

PREDIKSI PENGGUNAAN AIR BERSIH PERUSAHAAN UMUM DAERAH AIR MINUM MENGGUNAKAN ALGORITMA LONG SHORT TERM MEMORY

Hendri Kurniawan¹, Bagus Satrio Waluyo Poetro²

^{1,2}Universitas Islam Sultan Agung Semarang

anggihendri86@std.unissula.ac.id¹

bagusswp@unissula.ac.id²

Received: 08-08-2025

Revised: 20-08-2025

Approved: 27-08-2025

ABSTRAK

Perusahaan Umum Daerah Air Minum (PDAM) memiliki peran penting dalam menyediakan layanan air bersih bagi masyarakat. Tantangan utama yang dihadapi adalah bagaimana memprediksi kebutuhan air secara akurat agar distribusi berjalan optimal dan berkelanjutan. Penelitian ini bertujuan untuk membangun model prediksi penggunaan air bersih dengan menggunakan algoritma Long Short-Term Memory (LSTM), salah satu jenis Recurrent Neural Network (RNN) yang efektif dalam mempelajari pola data deret waktu. Data penelitian berupa catatan historis jumlah pelanggan dan volume pemakaian air bersih dari 15 wilayah layanan PDAM Kabupaten Kendal periode 2020-2025. Tahapan penelitian meliputi pra-pemrosesan data, perancangan dan pelatihan model, evaluasi performa, serta implementasi sistem prediksi berbasis web menggunakan Streamlit. Hasil evaluasi menunjukkan bahwa model LSTM mampu memberikan prediksi dengan tingkat akurasi yang baik. Rata-rata nilai Mean Absolute Percentage Error (MAPE) untuk prediksi jumlah pemakaian air berada pada kisaran 3-5%, sedangkan untuk jumlah pelanggan rata-rata di bawah 2%. Selain itu, nilai Root Mean Squared Error (RMSE) juga relatif rendah sehingga model dapat menyesuaikan dengan pola fluktuatif pada setiap wilayah. Implementasi ke dalam aplikasi web memudahkan PDAM melakukan analisis prediksi secara interaktif dan mendukung perencanaan distribusi air yang lebih efisien.

Kata kunci: LSTM, prediksi, PDAM, penggunaan air bersih, deret waktu.

PENDAHULUAN

Air termasuk kebutuhan pokok yang sangat esensial bagi kelangsungan hidup manusia dan makhluk hidup lainnya (Rachman dkk., 2021). Air selama ini sering dipandang sebagai sumber daya alam yang tidak akan habis dan selalu bisa diakses kapan pun diperlukan (Hartini 2020). Namun, pada kenyataannya air tidak selalu tersedia dalam jumlah cukup, karena pergerakannya mengikuti siklus alam yang cenderung tetap, pasokan air jadi terbatas dan sulit menyesuaikan dengan kebutuhan yang terus meningkat (Djana 2023). Air sangat penting dalam kebutuhan pokok yang diperlukan kehidupan makhluk hidup salah satunya manusia untuk sehari hari (Nurzhafirah, Wahyuni, and Asqia 2024). Peningkatan jumlah penduduk, perkembangan infrastruktur menimbulkan lonjakan permintaan air bersih yang tinggi kebutuhannya dari waktu ke waktu (Sutisna dkk., 2023). Keadaan tersebut memacu penyedia air bersih agar senantiasa mampu memenuhi permintaan masyarakat yang semakin bertambah. Salah satu institusi yang memiliki peran penting dalam hal ini adalah Perusahaan Umum Daerah (PERUMDA) Air Minum (Musyarrof dkk., 2023).

Air menjadi aspek krusial dalam menjaga kelangsungan hidup manusia, karena dipakai untuk berbagai aktivitas harian, termasuk konsumsi, kebersihan, memasak, dan kebutuhan lingkungan. Yang dimaksud dengan air bersih adalah air tawar yang memenuhi standar kualitas, dapat dikonsumsi secara aman, dan tidak berdampak buruk bagi kesehatan manusia (Pahude 2022). Namun, berbagai masalah seperti distribusi yang tidak merata, jumlah pengguna yang terus meningkat, serta ketidakakuratan dalam memprediksi penggunaan air sering kali menjadi kendala. (Pradilla dkk., 2025).

Salah satu strategi yang bisa dilakukan adalah melakukan prediksi kebutuhan air pada tahun-tahun berikutnya, agar PDAM dapat memastikan pasokan air yang memadai bagi masyarakat. Selain itu, hasil prediksi kebutuhan air juga dapat dimanfaatkan untuk mengatur alokasi distribusi air kepada pelanggan agar tidak terjadi kekurangan maupun pemborosan (Indro Antoro dkk., 2024). Prediksi yang akurat akan memberikan manfaat lebih lanjut, yaitu menekan potensi kerugian baik dalam bentuk kehilangan air maupun biaya operasional perusahaan (Fatimah dkk., 2018).

Seiring dengan kemajuan teknologi, algoritma machine learning seperti LSTM muncul sebagai salah satu opsi solusi yang efektif. LSTM memiliki kemampuan dalam menangani data serta memberikan kinerja prediksi yang luar biasa. LSTM mampu mengenali pola jangka panjang, menangani data non-linear, serta tidak memerlukan asumsi stasioner (K. Wang dkk., 2023). Hal ini sangat relevan dengan karakteristik data volume penggunaan air PDAM. Berbagai metode telah dikembangkan dan diuji dalam penelitian terkait prediksi, salah satunya adalah *Long Short-Term Memory* (LSTM). LSTM merupakan salah satu arsitektur dari *Recurrent Neural Network* (RNN) yang dirancang untuk mengolah data berurutan (*sequential data*). Kemampuan ini menjadikan LSTM sangat sesuai digunakan dalam pemodelan dan prediksi data deret waktu (*time series*) (Tita Lattifia dkk., 2022).

