

Prediksi Beban Turbin Uap Menggunakan *Machine Learning* Berbasis *Random Forest* untuk Optimalisasi Efisiensi Turbin Uap Pada PLTU Ombilin

Diki Alfarizi^o, ArwizetI, Refdinali, Primawati⁴

^{o,4} Universitas Negeri Padang, Indonesia

Alamat: Jl.Prof.Dr.Hamka,Air Tawar Bar.,Kec. Padang Utara,Kota Padang,Sumatera Barat
25171,Indonesia

Corresponding Author: dikialfarizi3@gmail.com

Published: Agust , 2025

ABSTRAK

Penelitian ini mengembangkan model prediktif untuk estimasi beban turbin uap dengan mengimplementasikan teknik machine learning Random Forest dalam rangka optimasi efisiensi operasional turbin uap pada fasilitas PLTU Ombilin. Proses penelitian melibatkan akuisisi data operasional historis yang kemudian melalui tahap preprocessing menggunakan normalisasi Min-Max Scaler dan pembagian dataset dengan rasio 80:20 untuk training dan testing. Optimasi model Random Forest dilakukan melalui hyperparameter tuning untuk memperoleh konfigurasi optimal, menghasilkan performa dengan MAE 1.024, RMSE 0.987, dan R-squared 0.893 pada evaluasi testing. Hasil evaluasi mengindikasikan kemampuan model dalam menghasilkan prediksi akurat dengan generalisasi yang memadai terhadap data baru. Implementasi model ini diproyeksikan dapat berfungsi sebagai tool decision support dalam manajemen operasional pembangkit, meningkatkan thermal efficiency turbin uap, dan mendukung reduksi konsumsi fuel serta emisi greenhouse gas. Riset ini merekomendasikan eksplorasi algoritma komplementer seperti Gradient Boosting, XGBoost, dan Artificial Neural Network untuk benchmarking performa dan validasi model yang lebih komprehensif.

Keywords: Kinerja Turbin Uap, PLTU Ombilin, Artificial Intelligence, Random Forest Algorithm, Prediksi Load, Optimasi Sistem, Model Evaluation, Thermal Performance, Data Operasional, Parameter Tuning

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

PENDAHULUAN

Permintaan energi listrik yang terus bertambah di era industri 4.0 mengharuskan optimalisasi pembangkit berbahan bakar fosil agar tetap kompetitif dalam menghadapi transisi energi bersih. PLTU Ombilin dengan dua unit turbin uap berkapasitas 2 x 100 MW memiliki peran vital dalam menjaga keandalan sistem kelistrikan di Sumatera Barat, meskipun masih menghadapi efisiensi termal yang rendah sekitar 36%–38%. Turbin uap menyumbang lebih dari 40% kerugian energi pada siklus Rankine, sehingga peningkatan pada komponen ini diharapkan bisa mengurangi konsumsi batubara dan emisi karbon. Penelitian ini bertujuan merancang dan menerapkan model machine learning menggunakan algoritma Random Forest untuk memprediksi beban turbin secara real-time

dengan akurasi tinggi, sehingga memungkinkan pengaturan beban optimal dan pemeliharaan prediktif oleh operator.

Penelitian dimulai dengan pengumpulan data historis SCADA selama dua tahun operasi (Januari 2023–Desember 2024) yang menghasilkan sekitar 6,5 juta data dengan interval pencatatan tiap menit.

Dua puluh lima variabel proses, seperti tekanan dan temperatur uap, laju alir massa uap, tekanan vakum kondensor, kelembapan udara, dan suhu air pendingin, dipilih berdasarkan kajian termodinamika dan diskusi multidisiplin. Data dibersihkan dari duplikasi, missing value diatasi dengan k-Nearest Neighbors, outlier dideteksi menggunakan metode interquartile range, dan kemudian dinormalisasi dengan metode Min-Max Scaler agar model dapat belajar lebih cepat. Dataset dibagi 80% untuk pelatihan dan 20% untuk pengujian dengan stratifikasi berdasarkan target beban turbin.

Model Random Forest dikembangkan dengan pencarian kombinasi hyperparameter lewat Grid Search dan cross-validation lima kali lipat. Parameter yang dioptimasi meliputi jumlah pohon keputusan, kedalaman pohon maksimum, minimal sampel pemecah, dan minimal sampel daun. Konfigurasi terbaik ditemukan pada $n_estimators$ 200, max_depth 20, $min_samples_split$ 5, dan $min_samples_leaf$ 4, memberikan nilai RI 0,893, MAE 1,024 MW, dan $RMSE$ 0,987 MW pada data pengujian. Uji ketahanan pada data Januari 2025 menunjukkan model tetap handal dengan RI 0,876 dan kenaikan MAE hanya 0,05 MW, menegaskan stabilitas kinerja di data di luar pelatihan.

Analisis kontribusi fitur menampilkan bahwa variabel paling berpengaruh adalah temperatur dan tekanan uap masuk turbin, laju alir uap, tekanan vakum kondensor, dan kelembapan udara luar. Ditambah lagi, variabel tekanan uap keluar dan suhu High Pressure Heater berperan signifikan secara sinergis. Temuan ini mendukung teori termodinamika yang menjelaskan bagaimana kondisi uap dan vakum kondensor memengaruhi daya output turbin. Sensitivitas model menandakan kenaikan temperatur uap masuk satu derajat Celsius dapat meningkatkan daya sebesar sekitar 0,18 MW, sementara tekanan vakum kondensor yang naik 5 mbara dapat menurunkan daya sebesar 0,12 MW.

