

PEMBANGUNAN MODEL PREDIKSI KEBUTUHAN ENERGI LISTRIK PT PLN (PERSERO) UIW NTB TAHUN 2025 DENGAN METODE RANDOM FOREST BERDASARKAN KELOMPOK TARIF

(Construction of an Electricity Demand Forecasting Model for PT PLN (Persero) UIW NTB for the Year 2025 Using the Random Forest Algorithm Based on Tariff Categories)

Ida Ayu Dewi Purnama Anjani^[1], Raphael Bianco Huwae^[1], Komang Yehuda Prawira^[2], Andy Hidayat Jatmika^[1]

^[1]Dept Informatics Engineering, Mataram University
Jl. Majapahit 62, Mataram, Lombok NTB, INDONESIA

^[2]PT PLN UIW NTB

Jl. Langko 25-27, Taman Sari, Kec. Ampenan, Mataram, Lombok NTB, INDONESIA

Email: dayudewi37@gmail.com, Raphael.bianco.huwae@unram.ac.id, komangyehudaprawira@gmail.com, andy@unram.ac.id

Abstrak

Kebutuhan akan energi listrik di Nusa Tenggara Barat semakin tinggi seiring dengan meningkatnya aktivitas masyarakat dan jumlah pelanggan yang semakin banyak. Oleh karena itu, memiliki sistem prediksi yang akurat sangat penting untuk mendukung perencanaan distribusi energi yang efisien. Penelitian ini bertujuan untuk mengembangkan model yang dapat meramalkan konsumsi listrik pada tahun 2025 berdasarkan kategori tarif dengan menggunakan algoritma Random Forest. Data yang digunakan adalah konsumsi listrik dari tahun 2016 hingga 2024 yang diperoleh dari PT PLN (Persero) UIW NTB. Proses pengolahan data dilakukan melalui pembersihan dan preparasi, diikuti dengan pelatihan model serta penyesuaian hyperparameter menggunakan metode Grid Search. Hasil evaluasi model menunjukkan kinerja yang baik dengan nilai R^2 mencapai 0,9057 dan MAE sebesar 0,48. Prediksi untuk tahun 2025 memperlihatkan bahwa kelompok tarif rumah tangga dengan daya 450 VA akan menjadi kelompok konsumen yang paling besar. Hasil ini dapat dijadikan sebagai acuan untuk merencanakan strategi distribusi energi oleh PLN.

Keywords: Random Forest, Prediksi Konsumsi Listrik, PLN, Kelompok Tarif, Machine Learning.

1. PENDAHULUAN

Energi listrik merupakan kebutuhan esensial dalam berbagai aspek kehidupan masyarakat modern, baik di sektor rumah tangga, industri, bisnis, sosial, maupun pemerintahan. Pemanfaatan energi listrik menjadi fondasi utama dalam mendukung aktivitas ekonomi, pelayanan publik, serta peningkatan kesejahteraan masyarakat. Berdasarkan data dari PT Perusahaan Listrik Negara (Persero) yang dirilis pada tanggal 3 Juni 2024, jumlah pelanggan listrik di Indonesia hingga akhir tahun 2023 mencapai 89.153.278 pelanggan, dengan detail sektor rumah tangga sebanyak 81.551.325 pelanggan, bisnis sebanyak 4.706.947 pelanggan, sosial sebanyak 1.993.101 pelanggan, pemerintahan sebanyak 595.941 pelanggan, dan industri sebanyak 206.770 pelanggan [1]. Melihat data tersebut, masing-masing sektor menunjukkan karakteristik konsumsi energi yang berbeda, sehingga sangat penting untuk melakukan prediksi kebutuhan listrik berdasarkan kelompok tarif pelanggan guna mendukung perencanaan distribusi energi yang efektif dan efisien.

PT PLN (Persero) Unit Induk Wilayah Nusa Tenggara Barat (UIW NTB) merupakan salah satu badan usaha milik negara yang memiliki tanggung jawab penting dalam memastikan ketersediaan dan keandalan pasokan listrik bagi masyarakat NTB. Untuk mendukung pengambilan keputusan dan perencanaan infrastruktur kelistrikan di wilayah tersebut, PLN UIW NTB selama ini menggunakan perangkat lunak Microsoft Excel untuk memprediksi kebutuhan energi listrik. Namun, metode konvensional ini memiliki sejumlah keterbatasan, terutama dalam hal efisiensi dan akurasi ketika menghadapi data historis yang besar, kompleks, dan terus berkembang. Dalam konteks kebutuhan yang semakin mendesak akan sistem prediksi yang handal, teknologi berbasis *machine learning* menawarkan solusi yang lebih tepat dan efisien [2].

Salah satu algoritma yang banyak digunakan dalam prediksi kebutuhan energi adalah *Random Forest*. Algoritma ini memiliki sejumlah keunggulan, seperti kemampuannya menangani data non-linear, mengurangi risiko *overfitting* melalui mekanisme *bagging*, serta menghasilkan prediksi yang relatif stabil pada berbagai kondisi data. *Random Forest* juga mampu memberikan informasi mengenai tingkat kepentingan variabel (*feature importance*), sehingga memudahkan PLN dalam mengidentifikasi faktor-faktor utama yang memengaruhi konsumsi energi, misalnya pola cuaca, musim, dan aktivitas ekonomi [3]. Dibandingkan dengan algoritma lain, seperti *Gradient Boosting*, *Random Forest* cenderung lebih tahan terhadap *noise* dan membutuhkan *tuning* parameter yang lebih sederhana, meskipun *Gradient Boosting* sering kali mencapai akurasi lebih tinggi jika parameter dioptimalkan dengan baik [4]. Sementara itu, algoritma *Neural Network*, khususnya LSTM, unggul dalam menangkap pola deret waktu jangka panjang, tetapi membutuhkan data yang sangat besar serta komputasi intensif [5].

Beberapa penelitian sebelumnya telah membuktikan efektivitas penggunaan algoritma *Random Forest* dalam memprediksi konsumsi energi listrik. Penelitian oleh Liu dan rekan-rekannya menggabungkan *fuzzy c-means clustering*, *Random Forest*, dan *Deep Neural Network* (DNN) untuk menghasilkan prediksi beban listrik jangka pendek yang akurat [6]. Sementara itu, penelitian oleh Dudek menunjukkan bahwa *Random Forest* dapat memberikan hasil prediksi yang baik dalam sistem energi kampus [7]. Selain itu, perkembangan metode explainable AI seperti SHAP dan LIME juga mendukung penerapan *Random Forest* dengan memberikan interpretasi yang lebih transparan terhadap hasil prediksi model [8] [9]. Dengan dasar tersebut, penelitian ini dilakukan untuk membangun model prediksi konsumsi energi listrik berdasarkan kelompok tarif pelanggan PLN UIW NTB untuk tahun 2025 menggunakan algoritma *Random Forest*.