Penggunaan algoritma LSTM untuk prediksi konsumsi air pelanggan masih jarang dieksplorasi, meskipun algoritma ini sudah banyak diterapkan di berbagai bidang (Kühnert dkk., 2021). Oleh sebab itu, penelitian ini berfokus pada pembangunan model prediksi air berbasis algoritma LSTM (D. Wang dkk., 2024). Diharapkan hasil penelitian ini dapat membantu PDAM dalam mengelola penyediaan dan penyaluran air secara lebih efektif serta mendukung pengambilan keputusan berdasarkan informasi data.

Studi terdahulu yang membahas penerapan metode *Support Vector Regression* dengan *Ant Colony Optimization* (SVR-ACO) pada PDAM Kota Malang membuktikan bahwa kombinasi metode tersebut mampu menghasilkan prediksi yang cukup optimal. Nilai akurasi NMSE yang dicapai sebesar 0,281 menunjukkan bahwa pendekatan optimasi metaheuristik mampu meningkatkan performa SVR dalam melakukan prediksi. Keberhasilan ini memperlihatkan bahwa metode berbasis *machine learning* dapat dimanfaatkan untuk memprediksi kebutuhan air secara lebih akurat, sehingga membantu perusahaan daerah air minum dalam menyusun strategi distribusi. Namun demikian, penelitian ini juga memiliki keterbatasan pada pemodelan jangka panjang karena performa SVR sangat dipengaruhi oleh pemilihan parameter kernel serta sifat data yang digunakan (Maghfiroh dkk., 2019).

Selain itu, penelitian mengenai prediksi produksi padi di Kota Lubuklinggau dengan algoritma *Long Short Term Memory* (LSTM) memberikan gambaran yang jelas mengenai keunggulan LSTM dalam mengolah data deret waktu (*time series*). Penelitian tersebut menggunakan data produksi padi dari Januari 2019 hingga November 2024 dengan total 71 bulan, yang dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian. Model terbaik diperoleh dengan konfigurasi 35 neuron pada *hidden layer*, batch size 12, dan jumlah epoch sebanyak 50. Hasil pengujian menunjukkan nilai MAPE sebesar 4,44%, yang menandakan tingkat kesalahan prediksi sangat rendah. Model juga mampu mempelajari pola musiman dengan baik sehingga hasil prediksi hampir sesuai dengan data aktual. Hal ini membuktikan bahwa LSTM unggul dalam mengenali pola data jangka panjang yang cenderung kompleks dan non-linear (Adhany dkk., 2025).

Berbeda dengan LSTM, penelitian yang mengkaji penggunaan *Extreme Learning Machine* (ELM) dalam memprediksi volume pemakaian air bulanan di Kota Batu

menunjukkan performa yang kurang optimal. Konfigurasi terbaik diperoleh dengan menggunakan 6 neuron pada lapisan input, 5 neuron pada *hidden layer*, serta pembagian data training dan testing sebesar 80%:20%. Hasilnya menunjukkan nilai error rata-rata sebesar 16.437,5 m³ berdasarkan perhitungan *Root Mean Square Error* (RMSE). Nilai error yang cukup tinggi ini menunjukkan adanya gap signifikan antara hasil prediksi dengan data aktual. Salah satu penyebabnya adalah keterbatasan data historis serta berkurangnya jumlah data akibat penambahan jumlah fitur, sehingga berdampak pada penurunan kualitas hasil prediksi. Dengan demikian, meskipun ELM relatif cepat dalam proses pelatihan, kinerjanya kurang sesuai untuk kasus prediksi volume air dengan pola data yang kompleks (Fahrizal dkk., 2021).

Pendekatan lain dilakukan di PDAM Tirta Dharma Kota Pasuruan dengan menerapkan Jaringan Syaraf Tiruan *Backpropagation*. Penelitian ini menggunakan arsitektur 4-2-1, yaitu 4 neuron pada lapisan input, 2 neuron pada *hidden layer*, dan 1 neuron pada *output layer*. Hasil penelitian menunjukkan performa prediksi yang cukup memuaskan. Pada prediksi tahun 2023 dengan jumlah iterasi sebanyak 999 epoch, model memperoleh nilai MSE sebesar 0,001888878 pada data latih dan 0,02720151 pada data uji dengan akurasi mencapai 100%. Selanjutnya, pada prediksi tahun 2024 dengan 578 epoch, model mencatat MSE 0,00099987 pada data training dan 0,00053177 pada data testing, dengan akurasi tetap 100%. Untuk prediksi tahun 2025, model juga berhasil mempertahankan akurasi 100% meskipun nilai MSE pada data latih dan uji menunjukkan variasi. Hasil ini menunjukkan bahwa *Backpropagation Neural Network* cukup efektif, tetapi nilai akurasi yang terlalu tinggi juga berpotensi mengindikasikan adanya *overfitting* (Agustina dkk., 2023).

