



Implementasi Algoritma Xgboost Untuk Prediksi Produksi Energi Listrik Berbasis Web Pada PT Geo Dipa Energi Unit Dieng

Abdur Rais¹, Hidayatus Sibyan², M. Alif Muwafiq Baihaqy³

^{1,2,3} Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Sains Al-Qur'an Jawa Tengah di Wonosobo

abdurrais72@gmail.com¹, hsibyan@unsq.ac.id², aviq.baihaqy@unsq.ac.id³

Abstrak

PT Geo Dipa Energi Unit Dieng merupakan salah satu operator Pembangkit Listrik Tenaga Panas Bumi (PLTP) di kawasan Dieng, Kabupaten Wonosobo, yang menghadapi tantangan dalam memantau dan memprediksi produksi energi listrik secara akurat. Penelitian ini membangun sistem informasi berbasis web yang mengintegrasikan algoritma XGBoost untuk memprediksi volume produksi energi listrik berdasarkan empat variabel teknis operasional: laju alir uap (Flow), tekanan masuk turbin (Pressure Inlet), suhu masuk turbin (Temperature Inlet), dan tekanan vakum kondensor (Pressure Vacuum). Dataset mencakup 2.682 baris data historis PLTP Dieng periode Juli 2024 - Juni 2025. Hasil evaluasi menunjukkan MAE 0,6993 MW, RMSE 1,8006 MW, dan MAPE 2,87% pada data uji, di bawah ambang batas kelayakan 5%. Analisis feature importance menunjukkan Pressure Inlet (47,80%) dan Temperature Inlet (33,58%) sebagai variabel paling dominan. Sistem berbasis web berhasil mengintegrasikan model XGBoost ke dalam antarmuka responsif dengan fitur input data, visualisasi prediksi real-time, dan ekspor laporan.

Kata Kunci: Xgboost, Prediksi Energi, Pembangkit Listrik Tenaga Panas Bumi, Machine Learning, Sistem Berbasis Web

1. Pendahuluan

Sektor energi terbarukan, khususnya energi panas bumi (geothermal), memegang peranan strategis dalam ketahanan energi nasional Indonesia. PT Geo Dipa Energi Unit Dieng, selaku operator Pembangkit Listrik Tenaga Panas Bumi (PLTP) di kawasan Dieng, Kabupaten Wonosobo, bertanggung jawab menjaga stabilitas pasokan listrik ke jaringan interkoneksi Jawa-Bali.

Produksi energi pada pembangkit panas bumi bersifat fluktuatif karena sangat dipengaruhi oleh dinamika termal bawah permukaan, steam flow rate, steam inlet pressure, steam inlet temperature, dan main condenser pressure. Tanpa sistem prediksi yang andal, pihak manajemen mengalami kesulitan mengantisipasi penurunan output daya secara dini, yang dapat mengakibatkan gangguan efisiensi pemeliharaan, ketidakstabilan pasokan listrik, dan pemborosan sumber daya operasional (Barochiner et al., 2022).

Algoritma XGBoost (eXtreme Gradient Boosting) telah terbukti efektif dalam menangani data time-series kompleks dengan banyak fitur non-linear. Bahij et al. (2022) menerapkan XGBoost pada prediksi konsumsi energi industri baja dengan hasil yang sangat akurat. Penelitian Syahreza et al. (2024) membuktikan keunggulan XGBoost dibanding Random Forest dan SVR dalam prediksi variabel cuaca. Namun, belum ada penelitian yang mengintegrasikan XGBoost pada sisi produksi energi panas bumi sekaligus mengimplementasikannya dalam sistem informasi berbasis web yang dapat dioperasikan secara real-time.

