

OPTIMISATION OF VARIABLE COMBINATIONS FOR HOUSEHOLD ELECTRICITY CONSUMPTION PREDICTION USING A MULTIVARIATE TIME SERIES MACHINE LEARNING APPROACH

OPTIMASI KOMBINASI VARIABEL UNTUK PREDIKSI KONSUMSI LISTRIK RUMAH TANGGA DENGAN PENDEKATAN MACHINE LEARNING MULTIVARIATE TIME SERIES

Akhmad Faeda Insani^{1*}, Ahmad Mushawir², Zainuddin³, Aditya Adiaksa⁴, Sparisoma Viridi⁵
Program Studi Sains Komputasi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi
Bandung, Indonesia
Email: faedainsani92@gmail.com

Abstract - Accurate household electricity consumption prediction is vital for effective energy planning in Indonesia, a nation facing rapid economic growth and technological advancements. Inaccurate predictions can lead to inefficiencies in resource allocation and energy shortages. Traditional methods like ARIMA struggle with non-linear patterns, long-term dependencies, and multivariate relationships critical in understanding electricity consumption dynamics. To address these challenges, this study employs the Long Short-Term Memory (LSTM) algorithm with a multivariate time series approach, chosen for its ability to capture complex patterns and long-term trends. The dataset comprises monthly electricity consumption data (2004–2023) from PT PLN, enriched with macroeconomic and environmental variables like Household Consumption GDP, inflation, and average temperature. The Denton-Chollete method was used to transform quarterly GDP data into monthly intervals, and correlation analysis identified Household Consumption GDP ($r=0.98$) and Power Contract Additions ($r=0.64$) as significant predictors. Testing 63 feature combinations, the best (Power Contract Additions, Household Consumption GDP, and Household Electricity Consumption) achieved a Mean Absolute Percentage Error (MAPE) of 3.54%. These results highlight LSTM's superiority in handling dynamic and complex electricity consumption patterns and provide a robust predictive tool for PT PLN. This study underscores the importance of exploring additional variables and advanced optimisation techniques to enhance predictive accuracy further.

Keywords: Household Electricity Consumption, LSTM, Multivariate, Machine Learning, Time Series.

Abstrak- Prediksi konsumsi listrik rumah tangga yang akurat sangat penting untuk perencanaan energi yang efektif di Indonesia, negara yang menghadapi pertumbuhan ekonomi dan kemajuan teknologi yang pesat. Prediksi yang tidak akurat dapat menyebabkan inefisiensi dalam alokasi sumber daya dan kekurangan energi. Metode tradisional seperti ARIMA berjuang dengan pola non-linier, ketergantungan jangka panjang, dan hubungan multivariat yang penting dalam memahami dinamika konsumsi listrik. Untuk mengatasi tantangan ini, studi ini menggunakan algoritma LSTM dengan pendekatan deret waktu multivariat, yang dipilih karena kemampuannya untuk menangkap pola kompleks dan tren jangka panjang. Kumpulan data tersebut terdiri dari data konsumsi listrik bulanan (2004–2023) dari PT PLN, diperkaya dengan variabel makroekonomi dan lingkungan seperti PDB Konsumsi Rumah Tangga, inflasi, dan suhu rata-rata. Metode Denton-Chollete digunakan untuk mengubah data PDB triwulanan menjadi interval bulanan, dan analisis korelasi mengidentifikasi PDB Konsumsi Rumah Tangga ($r=0,98$) dan Penambahan Kontrak Listrik ($r=0,64$) sebagai prediktor signifikan. Pengujian terhadap 63 kombinasi fitur, fitur terbaik (Penambahan Kontrak Daya, PDB Konsumsi Rumah Tangga, dan Konsumsi Listrik Rumah Tangga) mencapai MAPE sebesar 3,54%. Hasil ini menyoroti keunggulan LSTM dalam menangani pola konsumsi listrik yang dinamis dan kompleks serta menyediakan alat prediksi yang kuat bagi PT PLN. Studi ini menggarisbawahi pentingnya mengeksplorasi variabel tambahan dan teknik optimasi tingkat lanjut untuk lebih meningkatkan akurasi prediksi.

Kata Kunci: Konsumsi Listrik Rumah Tangga, LSTM, Multivariate, Machine Learning, Time Series.

I. PENDAHULUAN

Konsumsi listrik, yang mencapai 122,34 TWh (42,41% dari total energi nasional) pada 2023, mendukung pertumbuhan ekonomi Indonesia. Prediksi akurat diperlukan untuk memastikan pasokan stabil dan perencanaan efektif. Metode regresi linear cukup akurat namun terbatas pada fluktuasi non-linear. Machine learning seperti Random Forest dan Neural Network lebih akurat tetapi membutuhkan komputasi tinggi. Model ARIMA cocok untuk jangka pendek dengan data stasioner namun kurang efektif untuk pola jangka panjang atau non-linear. Mengatasi tantangan tersebut, penelitian ini menggunakan algoritma Long Short-Term Memory (LSTM), sebuah jenis jaringan saraf berulang (Recurrent Neural Network) yang dikenal karena kemampuannya menangkap dependensi jangka panjang dan hubungan non-linear yang kompleks. Tidak seperti metode konvensional, LSTM unggul dalam mengolah data berurutan dan hubungan multivariat, menjadikannya sangat cocok untuk peramalan pola konsumsi listrik yang dinamis dan kompleks. Penelitian ini berfokus pada konsumsi listrik rumah tangga di Indonesia dengan menggunakan data agregat nasional untuk periode Januari 2004 hingga Desember 2023. Dataset mencakup variabel internal dari PT PLN (misalnya, konsumsi listrik, penambahan pelanggan, dan penambahan daya tersambung) serta faktor eksternal seperti Produk Domestik Bruto (PDB) Konsumsi Rumah Tangga, inflasi, dan suhu rata-rata. Dengan pendekatan multivariate time series, penelitian ini bertujuan untuk menunjukkan keunggulan LSTM dibandingkan metode konvensional dan memberikan wawasan tentang dinamika permintaan energi.