Penerapan LSTM juga telah dieksplorasi di bidang keuangan. Penelitian mengenai prediksi harga saham menggunakan pendekatan *Recurrent Neural Network* (RNN) dengan arsitektur LSTM menunjukkan performa yang unggul dalam menghasilkan estimasi harga. Evaluasi menggunakan indikator *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) membuktikan bahwa semakin kecil nilai kedua indikator tersebut, semakin baik hasil prediksi. Pada pengujian awal dengan 10 epoch, diperoleh nilai MAE sebesar 0,0286 dan MAPE sebesar 0,0488. Hasil pengujian dengan 20 epoch memberikan performa yang lebih baik, dengan MAE 0,0150 dan MAPE 0,0257. Hal ini menegaskan bahwa LSTM mampu mempelajari pola pergerakan harga saham dengan baik serta menghasilkan prediksi yang mendekati nilai aktual (Hanafiah dkk., 2023).

Di bidang lingkungan, penelitian lain membandingkan performa LSTM dan *Support Vector Regression* (SVR) dalam memprediksi kualitas udara di Jakarta Pusat. Prediksi dilakukan terhadap rata-rata konsentrasi polutan PM₁₀, SO₂, CO, O₃, dan NO₂. Hasil penelitian menunjukkan bahwa LSTM lebih unggul dibandingkan SVR dengan nilai MAPE sebesar 12,15% dan RMSE 0,0941. Sebaliknya, SVR menghasilkan nilai MAPE 16,19% dan RMSE 0,1666, yang berarti tingkat kesalahannya lebih tinggi. Perbandingan ini memperlihatkan bahwa LSTM memiliki keunggulan dalam menangani data runtut waktu yang kompleks serta lebih mampu memberikan hasil peramalan yang mendekati data aktual (Wicaksono dkk., 2023).

METODE PENELITIAN

Penelitian ini dilaksanakan melalui beberapa tahapan metode penelitian sebagai berikut:

1. Pengumpulan data

Data penelitian ini berasal dari dokumentasi PDAM Kabupaten Kendal yang

mencatat konsumsi air pelanggan per periode (bulanan), sehingga dapat merepresentasikan pola penggunaan air.

REKAPITULASI TAGIHAN REKENING AIR TAHUN 2020							
NAMA CABANG	KODE WIL.	JANUARI			S/D BULAN INI		
		JUMLAH PELANGGAN (SR)	JUMLAH PEMAKAIAN AIR (M ³)	RATA-RATA PEMAKAIAN AIR (M ³)	JUMLAH PELANGGAN (SR)	JUMLAH PEMAKAIAN AIR (M ³)	RATA-RATA PEMAKAIAN AIR (M ³)
KENDAL BARAT	02.11	11.593	184.798	15.96	11.593	184.798	15.96
KENDAL TIMUR	03.14	8.340	134.690	16.15	8.340	134.690	16.15
PATEBON	04.42	4.921	70.280	14.28	4.921	70.280	14.28
BRANDOSONG	05.13	5.691	86.327	15.17	5.691	86.327	15.17
KALIWUNGU I	06.22	8.043	125.469	15.60	8.043	125.469	15.60
KALIWUNGU II	07.21	7.931	113.037	14.25	7.931	113.037	14.25
WELERI	08.31	6.426	93.031	14.48	6.426	93.031	14.48
ROWOSARI	08.32	1.335	17.018	12.75	1.335	17.018	12.75
CEPIRING	09.41	9.034	125.643	13.91	9.034	125.643	13.91
PEGANDON	10.44	3.741	52.844	14.13	3.741	52.844	14.13
SUKOREJO	11.51	4.812	70.858	14.73	4.812	70.858	14.73
PATEAN	11.52	1.360	19.088	14.04	1.360	19.088	14.04
PAGERUYUNG	12.53	1.235	14.061	11.39	1.235	14.061	11.39
BOJA	13.51	5.240	88.557	11.18	5.240	88.557	11.18
GEMUH	14.01	3.951	49.314	12.81	3.951	49.314	12.81
JUMLAH		83,540	1,215,015	14,54	83,540	1,215,015	14,54
JUMLAH	ABT		1,104,640			1,104,640	
JUMLAH	APT		110,375			110,375	

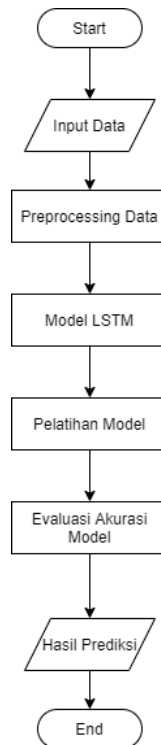
Gambar 1 Dataset PDAM

2. Preprocessing Data

Data mentah yang diperoleh dari PDAM tidak langsung dapat digunakan, sehingga diperlukan tahap pra-pemrosesan untuk meningkatkan kualitas data. Proses ini diawali dengan pembersihan data, yaitu memeriksa adanya data yang hilang atau tidak relevan, kemudian memperbaikinya atau menghapusnya agar tidak mengganggu hasil analisis. Selanjutnya dilakukan normalisasi data dengan metode *Min-Max Scaling*, yaitu menyesuaikan skala data ke dalam rentang tertentu (0–1) agar model lebih mudah mengenali pola. Setelah itu, data dibagi menjadi dua bagian melalui split data dengan rasio 80% untuk data latih dan 20% untuk data uji. Data latih digunakan untuk membangun dan melatih model, sementara data uji dipakai untuk mengevaluasi kinerja model pada data baru. Tahapan ini penting agar model tidak hanya mampu menghafal data, tetapi juga dapat menghasilkan prediksi yang akurat serta mengurangi risiko overfitting.

3. Model LSTM

Penelitian ini berfokus pada pembangunan model LSTM untuk memprediksi penggunaan air bersih PDAM dengan mengenali pola jangka panjang pada data time series konsumsi air.