2. TINJAUAN PUSTAKA

2.1 Machine Learning

Machine learning merupakan bagian dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan komputer untuk belajar dari data dan mengambil keputusan tanpa perlu diprogram secara eksplisit. Dalam dunia energi, khususnya pada prediksi konsumsi listrik, *machine learning* menjadi alat yang sangat bermanfaat karena mampu memanfaatkan data historis seperti catatan pemakaian listrik, kondisi cuaca, dan informasi pelanggan untuk memperkirakan kebutuhan listrik di masa mendatang dengan lebih akurat. Bagi perusahaan seperti PT PLN (Persero), pendekatan ini sangat membantu dalam perencanaan distribusi dan pengambilan keputusan operasional yang lebih efisien [10] [11].

Tidak hanya untuk prediksi, *machine learning* juga bisa digunakan untuk segmentasi pelanggan dan penentuan strategi layanan yang lebih tepat sasaran, seperti pemberian insentif atau program efisiensi energi. Hal ini dimungkinkan karena *machine learning* dibangun di atas dasar statistik, matematika, dan teknik *data mining* yang memungkinkan sistem belajar dan berkembang seiring bertambahnya data. Berbagai algoritma telah digunakan dalam konteks ini, salah satunya yang paling sering dipakai adalah *Random Forest*.

Beberapa penelitian sebelumnya memperkuat peran penting *machine learning*. Liu dan Dong [12], misalnya, menunjukkan bahwa *Random Forest* sangat efektif dalam memilih fitur terbaik untuk sistem prediksi berbasis CNN-BiGRU. Penelitian lainnya yang menggabungkan *multilayer perceptron* (MLP) dengan *Random Forest* juga terbukti mampu memprediksi beban listrik secara akurat di lingkungan kampus teknik [13]. Di sisi lain, kebutuhan untuk memahami bagaimana model membuat keputusan mendorong penggunaan metode explainable AI seperti SHAP dan LIME. Penelitian oleh Salih dan rekan-rekannya [14] dan penelitian dari Universitas Koblenz [15] menunjukkan bagaimana visualisasi SHAP dapat memberikan pemahaman yang lebih jelas tentang peran setiap fitur dalam model prediksi.

2.2 Algoritma Random Forest

Random Forest adalah salah satu algoritma populer dalam *machine learning* yang menggabungkan banyak pohon keputusan (*decision tree*) untuk membuat prediksi yang lebih stabil dan akurat [16]. Metode ini bekerja dengan membentuk sejumlah pohon dari sampel data yang dipilih secara acak, lalu hasil akhir diperoleh berdasarkan mayoritas prediksi dari pohon-pohon tersebut. Pendekatan ini membuat *Random Forest* lebih tahan terhadap *overfitting* dibanding *decision tree* tunggal dan mampu menangani data yang kompleks dan bervariasi [1].

Keunggulan utama *Random Forest* adalah kemampuannya dalam menghasilkan model yang handal tanpa perlu pengaturan parameter yang rumit. Selain itu, algoritma ini dapat menangani data kategorial dan numerik sekaligus, serta relatif tidak terpengaruh oleh nilai yang hilang atau tidak lengkap. Dalam praktiknya, *Random Forest* sangat cocok digunakan untuk memprediksi konsumsi energi karena dapat mengenali pola yang tersembunyi dan

memberikan estimasi yang realistis. Penelitian dari MDPI bahkan menunjukkan bahwa akurasi prediksi *Random Forest* bisa mencapai lebih dari 85% setelah dilakukan *tuning* terhadap *hyperparameter* [17].

Selain itu, *Random Forest* juga banyak digunakan di berbagai bidang seperti deteksi anomali, segmentasi pelanggan, dan analisis risiko. Teknik pemilihan fitur secara acak yang digunakan dalam setiap pohon menjadikan setiap model unik dan memperkuat stabilitas hasil prediksi. Kombinasi antara kesederhanaan, fleksibilitas, dan akurasi menjadikan *Random Forest* sebagai pilihan yang tepat untuk membangun model prediksi konsumsi listrik.

2.3 Evaluasi Model Prediksi dengan R^2

Evaluasi kinerja model prediksi adalah hal yang penting untuk memahami seberapa efektif model dalam menggambarkan hubungan antara fitur dan target dalam sebuah dataset. Salah satu cara yang umum digunakan untuk evaluasi model regresi adalah dengan menggunakan *R-squared* (R^2), atau disebut juga koefisien determinasi. Metrik ini mengukur seberapa besar proporsi variasi data aktual yang bisa dijelaskan oleh model. Nilai R^2 berkisar antara 0 hingga 1, di mana semakin mendekati 1 berarti model semakin baik dalam menjelaskan data [18]. R^2 dapat dihitung sebagai berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{\sum_{i=1}^n |y_i - \bar{y}|^2} \quad (1)$$

di mana y_i merupakan nilai *actual*, $\hat{y}_i$ merupakan nilai prediksi dari model, $\bar{y}$ merupakan rata-rata nilai *actual*, dan n merupakan jumlah data.

Secara sederhana, jika nilai prediksi dari model sangat dekat dengan nilai sebenarnya, maka kesalahan kuadrat yang dihasilkan akan kecil, dan nilai R^2 akan tinggi. Dalam konteks konsumsi energi, nilai R^2 yang tinggi menandakan bahwa model sudah mampu menangkap tren atau pola konsumsi dengan sangat baik.

2.4 Evaluasi Model Prediksi dengan MAE

Selain R^2 , ada juga metrik lain yang tidak kalah penting yaitu *Mean Absolute Error* (MAE). MAE mengukur rata-rata dari selisih absolut antara nilai prediksi dan nilai sebenarnya. Kelebihan MAE adalah hasil evaluasinya berada dalam satuan yang sama dengan target prediksi, misalnya kWh dalam kasus konsumsi energi, sehingga lebih mudah dipahami dan diinterpretasikan [16] [18]. MAE dapat dihitung dengan cara berikut :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2)$$

di mana n adalah jumlah total observasi, y_i adalah nilai aktual (observasi ke- i), dan $\hat{y}_i$ adalah nilai prediksi untuk observasi ke- i .

MAE cenderung lebih tahan terhadap nilai-nilai ekstrem dibandingkan MSE karena tidak mengkuadratkan nilai kesalahan. Oleh karena itu, MAE sangat cocok digunakan jika kita ingin mengetahui seberapa besar rata-rata kesalahan prediksi yang dilakukan oleh model secara umum. Semakin kecil nilai MAE, semakin baik performa model prediksi yang digunakan.