Analisis dalam penelitian ini membandingkan beberapa metode sebelumnya untuk menyoroti kelebihan dan kekurangannya, yang kemudian menjadi motivasi penggunaan LSTM. ARIMA efektif untuk data jangka pendek dan stasioner, tetapi tidak mampu menangani pola non-linear. Random Forest menawarkan akurasi tinggi tetapi membutuhkan sumber daya komputasi besar, sedangkan Neural Network, meskipun dapat menangkap pola non-linear, memiliki proses tuning yang kompleks. LSTM, di sisi lain, mengatasi keterbatasan ini dengan kemampuannya menangani data sekuensial dan hubungan antar variabel yang lebih kompleks. Dengan pendekatan multivariate time series ini, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam meningkatkan akurasi prediksi konsumsi listrik rumah tangga serta mendukung PT PLN dalam perencanaan energi dan pengelolaan infrastruktur listrik di masa depan. Misalnya[1] menggunakan regresi linear yang memberikan hasil cukup akurat namun metode ini memiliki keterbatasan dalam menangkap pola non-linear fluktuatif yang sering muncul[2]. mengevaluasi algoritma machine learning, menemukan bahwa random forest dan neural network adalah metode dengan akurasi terbaik tetapi membutuhkan waktu komputasi yang lebih tinggi serta tuning hyperparameter yang rumit[3]. menggunakan ARIMA, yang memberikan hasil baik namun model ini terbatas dalam menangani dependensi jangka panjang dan proses penyesuaian parameter yang rumit dan kurang fleksibel untuk data non-stasioner yang dinamis atau memiliki pola non-linear.

Penelitian terbaru oleh Yuli Astuti (2023)[4] menemukan bahwa model *Long Short-Term Memory* (LSTM) mampu memberikan prediksi yang lebih akurat dibandingkan dengan metode konvensional seperti ARIMA dan SARIMA, khususnya dalam menangkap pola musiman dan tren jangka panjang namun, penelitian ini masih terbatas pada penggunaan satu variabel data (*univariate*), sehingga belum mampu mengeksplorasi hubungan antar variabel yang relevan[4]. Penelitian ini berupaya untuk membuktikan bahwa pendekatan *multivariate* menggunakan LSTM memiliki potensi untuk menghasilkan prediksi yang lebih akurat dibandingkan metode konvensional, seperti ARIMA dan regresi linear. Dengan memanfaatkan hubungan kompleks antar variabel, model ini diharapkan mampu memberikan hasil yang lebih relevan untuk mendukung perencanaan energi yang efisien. Dalam penelitian ini, fokus utama adalah pada data konsumsi

listrik segmen rumah tangga di Indonesia. Data yang digunakan merupakan total agregat nasional yang mencerminkan kondisi konsumsi listrik rumah tangga secara keseluruhan di Indonesia. Periode waktu yang dianalisis mencakup Januari 2004 hingga Desember 2023 yang memberikan cakupan yang luas dan representatif untuk mengevaluasi tren jangka panjang dan fluktuatif bulanan dalam konsumsi listrik rumah tangga.

Penelitian ini menemukan bahwa model LSTM mampu memberikan prediksi yang akurat dibandingkan dengan metode konvensional seperti ARIMA dan SARIMA. Pada penelitian ini dilakukan pengujian model menggunakan data konsumsi energi listrik dari PT PLN (Persero) dan menemukan bahwa model LSTM memiliki akurasi yang lebih tinggi dalam menangkap pola musiman dan tren jangka panjang[4]. Dengan demikian, penelitian ini bertujuan untuk mengembangkan model prediksi konsumsi energi listrik segmen rumah tangga menggunakan machine learning *multivariate time series* yaitu dengan melakukan analisis pengaruh kombinasi antar variabel seperti Produk Domestik Bruto (PDB) Konsumsi Rumah Tangga, inflasi, suhu rata-rata, penambahan pelanggan dan penambahan daya tersambung terhadap akurasi model. Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam meningkatkan akurasi prediksi konsumsi energi listrik segmen rumah tangga di Indonesia, sehingga PT PLN (Persero) dapat lebih efektif dalam merencanakan dan mengelola kebutuhan energi listrik di masa depan.

II. SIGNIFIKASI STUDI

Related Work

Algoritma Long Short-Term Memory (LSTM) dipilih untuk mengatasi keterbatasan metode seperti ARIMA, yang efektif untuk peramalan jangka pendek tetapi kurang mampu menangani data non-stasioner, pola jangka panjang, dan hubungan non-linear (Gianika et al., 2024)[3]. Metode statistik tradisional, meskipun andal untuk data sederhana, juga tidak mampu menangkap kompleksitas hubungan antar variabel (Dwianto, 2020)[5]. LSTM telah terbukti unggul dalam menangkap pola musiman dan tren jangka panjang, sebagaimana dibahas oleh Astuti (2023)[4] dan Emsahagin et al. (2022)[6]. Namun, penelitian sebelumnya sering terbatas pada data univariate. Studi ini memperluas pendekatan dengan memanfaatkan data *multivariate*, seperti PDB Konsumsi Rumah Tangga, inflasi, suhu rata-rata, penambahan pelanggan, dan penambahan daya tersambung, untuk meningkatkan akurasi prediksi. Selain itu, seleksi fitur yang tepat memastikan model ini efisien dalam mendukung perencanaan energi dan infrastruktur ketenagalistrikan di Indonesia. Dibandingkan penelitian seperti Kristiana et al. (2015)[7] yang menggabungkan ARIMA dan ANFIS, serta studi berbasis jaringan saraf oleh Kurniawati (2019)[8] dan Rifais (2019)[9], pendekatan ini menawarkan cakupan lebih luas dengan penyesuaian variabel lokal. Hal ini memberikan keunggulan dibanding studi global seperti Patel (2021)[10].