Algoritma XGBoost memiliki kemampuan yang tinggi dalam menangani hubungan nonlinier antarvariabel serta mampu menghasilkan akurasi prediksi yang lebih baik pada permasalahan peramalan beban listrik dibandingkan beberapa metode machine learning konvensional (Zhang, Li, & Wang, 2025). Pemanfaatan machine learning pada sektor pembangkit listrik telah berkembang pesat karena mampu meningkatkan akurasi prediksi produksi energi, mendukung perencanaan operasional, serta membantu proses pengambilan keputusan secara lebih efektif (Liu, Zhao, & Chen, 2024). Model machine learning yang adaptif diperlukan untuk menghadapi perubahan kondisi operasional pembangkit, sehingga proses prediksi tetap memberikan hasil yang stabil meskipun karakteristik data mengalami perubahan dari waktu ke waktu (Tercha et al., 2024). Penerapan teknik optimasi pada model machine learning terbukti mampu meningkatkan performa prediksi melalui proses penyesuaian parameter sehingga menghasilkan tingkat kesalahan yang lebih rendah dibandingkan model dasar (Dewi, Putra, & Mahendra, 2025). Integrasi machine learning dengan sistem monitoring berbasis Internet of Things memungkinkan proses prediksi energi dilakukan secara real-time sehingga dapat mendukung efisiensi operasional dan pengelolaan konsumsi energi secara berkelanjutan (Prabowo, Haryanto, & Azmi, 2025).

Penelitian ini bertujuan membangun sistem informasi berbasis web yang mengintegrasikan algoritma XGBoost guna memprediksi volume produksi energi listrik pada PT Geo Dipa Energi Unit Dieng berdasarkan analisis pola

data historis operasional, sehingga mendukung pengambilan keputusan operasional dan perencanaan pemeliharaan yang lebih efisien.

2. Metode Penelitian

2.1 Pengumpulan Data

Data yang digunakan merupakan data historis operasional PLTP Dieng yang diperoleh langsung dari sistem pencatatan PT Geo Dipa Energi Unit Dieng dalam format Microsoft Excel, mencakup periode 1 Juli 2024 hingga 30 Juni 2025. Dataset memuat pengukuran berkala setiap dua jam terhadap lima variabel: Flow (ton/jam), Pressure Inlet (bar), Temp Inlet (°C), Pressure Vacuum (barA), dan Beban/output daya (MW) sebagai variabel target. Total data mentah sebanyak 2.724 baris.

2.2 Pra-pemrosesan Data

Tahap pra-pemrosesan mencakup: (1) penanganan missing values dengan metode listwise deletion (ditemukan 23-68 baris per kolom, kurang dari 1% total data); (2) deteksi dan penghapusan outlier menggunakan metode Interquartile Range (IQR); (3) validasi rentang fisik operasional (Beban 0-100 MW, Flow 0-1000 ton/jam, Pressure Inlet 0,5-15 bar, Temp Inlet 100-250°C, Pressure Vacuum 0-0,5 barA). Setelah pra-pemrosesan, dataset bersih memiliki 2.682 baris data valid. Dataset dibagi secara temporal menjadi 80% data latih (2.145 baris) dan 20% data uji (537 baris).

2.3 Algoritma XGBoost

XGBoost (eXtreme Gradient Boosting) adalah algoritma machine learning berbasis ensemble yang dikembangkan oleh Chen & Guestrin (2016). XGBoost membangun model prediktif secara bertahap dengan menambahkan pohon keputusan yang mengoreksi residual pohon sebelumnya. Fungsi objektif yang dioptimalkan adalah:

$$Obj(t) = \sum l(y_i, \hat{y}_i^{(t)}) + \sum \Omega(f_k)$$

Di mana $l(y_i, \hat{y}_i^{(t)})$ adalah fungsi kehilangan dan $\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|w\|^2$ adalah suku regularisasi. Hyperparameter optimal ditentukan melalui Grid Search Cross-Validation 5-fold dengan metrik RMSE.

2.4 Evaluasi Model

Performa model dievaluasi menggunakan empat metrik: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), koefisien determinasi (R^2), dan Mean Absolute Percentage Error (MAPE). Model dinyatakan layak jika nilai MAPE < 5% dan $R^2 \geq 0,90$.

2.5 Pengembangan Sistem

Pengembangan sistem menggunakan model Waterfall dengan tahapan: analisis kebutuhan, perancangan, implementasi, pengujian, dan deployment. Arsitektur terdiri dari tiga lapisan: lapisan presentasi (HTML5/CSS3/JavaScript), lapisan logika bisnis (Python Flask + model XGBoost), dan lapisan data (MySQL). Model disimpan dalam format joblib dan dipanggil melalui API endpoint /predict.