Gambar 2 Flowchart Perancangan Model

Penelitian ini diawali dengan pengumpulan data historis yang mencakup volume penggunaan air bersih dan jumlah pelanggan aktif setiap bulan. Data tersebut kemudian melalui tahap *preprocessing*, meliputi pembersihan dari data kosong dan duplikasi, normalisasi dengan *Min-Max Scaling*, serta penyusunan dalam bentuk urutan waktu sesuai format input LSTM. Selanjutnya, data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian agar model dapat dievaluasi secara objektif dan terhindar dari overfitting.

Model prediksi dibangun menggunakan algoritma Long Short-Term Memory (LSTM) dengan library TensorFlow dan Keras. Arsitektur model terdiri atas dua lapisan LSTM (64 unit dan 32 unit) yang dipadukan dengan dropout untuk mencegah overfitting, serta diakhiri dengan lapisan Dense beraktivasi linear sebagai output. Model dikompilasi menggunakan fungsi loss *Mean Squared Error* (MSE) dan dioptimasi dengan Adam Optimizer. Proses pelatihan dilakukan dengan menyesuaikan parameter seperti epoch, batch size, dan learning rate melalui serangkaian eksperimen.

Evaluasi model dilakukan menggunakan metrik *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). RMSE digunakan untuk mengukur besar kesalahan prediksi dalam satuan yang sama dengan data asli, sedangkan MAPE memberikan gambaran akurasi dalam bentuk persentase. Setelah pelatihan selesai, model dimanfaatkan untuk memprediksi volume pemakaian air bulan berikutnya.

4. Evaluasi Model

Model dievaluasi menggunakan metrik RMSE dan MAPE untuk mengukur efektivitas prediksi pemakaian air bersih PDAM. RMSE menekankan kesalahan besar sehingga efektif mendeteksi outlier, sedangkan MAPE menunjukkan tingkat akurasi rata-rata dalam bentuk persentase (Cahyani dkk., 2023).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_1 - \hat{y}_1)^2}$$

Dimana :

y_1 : nilai aktual

\hat{y}_1 : nilai hasil prediksi

Σ : menjumlahkan semua eror dari semua data

n : jumlah data yang di evaluasi

Penelitian ini juga menggunakan MAPE sebagai metrik evaluasi, yang mengukur rata-rata persentase perbedaan antara prediksi dan data aktual. Nilai MAPE yang rendah menunjukkan akurasi tinggi, sehingga metrik ini memberikan gambaran praktis dan intuitif bagi pengambil keputusan di PDAM (Simanjuntak dkk., 2025).

$$MAPE = \frac{100\%}{n} \left| \frac{y_1 - \hat{y}_1}{y_1} \right|$$

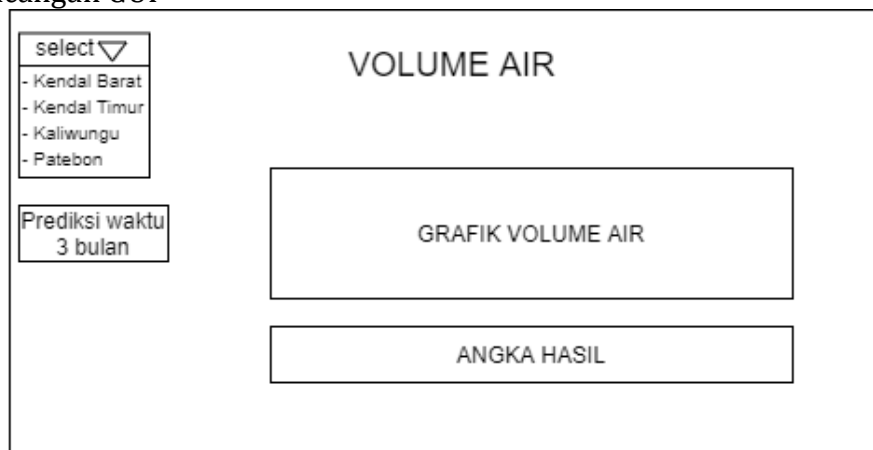
Dimana:

y_1 : nilai aktual

\hat{y}_1 : nilai hasil prediksi

n : jumlah data yang dievaluasi

5. Rancangan GUI



Gambar 3 Rancangan GUI

Rancangan antarmuka pengguna (GUI) dibuat untuk memudahkan analisis dan prediksi volume air pada beberapa wilayah. Antarmuka dilengkapi menu dropdown untuk memilih wilayah analisis, tombol prediksi tiga bulan ke depan, serta tampilan output berupa grafik tren volume air dan hasil numerik prediksi. Tata letak dirancang sederhana dan terstruktur, dengan area input di sisi kiri dan area output di bagian tengah, sehingga alur penggunaan jelas dan mudah dipahami.

HASIL DAN PEMBAHASAN

1) Hasil Pengumpulan Data

Tabel 1 Hasil Pengumpulan Data

bulan	jumlah_pelanggan	jumlah_pemakaian_air	wilayah
1-Jan-20	5240	58557	BOJA
1-Jan-20	5691	86327	BRANGSONG
1-Jan-20	9034	125643	CEPIRING
1-Jan-20	3851	49314	GEMUH
1-Jan-20	8043	125469	KALIWUNGU_I
1-Jan-20	7931	113037	KALIWUNGU_II

bulan	jumlah_pelanggan	jumlah_pemakaian_air	wilayah
1-Jan-20	11580	184798	KENDAL_BARAT
1-Jan-20	8340	134690	KENDAL_TIMUR
1-Jan-20	1235	14061	PAGERUYUNG
1-Jan-20	1360	19088	PATEAN
1-Jan-20	4921	70280	PATEBON
1-Jan-20	3741	52844	PEGANDON
1-Jan-20	1335	17018	ROWOSARI
1-Jan-20	4812	70858	SUKOREJO
1-Jan-20	6426	93031	WELERI

Data penelitian berasal dari catatan historis PDAM Kendal periode Januari 2020–Juni 2025, mencakup jumlah pelanggan dan volume pemakaian air bersih di 15 wilayah cabang. Dataset disusun dengan empat variabel utama, yaitu bulan, jumlah pelanggan, pemakaian air (m³), dan wilayah. Setiap cabang memiliki karakteristik penggunaan yang berbeda, sehingga data ini menjadi dasar pemodelan prediksi dengan metode LSTM.

