

# Efektivitas *Stacking Model Gradient Boosting* Dalam Prediksi Biaya Listrik: Data Bangunan Residensial

## *Effectiveness of Gradient Boosting Stacking Model in Predicting Electricity Costs: Residential Building Data*

Ulfatun Nadifa  
Prodi Teknik Komputer  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
\*ulfatun@ung.ac.id

Haeriani H.  
Prodi Teknik Informatika  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
haerianih@gmail.com

Syahrir Abdussamad  
Prodi Teknik Komputer  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
syahrirabdussamad@ung.ac.id

Ade Irawaty Tolago  
Prodi Teknik Elektro  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
adeirawaty75@ung.ac.id

Rahmat Deddy Rianto Dako  
Prodi Teknik Komputer  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
rahmatdeddy@ung.ac.id

Zainudin Bonok  
Prodi Teknik Elektro  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
zainudinbonok@ung.ac.id

Bambang Panji Asmara  
Teknik Komputer  
Universitas Negeri Gorontalo  
Gorontalo, Indonesia  
bambang@ung.ac.id

---

Diterima : Juli 2025  
Disetujui : Juli 2025  
Dipublikasi : Juli 2025

---

**Abstrak**— Prediksi biaya listrik yang akurat penting untuk mendukung efisiensi energi dan pengelolaan sumber daya, terutama pada bangunan residensial maupun komersial. Penelitian ini bertujuan untuk menguji efektivitas model *Gradient Boosting* dalam memprediksi biaya listrik bulanan. Model dibangun dengan menggunakan metode *Stacking Ensemble*, yaitu teknik penggabungan beberapa algoritma *Gradient Boosting* secara bertingkat untuk meningkatkan akurasi prediksi. Untuk meningkatkan kinerja model, digunakan pemilihan nilai parameter terbaik secara otomatis (*Optimasi Hyperparameter*) dengan bantuan *Optuna*. Tahapan awal dimulai dengan membangun *pipeline preprocessing* berbasis *Tree Model* untuk menangani variasi dan kompleksitas data. Model dievaluasi dengan menggunakan metode *K-Fold Cross Validation*, yaitu pembagian data menjadi beberapa bagian untuk pengujian yang lebih representatif, dan hasilnya diukur menggunakan metrik *Root Mean Squared Logarithmic Error (RMSLE)* untuk menilai ketepatan prediksi. Hasil evaluasi menunjukkan bahwa model mampu mencapai nilai *RMSLE* sebesar 0.22. Selain itu, waktu prediksi rata-rata adalah 0.00029 detik. Temuan ini menunjukkan bahwa meskipun model *Gradient Boosting* umumnya digunakan pada dataset berdimensi besar, pendekatan ini tetap efektif pada data berdimensi kecil. Kombinasi teknik *ensemble* dan *Optimasi Hyperparameter* mampu menghasilkan prediksi yang akurat dan efisien. Oleh karena itu, pendekatan ini dapat diterapkan dalam skenario nyata, seperti manajemen energi di kawasan perkotaan.

**Kata Kunci**— *Optuna; Gradient Boosting; Pipeline; Stacking Model; Biaya Listrik.*

*Accurate electricity cost prediction is essential to support energy efficiency and resource management, particularly in residential and commercial buildings. This study aims to evaluate the effectiveness of the Gradient Boosting model in predicting monthly electricity costs. The model is built using the Stacking Ensemble method, a technique that combines multiple Gradient Boosting algorithms in a layered manner to improve prediction accuracy. To enhance the model's performance, automatic selection of the best parameter values (Hyperparameter Optimization) is conducted using Optuna. The initial phase involves developing a tree-based preprocessing pipeline to address data variability and complexity. The model is evaluated using the K-Fold Cross Validation method, which divides the data into several subsets for more representative testing. The performance is assessed using the Root Mean Squared Logarithmic Error (RMSLE) metric to measure prediction accuracy. The evaluation results show that the model achieves an RMSLE score of 0.22, with an average prediction time of 0.00029 seconds. These findings suggest that although Gradient Boosting models are typically used on high-dimensional datasets, this approach remains effective for low-dimensional data. The combination of ensemble techniques and hyperparameter optimization yields accurate and efficient predictions. Therefore, this approach can be applied in real-world scenarios, such as urban energy management.*

**Keywords**— *Optuna; Gradient Boosting; Pipeline; Stacking Model; Electricity Costs.*

## I. PENDAHULUAN

Kebutuhan akan energi listrik meningkat secara signifikan seiring pertumbuhan populasi, perkembangan industri, dan urbanisasi. Dimana bangunan residensial maupun komersial menyumbang porsi konsumsi energi yang cukup besar. Menurut laporan International Energy Agency 2003, sekitar 30% dari total konsumsi energi global berasal dari sektor bangunan [1]. Oleh karena itu, prediksi biaya listrik yang akurat akan mendukung efisiensi energi, perencanaan anggaran, serta pengambilan kebijakan dalam pembangunan berkelanjutan seperti kota pintar (smart city).