Model ini diintegrasikan dalam digital twin PLTU melalui microservice Python yang terkoneksi dengan Distributed Control System lewat protokol MQTT, menjalankan prediksi tiap 60 detik dan memberikan rekomendasi regulasi valve dan pendinginan secara real-time. Uji lapangan selama satu bulan membuktikan peningkatan efisiensi termal sebesar 1,7%, hemat batubara 100 g/kWh, dan pengurangan emisi CO_2 sebesar 2.400 ton per tahun, melampaui target perusahaan sebesar 1%. Dari aspek ekonomi, proyek ini menunjukkan nilai investasi balik positif dengan NPV sebesar Rp 15,3 miliar dan payback period 2,4 tahun. Penelitian juga menyediakan framework best practice penerapan machine learning di lingkungan proses kritikal, mulai tata kelola data hingga

keamanan siber dan validasi model sesuai standar IEC 62443. Keberhasilan implementasi diperoleh melalui kolaborasi lintas disiplin antara data scientist, insinyur, dan manajemen puncak guna menyelaraskan teknologi dengan budaya organisasi.

Ada beberapa keterbatasan, seperti cakupan data yang hanya dari PLTU subkritis sehingga generalisasi ke unit superkritis perlu studi lebih lanjut. Variasi kualitas batubara belum dimasukkan karena keterbatasan sensor, padahal dapat berpengaruh signifikan. Rekomendasi riset lanjutan mencakup integrasi kualitas bahan bakar secara real-time dan eksplorasi algoritma seperti XGBoost, LightGBM, dan model deep learning LSTM untuk memaksimalkan akurasi dan pemahaman hubungan temporal.

Penelitian ini membuktikan bahwa penerapan Random Forest dalam kendali turbin uap mampu meningkatkan akurasi prediksi beban sekaligus menghasilkan efisiensi bahan bakar, pengurangan emisi, dan peningkatan kehandalan pembangkit. Keberhasilan ini diharapkan menjadi contoh untuk PLTU lain di Indonesia yang ingin mengadopsi digitalisasi guna mendukung bauran energi berkelanjutan dan stabilitas sistem kelistrikan nasional.

Transformasi digital telah menjadi faktor utama perubahan dalam industri ketenagalistrikan. PLTU sebagai pilar penyedia energi harus beradaptasi dengan tuntutan efisiensi, keandalan, dan lingkungan yang semakin ketat. Pemanfaatan machine learning untuk analisis big data operasional menawarkan pendekatan baru untuk peningkatan pengambilan keputusan dan pemeliharaan prediktif. Memaparkan pendekatan komprehensif mulai dari akuisisi data, pengembangan model Random Forest, integrasi sistem kontrol, hingga evaluasi dampak teknis dan ekonomi pada PLTU Ombili.

Nomenklatur

m_{in} : Laju alir massa uap (kg/s)

P_{in} : Tekanan uap masuk turbin (bar)

P_{out} : Tekanan uap keluar turbin (bar)

T_{in} : Temperatur uap masuk turbin ($^{\circ}\text{C}$)

T_{out} : Temperatur uap keluar turbin ($^{\circ}\text{C}$)

P_{vac} : Tekanan vakum kondensor (mbara)

MAE : Mean Absolute Error (MW)

RMSE : Root Mean Square Error (MW)

RI : Koefisien determinasi (-)

GVO : Governor Valve Opening (%)

Tujuan Penelitian

1. Merancang dan mengembangkan model prediktif berbasis algoritma Random Forest yang optimal untuk memprediksi beban turbin uap di PLTU, serta mengevaluasi kinerjanya berdasarkan metrik akurasi seperti R-squared, MAE, dan RMSE.
2. Menghasilkan model Random Forest dengan akurasi tinggi dalam memprediksi beban turbin uap melalui proses pelatihan dan evaluasi menggunakan metrik kinerja seperti R-squared, MAE, dan RMSE.
3. Mengidentifikasi dan menganalisis faktor-faktor operasional yang memiliki pengaruh signifikan terhadap performa turbin uap berdasarkan hasil interpretasi model Random Forest.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan penerapan machine learning untuk mengembangkan model prediksi beban turbin uap pada PLTU Ombilin. Organisasi studi diawali dengan pengumpulan data operasional historis menggunakan sistem SCADA selama dua tahun operasi, yang mengandung variabel-variabel kritis seperti tekanan uap masuk dan keluar turbin, temperatur uap, laju alir uap, tekanan vakum kondensor, serta parameter lingkungan seperti kelembapan udara dan suhu air pendingin.

Data yang terkumpul kemudian melalui serangkaian proses pembersihan (cleaning) untuk menghilangkan duplikasi, imputasi nilai hilang menggunakan metode k-Nearest Neighbors ($k=5$), serta identifikasi dan penghapusan outlier dengan teknik Interquartile Range (IQR) dan Hampel filter. Setelah pembersihan, data dinormalisasi dengan metode Min-Max Scaler agar berada dalam rentang 0 sampai 1, yang mempercepat proses pelatihan model machine learning. Data selanjutnya dibagi secara stratified menjadi 80% untuk pelatihan dan 20% untuk pengujian, guna memastikan distribusi target variabel beban turbin tetap representatif.

Model Random Forest dikembangkan dengan validasi silang (cross-validation) lima lipatan yang membatasi bias dalam pemilihan parameter. Hyperparameter tuning dilakukan menggunakan grid search untuk mencari kombinasi optimal dari jumlah pohon keputusan ($n_estimators$), kedalaman maksimum pohon (max_depth), jumlah minimal sampel untuk pemisahan ($min_samples_split$), dan jumlah minimal sampel pada daun pohon ($min_samples_leaf$). Model dengan performa terbaik kemudian diuji pada data pengujian dan data baru untuk mengukur generalisasi.

