

Comparison of LSTM and Random Forest for Hydrogen Production Prediction in Alkaline Water Electrolysis

Ulil Amri ^{1*}, Ardi Pujiyanta ^{2*}, Sunardi ^{3*}

^{1,2,3} Master of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55164, Indonesia
2507057005@webmail.uad.ac.id ¹, ardipujiyanta@tif.uad.ac.id ², sunardi@ee.uad.ac.id ³

Article Info

Article history:

Received 2026-06-09

Revised 2026-07-02

Accepted 2026-08-12

Keyword:

*Alkaline water Electrolysis,
Random Forest,
Long Short- Term Memory,
Hydrogen Production Prediction,
Machine Learning.*

ABSTRACT

Hydrogen has emerged as a promising clean energy carrier capable of supporting the transition toward sustainable energy systems. In alkaline water electrolysis systems, accurate prediction of hydrogen production is essential for improving system monitoring, operational efficiency, and future control strategies. This study aims to compare the performance of Random Forest (RF) and Long Short-Term Memory (LSTM) algorithms in predicting hydrogen production based on operational parameters of an alkaline water electrolysis system. The dataset used in this study consists of more than 100,000 operational data samples collected from laboratory-scale electrolysis experiments, including voltage, current, temperature, and hydrogen gas pressure measurements. Hydrogen production was calculated using the ideal gas law and used as the target variable for model development. Prior to model training, the dataset underwent preprocessing, including data cleaning, normalization, and train-test splitting. The predictive performance of both models was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Experimental results show that the Random Forest model achieved superior performance with an MAE of 0.0085, RMSE of 0.0106, and R^2 of 0.9905, while the LSTM model obtained an MAE of 0.0302, RMSE of 0.0361, and R^2 of 0.7907. The findings indicate that Random Forest is more effective than LSTM in modeling the relationship between operational parameters and hydrogen production in the investigated alkaline water electrolysis system. This study demonstrates the potential of machine learning approaches for accurate hydrogen production prediction and provides insights into the suitability of different predictive models for electrolysis-based hydrogen generation systems.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Peningkatan kebutuhan energi global yang diiringi dengan semakin menipisnya cadangan bahan bakar fosil telah mendorong berbagai negara untuk mempercepat transisi menuju sumber energi yang lebih bersih dan berkelanjutan. Salah satu teknologi yang saat ini mendapatkan perhatian besar adalah pemanfaatan hidrogen sebagai pembawa energi (*energy carrier*) karena memiliki densitas energi yang tinggi dan tidak menghasilkan emisi karbon selama proses pemanfaatannya. Hidrogen dipandang sebagai salah satu solusi potensial dalam

mendukung pencapaian target *net-zero emission* dan pembangunan sistem energi berkelanjutan di masa depan [1],[2]

Produksi hidrogen dapat dilakukan melalui berbagai metode, namun elektrolisis air menjadi salah satu pendekatan yang paling menjanjikan karena mampu menghasilkan hidrogen dengan tingkat kemurnian tinggi tanpa menghasilkan emisi karbon secara langsung. Selain itu, teknologi elektrolisis dapat diintegrasikan dengan sumber energi terbarukan seperti tenaga surya dan tenaga angin sehingga mendukung konsep *green hydrogen* yang saat ini menjadi fokus pengembangan

energi global[3] [2] [1]. Pada sistem elektrolisis air alkalin (*alkaline water electrolysis*), larutan elektrolit seperti kalium hidroksida (KOH) digunakan untuk meningkatkan konduktivitas listrik sehingga proses pemisahan molekul air menjadi hidrogen dan oksigen dapat berlangsung lebih efektif[4].

Meskipun memiliki prospek yang sangat baik, performa sistem elektrolisis air alkalin dipengaruhi oleh berbagai parameter operasi yang saling berinteraksi secara kompleks. Tegangan, arus, temperatur, tekanan, jenis elektroda, dan konsentrasi elektrolit merupakan beberapa faktor yang diketahui berpengaruh terhadap laju produksi hidrogen dan efisiensi energi sistem.

Berdasarkan Hukum Faraday, jumlah hidrogen yang dihasilkan berbanding lurus dengan jumlah muatan listrik yang mengalir selama proses elektrolisis. Semakin besar arus listrik yang diberikan, semakin banyak elektron yang terlibat dalam reaksi reduksi pada katoda sehingga jumlah hidrogen yang dihasilkan akan meningkat. Selain itu, temperatur elektrolit mempengaruhi mobilitas ion dan menurunkan resistansi larutan sehingga mampu meningkatkan kinetika reaksi elektrokimia. Tekanan dan temperatur gas juga menentukan jumlah mol hidrogen yang dihasilkan sesuai dengan hukum gas ideal. Oleh karena itu, tegangan, arus, temperatur, dan tekanan memiliki keterkaitan fisik yang kuat dalam menentukan laju produksi hidrogen pada sistem elektrolisis air alkalin.

Penelitian yang dilakukan oleh Kotijah menunjukkan bahwa peningkatan tegangan operasi mampu meningkatkan volume hidrogen yang dihasilkan, namun peningkatan konsentrasi elektrolit yang terlalu tinggi justru dapat menyebabkan penurunan efisiensi akibat kejenuhan ion dalam larutan [5]. Sementara itu, penelitian lain menunjukkan bahwa pemilihan jenis elektroda yang tepat mampu meningkatkan laju produksi hidrogen secara signifikan karena mempengaruhi karakteristik reaksi elektrokimia yang terjadi pada permukaan elektroda[6]

Selain parameter operasi, temperatur juga memiliki pengaruh penting terhadap kinetika reaksi elektrolisis. Peningkatan temperatur dapat meningkatkan mobilitas ion dalam larutan, menurunkan hambatan listrik, serta mempercepat laju transfer massa sehingga berkontribusi terhadap peningkatan efisiensi proses produksi hidrogen [7]. Berbagai penelitian terbaru juga menunjukkan bahwa pengembangan desain elektroda, optimasi kondisi operasi, dan pemodelan sistem menjadi faktor penting dalam meningkatkan performa teknologi elektrolisis air alkalin[3] [4], [8].

Ketersediaan data operasional dalam jumlah besar mendorong berkembangnya pemanfaatan metode *machine learning* untuk memodelkan dan memprediksi performa sistem energi. Berbeda dengan pendekatan matematis konvensional yang sering mengalami keterbatasan dalam memodelkan hubungan non-linear, algoritma *machine learning* mampu mempelajari pola hubungan yang kompleks dari data historis sehingga menghasilkan prediksi yang lebih adaptif terhadap perubahan kondisi operasi[9]. Dalam konteks produksi hidrogen, Maghfuri dan Kuku menunjukkan bahwa algoritma *machine learning* mampu memprediksi produksi hidrogen

berdasarkan parameter tegangan, arus, temperatur, dan konsentrasi elektrolit dengan tingkat akurasi yang tinggi[5]. Temuan tersebut menunjukkan bahwa pendekatan berbasis data (*data-driven*) memiliki potensi besar dalam mendukung optimasi sistem elektrolisis.

Salah satu algoritma yang banyak digunakan dalam pemodelan sistem energi adalah Random Forest. Algoritma ini merupakan metode *ensemble learning* yang dibangun dari kumpulan *decision tree* dan menghasilkan prediksi melalui agregasi hasil dari seluruh pohon keputusan. Random Forest memiliki kemampuan yang baik dalam menangani hubungan non-linear, data berdimensi tinggi, serta mampu mengurangi risiko *overfitting* sehingga banyak digunakan pada berbagai aplikasi prediksi dan optimasi sistem energi [10]. Selain itu, Random Forest juga menyediakan informasi *feature importance* yang dapat digunakan untuk mengidentifikasi parameter operasi yang paling berpengaruh terhadap performa sistem[10].