2.5 Hubungan Kelompok Tarif dengan Konsumsi Listrik

Dalam sektor energi, tarif listrik diberlakukan berbeda-beda sesuai dengan kategori pelanggan seperti rumah tangga, bisnis, industri, sosial, dan pemerintahan. Setiap kategori memiliki karakteristik konsumsi yang unik, sehingga penting untuk memahami bagaimana struktur tarif ini berhubungan dengan jumlah konsumsi listrik. Dengan mengkaji hubungan ini, PLN dapat merancang strategi distribusi yang lebih efisien dan kebijakan tarif yang lebih adil.

Dalam penelitian ini, variabel “TARIF” dikodekan menggunakan teknik *one-hot encoding* agar dapat dimanfaatkan dalam model *machine learning*. Data waktu dalam bentuk “Tahun” juga digunakan untuk menangkap perubahan tren dari tahun ke tahun. Sebelum data dilatih pada model, nilai target “*Total Consumption*” terlebih dahulu ditransformasi menggunakan logaritma natural untuk meredam pengaruh nilai ekstrem yang dapat mengganggu kinerja model.

Setelah model dilatih dan diuji, evaluasi menggunakan MAE dan R^2 menunjukkan bahwa prediksi model cukup akurat. Model kemudian digunakan untuk memproyeksikan konsumsi energi di tahun 2025 berdasarkan masing-masing kelompok tarif. Hasilnya menunjukkan bahwa tarif rumah tangga, khususnya R1 dengan daya 450 VA, memiliki konsumsi tertinggi. Analisis ini penting untuk menyusun kebijakan seperti subsidi energi, pengendalian beban, dan pengembangan infrastruktur yang lebih terfokus pada kelompok pengguna terbesar.

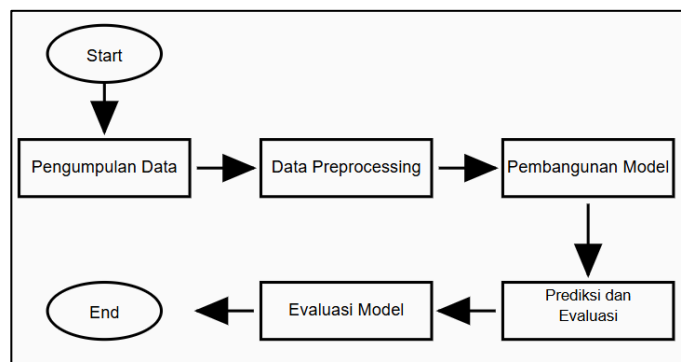
3. METODE PENGABDIAN MASYARAKAT

3.1 Metode Pengabdian

Kegiatan pengabdian kepada masyarakat ini dilaksanakan melalui penerapan teknologi pembelajaran mesin untuk mengembangkan model yang memprediksi penggunaan listrik dengan menggunakan algoritma *Random Forest* di PT PLN (Persero) Unit Induk Wilayah (UIW) Nusa Tenggara Barat (NTB). Tujuan utama dari program ini adalah untuk membantu PLN dalam meningkatkan efisiensi dan ketepatan prediksi kebutuhan energi listrik yang didasarkan pada data konsumsi pelanggan di masa lalu yang telah dikelompokkan berdasarkan jenis tarif. Dengan penerapan model prediktif ini, diharapkan proses perencanaan infrastruktur kelistrikan dapat lebih tepat sasaran dan peka terhadap kebutuhan masyarakat.

Beberapa kegiatan yang dilakukan selama pelaksanaan pengabdian antara lain:

1. Melakukan diskusi dengan pembimbing lapangan mengenai rancangan model prediksi konsumsi listrik yang akan dibuat.
2. Mengambil data konsumsi listrik.
3. Bekerja sama dengan tim untuk membersihkan data.
4. Menyaring data konsumsi listrik berdasarkan kelompok tarif sesuai format data.



Gambar 1 Metode pengabdian

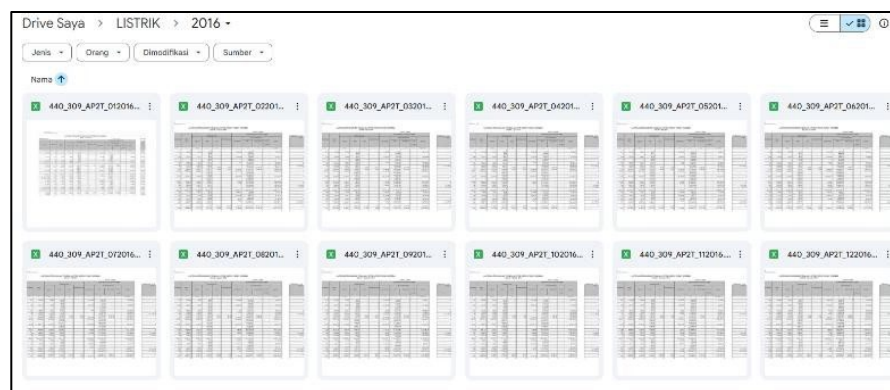
Pada Gambar 1, proses dimulai dengan mengumpulkan data dari tahun 2016 sampai 2024, kemudian dilanjutkan dengan membersihkan data untuk menghilangkan data yang tidak relevan dan data yang tidak digunakan dalam proses prediksi konsumsi listrik. Setelah itu, dilakukan *preprocessing data*, termasuk transformasi logaritma, *one-hot encoding*, dan normalisasi. Selanjutnya, model dibangun dengan menggunakan *Random Forest* dan dioptimalkan melalui penyesuaian parameter. Begitu model selesai dibuat, prediksi konsumsi listrik untuk tahun 2025 dilakukan, kemudian divisualisasikan dan dievaluasi dengan metrik R^2 dan MAE.

3.2 Pengumpulan Data

Untuk memperoleh data historis konsumsi listrik pelanggan PLN UIW NTB periode 2016–2024, langkah pertama yang dilakukan adalah mengajukan permohonan resmi kepada pihak PLN melalui divisi terkait, seperti Perencanaan atau Niaga, dengan melampirkan surat pengantar serta penjelasan tujuan penelitian. Permohonan ini mencakup detail jenis data yang dibutuhkan, antara lain catatan konsumsi listrik pelanggan per kategori tarif, jumlah pelanggan, serta data beban puncak. Setelah akses diberikan, proses pengumpulan data diperoleh dari sistem internal PLN seperti *Customer Information System* (CIS) dan laporan operasional. Data yang terkumpul disimpan dalam format Excel dan dalam struktur folder berdasarkan tahun, di mana masing-masing folder berisi 12 file bulanan yang mencerminkan data konsumsi listrik setiap bulannya. Setiap *file* Excel memiliki *sheet* “JUMLAH” yang berisi data agregasi berdasarkan kelompok tarif untuk bulan tersebut.