Peramalan Deret Waktu (Time Series Forecasting)

Peramalan deret waktu adalah teknik untuk memprediksi nilai masa depan berdasarkan data historis dalam urutan waktu. Teknik ini bermanfaat di berbagai bidang, termasuk ekonomi dan energi, seperti untuk PT PLN (Persero) dalam merencanakan kebutuhan energi listrik[4]. Dengan mengamati data pada waktu tertentu (t), nilai masa depan di waktu $t+1, t+2, t+3$ dapat diprediksi, membantu perencanaan dan optimalisasi[11]. *Lead time* (l) menunjukkan jangka waktu prediksi, misalnya konsumsi listrik bulan depan hingga 12 bulan mendatang menggunakan data historis[11].

Konsumsi Energi Listrik

Konsumsi energi listrik adalah jumlah listrik yang digunakan konsumen dalam kilowatt-jam (kWh), mencerminkan kebutuhan energi rumah tangga, bisnis, dan industri. Satu kWh setara dengan energi

yang digunakan perangkat 1 kW selama satu jam. Perhitungan konsumsi energi listrik dalam kWh adalah sebagai berikut:

$$\text{Energi (kWh)} = \text{Daya (kW)} \times \text{Waktu (jam)}$$

Dimana:

Daya (kW) adalah besarnya daya listrik yang digunakan oleh perangkat listrik. Daya ini biasanya tertera pada label perangkat dan diukur dalam kilowatt (1 kilowatt = 1000 watt). Waktu (jam) adalah lamanya waktu listrik digunakan.

Long Short-Term Memory (LSTM)

LSTM diperkenalkan oleh Hochreiter dan Schmidhuber (1997), adalah jaringan saraf berulang (RNN) yang mengelola urutan data dan dependensi jangka panjang. Dengan tiga gerbang utama—masukan, keluaran, dan lupa—LSTM mengatur aliran informasi dan mengatasi masalah gradien lenyap pada RNN tradisional[12].

1. Gerbang Masukan (*Input Gate*): Mengontrol informasi yang akan dimasukkan ke dalam sel memori.
2. Gerbang Keluaran (*Output Gate*): Mengatur informasi yang akan dikeluarkan dari sel memori.
3. Gerbang Lupa (*Forget Gate*): Memutuskan informasi mana yang akan dihapus dari sel memori.

Peramalan deret waktu memprediksi nilai masa depan dari data historis, membantu PT PLN (Persero) merencanakan kebutuhan listrik[4][11]. Lead time mencakup jangka prediksi, seperti konsumsi listrik hingga 12 bulan ke depan[13]

Metode Disagregasi Denton-Chollete

Metode *Denton-Chollete* digunakan untuk disagregasi data frekuensi rendah ke frekuensi tinggi, mempertahankan pola temporal dan meminimalkan perubahan ekstrem. Penelitian Irsyad (2020) menunjukkan bahwa metode ini menghasilkan data PDB bulanan yang konsisten dengan data triwulanan awal, dengan Indeks Produksi Manufaktur sebagai indikator andal dan validasi yang signifikan secara statistik.

Evaluasi Kinerja Model

Metrik evaluasi model, seperti RMSE, MAPE, MAE, dan MSE, digunakan untuk mengukur kinerja model prediksi, termasuk peramalan deret waktu. RMSE, sebagai akar rata-rata kuadrat selisih antara nilai prediksi dan aktual, sensitif terhadap kesalahan besar dan memberikan penalti lebih tinggi untuk kesalahan tersebut, sehingga sering digunakan untuk menggambarkan akurasi model secara umum[14].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

n: jumlah observasi

y_i : nilai aktual pada titik observasi ke- i

\hat{y}_i : nilai yang diprediksi pada titik observasi ke- i

Mean Absolute Percentage Error (MAPE) adalah rata-rata persentase kesalahan absolut antara nilai prediksi dan aktual. MAPE mudah dipahami karena hasilnya dalam bentuk persentase, namun dapat tidak stabil jika nilai aktual sangat kecil atau nol[15].