Meskipun demikian, sebagian besar penelitian yang memanfaatkan Random Forest pada sistem produksi hidrogen masih berfokus pada hubungan statis antarvariabel dan belum sepenuhnya mempertimbangkan karakteristik temporal data operasi. Padahal parameter seperti temperatur, tekanan, arus, dan tegangan berubah secara dinamis terhadap waktu sehingga membentuk pola deret waktu (*time series*) yang dapat dimanfaatkan untuk meningkatkan akurasi prediksi [5], [9]. Oleh karena itu, diperlukan algoritma yang mampu mempelajari ketergantungan temporal jangka panjang dari data historis.

Salah satu metode yang banyak digunakan untuk menangani permasalahan tersebut adalah Long Short-Term Memory (LSTM). LSTM merupakan pengembangan dari Recurrent Neural Network (RNN) yang dirancang untuk mempelajari pola temporal dan mempertahankan informasi historis dalam jangka panjang. Berbagai penelitian menunjukkan bahwa LSTM memiliki performa yang sangat baik dalam memprediksi data deret waktu pada sistem energi, termasuk pada prediksi produksi hidrogen dan kapasitas sistem berbasis energi terbarukan [9], [11]. Kemampuan ini menjadikan LSTM sebagai kandidat yang menjanjikan untuk diterapkan pada sistem elektrolisis yang memiliki karakteristik operasi dinamis.

Berdasarkan kajian literatur yang telah dilakukan, penelitian mengenai optimasi elektrolisis air alkalin, monitoring berbasis IoT, maupun penerapan *machine learning* telah banyak dilakukan secara terpisah. Namun penelitian yang mengintegrasikan sistem monitoring elektrolisis air alkalin berbasis IoT dengan algoritma Random Forest dan LSTM untuk memprediksi produksi hidrogen dan efisiensi energi secara real-time masih relatif terbatas. Selain itu, belum banyak penelitian yang secara khusus membandingkan kemampuan Random Forest dalam menangani hubungan non-linear dengan kemampuan LSTM dalam mempelajari pola temporal pada sistem elektrolisis air alkalin berbasis KOH.

Keunikan studi ini terletak pada tiga kontribusi utama. Pertama, studi ini memanfaatkan lebih dari 100.000 dataset operasional waktu nyata yang dikumpulkan dari sistem elektrolisis air alkalin skala laboratorium yang terintegrasi

dengan platform pemantauan berbasis mikrokontroler, sedangkan sebagian besar studi sebelumnya bergantung pada data simulasi atau dataset eksperimen kecil. Kedua, produksi hidrogen diperkirakan dari data tekanan dan suhu yang diukur menggunakan hukum gas ideal, memungkinkan pengembangan kerangka kerja prediksi berbasis data berdasarkan kondisi elektrolisis aktual. Ketiga, studi ini memberikan investigasi komparatif model Random Forest dan Long Short-Term Memory (LSTM) dalam kondisi operasional yang identik untuk mengevaluasi apakah model pembelajaran temporal memberikan keunggulan dibandingkan metode berbasis pohon ensemble dalam memprediksi produksi hidrogen dari sistem elektrolisis air alkali.

II. METODE

Penelitian ini bertujuan untuk mengembangkan model prediksi produksi hidrogen dan efisiensi energi pada sistem elektrolisis air alkalin menggunakan algoritma Random Forest dan Long Short-Term Memory (LSTM). Pengembangan model prediksi berbasis machine learning menjadi penting karena performa sistem elektrolisis dipengaruhi oleh berbagai parameter operasi yang saling berinteraksi secara kompleks, seperti tegangan, arus, temperatur, tekanan, karakteristik elektroda, dan konsentrasi elektrolit [1],[5]. Berbagai penelitian menunjukkan bahwa hubungan antarparameter tersebut bersifat non-linear dan dinamis sehingga memerlukan pendekatan berbasis data (*data-driven*) untuk memperoleh model prediksi yang lebih akurat dibandingkan pendekatan matematis konvensional [5],[6].

Sistem elektrolisis yang digunakan dalam penelitian ini merupakan elektroliser air alkalin skala laboratorium dengan larutan kalium hidroksida (KOH) sebagai elektrolit [12]. Pemilihan sistem elektrolisis alkalin didasarkan pada kemampuannya menghasilkan hidrogen dengan biaya operasional yang relatif rendah, teknologi yang telah matang, serta kemudahan integrasinya dengan sumber energi terbarukan [6]. Selain itu, berbagai penelitian terbaru menunjukkan bahwa optimasi parameter operasi pada sistem elektrolisis alkalin masih menjadi tantangan utama dalam meningkatkan efisiensi produksi hidrogen [13],[14].

Data operasional diperoleh melalui sistem yang dirancang untuk melakukan akuisisi data secara real-time selama proses elektrolisis berlangsung. Sistem monitoring terdiri atas sensor tegangan, sensor arus, sensor temperatur elektrolit, dan sensor tekanan gas hidrogen yang terhubung dengan mikrokontroler ESP32 [15],[16]. Data hasil pengukuran kemudian dikirimkan dan disimpan pada platform berbasis cloud untuk mendukung proses pemantauan dan pengolahan data secara berkelanjutan [17]. Pendekatan ini sejalan dengan berbagai penelitian yang menunjukkan bahwa integrasi IoT pada sistem produksi hidrogen mampu meningkatkan kemampuan monitoring, akuisisi data, serta mendukung pengembangan sistem prediksi berbasis kecerdasan buatan [18].

Parameter masukan (*input features*) yang digunakan dalam penelitian ini meliputi temperatur elektrolit ($^{\circ}\text{C}$), arus listrik (A), tegangan sistem (V), dan tekanan gas hidrogen (kPa). Pemilihan parameter tersebut didasarkan pada hasil penelitian

terdahulu yang menunjukkan bahwa variabel-variabel tersebut merupakan faktor dominan yang mempengaruhi laju produksi hidrogen dan efisiensi energi pada sistem elektrolisis air alkalin [6]. Sementara itu, variabel keluaran (*target variable*) yang digunakan adalah nilai produksi hidrogen yang dihitung berdasarkan data operasional hasil pengukuran selama proses elektrolisis berlangsung.

Untuk membangun model prediksi pertama, penelitian ini menggunakan algoritma Random Forest Regression. Random Forest merupakan metode *ensemble learning* yang dibangun dari sekumpulan *decision tree* dan menghasilkan prediksi akhir melalui agregasi hasil seluruh pohon keputusan [10]. Algoritma ini dipilih karena memiliki kemampuan yang baik dalam menangani hubungan non-linear antarvariabel, mengurangi risiko *overfitting*, serta mampu mengidentifikasi tingkat pengaruh masing-masing parameter terhadap hasil prediksi melalui analisis *feature importance* [10]. Keunggulan tersebut menjadikan Random Forest banyak digunakan dalam berbagai penelitian prediksi dan optimasi sistem energi maupun produksi hidrogen [19].

Selain Random Forest, penelitian ini juga menerapkan Long Short-Term Memory (LSTM) sebagai model pembanding. LSTM merupakan pengembangan dari Recurrent Neural Network (RNN) yang dirancang untuk menangani data deret waktu (*time-series*) dengan kemampuan menyimpan informasi historis dalam jangka panjang [11]. Penggunaan LSTM pada penelitian ini didasarkan pada karakteristik data elektrolisis yang diperoleh secara kontinu selama proses berlangsung sehingga memiliki ketergantungan temporal antar waktu pengamatan. Dengan kemampuan tersebut, LSTM diharapkan mampu menangkap pola perubahan parameter operasi dan produksi hidrogen secara lebih baik dibandingkan model yang hanya memanfaatkan hubungan statis antarvariabel [11].

Sebelum proses pelatihan model dilakukan, data hasil akuisisi terlebih dahulu melalui tahap pra-pemrosesan (*pre-processing*). Tahapan ini meliputi pemeriksaan data hilang (*missing values*), pembersihan data (*data cleaning*), normalisasi data, serta pembagian dataset menjadi data latih (*training data*) dan data uji (*testing data*). Proses normalisasi dilakukan untuk menyamakan rentang nilai antarvariabel sehingga proses pembelajaran model dapat berlangsung secara lebih optimal [9],[12]. Tahapan pra-pemrosesan menjadi penting karena kualitas data sangat mempengaruhi performa model machine learning yang dihasilkan.

