBSCAN

Algoritma DBSCAN mengidentifikasi *cluster* berdasarkan konsep pencapaian kepadatan. Suatu titik dianggap dapat dijangkau secara langsung kepadatan dari titik lain jika berada dalam jarak  $\epsilon$  dan titik lainnya adalah titik inti. Ini berarti bahwa *cluster* dibentuk dengan menghubungkan titik-titik inti dan titik-titik yang dapat dijangkau langsung (Cheng et al., 2021). Dalam pemodelan dengan DBSCAN Adapun hasil visualisasi *cluster* untuk metode DBSCAN seperti Gambar 6.

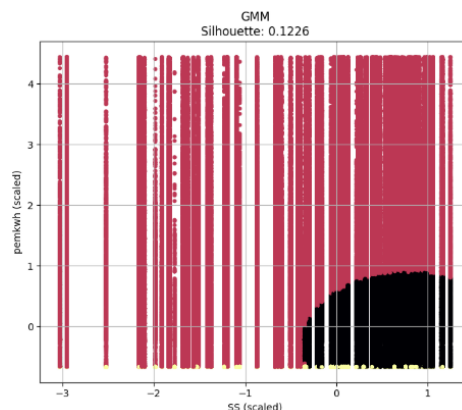


Gambar 6. Visualisasi Clustering dengan DBSCAN

Dari hasil visualisasi Gambar 6, terlihat bahwa metode DBSCAN hanya menghasilkan sebuah *cluster* yang besar dan padat. Ada 2 buah titik berwarna biru yang berbeda dari titik mayoritas, kemungkinan titik tersebut merupakan *outlier / noise*.

## 2.3. Clustering dengan GMM

Algoritma GMM adalah algoritma clustering berbasis model probabilistik, yang mengasumsikan bahwa data berasal dari campuran beberapa distribusi Gaussian (*normal distribution*). Tujuan utama GMM adalah mengelompokkan data dengan memodelkan setiap *cluster* sebagai distribusi Gaussian. GMM tidak hanya memodelkan data tetapi juga berfungsi sebagai sinyal pengawasan untuk meningkatkan representasi, meningkatkan kinerja pengelompokan secara keseluruhan (Wang & Jiang, 2021). Hasil visualisasi dari algoritma GMM seperti Gambar 7.



Gambar 7. Visualisasi Clustering dengan GMM

Dari hasil visualisasi Gambar 7 terlihat bahwa clustering kurang optimal dan dilihat distribusi *cluster* terlihat tidak natural.

#### 2.4. Hasil Clustering

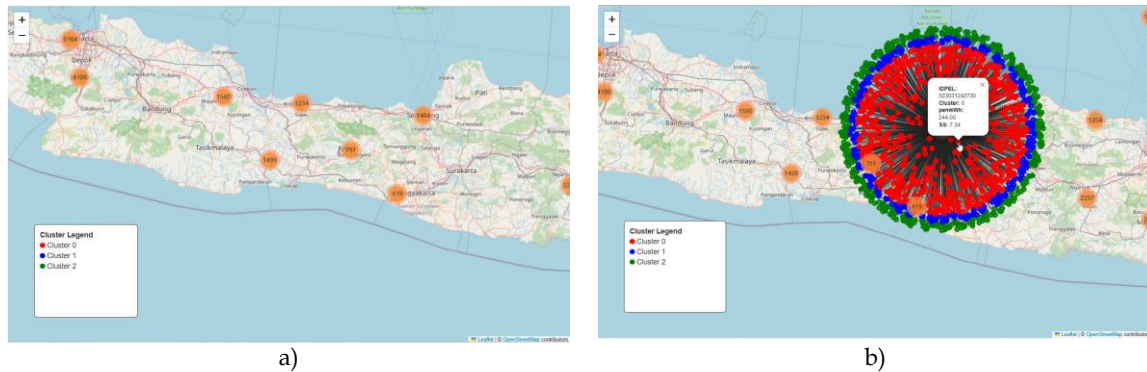
Dalam penelitian ini, dilakukan perbandingan tiga algoritma clustering KMeans, DBSCAN, dan GMM untuk segmentasi lokasi potensial pemasangan PLTS atap. Untuk hasil evaluasi menggunakan Silhouette Score dapat dilihat Tabel 2.