- 2) Hasil Reprocessing Data
 - a. Melihat Data Value

Tabel 2 Data Hilang

Nama	Data Hilang
Bulan	0
Jumlah_pelanggan	0
Jumlah_pemakaian_air	0
wilayah	0

Berdasarkan Tabel 4.2, seluruh kolom tercatat tanpa missing value (0), sehingga tidak diperlukan proses imputasi maupun penghapusan data.

- b. Duplikat Data

Tabel 3 Duplikat Data

Nama	Data duplikat
Bulan	0
Jumlah_pelanggan	0
Jumlah_pemakaian_air	0
wilayah	0

Berdasarkan Tabel 4.3, dataset tidak mengandung data duplikat sehingga seluruh baris bersifat unik dan tidak memerlukan proses penghapusan.

- c. Normalisasi Data

Sebelum membangun model, dilakukan proses normalisasi untuk menyeragamkan skala nilai antarvariabel sehingga tidak terjadi bias terhadap variabel dengan nilai lebih besar. Normalisasi diterapkan pada variabel jumlah pemakaian air dan jumlah pelanggan menggunakan metode *Min-Max Scaling*, yang mengubah seluruh data ke rentang 0–1 dengan bantuan fungsi *MinMaxScaler* dari *scikit-learn*. Prinsipnya adalah memetakan nilai berdasarkan proporsi terhadap nilai minimum dan maksimum variabel, sehingga seluruh data berada pada skala yang sama.

- d. Split Data

Pada tahap ini, data dibagi menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*). Data latih digunakan untuk mengajarkan model agar dapat mengenali pola-pola pada data, sedangkan data uji digunakan untuk melihat seberapa baik model bekerja pada data yang belum pernah dilihat sebelumnya. Dalam penelitian ini, pembagian dilakukan dengan perbandingan 80% untuk data latih dan 20% untuk data uji. Pembagian ini dilakukan dengan menghitung terlebih dahulu jumlah data latih, kemudian memisahkan data fitur dan label berdasarkan indeks tersebut. Data latih diambil dari awal hingga 80% jumlah data, sedangkan sisanya menjadi data uji. Pada penelitian ini, setiap wilayah memiliki 53 data latih dan 13 data uji, dengan total 15 wilayah yang digunakan.

3) Model LSTM

Pada tahap ini dilakukan perancangan model Long Short-Term Memory (LSTM) yang digunakan untuk memprediksi jumlah pemakaian air dan jumlah pelanggan di setiap wilayah. Model dibangun dengan memanfaatkan *Sequential API* dari Keras yang terdiri atas dua lapisan LSTM dengan masing-masing 64 neuron. Setiap lapisan LSTM dilengkapi dengan lapisan *Dropout* sebesar 0,2 untuk mengurangi risiko *overfitting*. Selanjutnya, ditambahkan lapisan keluaran berupa *Dense layer* dengan dua neuron yang menggunakan fungsi aktivasi linear. Proses kompilasi model dilakukan dengan menggunakan *optimizer Adam* dan fungsi *loss* berupa *Mean Squared Error (MSE)*.

Setelah model dirancang, tahap berikutnya adalah melakukan pelatihan dengan memanfaatkan data historis. Data diproses dalam bentuk urutan dengan jendela waktu tiga bulan, di mana tiga bulan terakhir digunakan untuk memprediksi bulan berikutnya. Dataset dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan dan 20% untuk data pengujian. Proses pelatihan model dijalankan dengan parameter 200 *epoch* dan *batch size* sebesar 16. Untuk menghindari terjadinya *overfitting*, digunakan teknik *Early Stopping* dengan *patience* selama 15 *epoch* serta opsi *restore_best_weights*. Pelatihan dilakukan secara terpisah untuk setiap wilayah sehingga model dapat menyesuaikan diri dengan pola konsumsi air dan jumlah pelanggan yang berbeda-beda di tiap wilayah.

4) Evaluasi Model

Evaluasi model dilakukan untuk mengukur kemampuan LSTM dalam memprediksi jumlah pemakaian air di tiap wilayah. Kinerja model dinilai menggunakan dua metrik utama, yakni *Root Mean Squared Error (RMSE)* untuk melihat besarnya kesalahan absolut, serta *Mean Absolute Percentage Error (MAPE)* untuk mengetahui tingkat kesalahan relatif dalam bentuk *persentase*.