Pada model LSTM, data yang telah diproses kemudian disusun dalam bentuk sekuens berdasarkan interval waktu tertentu sehingga mampu merepresentasikan kondisi sistem pada beberapa waktu sebelumnya. Penyusunan data dalam bentuk sekuens memungkinkan model mempelajari pola temporal yang terjadi selama proses elektrolisis berlangsung. Pendekatan ini banyak digunakan pada penelitian prediksi berbasis data deret waktu karena mampu meningkatkan kemampuan model dalam menangkap pola historis yang berpengaruh terhadap kondisi masa depan [11].

Evaluasi performa model dilakukan menggunakan beberapa metrik statistik yang umum digunakan pada permasalahan regresi, yaitu Mean Absolute Error (MAE), Root Mean Square

Error (RMSE), dan koefisien determinasi (R^2). Nilai MAE digunakan untuk mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual, sedangkan RMSE digunakan untuk mengevaluasi besarnya kesalahan prediksi dengan memberikan penalti yang lebih besar terhadap kesalahan yang ekstrem. Sementara itu, nilai R^2 digunakan untuk menunjukkan kemampuan model dalam menjelaskan variasi data aktual [10]. Penggunaan ketiga metrik tersebut diharapkan mampu memberikan gambaran yang komprehensif mengenai tingkat akurasi dan kemampuan generalisasi model Random Forest maupun LSTM dalam memprediksi produksi hidrogen[10].

Secara keseluruhan, penelitian ini mencakup perancangan sistem elektrolisis air alkalin, pengembangan sistem monitoring berbasis machine learning, akuisisi data operasional secara real-time dengan mikrokontroler esp32, proses pra-pemrosesan data, pembangunan model Random Forest dan LSTM, serta evaluasi performa model menggunakan parameter statistik yang telah ditentukan.

Alur penelitian secara lengkap ditunjukkan pada diagram alir penelitian yang disajikan pada Gambar berikut.



Gambar 2.1

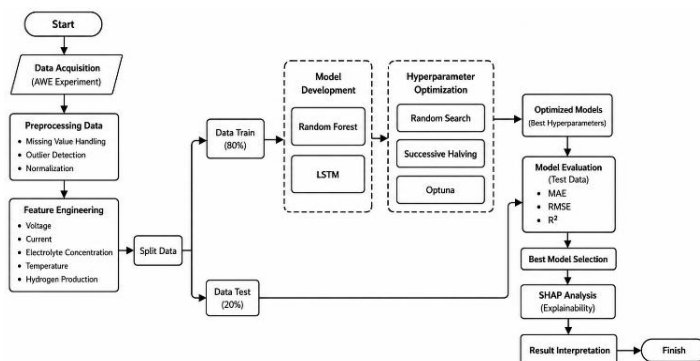


Figure 1. Analysis Procedures

Gambar 2.2

A. Data Acquisition

Sistem Data diperoleh dari sistem elektrolisis secara real-time menggunakan mikrokontroler sebagai perangkat akuisisi data. Parameter yang diukur meliputi tegangan listrik (V), arus listrik (A), suhu elektrolit ($^{\circ}\text{C}$), dan tekanan gas hydrogen. Dataset terdiri atas lebih dari 100.000 sampel yang diperoleh

dengan interval waktu setiap 2 detik selama 20 hari kurang lebih elektrolisis berlangsung selama 4 sampai 5 jam dalam sehari. Analisis statistik deskriptif dilakukan untuk mengetahui rentang operasi masing-masing parameter dan distribusi data target produksi hidrogen. Data kemudian disimpan dalam format digital (csv) untuk keperluan analisis dan pemodelan machine learning.

Tahapan penelitian secara keseluruhan ditunjukkan pada Gambar 3.1 dan dapat dijelaskan sebagai berikut:

Perancangan sistem elektrolisis tahap ini meliputi perancangan sel elektrolisis, pemilihan elektroda yaitu berbahan stainless steel, penentuan jenis elektrolit dengan bahan KoH konsentrasi 25%, serta penempatan sensor- sensor pengukuran. Akuisisi data dalam tahapan ini data diambil secara *real-time* dari sistem elektrolisis dengan menggunakan microcontroller. Data yang dikumpulkan meliputi suhu elektrolit, arus, tegangan, tekanan gas yang nantinya digunakan untuk mengukur laju produksi gas.

Meskipun data diperoleh secara berurutan, hasil analisis menunjukkan bahwa produksi hidrogen lebih dipengaruhi oleh

hubungan sesaat antar parameter operasional dibandingkan ketergantungan temporal jangka panjang. Oleh karena itu, pembagian data dilakukan menggunakan random split 80:20. Untuk menghindari data leakage, proses normalisasi hanya dilakukan menggunakan data latih dan parameter normalisasi yang diperoleh kemudian diterapkan pada data uji.

Preprocessing data merupakan tahapan data hasil akuisisi disimpan dalam bentuk dataset, kemudian dilakukan pembersihan data, normalisasi, dan pemilihan fitur yang relevan sebelum digunakan untuk pelatihan model. Pelatihan model machine learning yaitu model random forest dan LSTM dilatih menggunakan dataset historis untuk mempelajari hubungan antara parameter. Kinerja model dievaluasi menggunakan metrik seperti *Mean Absolute Error* (MAE) dan koefisien determinasi (R^2). Hasil prediksi dianalisis untuk mengetahui kondisi operasi yang memberikan produksi hydrogen secara optimal.

B. Dataset dan Variabel

Dataset disusun dalam bentuk data deret waktu (time-series) yang diperoleh dari hasil akuisisi data sistem elektrolisis. Variabel input yang digunakan meliputi tegangan, arus, suhu, dan tekanan, sedangkan variabel output (target) adalah laju produksi hidrogen.

$$n = PV/RT$$

dengan:

- n = jumlah mol hidrogen (mol)
- P = tekanan gas hidrogen (Pa)
- V = volume tabung penyimpanan (m^3)
- R = konstanta gas ideal ($8,314 \text{ J/mol} \cdot \text{K}$)
- T = temperatur absolut (K)

Jumlah hidrogen yang dihasilkan dihitung menggunakan hukum gas ideal. Nilai tekanan diperoleh dari sensor tekanan, sedangkan temperatur diperoleh dari sensor suhu yang terpasang pada sistem. Dengan mengetahui volume tabung hydrogen 2 liter, jumlah mol hidrogen dapat dihitung

berdasarkan hubungan antara tekanan, volume, dan temperatur.

Sebelum digunakan dalam pelatihan model machine learning, data terlebih dahulu melalui tahap preprocessing yang meliputi pembersihan data dari noise, penanganan data hilang, normalisasi nilai parameter, serta pembagian dataset menjadi data latih dan data uji.

C. Perancangan Model Random Forest

Model Random Forest digunakan untuk memprediksi produksi hidrogen berdasarkan parameter operasional sistem elektrolisis. Model ini bekerja dengan membangun sejumlah pohon keputusan dan menggabungkan hasil prediksi untuk meningkatkan akurasi dan mengurangi overfitting.

Parameter utama yang digunakan dalam perancangan model Random Forest meliputi jumlah pohon (*number of trees*), kedalaman maksimum pohon (*max depth*), dan jumlah fitur acak pada setiap pembelahan. Proses pelatihan model dilakukan menggunakan data latih, sedangkan evaluasi performa model dilakukan menggunakan data uji.