Gambar 2 Dokumentasi pengumpulan data



Gambar 3 Arsip laporan excel 2016-2025

Gambar 3 menampilkan isi dari salah satu folder, yaitu folder tahun 2016, yang berisi 12 *file* Excel. Masing-masing *file* merepresentasikan data konsumsi listrik untuk satu bulan. Ini menjadi representasi umum untuk semua tahun, sehingga jumlah *file* yang dikumpulkan secara keseluruhan mencapai 108 *file* dari 9 tahun.

3.3 Preprocessing Data

PEMBERSIHAN : PUSAT		STATUS : CUCI									
TARIF	PELANGGAN	DAYA (VA)	PENYALURAN KWH				RUMAH PENJUALAN TENAGA LISTRIK				
			JUMLAH	LRSP	WSP	KELUAR KWH	BAKAL BEBAN	KWH	KEL. KWH	TLR	JUMLAH
1	2	3	4	5	6	7	8	9	10	11	12=4+5+6+7+8+9+10+11
S 1 / 220 VA											
S 2 / 450 VA I	6.305	3.818.150	448.187	183.716	0	0	33.363.360	35.802.800	0		138.561.833
0	0	0	0	81.477	0	0	0	24.241.405	0		0
0	0	0	0	152.081	0	0	0	35.802.800	0		0
S 2 / 900 VA	6.351	5.882.800	513.550	268.712	0	0	33.667.650	39.626.321	0		226.510.203
0	0	0	0	51.886	0	0	0	22.682.255	0		0
0	0	0	0	132.267	0	0	0	8.246.128	0		0
S 2 / 1.300 VA	4.243	5.010.800	541.872	541.872	0	0	640.360	393.776.350	0		380.443.056
S 2 / 2.200 VA	960	2.173.600	258.364	258.364	0	0	0	45.970.136	0		185.870.136
S 2 / 3.500 VA s/d 200	1.160	19.266.400	1.658.527	1.658.527	0	0	0	1.437.346.251	0		1.973.366.371
S 3 / 200 KVA	12	7.262.000	760.760	626.120	10.640	13.321	0	637.864.658	12.687.825		670.652.473
0	18.687	18.180.800	2.371.327	2.288.620	10.640	13.321	37.776.360	2.222.107.670	12.687.825		2.294.516.814
0	383.373	179.446.715	26.263.433	1.775.530	10.640	13.321	1.271.627.835	4.578.533.358	12.687.825		17.146.100.877
0	0	0	0	5.636.415	0	0	0	2.033.532.464	0		0
0	0	0	0	8.665.477	0	0	0	4.884.337.335	0		0
R 1 / 450 VA	468.368	181.550.100	38.634.434	26.554.875	0	0	1.360.246.370	10.515.546.315	0		21.287.825.641
0	0	0	0	2.714.688	0	0	0	1.253.887.411	0		0
0	0	0	0	6.379.660	0	0	0	3.786.433.345	0		0
R 1 / 900 VA	86.247	127.321.100	13.628.687	13.628.687	0	0	3.023.680	17.873.468.776	0		17.873.468.776
0	14.051	19.266.400	1.658.527	1.658.527	0	0	1.863.171	2.973.408.528	0		2.513.472.658
R 3 / 1.500 VA s/d 5.500	4.258	17.862.400	2.031.872	2.031.872	0	0	0	2.945.382.052	0		2.945.382.052
S 2 / 6.600 VA ke atas I	772	8.626.800	1.108.872	1.108.872	0	0	0	1.578.862.312	0		1.578.862.312
0	0	0	0	0	0	0	0	0	0		0
JUMLAH R	868.633	760.621.350	81.545.641	81.545.641	0	0	2.610.812.187	66.787.684.387	0		67.408.427.384
0	2.070	1.388.000	160.861	82.121	0	0	17.648.252	37.847.758	0		86.536.371
0	0	0	0	107.565	0	0	0	12.761.258	0		0
R 1 / 900 VA	4.783	4.313.700	518.463	375.686	0	0	48.736.786	98.560.177	0		316.654.833
0	0	0	0	143.552	0	0	0	85.357.325	0		0
R 1 / 1.300 VA	8.783	10.631.400	1.159.242	1.159.242	0	0	1.202.560	1.118.878.268	0		1.120.284.831
R 1 / 2.200 VA s/d 5.500	11.265	31.482.800	2.866.778	2.866.778	0	0	8.118.000	8.288.832.215	0		8.608.644.215
S 2 / 6.600 VA ke atas I	2.647	763.762.400	1.108.872	1.108.872	0	0	0	20.353.788.180	0		20.353.788.180
0	0	0	0	0	0	0	0	0	0		0
S 3 / > 200 KVA	42	18.210.000	6.510.475	4.678.708	1.125.066	137.331	0	6.838.556.468	577.416.800		7.556.980.311
JUMLAH R	33.163	183.982.600	26.270.644	25.121.360	1.125.066	137.331	37.718.613	33.555.467.558	577.416.800		33.653.687.008
0	1	450	118	118	0	0	0	57.580	0		57.580
0	0	0	0	0	0	0	0	0	0		0
R 1 / 900 VA	1	900	118	72	0	0	28.352	22.668	0		66.230
0	0	0	0	0	0	0	0	15.986	0		0
R 1 / 1.300 VA	0	0	0	0	0	0	0	444.200	0		444.200
0	0	0	0	0	0	0	0	146.808	0		146.808
S 1 / 2.500 VA ke atas I	37	18.400	778	778	0	0	0	38.867.696	0		38.867.696
12.171 VA ke atas 200 KVA	1.603	18.836.000	2.383.187	2.684.761	268.496	148.877	0	2.156.633.660	155.248.868		2.611.632.828
S 2 / > 200 KVA	26	18.275.000	6.186.271	3.545.181	555.180	66.380	0	4.418.052.416	72.886.683		4.490.156.109
S 1 / 30.000 KVA ke atas I	3	0	0	0	0	0	0	0	0		0
JUMLAH	230	33.841.050	6.510.066	5.671.468	843.586	213.053	28.350	6.814.808.365	227.348.660		7.142.285.273
0	1.150	275	127	127	0	0	1.872.000	3.726.200	0		21.857.837
R 1 / 900 VA	681	922.400	37.400	37.400	0	0	7.348.052	37.621.214	0		45.667.266
R 1 / 1.300 VA	656	1.112.800	87.732	87.732	0	0	0	101.886.608	0		101.886.608
R 1 / 2.200 VA s/d 5.500	760	2.866.400	328.261	328.261	0	0	0	333.736.253	0		333.736.253
R 1 / 6.600 VA ke atas I	875	11.682.400	1.854.058	1.854.058	0	0	0	2.768.223.515	0		2.768.223.515
0	0	0	0	0	0	0	0	0	0		0
R 2 / > 200 KVA	0	3.845.000	358.700	268.062	50.000	8.000	0	388.175.581	6.684.505		394.260.086
JUMLAH R	1.986	13.687.600	4.770.811	4.770.811	0	0	0	6.788.264.683	6.788.264.683		8.758.264.683
0	5.640	27.871.812	7.600.400	7.552.801	66.600	8.000	11.821.463	46.278.127.481	6.684.505		10.783.152.008
0	0	0	0	0	0	0	0	0	0		0
0	0	0	0	0	0	0	0	0	0		0
0	19.255	182.350	182.350	182.350	0	0	0	271.136.656	0		271.136.656
JUMLAH	7.026.269	1.083.282.012	128.478.008	127.332.568	2.140.030	768.851	2.737.418.260	112.639.222.353	103.545.616		110.165.166.471