Pesatnya perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*) saat ini, diantaranya semakin banyak penerapan pendekatan berbasis *Machine Learning* untuk memodelkan dan memprediksi perilaku konsumsi energi. Berbagai studi telah menunjukkan efektivitas algoritma seperti *Gradient Boosting*, *Random Forest*, dan *Deep Learning* dalam memprediksi beban listrik dan biaya energi. Misalnya pada penelitian Khan 2022 membandingkan model *XGBoost*, *LightGBM*, dan *SVR* dalam prediksi konsumsi energi harian, dan menunjukkan bahwa model boosting memberikan hasil yang paling akurat [2]. Demikian pula penelitian yang dilakukan Li 2021 menunjukkan bahwa teknik *ensemble* memberikan akurasi yang lebih tinggi dibandingkan model tunggal dalam memprediksi biaya listrik disektor komersial [3].

Namun, tantangan utama dalam membangun model prediksi yang handal terletak pada pemilihan fitur yang optimal, pengolahan data yang kompleks, dan penyetelan parameter model (*hyperparameter tuning*). Proses tuning atau penentuan nilai parameter yang manual sering kali memakan waktu dan tidak efisien. Untuk mengatasi hal ini, digunakan pendekatan *Optimasi Hyperparameter* otomatis seperti *Optuna* yang memanfaatkan *Bayesian Optimization* dan *Tree-structured Parzen Estimator (TPE)*. Studi ini dilakukan juga oleh Akiba 2019 membuktikan bahwa *Optuna* mampu secara signifikan mempercepat proses tuning dan meningkatkan performa model [4].

Penelitian ini bertujuan mengembangkan model prediksi biaya listrik bulanan pada bangunan residensial dengan menerapkan *Optimasi Hyperparameter* otomatis menggunakan *Optuna*. Proses dimulai dengan membangun *Pipeline Preprocessing* berbasis *Tree Model* untuk menangani keragaman dan kompleksitas data. Selanjutnya, digunakan metode *Stacking Ensemble* dengan menggabungkan tiga algoritma *Gradient Boosting* populer: *CatBoost*, *LightGBM*, dan *XGBoost*, yang telah terbukti efektif dalam berbagai studi energi seperti yang dilakukan pada penelitian Zhou 2023 [5], Huang 2022 [6], dan Eissa 2022 [7].

Evaluasi model dilakukan menggunakan teknik *K-Fold Cross Validation* dan metrik *Root Mean Squared Logarithmic Error (RMSLE)* yang dinilai cocok untuk data energi yang cenderung skewed. Hasil dari beberapa studi seperti Zhang 2020 [8] dan Bianchi 2022 [9] juga mendukung bahwa penggunaan *RMSLE* memberikan keunggulan dalam menangani kesalahan prediksi yang bersifat proporsional.

Model yang dikembangkan dalam studi ini menunjukkan performa yang baik, dengan akurasi tinggi dan waktu prediksi yang cepat. Hal ini memperkuat temuan sebelumnya bahwa kombinasi *Pipeline* yang terstruktur,

metode *Ensemble*, dan *Optimasi Hyperparameter* yang tepat mampu menghasilkan sistem prediksi biaya listrik yang efisien, akurat, dan layak diterapkan dalam pengelolaan energi di lingkungan perkotaan modern.

## II. METODE

Berikut ini adalah metode penelitian yang digunakan sebagai pendekatan untuk mendapatkan model prediksi biaya listrik dari data bangunan residensial

### 1. Dataset

Data yang digunakan dalam penelitian ini diambil pada platform Kaggle, merupakan salah satu data yang di publikasikan sebagai data challenge. Didalam dataset ini terdapat 3 file: data training, testing, dan submission. Pada tabel 1 dijelaskan detail informasi dataset.