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x)$$

Keterangan:

- $\hat{y}$ = hasil prediksi produksi hidrogen
- N = jumlah pohon keputusan (*decision tree*)
- $T_i(x)$ = hasil prediksi pohon ke- i
- x = variabel input (tegangan, arus, suhu, tekanan)

Model Random Forest Regression pada penelitian ini dibangun menggunakan 100 pohon keputusan (*decision trees*) yang dikombinasikan melalui metode *ensemble learning*. Pemilihan jumlah pohon yang besar bertujuan untuk meningkatkan kemampuan model dalam menangkap hubungan non-linear antara variabel masukan berupa tegangan, arus, temperatur, dan tekanan terhadap variabel keluaran berupa produksi hidrogen. Setiap pohon keputusan dibangun menggunakan subset data dan subset fitur yang dipilih secara acak sehingga mampu mengurangi variansi model serta meminimalkan risiko *overfitting*. Nilai *random_state* sebesar 42 digunakan untuk memastikan konsistensi hasil pelatihan sehingga proses eksperimen dapat direplikasi pada kondisi yang sama. Seluruh proses pelatihan memanfaatkan seluruh inti prosesor yang tersedia melalui pengaturan parameter *n_jobs* = -1 guna meningkatkan efisiensi komputasi pada dataset berskala besar. Dengan konfigurasi tersebut, model Random Forest diharapkan mampu menghasilkan prediksi produksi hidrogen yang akurat serta mendekati data aktual hasil pengukuran sistem elektrolisis air alkalin. Pada algoritma Random Forest Regression, prediksi produksi hidrogen diperoleh dari rata-rata hasil prediksi seluruh pohon keputusan yang dibangun selama proses pelatihan. Setiap pohon menerima subset data dan subset fitur yang berbeda sehingga mampu mempelajari berbagai pola hubungan non-linear antara parameter operasi elektrolisis dan produksi hidrogen. Pada penelitian ini, model Random Forest dikonfigurasi menggunakan 100 pohon keputusan (*n_estimators* = 100) untuk meningkatkan

kemampuan model dalam mempelajari hubungan non-linear antara tegangan, arus, temperatur, tekanan, dan produksi hidrogen. Nilai *random_state* sebesar 42 digunakan untuk memastikan konsistensi hasil pelatihan sehingga eksperimen dapat direproduksi. Selain itu, parameter *n_jobs* diatur sebesar -1 untuk memanfaatkan seluruh inti prosesor yang tersedia selama proses komputasi dan mempercepat waktu pelatihan model. Random Forest dipilih karena mampu memodelkan hubungan non-linear antar parameter operasional elektrolisis tanpa memerlukan asumsi distribusi data tertentu serta relatif tahan terhadap overfitting pada dataset berukuran besar.

D. Perancangan Model LSTM

Model Long Short-Term Memory (LSTM) digunakan untuk memodelkan hubungan temporal pada data deret waktu sistem elektrolisis. Model ini mampu menangkap ketergantungan jangka pendek dan jangka panjang dalam data untuk meningkatkan akurasi prediksi produksi hidrogen.

Arsitektur LSTM pada penelitian ini terdiri dari lapisan LSTM, lapisan *dropout* untuk mengurangi *overfitting*, serta lapisan *dense* sebagai keluaran model. Model LSTM dilatih menggunakan data deret waktu yang telah dinormalisasi, dan performanya dievaluasi menggunakan metrik yang sama dengan model Random Forest.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Forget gate digunakan untuk menentukan informasi historis yang masih relevan dan perlu dipertahankan dalam memori model. Pada penelitian ini, informasi historis berupa data tegangan, arus, temperatur, dan tekanan dari waktu sebelumnya dievaluasi sebelum digunakan untuk memprediksi produksi hidrogen pada waktu berikutnya.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

Persamaan ini menggambarkan proses pembaruan memori (*cell state*) pada jaringan LSTM. Informasi lama yang dianggap penting dipertahankan melalui forget gate, sedangkan informasi baru yang diperoleh dari data operasional elektrolisis ditambahkan ke memori jaringan. Mekanisme ini memungkinkan LSTM mempelajari perubahan kondisi sistem elektrolisis yang terjadi secara bertahap terhadap waktu.

$$h_t = o_t \odot \tanh(C_t)$$

Hidden state yang dihasilkan digunakan sebagai representasi kondisi sistem pada waktu tertentu dan menjadi dasar dalam menghasilkan prediksi produksi hidrogen. Dengan memanfaatkan informasi historis yang tersimpan pada cell state, model dapat menghasilkan prediksi yang lebih mendekati kondisi aktual sistem elektrolisis.

Pada penelitian ini, model Long Short-Term Memory (LSTM) dibangun menggunakan dua lapisan tersembunyi (*hidden layers*) yang masing-masing terdiri atas 32 unit neuron. Arsitektur tersebut dipilih untuk memungkinkan model mempelajari hubungan temporal dari data operasional sistem elektrolisis yang terdiri atas parameter tegangan, arus, temperatur, dan tekanan gas hidrogen. Untuk meningkatkan kemampuan generalisasi model dan mengurangi risiko *overfitting*, pada setiap lapisan LSTM diterapkan teknik *dropout* sebesar 20%. Selanjutnya, hasil ekstraksi fitur temporal diteruskan ke lapisan *dense* yang terdiri atas 16

neuron sebelum menghasilkan satu nilai keluaran berupa prediksi produksi hidrogen.

Pelatihan model dilakukan menggunakan algoritma optimasi Adam dengan fungsi objektif Mean Squared Error (MSE). Jumlah epoch maksimum ditetapkan sebanyak 20 dengan ukuran *batch* sebesar 32 sampel pada setiap iterasi pelatihan. Sebanyak 20% dari data pelatihan digunakan sebagai data validasi untuk mengevaluasi performa model selama proses pembelajaran berlangsung. Selain itu, mekanisme *Early Stopping* diterapkan dengan nilai *patience* sebesar lima epoch untuk menghentikan proses pelatihan secara otomatis apabila tidak terjadi perbaikan pada nilai *validation loss*. Strategi ini digunakan untuk memperoleh model yang memiliki kemampuan generalisasi yang baik terhadap data baru dan menghindari proses pelatihan yang berlebihan. LSTM dipilih karena data operasional diperoleh secara kontinu selama proses elektrolisis berlangsung sehingga secara teoritis berpotensi memiliki karakteristik temporal. Model LSTM diharapkan mampu memanfaatkan informasi historis dari parameter operasi untuk meningkatkan akurasi prediksi produksi hidrogen.

E. Evaluation Metrics

Kinerja model dievaluasi menggunakan beberapa metrik, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). Evaluasi ini digunakan untuk membandingkan performa model Random Forest dan LSTM dalam memprediksi produksi hidrogen. Nilai RMSE digunakan sebagai indikator utama dalam membandingkan akurasi model karena sensitif terhadap kesalahan prediksi dalam skala besar.

F. Pengujian Sistem

Pengujian dilakukan dengan membandingkan hasil prediksi model terhadap data aktual produksi hidrogen. Performa model Random Forest dan LSTM dianalisis berdasarkan metrik evaluasi untuk yang paling akurat.

III. HASIL DAN PEMBAHASAN

A. Hasil Pengolahan Data dan Pelatihan Model

Penelitian ini bertujuan untuk menganalisis kinerja dua model machine learning, yaitu Random Forest dan Long Short-Term Memory (LSTM), dalam memprediksi produksi hidrogen pada sistem elektrolisis berbasis parameter operasional berupa tegangan (V), arus (A), dan suhu ($^{\circ}\text{C}$). Dataset dari hasil parameter operasional alat elektrolisis serta tekanan gas hidrogen pada rentang 0–1 (bar) sesuai dengan spesifikasi sensor gas yang digunakan.

Random Forest menunjukkan performa yang lebih baik karena hubungan antara arus, tegangan, temperatur, tekanan, dan produksi hidrogen cenderung bersifat non-linear namun tidak memiliki ketergantungan temporal yang kuat. Oleh karena itu, pendekatan ensemble berbasis decision tree lebih efektif dalam mempelajari pola hubungan antarvariabel dibandingkan LSTM. Meskipun LSTM dirancang untuk mempelajari pola temporal, karakteristik data pada penelitian ini tidak menunjukkan perubahan waktu yang kompleks. Akibatnya, keunggulan mekanisme memori pada LSTM tidak

dapat dimanfaatkan secara optimal sehingga menghasilkan performa yang lebih rendah dibandingkan Random Forest.