**Tabel 2.** Evaluasi Menggunakan Silhouette Score

| Metode Evaluasi  | KMeans | DBSCAN | GMM    |
|------------------|--------|--------|--------|
| Silhouette Score | 0.5566 | 0.5594 | 0.1226 |

Hasil evaluasi menggunakan Silhouette Score menunjukkan bahwa DBSCAN memiliki skor tertinggi sebesar 0.5594, diikuti oleh KMeans dengan 0.5566, dan GMM dengan skor terendah, yaitu 0.1226. Namun, jika dilihat dari visualisasi hasil clustering, KMeans justru menunjukkan pemisahan klaster yang paling jelas dan seimbang, menjadikannya lebih unggul dalam konteks interpretabilitas dan segmentasi yang bermakna.

Hal ini menunjukkan bahwa nilai Silhouette Score yang tinggi tidak selalu menjamin kualitas clustering yang lebih baik, terutama jika tidak didukung oleh visualisasi yang representatif (Blasilli et al., 2024). KMeans dinilai sebagai algoritma paling tepat untuk digunakan dalam segmentasi lokasi pemasangan PLTS atap, karena mampu membagi data secara logis, terpisah dengan baik, dan mendukung tujuan analisis dalam penelitian ini. Berikut visualisasi hasil clustering dengan menggunakan KMeans dengan menampilkan peta sebarannya dapat dilihat dari Gambar 8.



**Gambar 8.** Sebaran *Cluster*: a) Tampilan Peta Sebaran dan b) Tampilan *Cluster* di satu daerah

### 3. Regresi

#### 3.1. Menentukan Target Variabel

Dalam penelitian ini, target variabel pada regresi adalah estimasi penghematan energi bulanan (dalam kWh) yang dapat dicapai oleh pelanggan apabila PLTS atap berkapasitas 3 kWp dipasang (Elieser et al., 2021; Ardianto & Yulianto, 2021; Suardana & Yasa, 2022). Estimasi ini dihitung berdasarkan model teknis umum dari potensi produksi listrik PLTS (Ramadhani & Nufus, 2024), Rumus menghitung energi yang dihasilkan PLTS atap :

$$E_{hari} = P \times SS \times \eta \quad (1)$$

$$E_{bulan} = E_{hari} \times 30 \quad (2)$$

$$Penghematan\ Energi = P_{listrik} - E_{bulan} \quad (3)$$

Di mana:

- $P$  = Kapasitas PLTS Atap (dalam Watt)
- $SS$  = Lama Penyinaran (dalam jam)
- $\eta$  = Efisiensi Sistem Panel Surya
- $P_{listrik}$  = Pemakaian energi pelanggan (dalam kWh)

Pendekatan ini digunakan dalam studi teknis energi surya, dan hasil estimasi dijadikan target regresi untuk memodelkan hubungan antara karakteristik pelanggan (lokasi, konsumsi, penyinaran) dan potensi penghematan energi pelanggan.

#### 3.2. Pemodelan dan Evaluasi

Tiga algoritma regresi dibandingkan dalam penelitian ini: ElasticNet (Mohanasundaram & Rangaswamy, 2025), Random Forest Regressor, dan XGBoost Regressor. Pemodelan dilakukan menggunakan dataset hasil *feature engineering* dengan validasi silang 5-fold. Metrik evaluasi yang digunakan meliputi *Mean Squared Error* (MSE) dan koefisien determinasi ( $R^2$ ) (Gökçe & Duman, 2022).

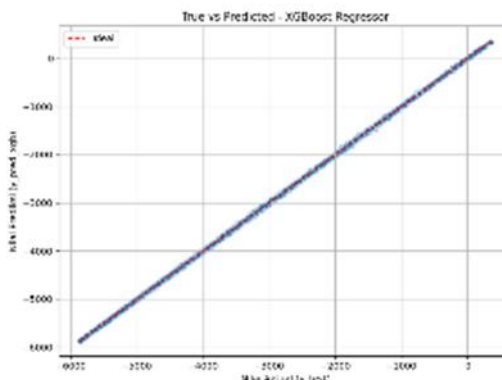
Dalam penelitian ini, seluruh model dilatih menggunakan data hasil *feature engineering* dengan proses validasi silang (*k-fold cross-validation*) menggunakan metrik MSE dan  $R^2$  sebagai acuan evaluasi utama (Hussain et al., 2025). Hasil evaluasi dari Pemodelan Regresi ditunjukkan pada Tabel 3.