TABEL 1. DETAIL INFORMASI DATASET

Jenis Data	Detail Data		
	Jumlah Kolom (Feature)	Banyak Data (Row)	Nama Kolom (Feature)
Data Training	9	11220	Site_Area, Structure_Type, Water_Consumption_Per_Building, Recycling_Rate, Utilization_Rate, Air_Quality_Index, Issue_Resolution_Time, Resident_Count, Electricity_Cost
Data Testing	9	300	
Data Submission	1	300	Electricity_Cost

Dataset yang digunakan dalam penelitian ini memiliki jumlah data yang relatif kecil yang di tunjukkan pada jumlah banyak data (Row) pada setiap jenis data. Kondisi ini dapat memengaruhi performa model dalam menghasilkan prediksi yang akurat dan stabil, karena data yang digunakan untuk model mempelajari pola dalam memprediksi data baru sedikit. Umumnya, model *machine learning* kompleks seperti *Gradient Boosting* dan teknik *stacking ensemble* menunjukkan performa optimal ketika diterapkan pada dataset berukuran besar dan kaya fitur. Oleh karena itu, dalam penelitian ini juga dianalisis bagaimana kinerja model tersebut saat diterapkan pada data berdimensi kecil.

Berikut ini adalah informasi nilai dari setiap feature:

1. *Site\_Area*: Mencakup data area dari setiap bangunan.
2. *Structure\_Type*: Mencakup tipe bangunan komersil, residens dll.
3. *Water\_Consumption\_Per\_Building*: Mencakup jumlah penggunaan air setiap bangunan.
4. *Recycling\_Rate*: Persentase efektivitas penggunaan listrik.
5. *Utilization\_Rate*: Persentase pemanfaatan Listrik.
6. *Air\_Quality\_Index*: Menggambarkan tingkat kebersihan atau pencemaran udara di suatu lokasi.
7. *Issue\_Resolution\_Time*: Lama waktu yang dibutuhkan untuk menyelesaikan masalah terkait sistem listrik atau layanan utilitas.
8. *Resident\_Count*: Jumlah penghuni/warga.
9. *Electricity\_Cost*: Biaya listrik.

Informasi dari setiap feature diatas, kemudian digunakan pada tahap selanjutnya, yaitu *Eksplorasi Data Analysis (EDA)*.

## 2. Proses Preprocessing Data

Dataset akan melalui tahapan preprocessing untuk mempersiapkan data sebelum dimasukkan ke dalam model prediksi. Beberapa tahapan utama sebagai berikut:

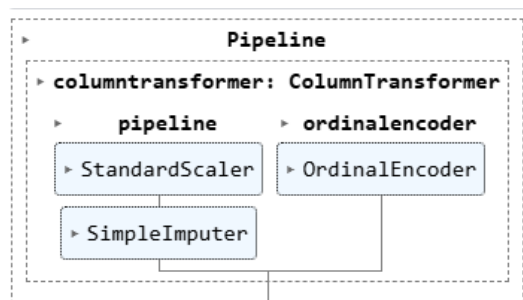
### a. Eksplorasi Data Analysis (EDA)

Tujuan *Eksplorasi Data Analysis (EDA)* untuk memahami kondisi data secara mendalam. Kemudian informasi hasil EDA akan digunakan untuk menentukan metode preprocessing yang tepat. Hasil dari EDA dapat digunakan untuk memilih metode yang tepat dalam mengatasi permasalahan data yang ditemukan, seperti data kosong dan data tidak berdistribusi normal. Pada penelitian ini hasil EDA akan berpengaruh langsung terhadap penyusunan arsitektur *Pipeline Preprocessing* yang dikembangkan.

### b. Pipeline Preprocessing

*Pipeline Preprocessing* berfungsi untuk memproses data dengan lebih efisien. Dengan arsitektur *Pipeline* maka Ketika ingin mengubah dan menambah metode baru serta parameternya dapat dilakukan dengan lebih terstruktur seperti pada gambar 1.

Pada penelitian ini berdasarkan hasil EDA, maka dipilih 3 metode seperti *Standard Scaler* untuk menormalisasi distribusi data, *Simple Imputer* untuk mengisi nilai-nilai yang hilang (missing values) dalam dataset menggunakan strategi tertentu seperti rata-rata (*mean*), *median*, *modus*, atau konstanta, dan *Ordinal Encoder* mengubah fitur kategori (bertipe string atau label) menjadi representasi numerik berurutan, yang kemudian akan di kemas didalam satu arsitektur *Pipeline Preprocessing*.