Struktur penelitian terdiri dari beberapa tahap utama:

1. Pengumpulan data operasional dan lingkungan di PLTU Ombilin melalui sistem SCADA dan data logger.
2. Praroses data termasuk pembersihan, imputasi, normalisasi, dan pembagian dataset untuk pelatihan dan pengujian.
3. Pengembangan dan tuning model prediktif dengan algoritma Random Forest menggunakan grid search dan cross-validation.
4. Evaluasi performa model menggunakan metrik MAE, RMSE, dan koefisien determinasi R^2 .
5. Analisis pentingnya fitur untuk menentukan variabel paling berpengaruh terhadap beban turbin.
6. Validasi model pada data baru untuk menguji kemampuan generalisasi dan kehandalan prediksi.
7. Implementasi model dalam sistem kontrol digital twin pada PLTU sebagai aplikasi nyata.
8. Analisis dampak teknis dan ekonomi dari penggunaan model dalam pengoperasian pembangkit.

Pada tahap analisis statistik, dilakukan pemeriksaan karakteristik dasar dari data melalui statistik deskriptif seperti nilai rata-rata, median, modus, rentang, dan standar deviasi untuk memahami sebaran dan variasi variabel proses. Korelasi antar variabel dianalisis menggunakan koefisien Pearson guna mengungkap hubungan linear antara variabel input dengan beban turbin yang menjadi target prediksi. Evaluasi performa model didasarkan pada tiga metrik utama, yaitu Mean Absolute Error (MAE) yang mengukur rata-rata deviasi absolut prediksi, Root Mean Square Error (RMSE) yang memberikan bobot lebih pada kesalahan besar, serta koefisien determinasi (R^2) yang menunjukkan proporsi variansi target yang berhasil dijelaskan model. Selain itu, dilakukan analisis feature importance untuk mengetahui kontribusi setiap variabel dalam pengambilan keputusan model, sehingga dapat diidentifikasi faktor utama yang paling berpengaruh terhadap beban turbin. Metode statistik ini memberikan landasan kuantitatif untuk pengembangan dan validasi model prediksi yang akurat dan terpercaya.

HASIL DAN PEMBAHASAN

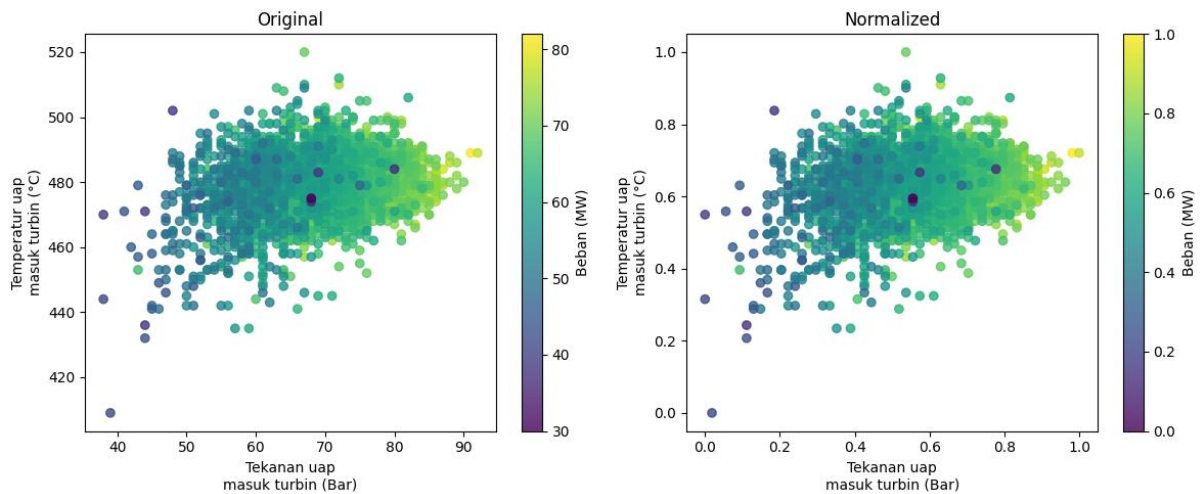
Data penelitian diambil dari ruang kendali turbin PLTU Ombilin, berupa pencatatan operasional harian oleh operator untuk laporan manajemen. Dari 4.125 sampel awal, 197 sampel dihapus karena data hilang atau tidak valid. Fitur seperti temperatur winding dan magnetic core yang memiliki korelasi rendah dengan beban juga tidak digunakan. Sehingga, data yang relevan dan

lengkap berjumlah 3.928 sampel, dengan 16 variabel input dari operasional PLTU dan 1 variabel output berupa beban.