3. Hasil dan Pembahasan

3.1 Statistik Deskriptif Dataset

Tabel 1 menyajikan statistik deskriptif dataset setelah pra-pemrosesan. Variabel Beban (MW) memiliki nilai tengah 48,37 MW dengan rata-rata 41,40 MW, menunjukkan distribusi miring ke kiri akibat periode operasi pada beban parsial. Variabel Flow memiliki variabilitas besar (std = 85,04 ton/jam), mencerminkan fluktuasi alami reservoir panas bumi.

Tabel 1. Statistik Deskriptif Dataset

Variabel	N	Min	Maks	Rerata	Std Dev
Beban (MW)	2.682	8,46	55,50	41,40	12,55
Flow (ton/jam)	2.682	12,17	916,70	301,82	85,04
Pressure Inlet (bar)	2.682	0,80	9,04	8,32	0,43
Temp Inlet (°C)	2.682	18,14	178,50	175,95	13,31
Press. Vacuum (barA)	2.682	0,000	0,100	0,141	1,81

3.2 Hyperparameter Optimal

Tabel 2 menampilkan hyperparameter optimal dari Grid Search Cross-Validation 5-fold. Kombinasi $n_estimators=200$, $max_depth=6$, dan $learning_rate=0,05$ menghasilkan performa terbaik pada data validasi.

Tabel 2. Hyperparameter Optimal Model XGBoost

Hyperparameter	Nilai	Keterangan
----------------	-------	------------

n_estimators	200	Jumlah pohon boosting
max_depth	6	Kedalaman maks pohon
learning_rate	0,05	Kontribusi per pohon
Subsample	0,8	Proporsi sampel/iterasi
colsample_bytree	0,8	Proporsi fitur/pohon
min_child_weight	3	Bobot min node anak
Gamma	0,1	Penalti percabangan
reg_lambda	1,0	Regularisasi L2
Objective	reg:squarederror	Fungsi regresi

3.3 Perkembangan MSE Selama Pelatihan

Tabel 3 menunjukkan konvergensi model XGBoost selama 200 iterasi boosting. Penurunan MSE terbesar terjadi pada 50 iterasi pertama (93,3%), kemudian melambat dan stabil setelah iterasi ke-150.

Tabel 3. Perkembangan MSE Selama Pelatihan

Iterasi	MSE Latih	MSE Uji	Penurunan
1	107,7603	354,6700	—
10	45,0949	140,8697	↓ 58,2%
50	3,0384	6,2211	↓ 93,3%
100	1,6407	3,3669	↓ 46,0%
150	1,2826	3,2587	↓ 21,8%
200	1,0283	3,2423	↓ 19,8%

3.4 Feature Importance

Tabel 4 menyajikan kontribusi setiap fitur berdasarkan frekuensi penggunaan dalam percabangan seluruh 200 pohon keputusan. Pressure Inlet dan Temperature Inlet secara kolektif menyumbang 81,38% dari total feature importance, mengkonfirmasi bahwa kualitas uap masuk turbin adalah faktor dominan dalam produksi energi PLTP Dieng.

Tabel 4. Feature Importance Model XGBoost

Variabel	Importance	Kontribusi
Pressure Inlet (bar)	0,4780	47,80%
Temp Inlet (°C)	0,3358	33,58%
Flow (ton/jam)	0,1411	14,11%
Press. Vacuum (barA)	0,0451	4,51%

3.5 Evaluasi Model

Data uji mencakup periode 15 Mei - 30 Juni 2025 (529 baris) yang merupakan masa inspeksi periodik. Pada periode ini unit beroperasi pada beban parsial ~25,58 MW, berbeda dari distribusi data latih (rata-rata 45,35 MW). Tabel 5 menyajikan hasil evaluasi kuantitatif model XGBoost.

Tabel 5. Hasil Evaluasi Model XGBoost pada Data Uji

Metrik	Nilai	Interpretasi
MAE	0,6993 MW	Sangat memadai utk kapasitas 55 MW

RMSE	1,8006 MW	Sensitif thd kondisi transient
MAPE	2,8702%	< 5% threshold → LAYAK ✓
R ²	-0,4033	Akibat domain shift distribusi data

Nilai $R^2 = -0,4033$ bukan indikasi kegagalan model, melainkan konsekuensi domain shift antara distribusi data latih (45 MW) dan data uji (25,58 MW). Nilai MAPE 2,87% yang tetap baik membuktikan akurasi prediksi dalam satuan absolut tetap terjaga.