Gambar 1. Arsitektur *Pipeline Preprocessing*

### c. Subset Feature

Setelah *Preprocessing* data selesai maka tahap selanjutnya adalah *subset Feature* untuk memilih fitur yang akan digunakan sebagai *feature* data *training* dan data *testing*. Pada penelitian ini *subset fitur* menggunakan metode *lasso* yang akan memilih *Feature* yang memiliki korelasi baik dengan label prediksi dalam hal ini *Electricity\_Cost* atau biaya listrik.

## 3. Hyperparameter Optuna

Nilai parameter memiliki peran penting dalam meningkatkan akurasi hasil prediksi model. Namun, jika proses penentuan *hyperparameter* dilakukan secara manual, nilai yang diperoleh cenderung bersifat subjektif dan tidak selalu dapat dijadikan sebagai standar yang optimal untuk semua kondisi. Pada penelitian ini, proses *optimasi hyperparameter* dilakukan secara otomatis menggunakan metode *Optuna*. Pendekatan ini memungkinkan pencarian nilai *hyperparameter* yang optimal secara sistematis dan efisien, sehingga menghasilkan model dengan kinerja yang lebih stabil dan dapat direproduksi.

## 4. Training dan Validasi Stacking Model

*Stacking Ensemble* model dikembangkan untuk melatih *meta-model* dari *based-model*. Teknik ini bertujuan untuk meningkatkan akurasi secara keseluruhan dengan menggabungkan berbagai perspektif yang diberikan oleh model-model dasar. Dimana *Meta-model* akan belajar bagaimana mengombinasikan hasil-hasil prediksi tersebut dengan cara yang optimal, sehingga menghasilkan prediksi yang lebih tepat.

## 5. Evaluasi Model

Evaluasi kinerja model menggunakan metode *K-Fold Cross Validation* serta mengukur nilai *Root Mean Squared Logarithmic Error (RMSLE)*. Metode ini membantu memastikan bahwa model diuji dengan cara yang lebih representatif dan stabil.

## III. HASIL DAN PEMBAHASAN

### A. Hasil Pipeline Preprocessing

Berdasarkan hasil Preprocessing Data diperoleh beberapa *feature* yang memiliki korelasi kuat dan berpengaruh langsung terhadap prediksi biaya listrik. Beberapa *feature* tersebut adalah :

1. *Site\_Area*
2. *Water\_Consumption\_Per\_Building*
3. *Recycling\_Rate*
4. *Recycling\_Rate*
5. *Air\_Quality\_Index*
6. *Issue\_Resolution\_Time*
7. *Resident\_Count*

Ketujuh *feature* diatas dipilih secara otomatis menggunakan metode *Lasso* dengan menetapkan nilai korelasi  $10^{-4}$  sebagai nilai korelasi terbaik yang mendekati 1.

### B. Hasil Optimasi Hyperparameter

Parameter yang ditentukan nilai *hyperparameter*nya pada penelitian ini adalah *learning-rate*, *n-estimator*, dan *max-depth*. *Learning rate* adalah salah satu parameter yang berpengaruh terhadap waktu komputasi, sedangkan parameter *n-estimator* dan *max-depth* adalah jumlah *tree* dan maksimal kedalam *tree model*, yang mana menggambarkan kompleksitas dari arsitektur model [10].

Pada penelitian ini dalam pemilihan *hyperparameter* digunakan beberapa kali percobaan (*trial*) yang dilakukan. Tujuan dari *trial* ini untuk menemukan *hyperparameter* terbaik serta sebagai analisis perbandingan. Berikut pada tabel 2 hasil nilai setiap *hyperparameter* dengan metode *optuna*.

TABEL 2. HASIL TRIAL HYPERPARAMETER METODE OPTUNA

No.	Trial	Hyperparameter			RMSLE
		Learning rate	n-estimator	Max-depth	
1.	50	0.015	958	9	0.25
2.	150	0.017	1124	12	0.21
3.	300	0.010	2998	16	0.19

Berdasarkan tabel 2, diketahui bahwa semakin kompleks arsitektur model yang ditandai dengan nilai  $n$ -estimator dan  $max$ -depth yang semakin tinggi dapat menghasilkan prediksi yang lebih baik berdasarkan nilai  $RMSLE$  semakin kecil, namun hal tersebut juga dapat diikuti dengan waktu komputasi yang akan semakin lama karena model terlalu kompleks, ditunjukkan dari nilai learning rate semakin kecil. Begitu juga sebaliknya jika waktu komputasi semakin lama dan model sangat sederhana, kemungkinan model akan mendapati *underfitting* atau model tidak dapat memprediksi data baru dengan tepat. Namun jika dilihat dari hasil nilai validasi  $RMSLE$  pada tabel 2, tidak ditemukan perbedaan nilai secara signifikan, dan juga tidak terdapat nilai anomaly dari setiap *hyperparameter*, hal ini menunjukkan model mampu memprediksi dengan cukup stabil.