Tabel 4 Evaluasi Pemakaian Air

Daerah	RMSE train	RMSE test	MAPE train	MAPE test
BOJA	4746.36 m ³	4694.70 m ³	5.02 %	4.22%
BRANGSONG	5564.61 m ³	5017.47 m ³	5.64%	4.21%
CEPIRING	6253.00 m ³	7749.77 m ³	3.96%	4.07%
GEMUH	3466.92 m ³	4699.57 m ³	4.75%	4.51%
KALIWUNGU I	5426.33 m ³	4713.20 m ³	3.08%	2.75%
KALIWUNGU II	6304.05 m ³	6049.30 m ³	3.94%	3.33%
KENDAL BARAT	7951.28 m ³	7089.39 m ³	3.21%	2.65%
KENDAL TIMUR	6584.46 m ³	6010.53 m ³	3.75%	3.37%
PAGERUYUNG	1064.89 m ³	1337.12 m ³	6.03%	8.25%
PATEAN	1927.16 m ³	2181.44 m ³	7.98%	7.72%
PATEBON	3364.48 m ³	4423.76 m ³	3.30%	4.29%

Daerah	RMSE train	RMSE test	MAPE train	MAPE test
PEGANDON	3438.68 m ³	5066.04 m ³	4.70%	6.32%
ROWOSARI	1365.31 m ³	1108.06 m ³	6.26%	4.38%
SUKOREJO	4378.64 m ³	5317.16 m ³	4.95%	5.64%
WELERI	4176.72 m ³	4293.69 m ³	3.60%	3.40%

Berdasarkan hasil evaluasi menggunakan metrik RMSE dan MAPE, model LSTM menunjukkan kinerja yang bervariasi pada tiap wilayah PDAM. Beberapa wilayah memberikan hasil prediksi yang sangat baik, seperti Kaliwungu I, Kendal Barat, Weleri, Kaliwungu II, Kendal Timur, dan Cepiring. Pada wilayah-wilayah tersebut nilai MAPE uji berada di bawah 4%, bahkan Kaliwungu I hanya sebesar 2,75% dan Kendal Barat 2,65%. Hal ini menunjukkan bahwa model mampu mengikuti pola konsumsi air dengan sangat akurat sehingga selisih antara data aktual dan hasil prediksi relatif kecil.

Sementara itu, terdapat wilayah dengan hasil prediksi yang tergolong cukup baik, yaitu Boja, Brangsong, Gemuh, Patebon, Rowosari, dan Sukorejo. Pada wilayah ini nilai MAPE uji berada pada kisaran 4% hingga 5,6%. Prediksi yang dihasilkan masih mampu mengikuti tren historis pemakaian air, namun terdapat deviasi yang lebih jelas terutama ketika terjadi lonjakan konsumsi yang tajam.

Adapun wilayah dengan hasil prediksi yang kurang baik adalah Pageruyung, Patean, dan Pegandon, di mana nilai MAPE uji lebih dari 6%. Pageruyung bahkan mencapai 8,25%, sedangkan Patean sebesar 7,72% dan Pegandon 6,32%. Hal ini menunjukkan bahwa model kesulitan menangkap pola fluktuatif pada wilayah tersebut, sehingga tingkat kesalahan relatif lebih tinggi dibanding wilayah lainnya.

Secara keseluruhan, mayoritas wilayah masih menunjukkan akurasi yang baik dengan MAPE rata-rata di bawah 5%, sehingga model LSTM dapat dikatakan efektif untuk memprediksi pemakaian air bersih. Perbedaan akurasi antarwilayah lebih dipengaruhi oleh karakteristik data historis, di mana wilayah dengan pola konsumsi stabil menghasilkan prediksi lebih akurat, sedangkan wilayah dengan pola yang sangat fluktuatif menghasilkan error yang lebih tinggi.

Tabel 5 Jumlah Pelanggan

DAERAH	RMSE train	RMSE test	MAPE train	MAPE test
BOJA	125.63	55.64	1.42%	0.57%
BRANGSONG	161.67	37.63	2.14%	0.46%
CEPIRING	446.31	102.68	3.57%	0.65%
GEMUH	141.48	28.51	2.29%	0.38%
KALIWUNGU I	93.40	30.97	0.75%	0.29%
KALIWUNGU II	160.98	56.91	1.47%	0.51%
KENDAL BARAT	123.79	89.03	0.67%	0.44%
KENDAL TIMUR	50.75	23.44	0.41%	0.20%
PAGERUYUNG	5.72	2.49	0.35%	0.16%
PATEAN	20.34	5.61	0.83%	0.24%
PATEBON	88.45	27.99	1.08%	0.36%
PEGANDON	71.50	57.73	1.49%	1.00%
ROWOSARI	14.51	6.77	0.80%	0.42%
SUKOREJO	18.22	12.21	0.29%	0.19%
WELERI	67.28	30.04	0.68%	0.32%

Hasil evaluasi model LSTM pada prediksi jumlah pelanggan menunjukkan performa yang sangat baik di seluruh wilayah. Nilai MAPE uji secara konsisten berada

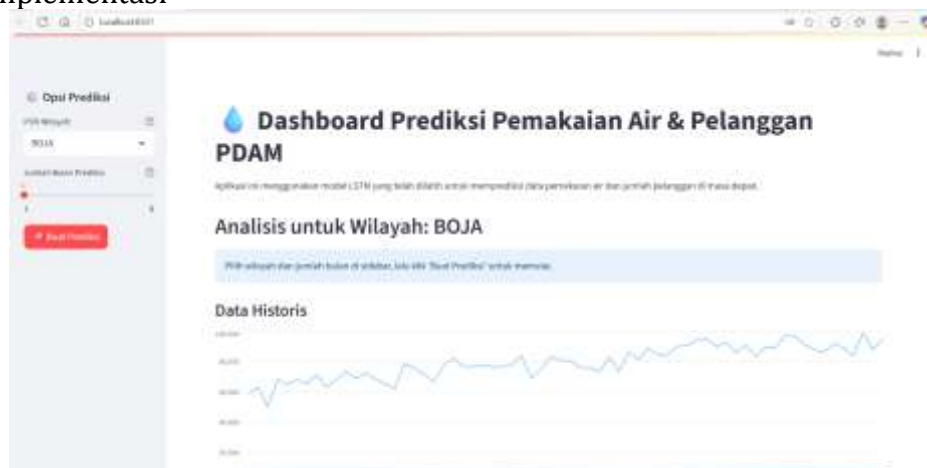
di bawah 2%, bahkan sebagian besar wilayah mencatatkan kesalahan di bawah 1%. Hal ini menandakan bahwa model mampu mengikuti pola pertumbuhan jumlah pelanggan dengan tingkat akurasi yang sangat tinggi.