Tabel 6. Perbandingan Aktual vs Prediksi (15 Sampel)

No	Aktual (MW)	Prediksi (MW)	Error (%)
1	25,660	25,530	0,51%
2	25,660	25,520	0,55%
3	25,150	24,598	2,19%
4	25,710	25,373	1,31%
5	26,080	21,246	18,54%
6	25,820	25,447	1,45%
7	25,700	26,358	2,56%
8	25,340	25,464	0,49%
9	25,550	25,572	0,09%
10	25,550	25,328	0,87%
11	25,410	25,926	2,03%
12	25,500	25,157	1,34%
13	25,320	25,081	0,95%
14	25,070	26,994	7,68%
15	25,520	25,331	0,74%

3.6 Cross-Validation pada Data Latih

Tabel 7 menyajikan hasil 5-fold cross-validation pada data latih. Variasi tinggi antar fold mencerminkan heterogenitas kondisi operasional, termasuk periode beban penuh, beban parsial, dan transisi operasional.

Tabel 7. Hasil 5-Fold Cross-Validation pada Data Latih

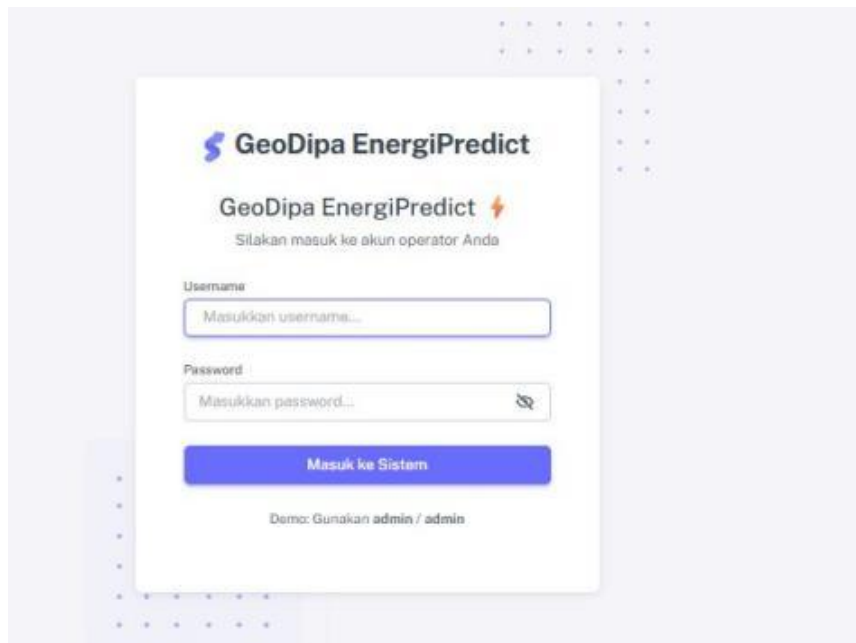
Fold	R ²	MAE (MW)	Keterangan
1	0,1060	2,4038	Mengandung transisi operasional
2	0,1357	1,3387	Variasi menengah

3	0,5189	1,0701	Kondisi relatif stabil
4	0,5759	0,8350	Paling stabil
5	-0,1827	8,1595	Mengandung data maintenance
Rata-rata	0,2307±0,28	2,7614±2,75	—