Tabel 3. Tabel hasil pemodelan regresi

| Model Regresi         | Mean Squared Error (MSE) | MSE Cross-Validation | Koefisien Determinasi (R <sup>2</sup> ) | Akurasi (%) |
|-----------------------|--------------------------|----------------------|---|-------------|
| ElasticNet Regression | 53,42                    | 53,294               | 0,965                                   | 96,50%      |
| Random Forest         | 174,05                   | 170,42               | 0,9998                                  | 99,98%      |
| XGBoost Regressor     | 32,25                    | 32,82                | 0,9999                                  | 99,99%      |

Tabel 3 menunjukkan bahwa XGBoost Regressor memiliki performa terbaik, dengan MSE terendah dan R<sup>2</sup> tertinggi. Keunggulan ini konsisten dengan hasil penelitian sebelumnya bahwa XGBoost efektif dalam memodelkan beban energi karena kemampuannya menangani non-linearitas, interaksi fitur, dan outlier dengan baik (Gökçe & Duman, 2022).

Untuk menunjukkan bahwa model mampu melakukan generalisasi dengan baik dan tidak mengalami bias terhadap kelompok data tertentu dapat dibuktikan visualisasi pada Gambar 9 terkait dengan perbandingan *true* dan *predicted* (Pagano et al., 2023):



Gambar 9. Akurasi *True* vs *Predicted*

### 3.3. Implikasi dan Keterkaitan dengan Tujuan Penelitian

Temuan regresi ini menjawab salah satu tujuan utama penelitian, yaitu menyediakan prediksi kuantitatif terkait potensi penghematan energi yang dapat digunakan PLN dalam merancang strategi pemasaran PLTS atap yang berbasis data. Dengan memanfaatkan model XGBoost yang sudah divalidasi, PLN dapat melakukan simulasi dampak penghematan energi dari pemasangan PLTS atap secara akurat untuk setiap pelanggan atau wilayah, yang sebelumnya hanya didasarkan pada estimasi kasar (PLN Icon Plus, 2022). Model ini dapat digunakan sebagai bagian dari *decision-support system* yang mendukung promosi berbasis estimasi keuntungan pelanggan (*value-based targeting*), terutama pada segmen rumah tangga menengah ke atas. Gambar 10 menunjukkan hasil prediksi penghematan energi dengan rekomendasi XGBoost Regressor sebagai analisis terbaik dari penelitian ini.

Hasil Prediksi (SS, pemkwh, Prediksi\_Penghematan) terbesar:

|         | SS       | pemkwh | Prediksi_Penghematan |
|---------|----------|--------|----------------------|
| 1390608 | 0.890323 | 999.0  | 333.563232           |
| 823184  | 0.890323 | 991.8  | 333.563232           |
| 823185  | 0.890323 | 984.3  | 333.563232           |
| 1899967 | 0.890323 | 937.0  | 291.803497           |
| 1260529 | 0.890323 | 924.0  | 272.635651           |

Gambar 10. Prediksi penghematan listrik PLTS atap dengan XGboost regressor

## KESIMPULAN DAN SARAN

Penelitian ini mengevaluasi tiga pendekatan *machine learning* secara independen untuk mendukung proses pengambilan keputusan dalam pemasangan PLTS atap di Pulau Jawa berdasarkan data radiasi matahari dan konsumsi listrik pelanggan. Tiga tugas utama yang diimplementasikan meliputi klasifikasi kelayakan pelanggan, segmentasi wilayah potensial, dan estimasi penghematan energi listrik.

Pada tugas klasifikasi, lima algoritma dibandingkan untuk menentukan kelayakan pelanggan dalam pemasangan PLTS atap, dengan model LightGBM menghasilkan performa terbaik dengan akurasi 87%. Segmentasi wilayah menggunakan tiga algoritma clustering menunjukkan bahwa meskipun DBSCAN memperoleh silhouette score tertinggi (0,5594), hasil visualisasi dari KMeans dinilai paling representatif dan interpretatif. Pada tugas regresi, prediksi penghematan energi bulanan menggunakan XGBoost Regressor menunjukkan akurasi tertinggi dengan nilai koefisien determinasi (R<sup>2</sup>) sebesar 0,9999 dan MSE terendah dibandingkan model lainnya.