Hasil pelatihan menunjukkan bahwa kedua model berhasil mempelajari pola dari dataset, namun dengan tingkat akurasi yang berbeda. Nilai loss pada model LSTM mengalami penurunan signifikan pada awal epoch dan kemudian stabil, yang mengindikasikan bahwa model telah mencapai konvergensi.

B. Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan tiga metrik utama, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2).

Hasil evaluasi ditampilkan pada Tabel 1.

MAE

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

RMSE

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

Koefisien Determinasi (R^2)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

Nilai MAE digunakan untuk mengukur rata-rata kesalahan absolut prediksi, RMSE digunakan untuk mengevaluasi besar kesalahan prediksi dengan memberikan penalti lebih tinggi pada kesalahan besar, sedangkan R^2 digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data aktual produksi hidrogen. Semakin kecil nilai MAE dan RMSE serta semakin mendekati satu nilai R^2 yang diperoleh, maka semakin baik performa model prediksi yang dihasilkan.

TABEL 3.1
PERBANDINGAN KINERJA MODEL

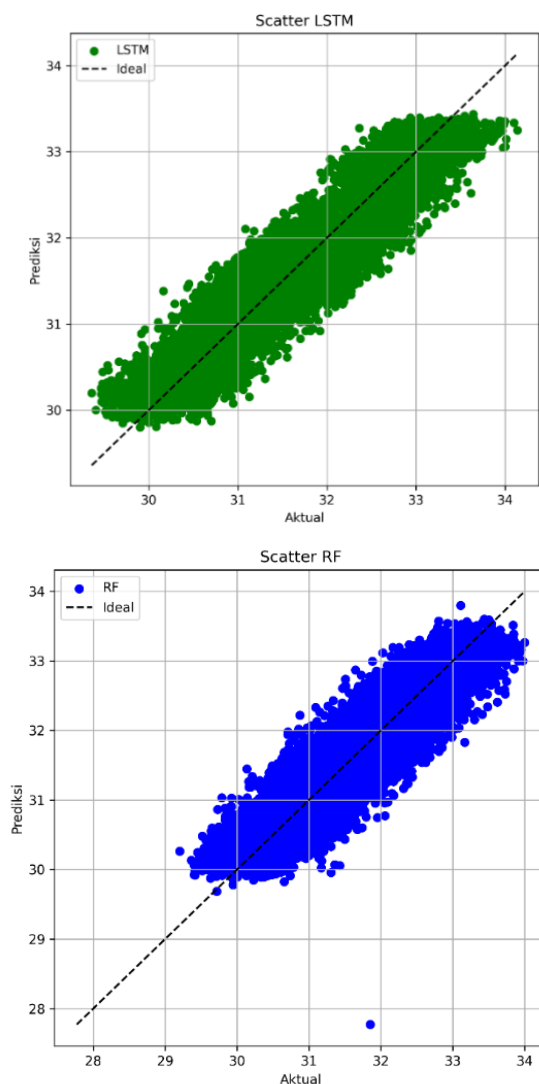
No	Model	MAE	RMSE	R^2
1.	Random Forest	0.0085	0.0106	0.9905
2.	LSTM	0.0302	0.0361	0.7907

Berdasarkan Tabel 1, model Random Forest menunjukkan performa yang jauh lebih baik dibandingkan LSTM. Nilai MAE dan RMSE yang lebih kecil menunjukkan bahwa kesalahan prediksi model Random Forest sangat rendah. Selain itu, nilai R^2 yang mendekati 1 menunjukkan bahwa model mampu menjelaskan hampir seluruh variabilitas data. Sebaliknya, model LSTM memiliki nilai error yang lebih tinggi serta nilai R^2 yang lebih rendah, yaitu sekitar 0.79. Hal ini menunjukkan bahwa meskipun model LSTM mampu menangkap sebagian pola dalam data, namun performanya masih jauh di bawah Random Forest.

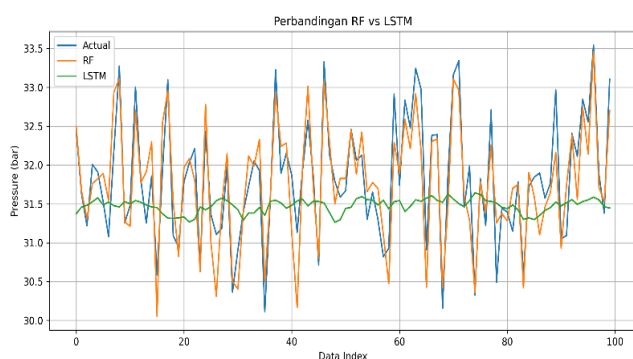
Hasil penelitian ini sejalan dengan penelitian Maghfuri et al. yang menunjukkan bahwa algoritma ensemble learning mampu menghasilkan akurasi tinggi pada prediksi produksi hidrogen. Nilai R^2 sebesar 0,9905 yang diperoleh pada penelitian ini menunjukkan bahwa Random Forest mampu

menjelaskan hampir seluruh variasi data produksi hidrogen. Temuan ini juga memperkuat hasil penelitian Yoon et al. yang melaporkan bahwa Random Forest efektif dalam memodelkan sistem produksi hidrogen yang memiliki hubungan non-linear antarparameter operasi.

C. Analisis Visualisasi Hasil Prediksi



Gambar 3.2 Scatter Random Forest



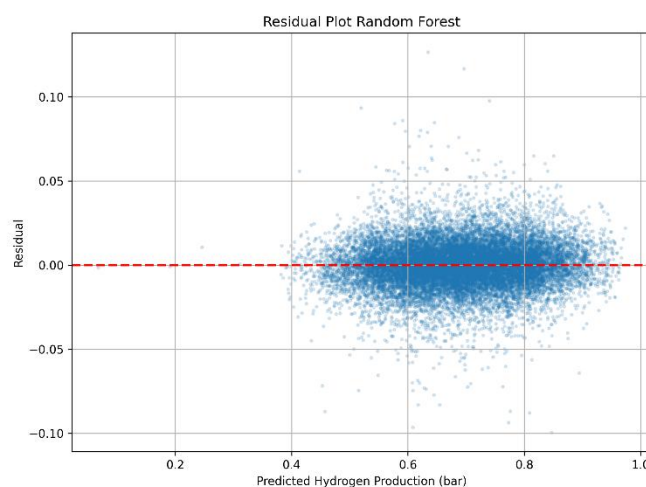
Gambar 3.3 Scatter LSTM

Visualisasi hasil prediksi ditunjukkan pada Gambar perbandingan antara nilai aktual, prediksi Random Forest, dan prediksi LSTM. Berdasarkan grafik tersebut, terlihat bahwa model Random Forest mampu mengikuti fluktuasi data aktual dengan sangat baik, termasuk pada perubahan nilai yang bersifat cepat dan tidak linear.

Sebaliknya, model LSTM cenderung menghasilkan prediksi yang lebih halus (smooth) dan tidak mampu mengikuti perubahan tajam pada data. Hal ini menunjukkan bahwa model LSTM lebih cocok untuk menangkap tren jangka panjang dibandingkan fluktuasi jangka pendek.

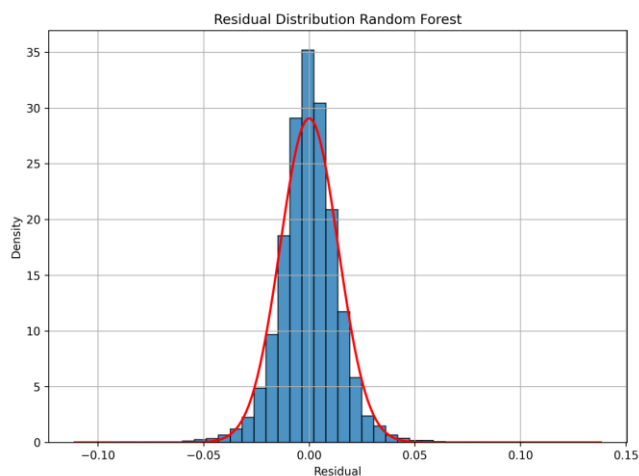
Selain itu, analisis scatter plot antara nilai aktual dan prediksi menunjukkan bahwa titik-titik hasil Random Forest lebih mendekati garis ideal ($y = x$), yang mengindikasikan tingkat akurasi yang tinggi. Sementara itu, distribusi titik pada model LSTM lebih menyebar, menunjukkan adanya deviasi yang lebih besar dari nilai aktual.