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

n: jumlah observasi

y_i : nilai aktual pada titik observasi ke- i

\hat{y}_i : nilai yang diprediksi pada titik observasi ke- i

Mean Absolute Error (MAE) adalah rata-rata selisih absolut antara nilai prediksi dan aktual. MAE memberikan gambaran tentang kesalahan rata-rata tanpa penalti lebih tinggi untuk kesalahan besar, mudah diinterpretasikan, dan tidak terlalu dipengaruhi oleh *outliers*[16].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

n: jumlah observasi

y_i : nilai aktual pada titik observasi ke-i

\hat{y}_i : nilai yang diprediksi pada titik observasi ke-i

Mean Squared Error (MSE) adalah rata-rata kuadrat selisih antara nilai prediksi dan aktual, sering digunakan karena penalti tinggi untuk kesalahan besar dan sifat matematisnya yang baik (Chai & Draxler, 2014).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

n: jumlah observasi

y_i : nilai aktual pada titik observasi ke-i

\hat{y}_i : nilai yang diprediksi pada titik observasi ke-i

III. HASIL DAN PEMBAHASAN

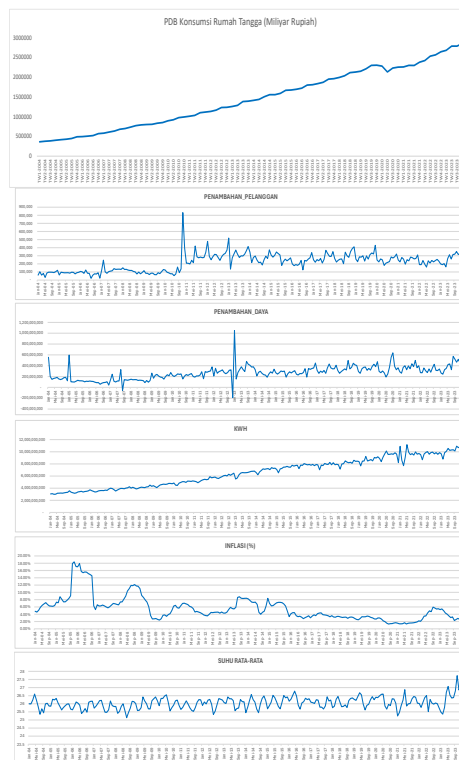
Metode Pengumpulan Data

Pengumpulan data diklasifikasikan menjadi data internal dan eksternal. Data internal merupakan data yang diterbitkan oleh PT PLN (Persero) sedangkan data eksternal merupakan data yang diterbitkan diluar PT PLN (Persero).

Tabel 1. Klasifikasi Data

Klasifikasi	Data	Periode Waktu	Rentang Waktu	Jumlah Dataset	Sumber
Data Internal	Pemakaian listrik (kWh) rumah tangga (KWH), penambahan jumlah pelanggan (DELTA_PELANGGAN) dan penambahan jumlah daya tersambung (DELTA_DAYA)	2004-2023	Bulanan	240	PT PLN (Persero)
Data Eksternal	Produk Domestik Bruto (PDB) konsumsi rumah tangga (PDB_RT)	2004-2023	Triwulan	80	Badan Pusat Statistik (BPS)
	Inflasi (INFLASI)	2004-2023	Bulanan	240	Bank Indonesia
	Suhu rata-rata (SUHU)	2004-2023	Bulanan	240	<i>Climate Change Knowledge Portal</i>

Berikut ini merupakan visualisasi awal dari data tersebut:



Gambar 1. Visualisasi Awal Data

Disagregasi Variabel PDB Konsumsi Rumah Tangga

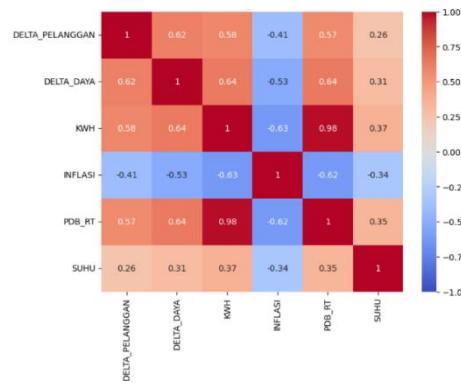
Peramalan deret waktu digunakan untuk memprediksi nilai masa depan berdasarkan data historis, penting di bidang ekonomi dan energi, seperti untuk PT PLN (Persero) dalam merencanakan kebutuhan listrik. Dengan mengamati data saat ini, nilai masa depan dapat diprediksi, membantu perencanaan dan optimalisasi. *Lead time* menunjukkan jangka waktu prediksi, misalnya konsumsi listrik bulan depan hingga 12 bulan mendatang.



Gambar 2. PDB Konsumsi Rumah Tangga Bulanan

Analisa Korelasi antar Variabel

Setelah seluruh variabel rentang waktunya sudah bulanan, perlu dilakukan analisa korelasi antar variabel sebelum melakukan pemodelan prediksi dengan hasil sebagai berikut.



Gambar 3. Hasil Correlation Matrix antar Variabel

PDB_RT memiliki korelasi kuat dengan konsumsi listrik (0,98), menjadikannya variabel utama dalam model prediksi. DELTA_DAYA dan DELTA_PELANGGAN memiliki korelasi moderat, sementara INFLASI dan SUHU memiliki korelasi lemah namun tetap relevan untuk hubungan yang lebih kompleks.

Pembagian Data

Pada penelitian ini, data dibagi menjadi dua kelompok utama: data pelatihan (*training*) dan data pengujian (*testing*). Pembagian data dilakukan berdasarkan rentang waktu untuk memastikan bahwa model dilatih pada data historis yang cukup panjang sebelum diuji pada data yang lebih baru.

Tabel 2. Pembagian Data

Parameter	Rentang Waktu	Jumlah Data	%
Training	Januari 2004 - Desember 2020	204	85%
Testing	Januari 2021 - Desember 2023	36	15%
Total	Januari 2004 - Desember 2023	240	

Data pelatihan mencakup periode 2004–2020 dengan 204 baris data (85%), digunakan untuk melatih model dalam memahami pola dan tren. Data pengujian mencakup periode 2021–2023 dengan 36 baris data (15%), digunakan untuk menguji performa model dalam memprediksi konsumsi listrik. Pembagian data ini memastikan evaluasi model yang objektif dan realistis dalam memprediksi konsumsi listrik masa depan.