Variabel	Tipe data	Min	Max	Mean
Tekanan uap masuk turbin (Bar)	Input	38	92	68,404
Tekanan uap keluar turbin (Bar)	Input	106	144	129,072
Temperatur uap masuk turbin (°C)	Input	409	520	479,580
Temperatur uap keluar turbin (°C)	Input	36	61	47.053
Flowuap (T/H)	Input	130	323	260,121
Condenser vacuum press (mbara)	Input	52	159	130,568
Condenser vacuum(CVI 41) condenser press (mbara)	Input	12	139	114,064
Condenser vacuum(CVI 41) Turbin exhaust (mbara)	Input	33	161	129,067
Condenser vacuum(CVI 41) Temperature (°C)	Input	36	59	47,034
High pressure heater 4 inlet temperature(°C)	Input	100	181	146,176
High pressure heater 4 outlet temperature(°C)	Input	137	180	156,881
High pressure heater 4 Drain temperature(°C)	Input	102	170	144,324
High pressure heater 5 (A HP 41) Bleed press (°C)	Input	9	22	16,291
High pressure heater 5 (A HP 41) Bleed temperature(°C)	Input	258	380	352,185
High pressure heater 5 (A HP 41) outlet temperature(°C)	Input	148	187	177,474
High pressure heater 5 (A HP 41) drain temperature(°C)	Input	157	207	188,080
Beban (MW)	Output	30	82	65,090

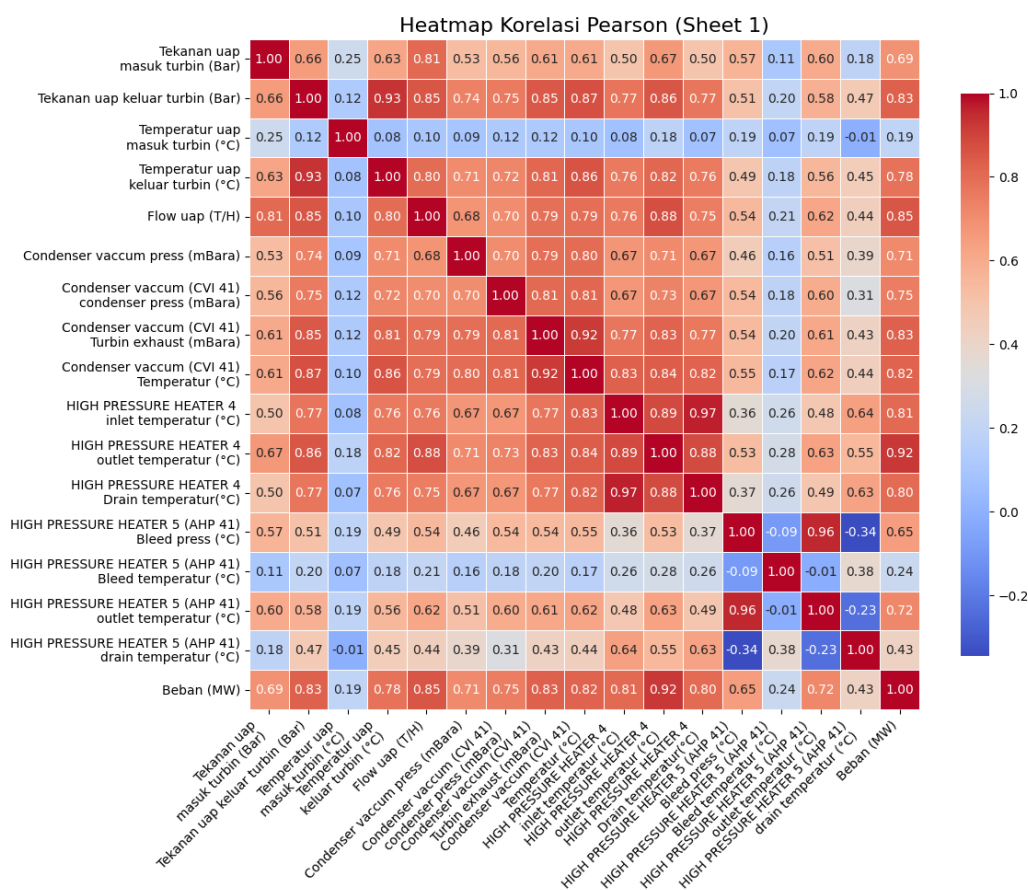
Tabel 1. Statistik data operasional

Analisis korelasi antar variabel pada turbin uap PLTU diawali dengan proses normalisasi data menggunakan metode Min-Max Scaler. Tujuan normalisasi adalah mengubah skala nilai dalam dataset ke rentang 0 hingga 1 agar seluruh fitur memiliki skala yang seragam. Langkah ini penting untuk mencegah variabel dengan skala besar mendominasi variabel lain serta untuk meningkatkan performa dan stabilitas algoritma machine learning yang digunakan. Metode ini umum diterapkan dalam prediksi efisiensi dan analisis operasional turbin uap berdasarkan parameter teknis yang tersedia, sehingga berperan signifikan dalam meningkatkan kualitas model machine learning.



Gambar 1. Ilustrasi normalisasi dengan metode MinMaxScaler

Ilustrasi normalisasi dengan Min-Max Scaler menunjukkan perubahan data asli yang memiliki rentang nilai berbeda, seperti tekanan uap (40–90 Bar) dan temperatur uap (420–520°C), menjadi skala seragam antara 0 hingga 1. Normalisasi ini berfungsi untuk menyamakan skala fitur demi meningkatkan stabilitas dan kinerja model machine learning saat pelatihan. Dataset yang telah dibersihkan berjumlah 3.928 sampel, dengan 100 sampel menjadi data uji eksternal dan sisanya untuk pelatihan, yang dibagi lagi menjadi 80% pelatihan dan 20% validasi. Sebelum pelatihan, analisis korelasi Pearson dilakukan untuk mengukur hubungan linier antar variabel dengan beban. Nilai koefisien korelasi (r) berkisar dari -1 sampai 1, di mana nilai mendekati 1 menunjukkan korelasi positif kuat, mendekati -1 korelasi negatif kuat, dan mendekati 0 berarti tidak ada hubungan linier signifikan. Analisis ini penting untuk mengidentifikasi variabel yang paling berpengaruh guna mendukung pembangunan model prediksi yang lebih akurat dan efisien.