D. Analisis Residual dan Distribusi Kesalahan Prediksi



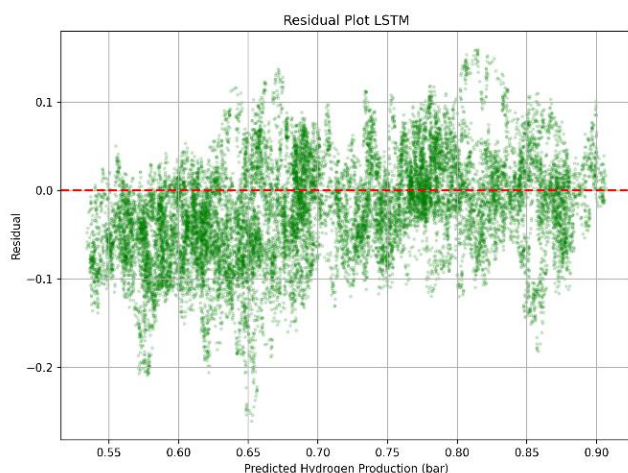
Gambar 3.4 Residual Plot Random Forest

Gambar 3.4 menunjukkan residual plot model Random Forest. Sebagian besar residual terdistribusi secara acak di sekitar garis nol dengan penyebaran yang relatif sempit dan tidak membentuk pola tertentu. Kondisi tersebut menunjukkan bahwa model Random Forest tidak mengalami bias prediksi yang signifikan dan mampu memodelkan hubungan non-linear antara parameter operasional elektrolisis dan produksi hidrogen dengan baik. Sebaran residual yang simetris di sekitar nol juga mengindikasikan bahwa kesalahan prediksi model cenderung kecil dan stabil pada berbagai rentang nilai produksi hidrogen.



Gambar 3.5 Residual Distribution Random Forest

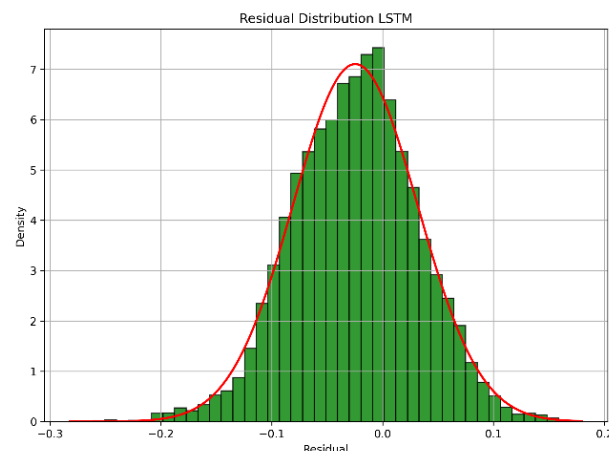
Gambar 3.5 menunjukkan distribusi residual model Random Forest yang terpusat di sekitar nilai nol dan memiliki penyebaran yang relatif sempit. Pola distribusi yang mendekati distribusi normal mengindikasikan bahwa model menghasilkan kesalahan prediksi yang kecil dan tidak menunjukkan kecenderungan overestimation maupun underestimation secara sistematis. Hasil ini mendukung nilai MAE dan RMSE yang rendah serta menunjukkan kemampuan generalisasi model yang baik.



Gambar 3.6 Residual Plot LSTM

Gambar 3.6 menunjukkan residual plot model LSTM dengan penyebaran residual yang lebih lebar dibandingkan Random Forest. Beberapa titik residual memiliki deviasi yang cukup besar dari garis nol dan menunjukkan variasi kesalahan prediksi yang lebih tinggi. Selain itu, terdapat indikasi pola residual yang tidak sepenuhnya acak, yang mengindikasikan bahwa model LSTM belum mampu merepresentasikan seluruh hubungan antara parameter operasional dan produksi hidrogen secara optimal. Kondisi ini menunjukkan bahwa karakteristik data yang digunakan

kemungkinan memiliki dependensi temporal yang lemah sehingga mekanisme memori pada LSTM tidak dapat dimanfaatkan secara maksimal.

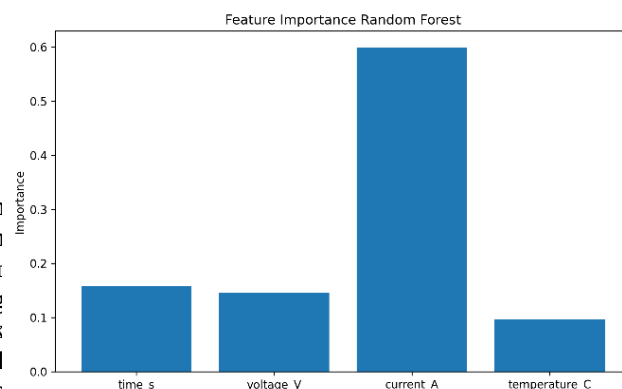


Gambar 3.7 Residual Distribution LSTM

Gambar 3.7 menunjukkan distribusi residual model LSTM yang memiliki rentang penyebaran lebih besar dibandingkan model Random Forest. Distribusi residual cenderung lebih lebar dan menunjukkan adanya beberapa kesalahan prediksi ekstrem. Kondisi tersebut mengindikasikan bahwa model LSTM menghasilkan variasi kesalahan prediksi yang lebih tinggi sehingga kemampuan generalisasi model menjadi lebih rendah. Hasil ini konsisten dengan nilai MAE, RMSE, dan R^2 yang diperoleh pada tahap evaluasi model.

Berdasarkan analisis residual dan distribusi kesalahan prediksi, model Random Forest menunjukkan performa yang lebih stabil dibandingkan model LSTM. Residual Random Forest cenderung terdistribusi secara acak di sekitar nol dengan penyebaran yang lebih kecil, sedangkan model LSTM menghasilkan residual yang lebih lebar dan beberapa kesalahan prediksi ekstrem. Temuan ini menunjukkan bahwa Random Forest lebih efektif dalam menangkap hubungan non-linear antarparameter operasional elektrolisis tanpa memerlukan dependensi temporal yang kuat.

E. Analisis Feature Importance



Gambar 3.8 Feature Importance Random Forest

Gambar 3.8 menunjukkan hasil analisis feature importance pada model Random Forest. Hasil analisis menunjukkan bahwa arus listrik merupakan parameter yang memberikan kontribusi terbesar terhadap prediksi produksi hidrogen, diikuti oleh temperatur elektrolit, tekanan gas hidrogen, dan tegangan sistem. Dominasi parameter arus sejalan dengan Hukum Faraday yang menyatakan bahwa jumlah hidrogen yang dihasilkan berbanding lurus dengan jumlah muatan listrik yang mengalir selama proses elektrolisis. Temperatur berpengaruh terhadap mobilitas ion dan resistansi elektrolit sehingga turut mempengaruhi laju reaksi elektrokimia. Sementara itu, tekanan dan temperatur gas menentukan jumlah mol hidrogen sesuai dengan hukum gas ideal. Kesesuaian antara hasil analisis machine learning dan teori elektrokimia menunjukkan bahwa model Random Forest tidak hanya menghasilkan akurasi prediksi yang tinggi, tetapi juga mampu memberikan interpretasi fisik yang dapat diterima secara ilmiah.

F. Analisis Karakteristik Data dan Model

Perbedaan performa antara Random Forest dan LSTM dapat dijelaskan berdasarkan karakteristik dataset dan sifat masing-masing algoritma. Dataset yang digunakan dalam penelitian ini memiliki hubungan yang lebih bersifat statis antara variabel input dan output, yaitu hubungan langsung antara tegangan, arus, suhu, dan produksi hidrogen.