Kontribusi utama dari penelitian ini adalah integrasi simultan data radiasi matahari dan konsumsi listrik pelanggan dalam satu pendekatan analitis yang komprehensif menggunakan teknik *machine learning*. Hal ini memberikan kerangka kerja baru untuk menggabungkan klasifikasi kelayakan, segmentasi wilayah, dan estimasi penghematan energi dalam

konteks pemasangan PLTS atap, yang sebelumnya umumnya diteliti secara terpisah.

Meskipun ketiga pendekatan dikembangkan dan dianalisis secara terpisah, hasilnya memberikan wawasan teknis yang signifikan untuk merumuskan strategi pemasangan PLTS atap yang berbasis data. Pendekatan ini dapat menjadi dasar pengembangan sistem pendukung keputusan (*decision-support system*) yang menggabungkan evaluasi kelayakan, pemetaan wilayah potensial, dan estimasi manfaat energi secara kuantitatif. Penelitian ini juga memperkuat pentingnya integrasi data spasial dan perilaku konsumsi dalam perencanaan energi terbarukan.

Keterbatasan utama penelitian ini terletak pada cakupan geografis yang hanya mencakup Pulau Jawa serta belum memperhitungkan variabel ekonomi pelanggan, seperti tingkat pengembalian investasi *Return On Investment* (ROI), struktur tarif listrik, dan insentif kebijakan. Penelitian lanjutan disarankan untuk memperluas cakupan wilayah, menambahkan parameter ekonomi dan regulasi, serta mengembangkan sistem berbasis AI yang terintegrasi secara *end-to-end*.

## UCAPAN TERIMAKASIH

Peneliti menyatakan bahwa seluruh proses penelitian, mulai dari perumusan masalah, pengumpulan dan analisis data, pengembangan model, hingga penulisan karya ilmiah ini, dilakukan secara mandiri dan orisinal. Semua data dan referensi yang digunakan telah dicantumkan secara jelas dan sesuai dengan kaidah akademik yang berlaku. Dengan ini ditegaskan bahwa tidak terdapat unsur plagiarisme dalam bentuk apa pun dalam naskah ini.

Peneliti mengucapkan terima kasih yang sebesar-besarnya kepada BMKG dan PLN yang telah memberikan data guna mendukung analisis dan eksperimen dalam penelitian ini. Sebagai bagian dari upaya peningkatan kualitas penelitian, penulis menggunakan Teknologi *Large Language Model* (LLM) yang berguna sebagai referensi pendukung untuk keperluan *literature review*, penyusunan struktur tata bahasa yang baik, serta memberikan rekomendasi dalam perbaikan redaksi naskah ilmiah secara efisien dan akurat.