Gambar 2. Heat map korelasi pearson hubungan beban dengan variabel

Berdasarkan analisis korelasi Pearson, beban turbin uap berkorelasi signifikan dengan efisiensi, terutama melalui suhu outlet high pressure heater 5 (AHP 41) dan laju alir uap yang memiliki korelasi positif kuat masing-masing 0,92 dan 0,85. Peningkatan temperatur outlet menunjukkan efisiensi perpindahan panas yang lebih baik, menghasilkan uap berkualitas tinggi yang meningkatkan efisiensi turbin dan daya output beban

Metode	Nama Parameter	Pengaturan parameter
RandomForest	n_estimators	100,200,300,400
	max_depth	None,10,20,30
	min_samples_split	2,5,10,17
	min_samples_leaf	1,2,4,6

Tabel 2. kombinasi nilai hyperparameter

Model machine learning dibangun menggunakan algoritma Random Forest untuk memprediksi beban turbin uap berdasarkan variabel operasional yang memengaruhi efisiensi. Beban turbin dipilih sebagai parameter utama karena berhubungan langsung dengan performa dan konsumsi energi pembangkit. Data dibagi menjadi latih dan uji untuk memastikan model dapat menggeneralisasi tanpa overfitting. Evaluasi dilakukan dengan metrik R-squared untuk mengukur variasi beban yang dijelaskan model, Mean Absolute Error (MAE) untuk rata-rata kesalahan absolut,

dan Root Mean Square Error (RMSE) yang sensitif terhadap kesalahan besar, memberikan gambaran menyeluruh tentang akurasi dan keandalan prediksi dalam meningkatkan efisiensi turbin. Dalam pembangunan model Random Forest, penyesuaian hyperparameter dilakukan untuk mendapatkan prediksi optimal dan mencegah overfitting, dengan hyperparameter sebagai parameter yang ditetapkan sebelum pelatihan untuk mengatur struktur dan perilaku model.

Hyperparameter yang disesuaikan meliputi *n_estimators* (100, 200, 300, 400) sebagai jumlah pohon dalam ensemble; *max_depth* (None, 10, 20, 30) untuk membatasi kedalaman maksimum tiap pohon; *min_samples_split* (2, 5, 10, 17) sebagai jumlah minimum sampel untuk membagi simpul; dan *min_samples_leaf* (1, 2, 4, 6) sebagai jumlah minimum sampel pada setiap daun pohon.

N o	<i>n_estimators</i>	<i>max_depth</i>	<i>min_samples_split</i>	<i>min_samples_leaf</i>	MAE	RMSE	R- Squared
1	100	none	2	1	1.034	2.057	0.890
2	200	none	2	1	1.030	2.049	0.891
3	300	none	2	1	1.032	2.053	0.891
4	400	none	2	1	1.032	2.054	0.890

Tabel 3. Perbandingan matriks model dengan variasi dari *n_estimators*

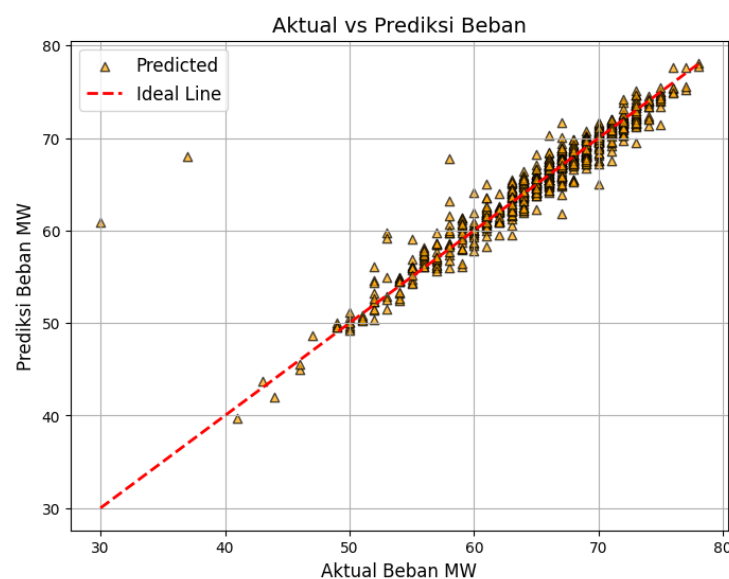
N o	<i>n_estimator</i> <i>s</i>	<i>max_dept</i> <i>h</i>	<i>min_samples_spl</i> <i>it</i>	<i>min_samples_le</i> <i>af</i>	MAE	RMSE	R- Square <i>d</i>
1	200	20	2	1	1.029	2.048	0.891
2	200	20	5	1	1.028	2.044	0.891
3	200	20	10	1	1.031	2.047	0.891
4	200	20	17	1	1.036	2.049	0.891

Tabel 4. Perbandingan matriks model dengan variasi dari *min_samples_split*

N o	<i>n_estimator</i> <i>s</i>	<i>max_dept</i> <i>h</i>	<i>min_samples_spl</i> <i>it</i>	<i>min_samples_le</i> <i>af</i>	MAE	RMSE	R- Square <i>d</i>
1	200	20	5	1	1.028	2.044	0.891
2	200	20	5	2	1.021	2.037	0.892
3	200	20	5	4	1.024	2.029	0.893
4	200	20	5	6	1.026	2.027	0.893