Pemodelan dan Pelatihan Prediksi

Penelitian ini menggunakan algoritma LSTM untuk memprediksi konsumsi listrik (kWh) dengan menguji berbagai skenario fitur dan konfigurasi hyperparameter untuk hasil terbaik. Pada penelitian ini memfokuskan pada menemukan skenario terbaik terhadap kombinasi fitur/variabel (*multivariate LSTM*) pada pemodelan prediksi sedangkan konfigurasi hyperparameter adalah sebagai berikut:

Tabel 3. Konfigurasi Hyperparameter LSTM

Indikator Hyperparameter	Nilai
LSTM Units	200
Dropout Rate	0,2
Learning Rate	0,001
Window Length	3

Evaluasi Hasil Model Prediksi

Terdapat 63 skenario kombinasi fitur yang mempengaruhi hasil evaluasi model LSTM yang menunjukkan performa prediksi yang berbeda untuk setiap skenario, dengan metrik utama yang digunakan adalah *Mean Squared Error (MSE)*, *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *Mean Absolute Percentage Error (MAPE)* pada tabel berikut:

Tabel 4. Hasil Evaluasi Model Prediksi

No	Kombinasi Fitur Variabel	Train_MSE	Train_MAE	Train_RMSE	Train_MAPE	Test_MSE	Test_MAE	Test_RMSE	Test_MAPE
1	['DELTA_PELANGGAN']	3.08E+18	1534222491	1755261832	32.20	8.26E+18	2822791827	2874193355	28.34
2	['DELTA_DAYA']	3.35E+18	1591898408	1829842486	34.08	7.36E+18	2664160651	2712272335	26.75
3	['INFLASI']	2.14E+18	1186166157	1461176180	23.37	6.19E+18	2201190609	2488012350	21.98
4	['PDB_RT']	9.41E+17	796486726	970255818	18.03	3.11E+17	406575610	557881476	4.11
5	['SUHU']	3.65E+18	1617506607	1910604050	34.52	8.59E+18	2795036524	2930201583	28.23
6	['KWH']	9.45E+17	796882281	972167248	18.11	5.04E+17	503273564	710246621	4.96
7	['DELTA_PELANGGAN', 'DELTA_DAYA']	2.98E+18	1500635678	1725244386	32.02	6.77E+18	2545597106	2602157636	25.55
8	['DELTA_PELANGGAN', 'INFLASI']	7.05E+17	652029720	839532710	12.06	4.46E+17	1774170715	2111490397	17.68
9	['DELTA_PELANGGAN', 'PDB_RT']	1.87E+17	351706891	432319088	6.47	3.70E+17	450723200	607996237	4.56
10	['DELTA_PELANGGAN', 'SUHU']	3.12E+18	1509156349	1766479415	32.38	7.17E+18	2560691813	2678262967	25.84
11	['DELTA_PELANGGAN', 'KWH']	1.04E+17	212929023	321800670	3.62	4.59E+17	450245573	677211743	4.52
12	['DELTA_DAYA', 'INFLASI']	1.50E+18	982441313	1225396868	19.31	4.47E+18	1871265691	2114663587	18.74
13	['DELTA_DAYA', 'PDB_RT']	2.97E+17	462604328	544859590	9.73	3.30E+17	417595189	574627432	4.22
14	['DELTA_DAYA', 'SUHU']	3.43E+18	1556082980	1853322358	34.06	6.49E+18	2433103771	2547903137	24.54
15	['DELTA_DAYA', 'KWH']	1.55E+18	102050820	1245207469	23.17	1.06E+18	883033241	1029610414	8.73
16	['INFLASI', 'PDB_RT']	2.14E+17	396073188	462207694	7.77	3.42E+17	431536947	584984742	4.36
17	['INFLASI', 'SUHU']	2.19E+18	1164243724	1478633169	22.89	7.06E+18	2261466307	2656946398	22.83
18	['INFLASI', 'KWH']	1.29E+17	270203266	359487779	4.98	4.23E+17	437460312	650067877	4.41
19	['PDB_RT', 'SUHU']	1.99E+17	376409131	446361243	7.42	5.54E+17	585816001	744608999	5.85
20	['PDB_RT', 'KWH']	5.92E+17	630177659	769563076	14.20	3.10E+17	363534385	556355371	3.67
21	['SUHU', 'KWH']	1.18E+17	259395123	343656409	4.56	5.41E+17	512491529	735627451	5.10
22	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI']	1.26E+18	937449194	1122062274	18.25	5.03E+18	2056229259	2241748125	20.63
23	['DELTA_PELANGGAN', 'DELTA_DAYA', 'PDB_RT']	1.62E+18	1086178513	1271235700	23.69	8.08E+17	782014787	898942536	7.80
24	['DELTA_PELANGGAN', 'DELTA_DAYA', 'SUHU']	3.05E+18	1481454474	1745185260	32.17	6.06E+18	2327959544	2460723501	23.49
25	['DELTA_PELANGGAN', 'DELTA_DAYA', 'KWH']	1.11E+17	224215198	333430286	3.97	4.48E+17	450938521	669023651	4.53
26	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT']	2.08E+17	361938886	456138560	6.12	2.98E+17	388943032	545932553	3.92
27	['DELTA_PELANGGAN', 'INFLASI', 'SUHU']	6.62E+17	633857538	813675811	11.41	4.01E+17	1656740209	2002963804	16.68
28	['DELTA_PELANGGAN', 'INFLASI', 'KWH']	1.45E+17	288344050	381367421	5.21	4.24E+17	438911059	650778271	4.42
29	['DELTA_PELANGGAN', 'PDB_RT', 'SUHU']	1.78E+17	318260072	421810278	5.58	5.90E+17	607477641	768231102	6.05
30	['DELTA_PELANGGAN', 'PDB_RT', 'KWH']	1.13E+17	966291558	1060781606	19.89	3.64E+17	454092823	603239285	4.63