## DAFTAR PUSTAKA

- Abdelsattar, M., AbdelMoety, A., & Emad-Eldeen, A. (2025). Comparative analysis of machine learning techniques for temperature and humidity prediction in photovoltaic environments. *Scientific Reports*, 15, 15650.
- Abdullah, M. R., Lu, Q.-C., Hussain, A., Tripura, S., Xu, P.-C., & Wang, S. (2024). Location optimization of EV charging stations: A custom K-means cluster algorithm approach. *Energy Reports*, 12, 5367-5382.
- Ardianto, D., & Yulianto, M. (2021). Simulation and experimental results of a 3 kWp rooftop PV system in Surabaya. *Proceedings of the 5th International Conference on Vocational Education and Electrical Engineering (ICVEE)*.
- Asian Development Bank. (2021). Indonesia energy sector assessment, strategy, and road map. <https://www.adb.org/sites/default/files/institutional-document/666741/indonesia-energy-asr-update.pdf>
- Berlanger, N., van Ophoven, N., Verdonck, T., & Wilms, I. (2023). Tree-based forecasting of day-ahead solar power generation from granular meteorological features. *arXiv preprint arXiv:2312.00090*.
- Blasilli, G., Kerrigan, D., Bertini, E., & Santucci, G. (2024, October). Towards a visual perception-based analysis of clustering quality metrics. *IEEE Visualization in Data Science (VDS)*, 15–24. <https://doi.org/10.1109/vds63897.2024.00007>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Cheng, D., Yue, G., Pei, T., & Wu, M. (2021). Clustering indoor positioning data using E-DBSCAN. *International Journal of Geo-Information*, 10(10), 669. <https://doi.org/10.3390/ijgi10100669>
- Citraningrum, M., & Tumiwa, F. (2019). Market potential of rooftop solar PV in Surabaya: A report. Institute for Essential Services Reform (IESR). <https://www.iesr.or.id>
- Climate Transparency. (2024). Implementation check: Renewable energy development in Indonesia. <https://www.climate-transparency.org/wp-content/uploads/2024/01/Implementation-Check-Renewable-Energy-Development-in-Indonesia-2024.pdf>
- de Luis-Ruiz, J. M., Salas-Menocal, B. R., Pereda-García, R., Pérez-Álvarez, R., Sedano-Cibrián, J., & Ruiz-Fernández, C. (2024). Optimal location of solar photovoltaic plants using geographic information systems and multi-criteria analysis. *Sustainability*, 16(7), 2895. <https://doi.org/10.3390/su16072895>
- Elieser, R., Sudibyo, H., & Widodo, T. (2021). Simulasi sistem PLTS atap dan harga satuan energi listrik untuk skala rumah tangga di Surabaya. *Jurnal Rekayasa ElektriKa*, 18(2), 100–107. [https://repository.ubaya.ac.id/42271/3/JRE%20Vol%2018%20No%202%20-%20Elieser\\_Rev.pdf](https://repository.ubaya.ac.id/42271/3/JRE%20Vol%2018%20No%202%20-%20Elieser_Rev.pdf)
- Fakhraian, E., Formenta, M. A., Valls Dalmau, F., Nameni, A., & Casañ Guerrero, M. J. (2021). Determination of the urban rooftop photovoltaic potential: A state of the art. *Energy Reports*, 7, 907–920. <https://doi.org/10.1016/j.egy.2021.06.031>
- Gökçe, M. M., & Duman, E. (2022). Performance comparison of simple regression, random forest and XGBoost algorithms for forecasting electricity demand. 2022 3rd International Informatics and Software Engineering Conference (IISEC), 1–6. <https://doi.org/10.1109/IISEC56263.2022.9998213>
- Hennig, C., & Viroli, C. (2013). Quantile-based classifiers. *arXiv preprint arXiv:1303.1282*.
- Hussain, I., Ching, K. B., Uttraphan, C., Tay, K. G., & Noor, A. (2025). Evaluating machine learning algorithms for energy consumption prediction in electric vehicles: A comparative study. *Scientific Reports*, 15.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, 3146–3154.
- Kementerian Energi dan Sumber Daya Mineral. (n.d.). Matahari untuk PLTS di Indonesia. Diakses 2 April 2025, dari <https://www.esdm.go.id/id/media-center/arsip-berita/matahari-untuk-plts-di-Indonesia>