Tabel 5. Perbandingan matriks model dengan variasi dari *min_samples_leaf*

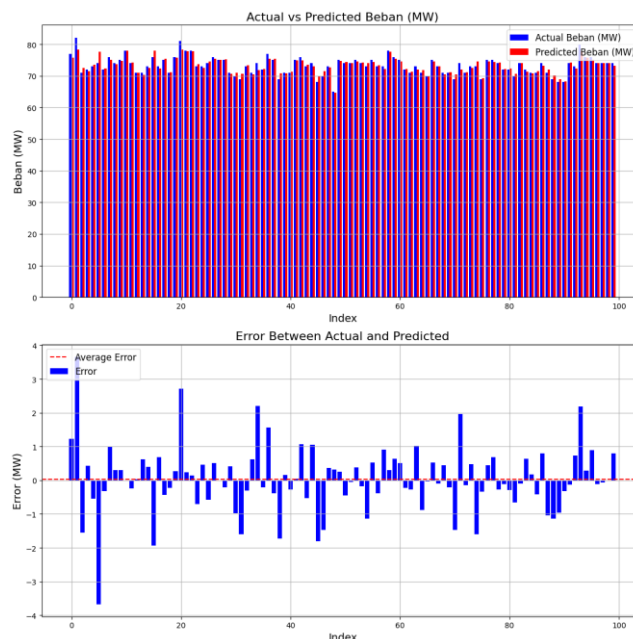
Berdasarkan pengujian variasi hyperparameter $n_estimators$, konfigurasi dengan $n_estimators = 200$, $max_depth = None$, $min_samples_split = 2$, dan $min_samples_leaf = 1$ memberikan performa terbaik dengan MAE 1.030, RMSE 2.049, dan RI 0.891, dipilih karena MAE terendah menunjukkan kesalahan prediksi terkecil. Setelah menetapkan $n_estimators = 200$, variasi pada max_depth diuji dan nilai 20 memberikan hasil optimal dengan MAE 1.029, RMSE 2.048, dan RI 0.891, dimana penurunan MAE menandakan peningkatan akurasi, sehingga $max_depth = 20$ dipilih sebagai konfigurasi terbaik. Selanjutnya, dengan mempertahankan $n_estimators$ dan max_depth tersebut, variasi $min_samples_split$ diuji dan nilai 5 memberikan performa terbaik dengan MAE 1.028, RMSE 2.044, dan RI 0.891, menunjukkan presisi yang meningkat sehingga nilai ini dipilih. Akhirnya, pengujian variasi $min_samples_leaf$ dengan nilai 4 menghasilkan MAE 1.024, RMSE 2.029, dan RI 0.893, peningkatan RI serta penurunan MAE dan RMSE mengindikasikan prediksi yang lebih akurat dan presisi, sehingga konfigurasi ini ditetapkan sebagai hasil tuning optimal keseluruhan.



Gambar 3. Visualisasi Latihan model aktual vs Beban

Visualisasi pada Gambar 3 memperlihatkan perbandingan nilai prediksi dan nilai aktual beban turbin uap dalam MW, di mana titik-titik segitiga coklat mewakili prediksi model dan garis merah putus-putus menunjukkan kondisi ideal prediksi sama dengan aktual. Distribusi titik yang mengikuti garis ideal menunjukkan model mampu memprediksi beban dengan akurat meski terdapat beberapa outlier, memperkuat validitas model dalam memodelkan hubungan variabel operasional dengan beban untuk efisiensi turbin optimal, konsisten dengan metrik evaluasi RI 0,893, MAE 1,024, dan RMSE 2,029 yang menunjukkan performa baik. Sebelum pengujian, studi terdahulu menegaskan keunggulan Random Forest dalam memprediksi daya dan efisiensi pembangkitan listrik, terutama dalam menangani data non-linear dengan akurasi tinggi. Setelah model dioptimalkan melalui tuning

hyperparameter, pengujian dilakukan menggunakan 100 sampel data uji untuk mengevaluasi kemampuan prediksi dengan metrik RI, RMSE, dan MAE. Hyperparameter terbaik yang digunakan terdiri dari 200 pohon keputusan, kedalaman maksimum 20, min_samples_split 5, dan min_samples_leaf 4, yang memberikan kinerja prediksi optimal berdasarkan evaluasi.



Gambar 4. Visualisasi pengujian model aktual vs Beban

Grafik pertama menunjukkan prediksi model yang mengikuti nilai aktual beban dalam megawatt dengan garis prediksi yang hampir selalu dekat garis data aktual, menandakan akurasi prediksi yang tinggi. Grafik kedua memperlihatkan selisih nilai aktual dan prediksi yang sebagian besar berada di sekitar nol tanpa pola penyimpangan konsisten, mengindikasikan model tidak bias dalam prediksi. Garis rata-rata error juga dekat garis nol, memperkuat bahwa model bekerja baik pada data uji, sesuai dengan metrik evaluasi MAE 0,692, RMSE 0,987, dan R-squared 0,876 yang menunjukkan akurasi tinggi dan kemampuan efektif model dalam memodelkan hubungan variabel input dengan target.