No	Kombinasi Fitur Variabel	Train_MSE	Train_MAE	Train_RMSE	Train_MAPE	Test_MSE	Test_MAE	Test_R MSE	Test_MAPE
3	['DELTA_PELANGGAN', 'SUHU', 'KWH']	1.16E+17	250590	340764	4.27	5.81E+17	53948	761967	5.37
1		17	453	727		+17	6155	975	
3	['DELTA_DAYA', 'INFLASI', 'PDB_RT']	2.60E+17	432441	510293	8.69	3.10E+17	39861	557165	4.01
2		17	721	772		+17	9017	821	
3	['DELTA_DAYA', 'INFLASI', 'SUHU']	1.70E+18	104293	130311	20.87	5.28E+18	19031	229752	19.18
3		18	8041	9855		+18	00505	9879	
3	['DELTA_DAYA', 'INFLASI', 'KWH']	1.52E+17	289187	389783	5.39	4.22E+17	44135	649895	4.44
4		17	528	196		+17	5532	267	
3	['DELTA_DAYA', 'PDB_RT', 'SUHU']	2.35E+17	406911	484284	8.05	5.95E+17	61025	771140	6.08
5		17	542	648		+17	1234	422	
3	['DELTA_DAYA', 'PDB_RT', 'KWH']	8.14E+17	752079	902321	16.81	3.13E+17	35438	559083	3.54
6		17	836	250		+17	8758	135	
3	['DELTA_DAYA', 'SUHU', 'KWH']	1.29E+17	266470	358801	4.87	5.79E+17	53014	761129	5.27
7		17	856	965		+17	8441	651	
3	['INFLASI', 'PDB_RT', 'SUHU']	1.25E+17	260708	353497	4.17	5.18E+17	56147	719757	5.61
8		17	918	290		+17	2930	862	
3	['INFLASI', 'PDB_RT', 'KWH']	1.52E+17	320154	390056	6.59	3.67E+17	40508	605604	4.09
9		17	723	907		+17	3127	889	
4	['INFLASI', 'SUHU', 'KWH']	8.57E+16	203866	292808	3.52	4.87E+16	48035	697799	4.79
0		16	351	047		+16	2310	189	
4	['PDB_RT', 'SUHU', 'KWH']	1.83E+17	348568	427796	7.27	5.07E+17	52014	712189	5.17
1		17	437	739		+17	2892	643	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT']	2.13E+17	366820	461910	6.25	2.99E+17	39674	546961	3.98
2		17	412	404		+17	9567	489	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'SUHU']	3.08E+18	150677	175439	32.26	7.32E+18	26013	270501	26.25
3		18	1296	9976		+18	85883	3494	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'KWH']	1.21E+17	244004	347230	4.32	4.53E+17	45525	672855	4.56
4		17	712	848		+17	4813	544	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'PDB_RT', 'SUHU']	2.05E+17	365952	453016	6.97	5.67E+17	59435	752728	5.92
5		17	755	053		+17	7565	755	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'PDB_RT', 'KWH']	1.21E+17	265425	347710	4.83	3.51E+17	40633	592163	4.08
6		17	051	251		+17	3771	882	
4	['DELTA_PELANGGAN', 'DELTA_DAYA', 'SUHU', 'KWH']	1.97E+18	118697	140496	26.12	2.53E+18	13935	159003	14.00
7		18	8942	4139		+18	80086	4605	
4	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT', 'SUHU']	1.29E+17	258918	358867	4.09	5.32E+17	57188	729592	5.71
8		17	353	434		+17	6271	945	
4	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT', 'KWH']	1.50E+17	294567	386880	5.46	3.26E+17	37107	570958	3.72
9		17	930	678		+17	9004	788	
5	['DELTA_PELANGGAN', 'INFLASI', 'SUHU', 'KWH']	9.46E+16	217955	307511	3.76	5.34E+16	49687	730697	4.95
0		16	125	235		+16	7579	655	
5	['DELTA_PELANGGAN', 'PDB_RT', 'SUHU', 'KWH']	1.21E+17	274806	347399	5.11	4.29E+17	46674	655218	4.65
1		17	338	074		+17	8668	752	
5	['DELTA_DAYA', 'INFLASI', 'PDB_RT', 'SUHU']	1.98E+17	348415	445155	6.02	4.88E+17	53980	698609	5.36
2		17	721	473		+17	7961	541	
5	['DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	1.18E+17	262005	343758	4.49	3.34E+17	38099	577891	3.82
3		17	832	133		+17	7332	492	
5	['DELTA_DAYA', 'INFLASI', 'SUHU', 'KWH']	9.03E+16	211077	300462	3.58	6.24E+16	56262	790233	5.59
4		16	962	932		+16	4182	908	
5	['DELTA_DAYA', 'PDB_RT', 'SUHU', 'KWH']	1.07E+17	253485	326498	4.43	4.08E+17	45326	638431	4.51
5		17	853	400		+17	2117	452	
5	['INFLASI', 'PDB_RT', 'SUHU', 'KWH']	8.31E+16	211849	288282	3.46	4.01E+16	45273	633634	4.53
6		16	510	705		+16	2933	375	
5	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT', 'SUHU']	1.69E+17	311694	410506	5.17	4.54E+17	51912	673502	5.16
7		17	573	259		+17	1286	894	
5	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	1.42E+17	284419	377366	5.09	3.21E+17	38482	566370	3.86
8		17	811	448		+17	6341	359	
5	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'SUHU', 'KWH']	1.32E+17	265227	362962	4.72	5.33E+17	52955	729957	5.28
9		17	527	829		+17	4291	897	
6	['DELTA_PELANGGAN', 'DELTA_DAYA', 'PDB_RT', 'SUHU', 'KWH']	1.16E+17	269572	340279	4.67	3.90E+17	44998	624322	4.50
0		17	768	944		+17	7079	129	
6	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT', 'SUHU', 'KWH']	7.50E+16	194521	273919	3.16	4.08E+16	44972	639046	4.49
1		16	792	478		+16	7886	421	
6	['DELTA_DAYA', 'INFLASI', 'PDB_RT', 'SUHU', 'KWH']	8.28E+16	208236	287829	3.42	3.92E+16	44063	626134	4.39
2		16	415	766		+16	6277	830	
6	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT', 'SUHU', 'KWH']	8.49E+16	224560	291454	3.67	4.10E+16	45856	640691	4.58
3		16	564	012		+16	5474	544	