Gambar 4 Tampilan data konsumsi listrik dalam satu bulan

Pembersihan data dilakukan untuk menghapus data yang tidak digunakan dan meningkatkan kualitas data. Pembersihan data ini sangat penting agar model yang dibuat tidak condong pada satu sisi dan dapat belajar dari data yang valid. Dari data tersebut, yang akan digunakan dalam proses prediksi adalah kolom tarif dan kolom jumlah saja, sehingga akan dibuat *sheet* baru yang hanya berisi kolom tarif dan kolom jumlah.

TABEL 1. HASIL PEMBERSIHAN DATA DALAM SATU BULAN

Tarif	Jumlah
S 1 / 220 VA	0.00
S 2 / 450 VA I	440.20
S 2 / 900 VA I	513.57
S 2 / 1.300 VA	543.69
S 2 / 2.200 VA	256.31
S 2 / 3.500 VA s/d 200	1.859.527
S 3 / 200 KVA	760.76
R 1 / 450 VA I	29.260.422
R 1 / 900 VA I	35.634.134
R 1 / 1.300 VA	12.629.697
R 1 / 2.200 VA	3.572.549
R 2 / 3.500 VA s/d 5.500	2.034.670
R 3 / 6.600 VA ke atas I	1.108.872
B 1 / 450 VA I	154.96
B 1 / 900 VA I	519.46
B 1 / 1.300 VA	1.159.242
B 1 / 2.200 VA s/d 5.500	3.994.475
B 2 / 6.600 VA s/d 200 k	14.389.407
B 3 / > 200 KVA	6.012.475
I 1 / 450 VA I	119.00
I 1 / 900 VA I	110.00
I 1 / 1.300 VA	478.00

Tarif	Jumlah
I 1 / 2.200 VA	779.00
I 1 / 3.500 s/d 14 KVA	35.04
I 2 / > 14 KVA s/d 200 KV	2.383.167
I 3 / > 200 KVA	4.095.374
I 4 / 30.000 KVA ke atas	0.00
P 1 / 450 VA	36.37
P 1 / 900 VA	57.40
P 1 / 1.300 VA	97.73
P 1 / 2.200 VA s/d 5.500	326.20
P 1 / 6.600 VA s/d 200 k	1.954.059
P 2 / > 200 KVA	359.70
P 3	4.770.941
T / > 200 KVA	0.00
C / TM > 200 KVA	0.00
L	168.39

Tabel di atas merupakan hasil pembersihan data, di mana data yang digunakan adalah hanya kolom tarif dan kolom jumlah yang diletakkan pada *sheet* baru bernama “JUMLAH” untuk memudahkan proses prediksi menggunakan *machine learning*.

Tahapan selanjutnya setiap *file* Januari – Desember diproses dengan cara membaca *sheet* bertajuk “JUMLAH”, membersihkan data, menghapus nilai tarif yang tidak valid, mengubah format angka, dan menambah informasi tahun. Data kemudian dikategorikan berdasarkan tahun dan tarif, digabungkan menjadi satu *DataFrame* yang disebut “final_df”, dan disimpan dalam *file* Excel baru sebagai hasil akhir dari pengumpulan dan pembersihan data. Berikut hasil pengolahannya.

TABEL II. HASIL TOTAL PERSIAPAN DATA JANUARI – DESEMBER 2016

Year	Tarif	Jumlah
2016	II	0
2016	III	-
2016	B 1 / 1.300 VA	10,660,517
2016	B 1 / 2.200 VA s/d 5.500	36,265,775
2016	B 1 / 450 VA I	1,759,684
2016	B 1 / 900 VA I	4,803,827
2016	B 2 / 6.600 VA s/d 200 k	13,431,172
2016	B 3 / > 200 KVA	61,064,312
2016	C / TM > 200 KVA	-
2016	I 1 / 1.300 VA	7,473
2016	I 1 / 2.200 VA	9,409
2016	I 1 / 3.500 s/d 14 KVA	316,142
2016	I 1 / 450 VA I	895
2016	I 1 / 900 VA I	1,216
2016	I 2 / > 14 KVA s/d 200 KV	23,716,505
2016	I 3 / > 200 KVA	33,501,400
2016	I 4 / 30.000 KVA ke atas	-
2016	JUMLAH I	14,314,469
2016	JUMLAH	25,224,458
2016	L	1,743,379
2016	P 1 / 1.300 VA	2,406,305
2016	P 1 / 2.200 VA s/d 5.500	3,040,623
2016	P 1 / 450 VA	341,039
2016	P 1 / 6.600 VA s/d 200 k	18,555,968
2016	P 1 / 900 VA	1,929,648
2016	P 2 / > 200 KVA	3,385,387
2016	P 3	45,563,221

Year	Tarif	Jumlah
2016	R 1 / 1.300 VA	11,755,439
2016	R 1 / 2.200 VA	34,522,651

Setelah data terkumpul, dilakukan *preprocessing* menggunakan bahasa pemrograman *Python*. Proses ini meliputi pembacaan semua *file* dari *sheet* “JUMLAH”, penghapusan baris kosong atau tidak valid, penghilangan entri dengan nilai TARIF yang tidak digunakan seperti “I”, “II”, dan “III”, serta konversi format angka agar seragam dan siap diolah. Setelah itu, seluruh data dari semua *file* digabungkan ke dalam satu data *frame* utama agar bisa dianalisis secara menyeluruh.

Selanjutnya, dilakukan transformasi pada variabel target (*Total Consumption*) dengan logaritma natural menggunakan fungsi $\log_{1p}$ untuk mengurangi pengaruh outlier. Variabel kategorikal seperti “TARIF” juga diubah menjadi bentuk numerik melalui *one-hot encoding* agar bisa digunakan sebagai *input* ke dalam model. Semua fitur numerik kemudian dinormalisasi menggunakan “StandardScaler” untuk memastikan bahwa data berada dalam skala yang sebanding, yang bertujuan untuk meningkatkan performa model.