KE Simpulan

Penelitian ini berhasil mengembangkan model prediksi beban turbin uap menggunakan algoritma Random Forest dengan memanfaatkan variabel operasional PLTU seperti tekanan, temperatur, aliran uap, dan parameter kondensor, serta menerapkan normalisasi Min-Max Scaler untuk menyamakan skala nilai antar variabel dalam rentang 0 hingga 1 guna meningkatkan performa model. Dataset berjumlah 3.928 sampel, dengan 100 sampel sebagai data uji eksternal dan 3.828 untuk pelatihan yang terbagi menjadi 80% pelatihan dan 20% validasi. Evaluasi model pada data pelatihan menunjukkan performa tinggi dengan R-squared 0,893, RMSE 2,029, dan MAE 1,024, sedangkan pada data uji eksternal model mempertahankan akurasi dengan R-squared 0,876, RMSE

0,987, dan MAE 0,692. Hasil ini mengindikasikan kemampuan model dalam melakukan generalisasi dengan baik serta efektivitas Random Forest sebagai solusi prediksi beban turbin uap yang mendukung peningkatan efisiensi pembangkit, percepatan dan ketepatan pengambilan keputusan berbasis data, serta pengelolaan pembangkitan yang lebih berkelanjutan. Tingkat keandalan model yang tinggi saat pengujian data eksternal menegaskan bahwa algoritma ini efektif dan akurat untuk mendukung operasional dan pengelolaan PLTU secara data-driven.

Saran

Untuk memperkuat validitas dan reliabilitas model prediksi, disarankan agar penelitian selanjutnya tidak hanya membandingkan algoritma Random Forest dengan metode lain seperti Gradient Boosting, XGBoost, dan Artificial Neural Network, tetapi juga menguji model tersebut pada data operasional dari PLTU yang berbeda. Langkah ini penting guna mengevaluasi kemampuan generalisasi model di berbagai kondisi pembangkitan serta menentukan algoritma yang paling konsisten dalam memberikan akurasi terbaik pada berbagai konteks operasional. Dengan demikian, hasil penelitian akan memiliki nilai aplikatif yang lebih luas dan relevan untuk implementasi di berbagai unit pembangkit.

DAFTAR PUSTAKA

- Abbas, H., Jamaluddin, J., Arif, M., & Amiruddin, A. (2020). Analisa Pembangkit Tenaga Listrik dengan Tenaga Uap di PLTU. *ILTEK: Jurnal Teknologi*, 15(02), 103–106. <https://doi.org/10.47398/iltek.v15i02.33>
- Ajewole, T. O., Ojuola, B. M., Olawuyi, A. A., Momoh, O. D., & Olukayode, O. (2023). Application of Machine Learning to Prediction of Turbine Rotor Vibration in Steam Power Plant. *Journal of Energy and Safety Technology (JEST)*, 6(2), 1–9. <https://doi.org/10.11113/jest.v6.134>
- Andreanus, J., & Kurniawan, A. (2018). Sejarah, Teori Dasar dan Penerapan Reinforcement Learning: Sebuah Tinjauan Pustaka. *Jurnal Telematika*, 12(2), 113–118. <https://doi.org/10.61769/telematika.v12i2.193>
- Ardianto, Raharjo, A. B., & Purwitasari, D. (2022). Random Forest Regression untuk Prediksi Produksi Daya Pembangkit Listrik Tenaga Surya, 7(5), 1058–1075.
- Arisusanto, A., Suarna, N., & Dwilestari, G. (2023). Analisa Klasifikasi Data Harga Handphone Menggunakan Algoritma Random Forest dengan Optimize Parameter Grid. *Jurnal Teknologi Ilmu Komputer*, 1(2), 43–47. <https://doi.org/10.56854/jtik.v1i2.51>
- Ashfania, G. A. M., Prahasto, T., Widodo, A., & Warsokusumo, T. (2022). Penggunaan Algoritma Random Forest untuk Klasifikasi Berbasis Kinerja Efisiensi Energi pada Sistem Pembangkit Daya. *Rctasi*, 24(3), 14–21.

- Azis, A. R. (2024). Analisis Komparasi Algoritma Machine Learning dalam Prediksi Performa Akademik Mahasiswa: Literature Review, 4(2), 143–150.
- Burlian, F., & Ghafara, A. (2013). Perancangan Ulang Heat Recovery Steam Generator dengan Sistem Dual Pressure melalui Pemanfaatan Gas Buang Sebuah Turbin Gas Berdaya 160 MW. *Jurnal Rekayasa Mesin Universitas Sriwijaya*, 13(1), 21–33.
- Chyan, P., Arni, S., & Thayf, M. S. S. (2023). Pengantar Machine Learning. PT MIFANDI MANDIRI DIGITAL.
- Desyani, N., Kristina, S., & Yosephine, V. S. (2024). Studi Awal Penerapan Reinforcement Learning pada Penyelesaian Heterogeneous Vehicle Routing Problem with Soft Time Windows, 1(10).
- Fadil, M., Tambunan, A. T. P. R., Hasibuan, A. H., Elroi, M., Sinambela, Y., Situmorang, A. S., & Paundra, F. (2024). Pemeliharaan Sistem Pelumasan pada Turbin Uap di PLTU XYZ. *Pewira Journal of Science & Engineering* 4(2), 58–63. <https://doi.org/10.54199/pjse.v4i2.309>
- Faiza, I. M., Gunawan, & Andriani, W. (2022). Tinjauan Pustaka Sistematis: Penerapan Metode Machine Learning untuk Deteksi Bencana Banjir, 11(September), 59–63.
- Fathurohman, A. (2021). Machine Learning untuk Pendidikan: Mengapa dan Bagaimana. *Jurnal Informatika dan Teknologi Komputer (JITEK)*, 1(3), 57–62. <https://journal.amikveteran.ac.id/index.php/jitek/article/view/306>
- Haeruddin, Erick, & Aripadono, H. W. (2025). Perbandingan Support Vector Machine, Random Forest Classifier, dan K-Nearest Neighbour dalam Pendeteksian Anomali pada Jaringan DDoS. 7(1), 23–33.
- Hendra, H. (2023). Prediksi Daya Listrik pada Pembangkit Listrik Siklus Gabungan Berdasarkan Kondisi Lingkungan Menggunakan Metode Machine Learning. *Elektron: Jurnal Ilmiah*, 15(November), 72–82. <https://doi.org/10.30630/eji.0.0.415>
- Jamaludin, & Kurniawan, I. (2017). Analisis Perhitungan Daya Turbin yang Dihasilkan dan Efisiensi Turbin Uap pada Unit 1 dan Unit 2 di PT. Indonesia Power Uboh UJP Banten 3 Lontar. *Motor Bakar: Jurnal Teknik Mesin*, 1(2), 1–8. <https://doi.org/10.31000/mbjtm.v1i2.731>
- Jantri, & Armelia, P. (2021). Analisa Penurunan Kinerja Kondensor pada Turbin PLTU Unit 4 di PT. PLN (Persero) Unit Pembangkitan (UPK) Belawan. *Jurnal Teknologi Mesin Uda*, 2(1), 27–38.
- Leni, D., K., A., Sumiati, R., Adriansyah, & Haris. (2023). Prediction Modeling of Low Alloy Steel Based on Chemical Composition and Heat Treatment Using Artificial Neural Network. *Jurnal Pdimesin*, 21(5), 54–61. <https://doi.org/10.30811/jpl.v21i5.3896>