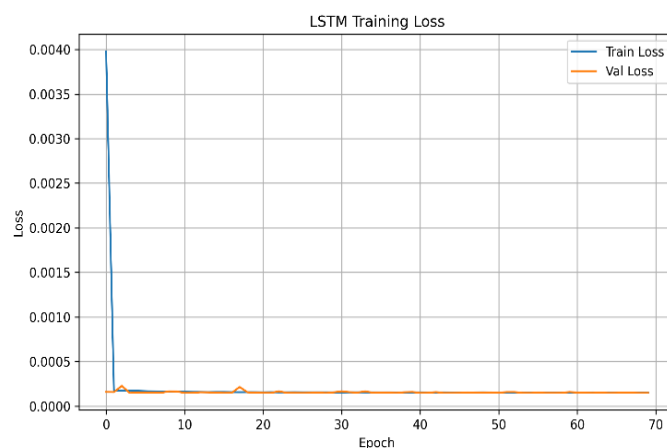
Dalam sistem elektrolisis, produksi hidrogen secara teoritis sangat dipengaruhi oleh arus listrik yang mengalir, sesuai dengan hukum Faraday. Hal ini menunjukkan bahwa hubungan antar variabel cenderung bersifat deterministik dan tidak memiliki dependensi waktu yang kompleks.

Model Random Forest sangat efektif dalam menangkap hubungan non-linear seperti ini karena menggunakan pendekatan ensemble berbasis decision tree. Model ini mampu mengidentifikasi pola kompleks tanpa memerlukan informasi urutan waktu.

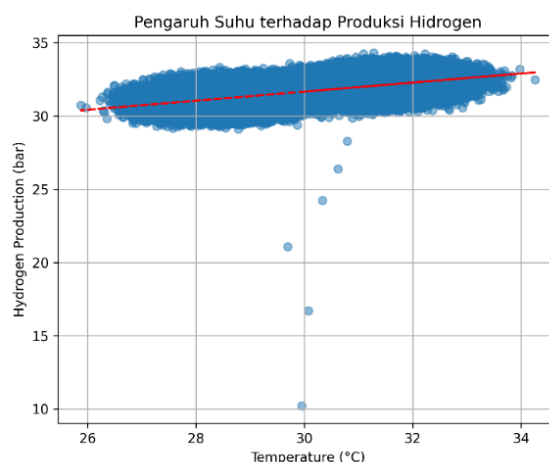
Sebaliknya, model LSTM dirancang untuk menangkap dependensi temporal dalam data time-series. Karena dataset yang digunakan tidak memiliki pola temporal yang kuat, maka LSTM tidak dapat bekerja secara optimal. Hal ini menyebabkan model cenderung menghasilkan prediksi yang mendekati nilai rata-rata, sehingga performanya lebih rendah.

Performa LSTM yang lebih rendah menunjukkan bahwa dataset yang digunakan tidak memiliki dependensi temporal yang kuat. Produksi hidrogen pada sistem elektrolisis lebih ditentukan oleh hubungan non-linear antar parameter operasi saat itu, khususnya arus, temperatur, dan tekanan, dibandingkan informasi historis jangka panjang. Kondisi tersebut menyebabkan mekanisme memori pada LSTM tidak dapat dimanfaatkan secara optimal dan cenderung menghasilkan prediksi yang lebih halus (over-smoothed prediction). Sebaliknya, Random Forest mampu menangkap hubungan non-linear tersebut secara lebih efektif. Meskipun memiliki beberapa keterbatasan, hasil penelitian ini menunjukkan bahwa pendekatan machine learning berbasis Random Forest dan LSTM memiliki potensi untuk

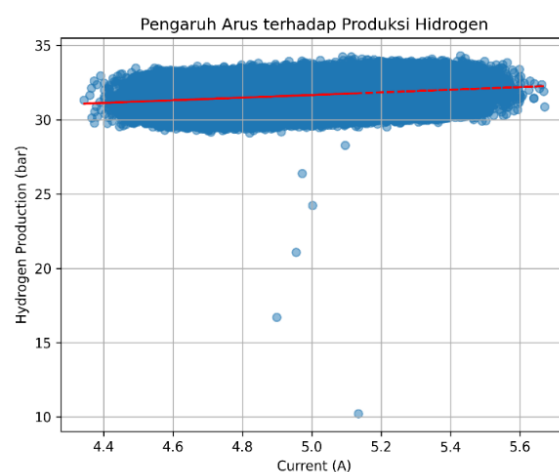
mendukung pengembangan sistem prediksi produksi hidrogen pada proses elektrolisis air alkalin dan dapat menjadi dasar bagi penelitian lanjutan menuju implementasi pada sistem berskala lebih besar.



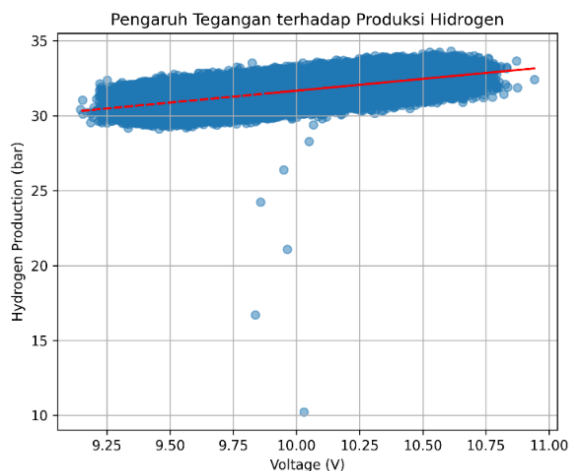
Gambar 3.9 LSTM Training Lost



Gambar 3.10 Pengaruh Suhu Terhadap Produksi Hidrogen



Gambar 3.11 Pengaruh Arus Terhadap Produksi Hidrogen



Gambar 3.12 Pengaruh Tegangan Terhadap Produksi Hidrogen

G. Implikasi terhadap Sistem Elektrolisis Real-Time

Hasil penelitian ini menunjukkan bahwa model Random Forest memiliki potensi yang sangat baik untuk diterapkan dalam sistem elektrolisis secara real-time. Keunggulan utama model ini adalah kecepatan prediksi yang tinggi, akurasi yang sangat baik, serta kompleksitas komputasi yang relatif rendah. Dalam implementasi nyata, model ini dapat digunakan sebagai bagian dari sistem monitoring cerdas untuk memprediksi produksi hidrogen berdasarkan parameter operasional yang diukur secara langsung menggunakan sensor.

Sementara itu, penggunaan LSTM dalam sistem ini memerlukan pertimbangan lebih lanjut, terutama terkait kebutuhan data time-series yang lebih kompleks serta beban komputasi yang lebih tinggi. Oleh karena itu, dalam konteks sistem elektrolisis yang memiliki karakteristik statis, model Random Forest lebih direkomendasikan.

Penelitian ini masih memiliki beberapa keterbatasan. Data yang digunakan diperoleh dari satu sistem elektrolisis air alkalin dengan konsentrasi elektrolit dan konfigurasi elektroda yang tetap, sehingga kemampuan generalisasi model terhadap kondisi operasi yang berbeda masih perlu diuji lebih lanjut. Selain itu, variabel masukan yang digunakan terbatas pada tegangan, arus, temperatur, dan tekanan gas hidrogen. Penelitian selanjutnya dapat menambahkan parameter lain seperti konsentrasi elektrolit, laju aliran elektrolit, serta karakteristik elektroda untuk meningkatkan akurasi dan kemampuan generalisasi model prediksi.

H. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil yang diperoleh. Pertama, dataset yang digunakan diperoleh dari satu sistem elektrolisis air alkalin (alkaline water electrolysis) skala laboratorium dengan konfigurasi elektroda dan konsentrasi elektrolit yang tetap. Kondisi tersebut menyebabkan kemampuan generalisasi model terhadap sistem elektrolisis

dengan konfigurasi dan kondisi operasi yang berbeda masih perlu dikaji lebih lanjut. Kedua, variabel masukan yang digunakan dalam penelitian ini terbatas pada tegangan, arus, temperatur, dan tekanan gas hidrogen sehingga belum sepenuhnya merepresentasikan seluruh fenomena fisik dan elektrokimia yang terjadi selama proses elektrolisis.