Year	B.1 / 1.300 VA	B.1 / 2.200 VA S/D 5.500 VA	B.1 / 2.200 VA S/D 5.500 VA	B.1 / 450 VA I	B.1 / 900 VA I	B.2 / 6.600 VA S/D 200 KVA I	B.2 / 6.600 VA S/D 200 KVA I	B.3 / 30.000 KVA	B.3 / > 200 KVA	B.3 / > 200 KVA	C / TR > 200 KVA S.D
0 2016	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1 2016	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2 2016	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3 2016	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4 2016	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

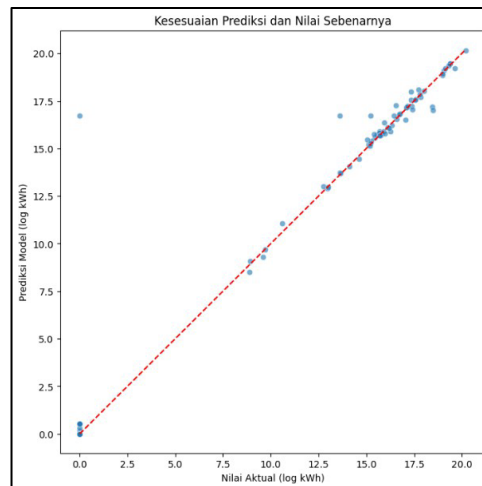
Gambar 5 *Preprocessing* data

Gambar 5 merupakan contoh dilakukannya *One-Hot Encoding*. *One-Hot Encoding* dilakukan untuk mengubah format data kategori tarif yang sebelumnya berupa teks menjadi format numerik biner agar dapat digunakan oleh *machine learning*.

3.4 Pembangunan Model

Model prediksi dibangun menggunakan algoritma *Random Forest Regressor* yang diimplementasikan melalui *library Scikit-learn* di *Python*. Data yang telah diproses dibagi menjadi dua bagian, yaitu data pelatihan sebesar 80% dan data pengujian sebesar 20%. Proses pelatihan dilakukan untuk membangun model berdasarkan data historis, sedangkan pengujian dilakukan untuk mengevaluasi performa model terhadap data yang belum pernah dilihat sebelumnya.

Untuk meningkatkan akurasi prediksi, dilakukan pencarian kombinasi parameter terbaik melalui proses *hyperparameter tuning* menggunakan *Grid Search* dengan *Cross Validation* sebanyak 5 *fold*. Parameter yang diatur mencakup jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), jumlah minimum sampel untuk split dan daun, serta metode pemilihan fitur terbaik.

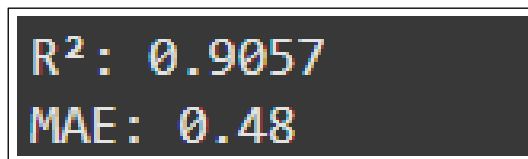


Gambar 6 Kesesuaian prediksi model dengan nilai sebenarnya

Hasil dari pelatihan mengindikasikan bahwa model dapat memahami pola penggunaan listrik berdasarkan kelompok tarif dengan sangat baik. Jika titik biru mendekati garis merah menunjukkan semakin sedikit selisih antara nilai asli dengan prediksi konsumsi Listrik

3.5 Evaluasi Model

Tahap evaluasi dilakukan untuk mengukur kinerja model *Random Forest* yang telah dikembangkan dengan memanfaatkan data uji. Dua metrik yang digunakan untuk analisis adalah *R-squared* (R^2) dan *Mean Absolute Error* (MAE). Nilai R^2 digunakan untuk menilai seberapa baik model mampu menjelaskan variasi dalam data, sedangkan MAE mencerminkan rata-rata kesalahan absolut antara nilai yang diprediksi dan nilai sebenarnya.



Gambar 7 Hasil evaluasi model

Evaluasi model dilakukan menggunakan dua metrik utama, yaitu *R-squared* (R^2) dan *Mean Absolute Error* (MAE). Nilai R^2 yang diperoleh sebesar 0.9057 menunjukkan bahwa model dapat menjelaskan sekitar 90.57% variasi dalam data konsumsi listrik, yang berarti model memiliki performa yang sangat baik dalam mengenali pola dari data historis. Sementara itu, nilai MAE sebesar 0.48 menunjukkan bahwa rata-rata kesalahan prediksi berada pada angka yang sangat rendah (dalam satuan logaritmik), yang memperkuat bukti bahwa model memiliki akurasi tinggi.