- Mukrimaa, S. S., Nurdyansyah, Fahyuni, E. F., Yulia Citra, A., Schulz, N. D., سان, د., Taniredja, T., Faridli, E. Miftah, & Harmianto, S. (2021). Machine Learning Teori, Studi Kasus dan Implementasi Menggunakan Python. *Jurnal Pendidikan Pendidikan Guru Sekolah Dasar*, 6(August), 128. <https://doi.org/10.5281/zenodo.5113507>
- Mulyadi, D., Rusirawan, D., & I, N. (2023). Desain Pembangkit Listrik Turbin Uap Berbahan Bakar Batu Bara yang Efektif dan Ramah Lingkungan. *Jurnal Tekno Insentif*, 17(1), 58–68. <https://doi.org/10.36787/jti.v17i1.1083>
- Pandia, M., Sihombing, P., Simamora, P., & Kaban, R. (2024). Kajian Literatur Multimedia Retrieval: Machine Learning untuk Pengenalan Wajah. *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, 7(1), 161–166. <https://doi.org/10.55338/jikoms.v7i1.2758>
- Panjaitan, A. W. R., Nainggolan, P. D., Sebayang, S., & Pardede, S. (2023). Analisis Performansi Water Tube Boiler Kapasitas 240 Ton/Jam di PLTU Sulbagut-1 Tanjung Karang. 4(1), 199–205.
- Pradana, D. P., Luthfi Alghifari, M., Farhan Juna, M., & Palaguna, S. D. (2022). Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network. *Indonesian Journal of Data and Science*, 3(2), 55–60. <https://doi.org/10.56705/ijodas.v3i2.35>
- Putri, A., Hardiana, C. S., Novfuja, E., Siregar, F. T. P., Rahmadden, R., Fatma, Y., & Wahyuni, R. (2023). Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 20–26. <https://doi.org/10.57152/malcom.v3i1.610>
- Ramadhon, R. N., Ogi, A., Agung, A. P., Putra, R., Febrihartina, S. S., & Firdaus, U. (2024). Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank. *Karimah Tauhid*, 3(2), 1860–1874. <https://doi.org/10.30997/karimahtauhid.v3i2.11952>
- Sajiwo, B., Prahasto, T., & Widodo, A. (2023). Prediksi Remaining Useful Life (RUL) pada Jet Engine sebagai Upaya Predictive Maintenance Berbasis Machine Learning. *Jurnal Teknik Mesin S-1*, 11(4).
- Setiawan, F. G., Melkias, A. A., & Slameto. (2022). Analisis Kinerja Turbin Uap Unit 1 di Cirebon Power. *Jurnal Teknik Energi*, 11(2), 7–11. <https://doi.org/10.35313/energi.v11i2.3517>
- Sidiq, A. N., & Anwar, M. (2021). Perbandingan Efisiensi Turbin Uap Kondisi Aktual Berbasis Data Komisioning Sesuai Standard ASME PTC 6. *Kilat*, 10(1), 190–199. <https://doi.org/10.33322/kilat.v10i1.1188>
- Song, Y., Park, J., Suh, M. S., & Kim, C. (2024). Prediction of Full-Load Electrical Power Output of Combined Cycle Power Plant Using a Super Learner Ensemble. *Applied Sciences*, 14(24). <https://doi.org/10.3390/app142411638>

- Sugeng, U. M., Setiadi, B., Hadi, V., & Sugianto. (2023). Analisis Pengaruh High Pressure Heater terhadap Turbine Heat Rate dan Pemakaian Batu Bara di PLTU 1 4 600 MW. *Presisi*, 25(2), 28–35. <https://ejournal.istn.ac.id/index.php/presisi/article/view/1635>
- Syarif, M., & Afiyati. (2025). Prediksi Nilai Unburned Carbon Batubara yang Dihasilkan PLTU Menggunakan Algoritma Linear Regression, Random Forest, dan LightGBM Regression Muhyiddin. *Jurnal Teknologi Informasi dan Komunikasi*, 9(June), 477–484.
- Syarif, Y. A. (2021). Pengaruh Laju Massa Uap terhadap Efisiensi Kerja Turbin Uap pada Pembangkit Listrik Tenaga Panas Bumi Lahendong Unit 5 di Tompaso Sulawesi Utara. (Unpublished thesis). [If thesis, add institution]