Beberapa wilayah dengan hasil prediksi terbaik antara lain Kendal Timur, Pageruyung, dan Sukorejo yang masing-masing memiliki MAPE uji hanya sebesar 0,20%, 0,16%, dan 0,19%. Nilai ini menunjukkan bahwa prediksi hampir berhimpit dengan data aktual. Selain itu, wilayah lain seperti Kaliwungu I, Gemuh, Boja, Brangsong, Patean, Rowosari, Patebon, dan Weleri juga mencatatkan hasil yang sangat baik dengan MAPE uji rata-rata kurang dari 0,5%.

Adapun wilayah dengan nilai kesalahan sedikit lebih tinggi adalah Pegandon dengan MAPE uji 1,00% dan Cepiring dengan MAPE uji 0,65%. Meski demikian, angka ini tetap tergolong sangat rendah sehingga masih menunjukkan akurasi prediksi yang sangat baik.

Secara keseluruhan, dapat disimpulkan bahwa model LSTM sangat efektif dalam memprediksi jumlah pelanggan PDAM. Tingkat kesalahan yang sangat kecil mengindikasikan bahwa pola pertumbuhan pelanggan relatif stabil dan mudah dipelajari oleh model. Hal ini menjadikan hasil prediksi jumlah pelanggan lebih akurat dibanding prediksi pemakaian air, yang cenderung lebih fluktuatif.

4. Hasil Implementasi



Gambar 4 Tampilan Awal Sistem

Gambar tersebut memperlihatkan antarmuka dasbor aplikasi web PDAM yang menggunakan model LSTM untuk memprediksi pemakaian air dan jumlah pelanggan. Di sisi kiri terdapat panel Opsi Prediksi untuk memilih wilayah (contoh: Boja) dan menentukan horizon waktu prediksi, kemudian pengguna menekan tombol Buat Prediksi. Bagian utama menampilkan instruksi awal serta grafik data historis yang menunjukkan tren pemakaian air dengan fluktuasi dari waktu ke waktu. Data historis ini menjadi dasar bagi model LSTM dalam menghasilkan prediksi, yang hasilnya akan ditampilkan pada bagian Analisis untuk Wilayah setelah proses dijalankan.



Gambar 5 Hasil Prediksi

Gambar ini menampilkan dashboard prediksi PDAM untuk wilayah Kendal Barat dengan horizon tiga bulan ke depan. Setelah parameter dipilih, sistem menampilkan hasil prediksi pada Juli 2025 pemakaian air diperkirakan 94.115 m³ dengan 8.016 pelanggan, Agustus turun sedikit menjadi 93.322 m³ dengan 7.990 pelanggan, lalu September naik kembali menjadi 94.466 m³ dengan 8.066 pelanggan. Hasil ini menunjukkan pemakaian air relatif stabil, sementara jumlah pelanggan cenderung meningkat.

KESIMPULAN

Hasil penelitian menunjukkan bahwa algoritma Long Short-Term Memory (LSTM) dapat memberikan prediksi yang akurat mengenai pemakaian air bersih. Hasil evaluasi model menunjukkan nilai RMSE dan MAPE yang cenderung rendah, yang mencerminkan tingkat akurasi prediksi yang tinggi, meskipun di beberapa daerah masih ada hasil yang tidak optimal. Ini menunjukkan bahwa penerapan LSTM efektif dan dapat menjadi pilihan solusi dalam sistem prediksi kebutuhan air bersih. Selanjutnya, penelitian ini memberikan kontribusi signifikan dalam mendukung manajemen dan perencanaan distribusi air bersih yang lebih efisien, serta dapat menjadi landasan untuk pengembangan sistem prediksi yang lebih adaptif dan menyeluruh di masa yang akan datang.

DAFTAR PUSTAKA

- Adhany, Putri Cheria, Cindi Wulandari, Bunga Intan, and Budi Santoso. 2025. "Prediksi Padi Menggunakan Algoritma Long Short Term Memory." 5(2): 120–27.
- Agustina, Dwi, Moh. Hafiyusholeh, Aris Fanani, and Dono Prasetyo. 2023. "Prediksi Distribusi Air Perusahaan Daerah Air Minum (PDAM) Tirta Dharma Kota Pasuruan Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation." *Jurnal PROCESSOR* 18(1): 8–16.
- Cahyani, Jumina, Syamsul Mujahidin, and Tegar Palyus Fiqar. 2023. "Implementasi Metode Long Short Term Memory (LSTM) Untuk Memprediksi Harga Bahan Pokok Nasional." *Jurnal Sistem dan Teknologi Informasi (JustIN)* 11(2): 346.
- Djana, Miftahul. 2023. "Analisis Kualitas Air Dalam Pemenuhan Kebutuhan Air Bersih Di Kecamatan Natar Hajimena Lampung Selatan." *Jurnal Redoks* 8(1): 81–87.
- Fahrizal, M A, S Adinugroho, and ... 2021. "Prediksi Volume Penggunaan Air Bulanan Kota Batu Menggunakan Metode Extreme Learning Machine (ELM)." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 5(7): 3078–86. <http://j->