- Kementerian Energi dan Sumber Daya Mineral Republik Indonesia. (2021). Peraturan Menteri Energi dan Sumber Daya Mineral Nomor 11 Tahun 2021 tentang Pelaksanaan Usaha Ketenagalistrikan. Berita Negara Republik Indonesia, No. 671.
- Kementerian Energi dan Sumber Daya Mineral Republik Indonesia. (2024). Peraturan Menteri ESDM Nomor 2 Tahun 2024 tentang Pembangkit Listrik Tenaga Surya (PLTS) Atap pada Jaringan Tenaga Listrik Pemegang Izin Usaha Penyedia Tenaga Listrik Untuk Kepentingan Umum. JDIIH ESDM. <https://jdih.esdm.go.id>
- Kompas. (2023, 15 Juni). Potensi PLTS atap Indonesia tembus 32,5 gigawatt. [lestari.kompas.com](https://lestari.kompas.com/read/2023/06/15/080000086/potensi-plts-atap-indonesia-tembus-32-5-gigawatt). <https://lestari.kompas.com/read/2023/06/15/080000086/potensi-plts-atap-indonesia-tembus-32-5-gigawatt>
- Levent, İ., Şahin, G., Işık, G., & van Sark, W. G. J. H. M. (2025). Comparative analysis of advanced machine learning regression models with advanced artificial intelligence techniques to predict rooftop PV solar power plant efficiency using indoor solar panel parameters. *Applied Sciences*, 15(6), 3320. <https://doi.org/10.3390/app15063320>
- McKinsey & Company. (2023). Indonesia's green powerhouse promise: Ten big bets that could pay off. <https://www.mckinsey.com/id/our-insights/indonesias-green-powerhouse-promise-ten-big-bets-that-could-pay-off>
- Mohanasundaram, V., & Rangaswamy, B. (2025). Elastic net with Bayesian density estimation model for feature selection for photovoltaic energy prediction. *Scientific Reports*, 15, 8736. <https://doi.org/10.1038/s41598-025-92633-1>
- Mora-Gaona, M., Neumann, U., Vargas-Canas, R., & López, D. M. (2021). Evaluating the impact of multivariate imputation by MICE in feature selection. *PLOS ONE*, 16(7), e0254720. <https://doi.org/10.1371/journal.pone.0254720>
- Pagano, T. P., Loureiro, R. B., Lisboa, F. V. N., Peixoto, R. M., Guimarães, G. A. S., Cruz, G. O. R., Araujo, M. M., Santos, L. L., Cruz, M. A. S., & Oliveira, E. L. S. (2023). Bias and unfairness in machine learning models: A systematic review on datasets, tools, fairness metrics, and identification and mitigation methods. *Big Data and Cognitive Computing*, 7(1), 15. <https://doi.org/10.3390/bdcc7010015>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- PLN Icon Plus. (2022, 21 September). Customer sales kit PV rooftop PLN Icon Plus [Slide PowerPoint]. PT PLN (Persero) & ICON+.
- Prashant. (2023). Solar power prediction using machine learning - Random Forest Regressor. Medium. <https://prashant-one4all.medium.com/solar-power-prediction-using-machine-learning-randomforestregressor-10babe26c8e1>
- Ramadhani, A., & Nufus, T. H. (2024). Evaluasi sistem PLTS grid-connected 21.60 kWp di Politeknik Negeri Jakarta dengan metode Failure Mode Effect Analysis (FMEA). *Jurnal Mekanik Terapan*, 5(2), 103-112.
- Rishitha, N., Muthu Reshmi, K., Gulecha, S., & Vani, K. (2024). Machine learning-driven solar panel site selection and rooftop potential estimation for sustainable development goals. *Proc. Asian Conf. Remote Sensing*, 1–8. <https://doi.org/10.1109/ACRS53953.2024.1014856>
- Santos, C., Pires, J. S., & Soares, J. (2025). Machine Learning for identifying potential photovoltaic installations in urban parking areas. *Proceedings of the 14th International Conference on Smart Cities and Green ICT Systems*.
- Saqib, M., Zhang, J., Iqbal, M. W., Irshad, A., & Hussain, S. (2023). Geographic information system and machine learning approach for solar photovoltaic site selection: A case study in Pakistan. *Processes*, 11(4), 981.
- Seidu, J. M., Ewusi, A., Kuma, J. S., Ziggah, Y. Y., & Voigt, H.-J. (2022). Impact of data partitioning in groundwater level prediction using artificial neural network for multiple wells. *International Journal of River Basin Management*, 1–12. <https://doi.org/10.1080/15715124.2022.2079653>
- Suardana, I. N. A., & Yasa, I. N. A. S. (2022). Studi performance PLTS rooftop 3kWp frameless dengan sistem on-grid di lingkungan perumahan Kori Nuansa Jimbaran. *Jurnal Teknik ITS*, 11(1), A90–A95. <https://www.researchgate.net/publication/363339238>
- Wang, J., & Jiang, J. (2021). Unsupervised deep clustering via adaptive GMM modeling and optimization. *Neurocomputing*, 433, 199–211. <https://doi.org/10.1016/j.neucom.2020.12.082>
- World Bank Group. (n.d.). Global solar atlas - Indonesia. Diakses 7 April 2025, dari <https://globalsolaratlas.info/map?c=-2.661188,118.1042,4&r=IDN>
- Zhang, Y., Wang, D., & Liu, W. (2020). Multistep-ahead solar radiation forecasting scheme based on the Light Gradient Boosting Machine. *Remote Sensing*, 12(14), 2271.