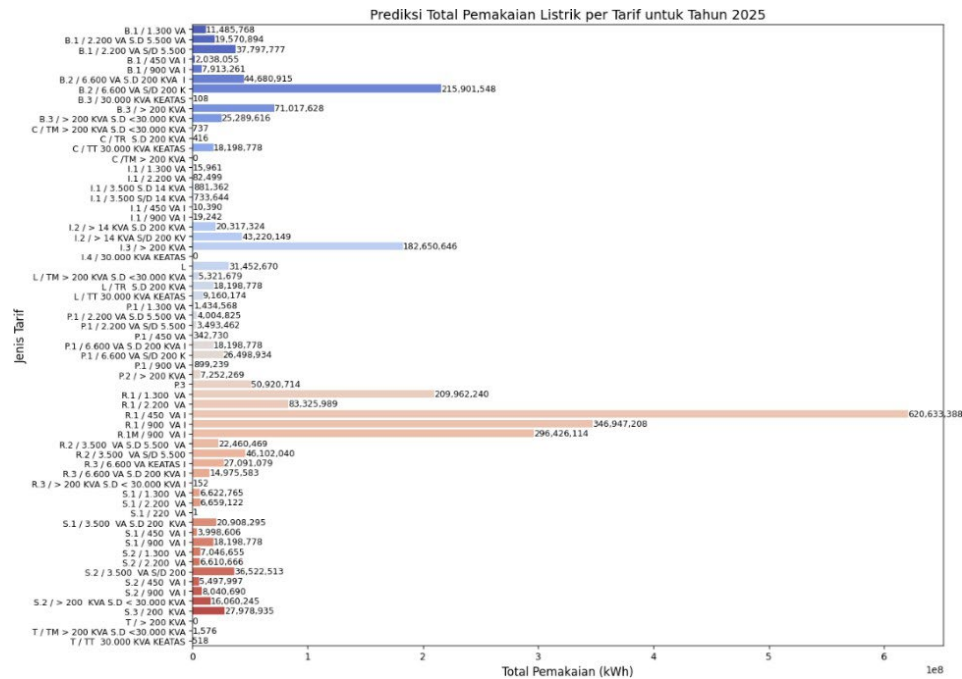
4. HASIL DAN PEMBAHASAN

4.1 Hasil dan Visualisasi Prediksi Konsumsi Listrik 2025 Berdasarkan Kelompok Tarif

Total Prediksi Konsumsi Listrik Tahun 2025: 2,916,380,608 kWh

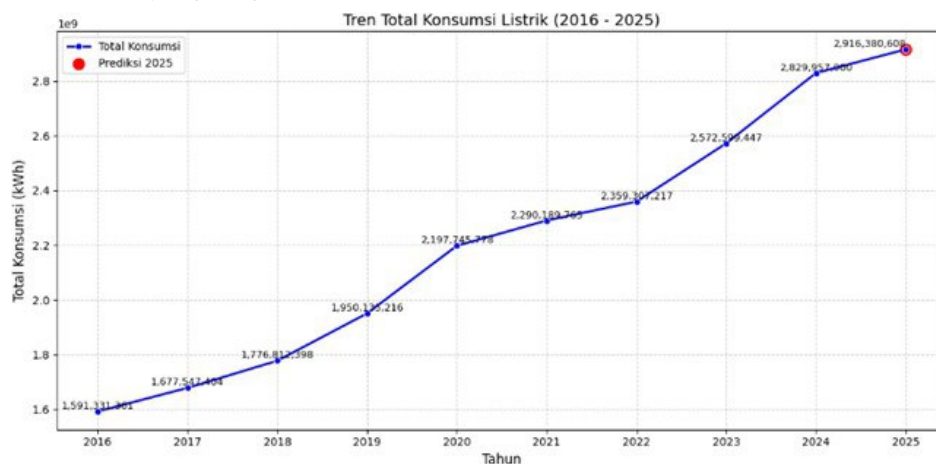
Gambar 8 Hasil prediksi konsumsi listrik per tarif tahun 2025

Model yang telah dilatih kemudian digunakan untuk melakukan prediksi konsumsi listrik di tahun 2025 berdasarkan masing-masing kelompok tarif. Total prediksi konsumsi listrik untuk semua kategori tarif di tahun 2025 diperkirakan mencapai 2,916,569,738 kWh (dua miliar sembilan ratus enam belas juta lima ratus enam puluh sembilan ribu tujuh ratus tiga puluh delapan *kilowatt-hour*). Hasil prediksi ditampilkan dalam bentuk tabel dan grafik visualisasi agar lebih mudah dipahami oleh pihak terkait dalam pengambilan keputusan.



Gambar 9 Visualisasi prediksi konsumsi listrik per tarif tahun 2025

Gambar di atas memperlihatkan grafik batang hasil prediksi konsumsi listrik berdasarkan kelompok tarif pada tahun 2025. Kelompok tarif R1 450 VA menunjukkan nilai tertinggi dalam konsumsi listrik, sementara beberapa kategori lain memiliki nilai yang sangat rendah atau bahkan nol.



Gambar 10 Visualisasi Tren Total Konsumsi Listrik Tahun 2016-2025

Gambar 10 menampilkan tren konsumsi listrik dari tahun 2016 hingga 2025. Grafik ini menggambarkan pola peningkatan konsumsi dari tahun ke tahun, yang menjadi dasar penting dalam perencanaan kebutuhan energi jangka panjang oleh PLN.

4.2 Analisis Hasil

Model prediksi yang dibuat dengan menggunakan algoritma *Random Forest* menunjukkan hasil yang sangat memuaskan, dengan nilai koefisien determinasi (R^2) mencapai 0,9057 dan nilai *Mean Absolute Error* (MAE) sebesar 0,48. Ini berarti model tersebut mampu menjelaskan lebih dari 90% variasi dalam data historis konsumsi listrik, dengan tingkat kesalahan yang cukup kecil. Tingkat akurasi ini membuat model tersebut cukup bisa diandalkan untuk perencanaan distribusi listrik, baik untuk jangka waktu pendek maupun menengah.

Dari hasil ramalan untuk tahun 2025, diprediksi bahwa kelompok rumah tangga dengan daya 450 VA akan menjadi konsumen terbesar dengan total konsumsi sekitar 620 juta kWh. Hal ini menunjukkan banyaknya pelanggan rumah tangga di wilayah NTB, sehingga perlu perhatian lebih terhadap efisiensi dan subsidi bagi kelompok ini.

Selain itu, sektor industri dan bisnis juga berkontribusi signifikan dengan konsumsi masing-masing sekitar 182 juta kWh dan 215 juta kWh.

Tren konsumsi listrik antara tahun 2016 hingga ramalan tahun 2025 menunjukkan rata-rata kenaikan sekitar 4–5% per tahun, dengan lonjakan tajam hingga 12% pada tahun 2020 saat pandemi COVID-19. Kenaikan ini menunjukkan bahwa perubahan dalam perilaku masyarakat dan kondisi sosial memiliki pengaruh besar terhadap permintaan energi. Pada penelitian John Garcia [19], juga mengungkapkan bahwa sektor residensial meningkat signifikan elektrifikasinya (+ ~16-17%) selama pandemi, sementara sektor lain seperti komersial atau industri mengalami penurunan tergantung kebijakan *lockdown*. Ini mendukung ide bahwa kondisi sosial / perilaku masyarakat sangat mempengaruhi permintaan energi. Ini menunjukkan adanya pertumbuhan ekonomi yang perlu didukung dengan pasokan energi yang cukup dan stabil. Diperkirakan total konsumsi listrik pada tahun 2025 akan mencapai sekitar 2,92 miliar kWh. Hal ini menjadi tantangan dalam menjaga keseimbangan antara permintaan dan pasokan energi di masa depan.