Selain itu, penelitian ini belum mempertimbangkan pengaruh parameter lain seperti konsentrasi KOH, konduktivitas elektrolit, karakteristik elektroda, efisiensi Faraday, laju aliran gas, serta degradasi elektroda yang berpotensi memengaruhi produksi hidrogen. Model yang dikembangkan juga belum divalidasi pada sistem elektrolisis berskala industri maupun pada variasi kondisi operasi yang lebih luas, seperti perubahan beban, perubahan temperatur lingkungan, dan kondisi operasi dinamis. Oleh karena itu,

penelitian selanjutnya perlu mengembangkan dataset yang lebih beragam, melibatkan sistem elektrolisis dengan skala yang lebih besar, serta mengevaluasi algoritma machine learning pada berbagai kondisi operasi untuk meningkatkan kemampuan generalisasi dan potensi implementasi model pada sistem produksi hidrogen secara nyata.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem prediksi produksi hidrogen berbasis data operasional nyata pada sistem elektrolisis air alkalin menggunakan algoritma Random Forest dan Long Short-Term Memory (LSTM). Berdasarkan hasil pengujian, Random Forest menunjukkan performa terbaik dengan nilai MAE sebesar 0,0085, RMSE sebesar 0,0106, dan R^2 sebesar 0,9905. Sementara itu, LSTM memperoleh nilai MAE sebesar 0,0302, RMSE sebesar 0,0361, dan R^2 sebesar 0,7907. Hasil tersebut menunjukkan bahwa Random Forest lebih efektif dalam memodelkan hubungan non-linear antarparameter operasi elektrolisis dibandingkan LSTM pada dataset yang digunakan. Penelitian selanjutnya dapat difokuskan pada pengujian berbagai konsentrasi elektrolit, pengembangan model ensemble yang lebih kompleks, serta penerapan sistem prediksi secara langsung pada platform IoT untuk mendukung monitoring hidrogen secara real-time.

DAFTAR PUSTAKA

- [1] H. Tüysüz, "Alkaline Water Electrolysis for Green Hydrogen Production," *Acc. Chem. Res.*, Feb. 2024, doi: 10.1021/acs.accounts.3c00709.
- [2] S. Dash *et al.*, "Advances in green hydrogen production through alkaline water electrolysis: A comprehensive review," *Int. J. Hydrogen Energy*, vol. 83, pp. 614–629, Sep. 2024, doi: 10.1016/j.ijhydene.2024.08.157.
- [3] L. P. Campo Schneider *et al.*, "Advancements in hydrogen production using alkaline electrolysis systems: A short review on experimental and simulation studies," *Curr. Opin. Electrochem.*, vol. 47, p. 101552, Oct. 2024, doi: 10.1016/j.coelec.2024.101552.
- [4] Z. Ibnu Faqih, E. Kurniawan, and U. Kurniawan Usman, "SYNERGY Jurnal Ilmiah Multidisiplin Penerapan Iot Sebagai Monitoring Jarak Jauh Untuk Mengetahui Proses Elektrolisis Air Mineral," vol. 1, no. 2, pp. 70–74, 2023, [Online]. Available: <https://ejournal.naurendigiton.com/index.php/sjim>
- [5] A. M. Maghfuri and M. Kuku, "Optimized hydrogen production through

- machine learning: comparative analysis of electrolyzer technologies using hybrid renewable energy,” *Appl. Therm. Eng.*, vol. 276, Oct. 2025, doi: 10.1016/j.applthermaleng.2025.126923.
- [6] R. Zhang, A. Xie, L. Cheng, Z. Bai, Y. Tang, and P. Wan, “Hydrogen production by traditional and novel alkaline water electrolysis on nickel or iron based electrocatalysts,” *Chemical Communications*, vol. 59, no. 53, pp. 8205–8221, 2023, doi: 10.1039/D3CC00996C.
- [7] M. Muthiah, M. Elnashar, W. Afzal, and H. Tan, “Safety assessment of hydrogen production using alkaline water electrolysis,” *Int. J. Hydrogen Energy*, vol. 84, pp. 803–821, Sep. 2024, doi: 10.1016/j.ijhydene.2024.08.237.
- [8] Y. Zuo *et al.*, “High-performance alkaline water electrolyzers based on Ru-perturbed Cu nanoplatelets cathode,” *Nat. Commun.*, vol. 14, no. 1, p. 4680, Aug. 2023, doi: 10.1038/s41467-023-40319-5.
- [9] Q. He, M. Zhao, S. Li, X. Li, and Z. Wang, “Machine Learning Prediction of Photovoltaic Hydrogen Production Capacity Using Long Short-Term Memory Model,” *Energies (Basel)*, vol. 18, no. 3, p. 543, Jan. 2025, doi: 10.3390/en18030543.
- [10] W. Xu, J. Tu, N. Xu, and Z. Liu, “Predicting daily heating energy consumption in residential buildings through integration of random forest model and meta-heuristic algorithms,” *Energy*, vol. 301, p. 131726, Aug. 2024, doi: 10.1016/j.energy.2024.131726.
- [11] H. Yadav and A. Thakkar, “NOA-LSTM: An efficient LSTM cell architecture for time series forecasting,” *Expert Syst. Appl.*, vol. 238, p. 122333, Mar. 2024, doi: 10.1016/j.eswa.2023.122333.
- [12] N. Hamed Al Hajri, R. Naji Al Harthi, G. K. Pasam, and R. Natarajan, “IoT and Machine Learning based Green Energy Generation using Hybrid Renewable Energy Sources of Solar, Wind and Hydrogen Fuel Cells,” *E3S Web of Conferences*, vol. 472, p. 01008, Jan. 2024, doi: 10.1051/e3sconf/202447201008.
- [13] G. S. Cassol *et al.*, “Ultra-fast green hydrogen production from municipal wastewater by an integrated forward osmosis-alkaline water electrolysis system,” *Nat. Commun.*, vol. 15, no. 1, p. 2617, Mar. 2024, doi: 10.1038/s41467-024-46964-8.
- [14] W. Luo *et al.*, “Research on the method to improve the efficiency of ScCO₂ fracturing and the mechanism of proppant transport and drag-reducing,” *Energy*, vol. 319, p. 135012, Mar. 2025, doi: 10.1016/j.energy.2025.135012.
- [15] S. Ali Akbar, M. Azis Sidiq, A. Yudhana, and R. Rian Hidayat, “Integrasi Sensor Pintar dengan Sistem IoT untuk Memantau Kadar Gas Nitrogen pada Reaktor Biogas,” *Jurnal Pekommas*, vol. 10, no. 1, Jun. 2025, doi: 10.56873/jpkm.v9i1.5918.
- [16] T. Sutikno, H. S. Purnama, A. Pamungkas, A. Fadlil, I. M. Alsofyani, and M. H. Jopri, “Internet of things-based photovoltaics parameter monitoring system using NodeMCU ESP8266,” *International Journal of Electrical and Computer Engineering*, vol. 11, no. 6, pp. 5578–5587, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5578-5587.
- [17] “Internet of Things and Cloud-Based Monitoring of an Electrolysis and Fuel Cell System,” *J. Sci. Ind. Res. (India)*, vol. 84, no. 7, Jul. 2025, doi: 10.56042/jsir.v84i7.14515.
- [18] S. Hari mukti, A. Pujiyanta, and A. Fadlil, “Klasifikasi Berat Itik Pedaging Berbasis Convolutional Neural Network dan Internet of Things,” *Jurnal Riset Rekayasa Elektro*, vol. 7, no. 2, pp. 223–236, Dec. 2025, doi: 10.30595/jrre.v7i2.27847.
- [19] J. Yoon, D.-Y. Cheong, and G. Baek, “Predicting current and hydrogen productions from microbial electrolysis cells using random forest model,” *Appl. Energy*, vol. 371, p. 123641, Oct. 2024, doi: 10.1016/j.apenergy.2024.123641.