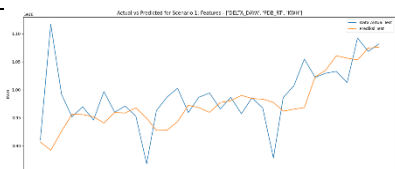
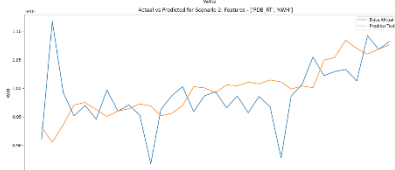

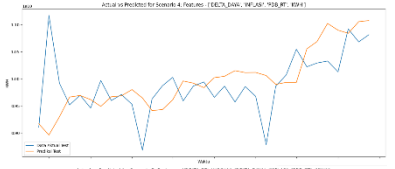

Dari hasil analisa perbandingan akurasi sesuai dengan skenario kombinasi variabel sesuai tabel di atas didapatkan bahwa: 5 Skenario kombinasi variabel berdasarkan nilai MAPE dengan akurasi terbaik pada data uji (testing) meliputi:

Tabel 5. Skenario Kombinasi Variabel berdasarkan Nilai MAPE dengan Akurasi Terbaik

No	Kombinasi Fitur Variabel	Train_M SE	Train_M AE	Train_RM SE	Train_MA PE	Test_M SE	Test_M AE	Test_RM SE	Test_MA PE
1	['DELTA_DAYA', 'PDB_RT', 'KWH']	8.14E+17	75207983 6	902321250	16.81	3.13E+1 7	3543887 58	55908313 5	3.54
2	['PDB_RT', 'KWH']	5.92E+17	63017765 9	769563076	14.20	3.10E+1 7	3635343 85	55635537 1	3.67
3	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT', 'KWH']	1.50E+17	29456793 0	386880678	5.46	3.26E+1 7	3710790 04	57095878 8	3.72
4	['DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	1.18E+17	26200583 2	343758133	4.49	3.34E+1 7	3809973 32	57789149 2	3.82
5	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	1.42E+17	28441981 1	377366448	5.09	3.21E+1 7	3848263 41	56637035 9	3.86

Kombinasi fitur DELTA_DAYA, PDB_RT, dan KWH memberikan MAPE terkecil sebesar 3,54%, menunjukkan prediksi dengan kesalahan sangat rendah. PDB_RT dan KWH konsisten berperan penting dalam model, meskipun kombinasi lain tetap kompetitif dengan akurasi sedikit lebih rendah. Berikut ini merupakan hasil visualisasi pada data testing dari 5 skenario terbaik dari model prediksi tersebut.

Tabel 6. Hasil Visualisasi pada Data Testing

No	Kombinasi Fitur Variabel	Visualisasi
1	['DELTA_DAYA', 'PDB_RT', 'KWH']	
2	['PDB_RT', 'KWH']	
3	['DELTA_PELANGGAN', 'INFLASI', 'PDB_RT', 'KWH']	
4	['DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	
5	['DELTA_PELANGGAN', 'DELTA_DAYA', 'INFLASI', 'PDB_RT', 'KWH']	

Hasil penelitian ini menunjukkan bahwa algoritma LSTM dengan pendekatan *multivariate* time series mampu memberikan prediksi konsumsi listrik rumah tangga yang akurat. Kombinasi variabel terbaik, yaitu DELTA_DAYA, PDB_RT, dan KWH, menghasilkan nilai MAPE sebesar 3,54% pada data uji, yang menunjukkan kesalahan relatif yang sangat rendah dalam prediksi. Hasil ini memperkuat pemahaman pola konsumsi listrik rumah tangga di Indonesia. Variabel PDB_RT memiliki korelasi tinggi dengan konsumsi listrik, menunjukkan hubungan erat antara pertumbuhan ekonomi dan kebutuhan energi. Konsumsi listrik berfungsi sebagai indikator ekonomi, mencerminkan peningkatan pendapatan, daya beli, dan adopsi teknologi. Variabel DELTA_DAYA juga signifikan, mencerminkan peningkatan permintaan kapasitas listrik akibat pertumbuhan pelanggan atau penambahan daya. Disamping itu, hasil ini juga mengindikasikan bahwa pengelolaan variabel ekonomi makro, seperti PDB_RT (PDB Konsumsi Rumah Tangga), dapat berfungsi sebagai alat perencanaan strategis untuk mengantisipasi kebutuhan energi di masa depan. Selain itu, hubungan variabel seperti inflasi dan suhu rata-rata, meskipun korelasinya lebih rendah, tetap relevan dalam menjelaskan variasi musiman dan fluktuasi bulanan dalam pola konsumsi listrik. Inflasi juga dapat memengaruhi daya beli masyarakat terhadap pemakaian listrik, sementara suhu rata-rata berkaitan dengan penggunaan pendingin udara atau pemanas, yang dapat memengaruhi konsumsi listrik secara langsung.