- ptiik.ub.ac.id/index.php/j-ptiik/article/download/9478/4282.
- Fatimah, Fajar, Andi Tejawati, and Novianti Puspitasari. 2018. "Prediksi Pemakaian Air PDAM Menggunakan Metode Simple Moving Average." *Jurnal Rekayasa Teknologi Informasi (JURTI)* 2(1): 55.
- Hanafiah, Anggi, Yudhi Arta, Hafiza Oktasia Nasution, and Yuyun Dwi Lestari. 2023. "Penerapan Metode Recurrent Neural Network Dengan Pendekatan Long Short-Term Memory (LSTM) Untuk Prediksi Harga Saham." *Bulletin of Computer Science Research* 4(1): 27–33.
- Hartini. 2020. "Analisa Kebutuhan Dan Ketersediaan Air." 10(26): 1–12.
- Indro Antoro, Achmad Sudrajad, Iskahar Iskahar, and Teguh Marhendi. 2024. "Analisis Kehilangan Air Pada Jaringan Distribusi Pdam Wonosobo." *CIVeng: Jurnal Teknik Sipil dan Lingkungan* 5(1): 1.
- Kühnert, Christian et al. 2021. "Application of LSTM Networks for Water Demand Prediction in Optimal Pump Control." *Water (Switzerland)* 13(5): 1–19.
- Maghfiroh, A, A W Widodo, and Y A Sari. 2019. "Prakiraan Penggunaan Volume Air PDAM Kota Malang Menggunakan Metode Support Vector Regression Dengan Ant Colony Optimization." *Pengembangan Teknologi Informasi dan Ilmu Komputer* 3(10): 10344–52. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/6676/3222>.
- Musyarrof, Muhammad Alfi, and Aries Susanty. 2023. "Peramalan Volume Produksi Air Bersih Menggunakan Metode Time Series (Studi Kasus: PERUMDAM Purwa Tirta Dharma Kabupaten Grobogan)." *Industrial Engineering Online Journal* 12(4). <https://ejournal3.undip.ac.id/index.php/ieoj/article/view/41339>.
- Nurzahfirah, Nurzahfirah, Sri Wahyuni, and Nurul Asqia. 2024. "Pentingnya Konsumsi Air Putih Daripada Minuman Manis Untuk Anak Usia Dini." *Kumaracitta : Jurnal Pendidikan Anak Usia Dini* 2(01): 35–47.
- Pahude, Mansur S. 2022. "Analisis Kebutuhan Air Bersih Di Desa Santigi Kecamatan Tolitoli Utara Kabupaten Tolitoli." *Jurnal Inovasi Penelitian* 03(02): 4801–10. https://www.google.com/url?sa=i&url=https%3A%2F%2Fstp-mataram.e-journal.id%2FJIP%2Farticle%2Fdownload%2F1744%2F1345&psig=AOvVaw1amkE6LLiImG7Y3YJOPjE&ust=1711081233630000&source=images&cd=vfe&opi=89978449&ved=0CAUQn5wMahcKEwjwp_rzwYSFAxUAAAAAHQAAAAAQBA.
- Pradilla, Diah Vivin, and Agus Bambang Irawan. 2025. "Analisis Tingkat Kehilangan Air Akibat Kebocoran Pipa PERUMDAM Tirta Projotamansari, Sub Unit Pulutan." *Prosiding Seminar Nasional Teknik Lingkungan Kebumihan SATU BUMI* 6(1): 224–29.
- Rachman, Delli Noviarti, and Susi Riwayati. 2021. "Perhitungan Kebutuhan Air Bersih Dan Perencanaan Kebutuhan Infrastruktur Perpipaan PDAM Di Kel Sako, Sako Baru, Sialang Dan Sukamaju Kec. Sako Kota Palembang." *Jurnal Tekno Global UIGM Fakultas Teknik* 10(1): 1–8.
- Simanjuntak, Dina Febrianti et al. 2025. "IMPLEMENTASI LONG SHORT-TERM MEMORY DALAM PERAMALAN PERMINTAAN AIR BERSIH DI KOTA PONTIANAK." 9(2): 352–58.
- Sutisna, and Nazar Mirsandi Yuniar. 2023. "Klasifikasi Kualitas Air Bersih Menggunakan Metode Naïve Baiyes." *Jurnal Sains dan Teknologi* 5(1): 243–46. <https://doi.org/10.55338/saintek.v5i1.1383>.
- Tita Lattifia, Putu Wira Buana, and NI Kadek Dwi Rusjyanthi. 2022. "Model Prediksi Cuaca Menggunakan Metode LSTM." *JITTER-Jurnal Ilmiah Teknologi dan Komputer* 3(1).
- Wang, Dingtong, Yanning Li, Benwei Hou, and Shan Wu. 2024. "Short-Term Water

- Demand Forecasting Based on LSTM Using Multi-Input Data †." *Engineering Proceedings* 69(1): 2–5.
- Wang, Ke et al. 2023. "MACLA-LSTM: A Novel Approach for Forecasting Water Demand." *Sustainability (Switzerland)* 15(4).
- Wicaksono, Mohammad Ernico Suryo, Gusti Made Arya Sasmita, and I Putu Agus Eka Pratama. 2023. "Peramalan Kualitas Udara Di Kota Jakarta Pusat Dengan Metode Long Short-Term Memory Dan Support-Vector Regression." *JITTER: Jurnal Ilmiah Teknologi dan Komputer* 4(1): 1576.