Pada rencana usaha penyediaan tenaga listrik (RUPTL) PLN 2025-2034, PLN sudah memproyeksikan bahwa penjualan listrik nasional akan mencapai 511 TWh pada tahun 2034, naik dari 306 TWh secara nasional dari tahun 2024 [20]. Berdasarkan kebutuhan tersebut, beberapa rekomendasi dari hasil prediksi bisa diusulkan. Pertama, diperlukan peningkatan kapasitas dan keandalan sistem distribusi listrik, terutama di daerah yang mengalami pertumbuhan jumlah pelanggan rumah tangga. Kedua, perluasan program pendidikan dan efisiensi energi bagi pelanggan bersubsidi agar pemakaian energi lebih efisien. Terakhir, kerja sama dengan sektor industri untuk menerapkan sistem manajemen konsumsi yang berbasis waktu bisa membantu menyeimbangkan beban dan mencegah lonjakan konsumsi pada waktu-waktu tertentu.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kegiatan pengabdian yang dilaksanakan di PT PLN (Persero) UIW Nusa Tenggara Barat berhasil menerapkan model prediksi konsumsi energi listrik dengan menggunakan algoritma *Random Forest*. Model ini dibuat berdasarkan data konsumsi listrik dari tahun 2016 hingga 2024 dengan jumlah *file* yang dikumpulkan secara keseluruhan mencapai 108 *file* selama 9 tahun. Hasil dari model menunjukkan kinerja yang sangat baik dengan nilai R^2 sebesar 0,9057 dan MAE sebesar 0,48. Berdasarkan prediksi diperkirakan bahwa konsumsi listrik di tahun 2025 akan mencapai sekitar 2,91 miliar kWh, dengan kelompok pelanggan rumah tangga yang memiliki daya 450 VA sebagai pengguna terbesar.

5.2 Saran

Adapun saran yang dapat diberikan agar prediksi menjadi lebih tepat dan sesuai, disarankan bagi PLN untuk secara teratur memperbarui informasi mengenai penggunaan listrik. Selain itu, peningkatan kemampuan dalam menganalisis data serta menggunakan algoritma prediksi yang didasarkan pada teknologi *machine learning* sangat penting untuk mendukung pengambilan keputusan yang efisien dan tepat di masa yang akan datang.

UCAPAN TERIMA KASIH

Penulis ingin menyampaikan rasa terima kasih kepada PT PLN (Persero) Unit Induk Wilayah Nusa Tenggara Barat atas peluang dan bantuan yang diberikan sepanjang pelaksanaan kegiatan pengabdian hingga penyusunan artikel ini. Ucapan terima kasih juga ditujukan semua pihak yang telah berkontribusi melalui arahan ilmiah dan saran teknis yang sangat membantu dalam perencanaan dan pelaksanaan kegiatan ini.

DAFTAR PUSTAKA

- [1] R. H. Nugraha, Yuwono, P. L., Y. B. Arief and a. H. Patria, "Analisis konsumsi energi listrik pelanggan dan biaya pokok produksi penyediaan energi listrik dengan machine learning," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, Vols. vol. 6, no. 1, 2022.
- [2] F. Ericson and A. H. Saputra, "Optimasi skema transportasi LNG untuk pembangkit listrik di Wilayah Nusa Tenggara," Fakultas Teknik Universitas Indonesia, Jakarta, 2018.
- [3] Y.-T. Chen, J. Eduardo Piedad and K. Cheng-Chien, "Energy Consumption Load Forecasting Using a Level- Based Random Forest Classifier," *Symmetry*, vol. 11, no. 8, 2019.

- [4] geeksforgeek, "Gradient Boosting vs Random Forest," 9 4 2024. [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/gradient-boosting-vs-random-forest>. [Accessed 15 5 2025].
- [5] K. Olcay, S. G. Tunca and M. A. Ozgur, "Forecasting and Performance Analysis of Energy Production in Solar Power Plants Using Long Short-Term Memory (LSTM) and Random Forest Models," *IEEE Access*, vol. 12, p. 103299 – 103312, 2024.
- [6] Y. Liu, H. Zhang and a. X. Chen, "A Hybrid Short-Term Load Forecasting Model Based on Improved Fuzzy C-means Clustering, Randonom Forest and Deep Neural Networks," *Int. J. Electr. Power Energy Syst*, vol. 125, 2021.
- [7] G. Dudek, "A comprehensive study of Random Forest for Short-Term Load Forecasting," *Energies*, vol. 15, no. 20, 2022.
- [8] K. M. Salih, A. A. Mohammed and a. M. Yusof, "A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME," *Appl. Energy*, vol. 345, 2023.
- [9] S. Sun, X. Huang and a. Y. Chen, "Enhanced Random Forest Model for Robust Short-Term Photovoltaic Power Forecasting Using Weather Measurements," *Energies*, vol. 14, 2021.
- [10] R. G. Wardhana, G. Wang and a. F. Sibuea, "Penerapan machine learning dalam prediksi tingkat kasus penyakit di Indonesia," *Journal of Information System Management*, vol. 5, no. 1, 2023.
- [11] C. Bianca, R. Chifu, C. Cordea, E. S. Chifu and O. Barsan, "Forecasting the short-term energy consumption using Random Forests and Gradient Boosting," *Energy*, vol. 239, 2022.
- [12] C. Cui, X. Zhang, W. Ma and a. X. Peng, "Advanced short-term load forecasting with XGBoost-RF feature selection and CNN-GRU," *Processes*, vol. 12, no. 11, 2024.
- [13] Singh, S. Sharma and a. R., "Hybrid MLP–Random Forest model for electrical consumption prediction: A case study in technical campus," *Int. J. Energy Res.*, vol. 42, no. 12, 2018.
- [14] Lee, S.-M. Lundberg and a. S.-I., "A Unified Approach to Interpreting Model Predictions," in *Neural Information Processing Systems*, 2017.
- [15] T. Clement, H. T. T. Nguyen, N. Kemmerzell, M. Abdelaal and a. D. Stjelja, "Beyond Explaining: XAI-based Adaptive Learning with SHAP Clustering for Energy Consumption Prediction," *arXiv*, p. 2, 2024.
- [16] G. A. M. Ashfania, T. Prahasto, A. Widodo, T. Warsokusumo and J. Soedarto, "Penggunaan algoritma Random Forest untuk klasifikasi berbasis kinerja efisiensi energi pada sistem pembangkit daya," *Jurnal Teknik Elektro dan Komputer*, vol. 24, no. 3, 2022.
- [17] P. B. J. P. M. d. R. C. a. S. M. B. Magalhães, "Short-term load forecasting based on optimized Random Forest and optimal feature selection," *Energies*, vol. 17, no. 8, 2024.
- [18] M. Ardian, S. Khomsah and a. R. Pandiya, "Perbandingan model regresi untuk memprediksi harga jual," *Jurnal Teknologi Informasi dan Komputer*, vol. 8, no. 2, 2023.
- [19] J. Garcia-Rendon, F. R. Londoño, L. J. A. Restrepo and S. B. Correa, "Sectoral analysis of electricity consumption during the COVID-19 pandemic: Evidence for unregulated and regulated markets in Colombia," *Energy*, vol. 268, 2023.
- [20] S. C. Arini, "PLN Bidik Penjualan Listrik Naik Jadi 511 TWh hingga 2034," *detikFinance*, 2 6 2025. [Online]. Available: <https://finance.detik.com/energi/d-7944061/pln-bidik-penjualan-listrik-naik-jadi-511-twh-hingga-2034>. [Accessed 5 7 2025].