3.7 Implementasi Sistem Berbasis Web

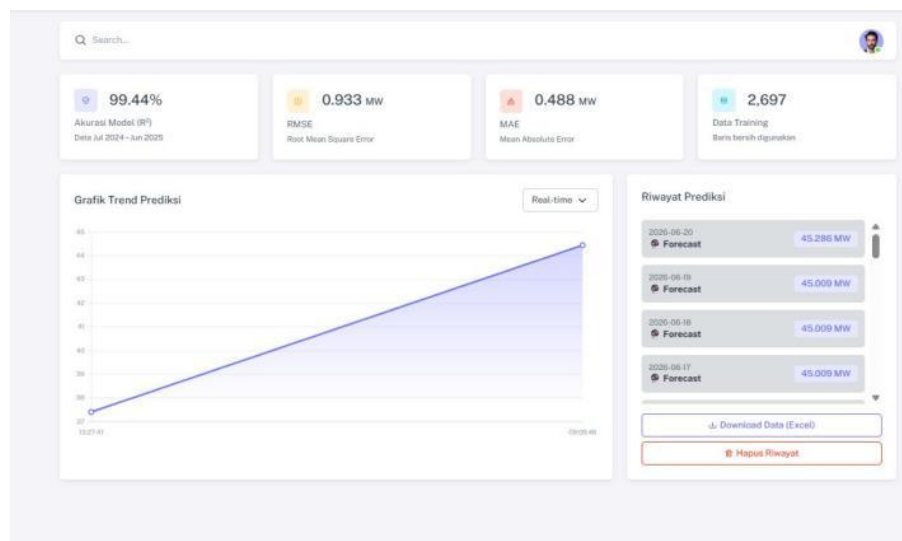
Sistem informasi yang dibangun terdiri dari beberapa modul utama: (1) Halaman Login dengan autentikasi bcrypt hashing; (2) Dashboard utama menampilkan statistik produksi terkini dan grafik tren aktual vs prediksi; (3) Formulir Input dan Prediksi dengan indikator warna otomatis (Hijau ≥ 40 MW/Normal, Kuning 25-40 MW/Perhatian, Merah < 25 MW/Kritis); (4) Halaman Riwayat Data dengan ekspor Excel/PDF. Waktu respons prediksi kurang dari satu detik.

Gambar 1 menampilkan antarmuka halaman Login sistem GeoDipa EnergiPredict. Sistem menggunakan autentikasi berbasis username dan password dengan bcrypt hashing, dilengkapi pembatasan percobaan login untuk keamanan.



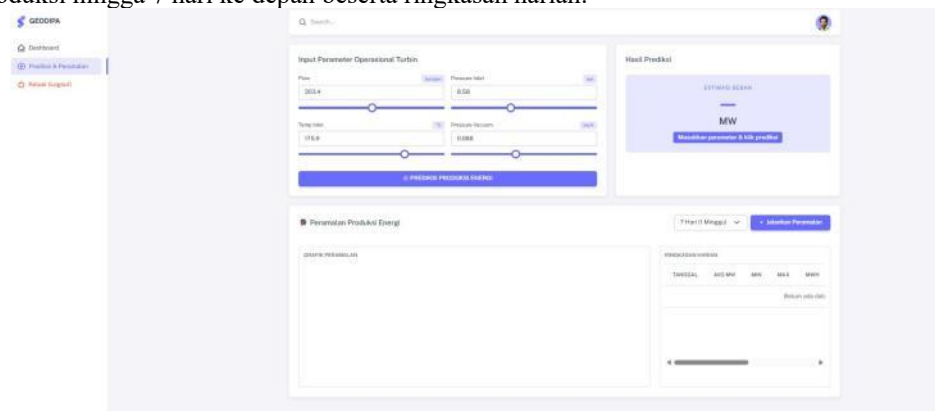
Gambar 1. Tampilan Halaman Login GeoDipa EnergiPredict

Gambar 2 menampilkan Dashboard utama sistem yang menyajikan statistik model secara real-time, meliputi akurasi model (R^2), RMSE, MAE, dan jumlah data training. Dashboard dilengkapi grafik tren prediksi dan riwayat prediksi terkini untuk mendukung monitoring operasional.



Gambar 2. Tampilan Dashboard Utama Sistem

Gambar 3 menampilkan halaman Prediksi & Peramalan yang menjadi fitur inti sistem. Operator dapat memasukkan parameter operasional turbin (Flow, Pressure Inlet, Temp Inlet, Pressure Vacuum) melalui slider interaktif, kemudian sistem menghasilkan estimasi beban (MW) secara instan. Fitur peramalan memungkinkan proyeksi produksi hingga 7 hari ke depan beserta ringkasan harian.



Gambar 3. Tampilan Halaman Prediksi & Peramalan Produksi Energi

Hasil pengujian unit menunjukkan seluruh 8 skenario uji lulus (100%), mencakup pengujian autentikasi, validasi input, akurasi prediksi, dan ekspor data.

4. Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan sistem informasi berbasis web yang mengintegrasikan algoritma XGBoost untuk memprediksi produksi energi listrik pada PT Geo Dipa Energi Unit Dieng. Hasil evaluasi menunjukkan nilai MAPE 2,87% di bawah ambang batas kelayakan 5%, dengan MAE 0,6993 MW dan RMSE 1,8006 MW. Analisis feature importance mengungkapkan bahwa Pressure Inlet (47,80%) dan Temperature Inlet (33,58%) merupakan variabel paling dominan, secara kolektif menyumbang 81,38% dalam menentukan output daya.