### C. Hasil Stacking Model

Prediksi biaya listrik dilakukan dengan menggunakan data *testing*. *Stacking model* pada penelitian ini menggunakan dua algoritma dasar yaitu *CatBoost* dan *LGBM* [11]. Berdasarkan hasil trial diketahui bahwa algoritma dasar sebagai model terbaik adalah *CatBoost* dengan nilai  $RMSLE$  lebih rendah dibandingkan dengan *LGBM*, walaupun perbedaan tersebut tidak cukup signifikan [12].

Adapun meta-model yang digunakan sebagai final estimator adalah *Gradient Boosting Machine (GBM)* [13]. Hasil prediksi biaya Listrik dengan *stacking model* menunjukkan nilai  $RMSLE$  rata-rata 0.22 dengan waktu prediksi sekitar 0.00029 detik. Berdasarkan waktu prediksi diketahui model mampu memprediksi data dengan waktu prediksi yang sangat baik.

### D. Evaluasi Model

Berdasarkan nilai  $RMSLE$  hasil prediksi biaya Listrik pada data bangunan residensial, kemudian dilakukan analisis secara manual untuk menghitung nilai prediksi sebenarnya yang akan digunakan untuk menganalisis presentase kesalahan prediksi model. Berikut ini adalah rumus  $RMSLE$  [14].

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(p_i + 1) - \log(a_i + 1))^2}$$

Misalkan jika prediksi biaya Listrik sebesar  $p_i = 15.000$ , dengan  $RMSLE$  0.22 maka prediksi aktual atau sebenarnya ( $a_i$ ) adalah berkisar antara:

$$a_i = 12.298 \text{ (Jika prediksi terlalu tinggi)}$$

$$a_i = 18.292 \text{ (Jika prediksi terlalu rendah)}$$

Sedangkan untuk menghitung persentase kesalahan dari nilai aktual dan prediksi adalah:

$$\text{Persentase Kesalahan} = \left| \frac{a_i - p_i}{a_i} \right|$$

Sehingga nilai persentase kesalahannya:

$$a_i = 22\% \text{ (Jika prediksi terlalu tinggi)}$$

$$a_i = 18\% \text{ (Jika prediksi terlalu rendah)}$$

Nilai persentase kesalahan dalam prediksi diharapkan semakin kecil untuk akurasi prediksi yang baik [15]. Berdasarkan persentase kesalahan model dalam memprediksi Biaya Listrik diatas menunjukkan nilai yang masih cukup besar, hal ini dikarenakan jumlah data training model yang sangat kecil, sehingga informasi pengenalan pola oleh model juga terbatas. Namun, jika dibandingkan dengan kondisi dataset yang ada, persentase kesalahan prediksi model tersebut masih tergolong cukup baik. Sehingga untuk penelitian selanjutnya dapat menggunakan dimensi data yang lebih banyak.

Meskipun terdapat perbedaan antara nilai prediksi dan nilai aktual, kesalahan ini tidak dianggap signifikan atau fatal, terutama dalam konteks prediksi Biaya Listrik, karena selisih ini relatif kecil dan tidak memberikan dampak besar dalam pengambilan keputusan terkait biaya operasional listrik. Namun masih perlu dilakukan analisis lebih dalam dengan dataset yang lebih kompleks untuk penerapan dalam skala industri.