Hasil penelitian ini memberikan wawasan bagi pemerintah dan PT PLN dalam merumuskan kebijakan energi dan infrastruktur listrik jangka panjang. Pemahaman hubungan variabel ekonomi dan konsumsi listrik memungkinkan perencanaan yang lebih presisi untuk stabilitas pasokan energi. Penelitian ini menekankan pentingnya data *multivariate* untuk meningkatkan keandalan model prediksi serta membuka peluang eksplorasi data granular dan variabel tambahan, seperti harga listrik dan penetrasi teknologi energi terbarukan, guna memahami dinamika konsumsi listrik secara lebih spesifik.

IV. KESIMPULAN

Penelitian ini mengembangkan model prediksi konsumsi listrik rumah tangga menggunakan algoritma LSTM dengan seleksi fitur berdasarkan analisis korelasi. Variabel utama seperti PDB Konsumsi Rumah Tangga (PDB_RT) dan Penambahan Daya Tersambung (DELTA_DAYA) memberikan kontribusi terbesar dengan MAPE 3,54%, sedangkan fitur tambahan cenderung menurunkan akurasi. Hasil penelitian menegaskan hubungan erat antara konsumsi listrik dengan faktor ekonomi makro seperti PDB_RT, suhu, dan inflasi. Namun, model ini belum mencakup variabel eksternal lain, seperti tarif listrik, tingkat suku bunga, atau penetrasi energi terbarukan, yang berpotensi memengaruhi konsumsi listrik. Penelitian lanjutan disarankan untuk memasukkan variabel ini serta menggunakan data yang lebih granular untuk meningkatkan akurasi prediksi. Dengan hasil ini, penelitian diharapkan memberikan kontribusi signifikan bagi PT PLN dan pemerintah dalam perencanaan energi yang lebih efektif, efisien, dan berkelanjutan.

REFERENSI

- [1] Johnson, "Linear regression model for predicting electricity consumption in Indonesia," *J. Appl. Energy*, 2018.
- [2] S. Lee, "Evaluation of machine learning algorithms for electricity consumption prediction," *Int. J. Energy Res.*, 2019.
- [3] G. R. Sosa, M. Z. Falah, D. F. L, A. P. Wibawa, A. N. Handayani, and J. A. H. Hammad, "Forecasting electrical power consumption using ARIMA method based on kWh of sold energy," *Sci. Inf. Technol. Lett.*, vol. 2, no. 1, pp. 9–15, 2021, doi: 10.31763/sitech.v2i1.637.
- [4] Y. Astuti, "Prediksi Konsumsi Energi Listrik Menggunakan LSTM: Studi Kasus di PT PLN (Persero)," *Univ. Indones.*, 2023.
- [5] D. A, "Statistical methods for energy demand forecasting," *J. Stat. Plan. Inference*, 2020.
- [6] S. Emshagin, W. K. Halim, and R. Kashef, "Short-term Prediction of Household Electricity Consumption Using Customized LSTM and GRU Models," 2022, doi: <http://dx.doi.org/10.48550/arXiv.2212.08757>.
- [7] A. Kristiana, Y. Wilandari, and A. Prahutama, "Peramalan Beban Puncak Pemakaian Listrik Di Area Semarang Dengan Metode Hybrid Arima (Autoregressive Integrated Moving Average)-Anfis (Adaptive Neuro Fuzzy Inference System) (Studi Kasus Di Pt Pln (Persero) Distribusi Jawa Tengah Dan Diy)," *None*, vol. 4, no. 4, pp. 714–723, 2015.
- [8] Kurniawati, "Prediksi Energi Listrik dengan Metode Jaringan Syaraf Tiruan Backpropagation Menggunakan Matlab untuk Kota Semarang Tahun 2019–2024 i," 2019.
- [9] A. Rifais, "Prediksi Konsumsi Energi Listrik Menggunakan Metode Jaringan Syaraf Tiruan Recurrent di PLN APJ Salatiga," 2019.
- [10] M. Patel, "Neural network models for electricity consumption prediction," *J. Energy Syst.*, 2021.
- [11] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Prentice Hall, Englewood Cliffs, N.J., 1994, 2015.
- [12] S. Hochreiter and J. Schmidhuber, *Long Short-Term Memory*. Neural Computation, 1997. doi: 10.1162/neco.1997.9.8.1735.
- [13] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of Recurrent Network architectures," *32nd Int. Conf. Mach. Learn. ICML 2015*, vol. 3, pp. 2332–2340, 2015.
- [14] R. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.
- [15] K. H and K. C, "A review of time series prediction techniques," *J. Stat. Res.*, 2016.
- [16] Willmott and Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, pp. 79–8, 2005.