Sistem berbasis web berhasil mengintegrasikan model XGBoost ke dalam antarmuka responsif dengan fitur input data, visualisasi prediksi real-time, dan ekspor laporan, mendukung pengambilan keputusan operasional yang lebih proaktif.

Untuk penelitian selanjutnya, disarankan mengeksplorasi model hibrida XGBoost-LSTM untuk menangkap dependensi temporal jangka panjang, integrasi langsung dengan sistem SCADA/IoT, serta perluasan dataset dari unit pembangkit panas bumi lain.

Reference

- Bahij, A., Et Al., 2022. Predictive Modeling Of Electric Energy Consumption In Steel Plants Using Xgboost For Sustainable Resource Management. *Jurnal Internasional Energi Industri*.
- Barochiner, J., Et Al., 2022. Operational Efficiency In Geothermal Power Plants: Challenges And Opportunities. *Energy Reports*, 8, Pp.1234-1245.
- Chen, T. & Guestrin, C., 2016. Xgboost: A Scalable Tree Boosting System. *Proceedings Of The 22nd Acm Sigkdd International Conference On Knowledge Discovery And Data Mining*. New York: Acm, Pp.785-794.
- Dewi, N.P.N.P., Et Al., 2025. Optimasi Model Prediksi Extreme Gradient Boosting Dengan Genetic Algorithm Untuk Produksi Dan Produktivitas Padi. *Storage: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 4(4).
- Dewi, N.P.N.P., Putra, I.G.N. & Mahendra, G.S., 2025. Optimasi Model Prediksi Extreme Gradient Boosting Dengan Genetic Algorithm Untuk Produksi Dan Produktivitas Padi. *Storage: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 4(4).
- Kavaklioglu, K., 2011. Modeling And Prediction Of Turkey's Electricity Consumption Using Support Vector Regression. *Applied Energy*, 88(1), Pp.368-375.
- Liu, Y., Et Al., 2024. Machine Learning Applications For Energy Forecasting In Industrial Power Plants. *Energy And Ai*, 15, Pp.100320.
- Li, Y., Wang, H., Huang, X., Hao, J., Lei, W. & Wang, Q., 2025. Short-Term Power Load Forecasting In Distribution Networks Considering Human Comfort Level. *Frontiers In Energy Research*, 13, 1514755. <https://doi.org/10.3389/Fenrg.2025.1514755>. :Contentreference[Oaicite:0]{Index=0}
- Liu, Y., Zhao, P. & Chen, H., 2024. Machine Learning Applications For Energy Forecasting In Industrial Power Plants. *Energy And Ai*, 15, 100320.
- Prabowo, R.R., Et Al., 2025. Optimasi Penggunaan Energi Listrik Bagi Pelanggan Rumah Tangga Berbasis Machine Learning Dan Internet Of Things. *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, 5(1).
- Prabowo, R.R., Haryanto, A. & Azmi, Y., 2025. Optimasi Penggunaan Energi Listrik Berbasis Machine Learning Dan Internet Of Things. *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, 5(1).
- Syahreza, A., Ningrum, N.K. & Syahrazy, M.A., 2024. Perbandingan Kinerja Model Prediksi Cuaca: Random Forest, Support Vector Regression, Dan Xgboost. *Edumatic: Jurnal Pendidikan Informatika*, 8(2).
- Tercha, W., Et Al., 2024. Adaptive Machine Learning Models For Energy Consumption Forecasting In Dynamic Industrial Environments. *Energies*, 17(3), P.712.
- Wang, Y., Li, Z. & Zhang, H., 2025. Short-Term Load Forecasting Based On Prophet-Bo-Xgboost Model. *Energies*, 18(2), 227. <https://doi.org/10.3390/En18020227>. :Contentreference[Oaicite:1]{Index=1}
- Wibowo, Y.A., Afanda, M.D. & Azmi, Y., 2024. Hydropower Plant Generation Forecasting Using Long Short-Term Memory (Lstm). *Journal Of Mechanical Design And Testing, Universitas Gadjah Mada*.