## IV. KESIMPULAN

Hasil komposisi model *stacking* yang terdiri dari *CatBoost*, *LightGBM (LGBM)*, dan *Gradient Boosting Machine (GBM)* sebagai *meta-model* menghasilkan nilai  $RMSLE$  rata-rata sebesar 0.22 dan waktu prediksi hanya 0.00029 detik, mengindikasikan efisiensi pemrosesan yang sangat baik. Evaluasi manual mengonfirmasi bahwa tingkat kesalahan prediksi berkisar antara 18–22%, yang masih dapat diterima dalam konteks dataset terbatas. Hasil ini menunjukkan bahwa meskipun *Gradient Boosting* umumnya digunakan pada dataset besar, kombinasinya dengan *Optuna* mampu memberikan performa stabil dan cukup akurat bahkan pada data skala kecil. Model ini memiliki potensi nyata untuk diterapkan dalam sistem manajemen energi bangunan cerdas, khususnya dalam pemantauan konsumsi listrik dan pengendalian biaya secara otomatis di kawasan urban. Untuk pengembangan selanjutnya, disarankan agar: dataset diperluas dengan menambahkan dimensi spasial (lokasi geografis), temporal (musiman), serta fitur lingkungan seperti indeks kualitas udara dan suhu rata-rata harian, validasi model dilakukan pada bangunan komersial atau industri, guna menilai skalabilitas dan generalisasi model pada profil konsumsi energi yang lebih kompleks. Integrasi dengan sistem IoT dan smart meter dapat diuji untuk memungkinkan prediksi real-time dan pengambilan keputusan otomatis dalam manajemen energi. Dengan pendekatan ini, model tidak hanya berfungsi sebagai alat prediksi, tetapi juga dapat menjadi komponen penting dalam sistem energi berkelanjutan berbasis kecerdasan buatan.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh teman-teman di Jurusan Teknik Elektro dan Komputer, Universitas Negeri Gorontalo, yang telah memberikan dukungan dan motivasi selama proses penelitian ini.

## REFERENSI

- [1] IEA, "Energy Efficiency 2023: Analysis and Outlook to 2030". International Energy Agency, 2023.
- [2] A. H. Khan, A. Ahmed, M. Asif, and L. Wang, "Comparison of machine learning techniques for daily residential energy consumption forecasting," *Sustainable Cities and Society*, vol. 78, p. 103591, 2022, doi: 10.1109/IC3I59117.2023.10397623.
- [3] Y. Li, J. Zhang, and K. Wang, "Electricity price forecasting based on ensemble learning: A case study in the commercial sector," *Energy Reports*, vol. 7, pp. 2034–2045, 2021.
- [4] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, Anchorage, AK, USA, 2019.
- [5] X. Zhou, M. Zhao, J. Liu, and H. Lin, "Stacking ensemble learning for smart grid load forecasting," *IEEE Access*, vol. 11, pp. 15920–15930, 2023.
- [6] Y. Huang, Q. Chen, W. Zhang, and R. Wang, "A comparative study of CatBoost, LightGBM, and XGBoost for predicting building energy consumption," *J. Build. Eng.*, vol. 45, p. 103462, 2022.
- [7] N. Barboza, R. Puentes, J. C. Perez, and S. Caicedo, "Stacking ensemble tree models to predict energy performance in residential buildings," *Sustainability*, vol. 12, no. 22, p. 9366, 2020.
- [8] J. Zhang, Z. Wang, and J. Wang, "Data-driven methods for electricity price and load forecasting: A review," *Renew. Sustain. Energy Rev.*, vol. 122, p. 109766, 2020.
- [9] F. M. Bianchi, D. Grassi, E. L. Livi, and M. Alippi, "Machine learning for energy forecasting: Recent advances and future challenges," *Energy and AI*, vol. 7, p. 100135, 2022.
- [10] J. Brownlee, "Performance metrics for regression in machine learning," *Machine Learning Mastery*, 2020. Available: <https://machinelearningmastery.com>.
- [11] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 31, pp. 6638–6648, 2018.
- [12] D. Liu, H. Wang, and M. Zhao, "An ensemble model for the energy consumption prediction of residential buildings," *Energy*, vol. 304, 2024, Art. no. 124935. Available: <https://www.sciencedirect.com/science/article/pii/S0360544224040337>.
- [13] J. Liu and S. Hu, "Prediction of home energy consumption based on gradient boosting regression tree algorithm," *Energy Reports*, vol. 7, pp. 152–160, 2021.
- [14] M. Kiran, "Electricity load forecasting using CNN-LSTM with Optuna optimization," NCI Master's Thesis, 2021. [Online]. Available: <https://norma.ncirl.ie/4440/>.
- [15] H. B. Eissa, "A Stacking Multi-Learning Ensemble Model for Predicting Near Real-Time Energy Consumption Demand of Residential Buildings," *Academia.edu*, 2020. Available: <https://www.academia.edu/41267227>.