

Prediksi Nilai *Unburned Carbon* Batubara yang Dihasilkan PLTU Menggunakan Algoritma *Linear Regression*, *Random Forest*, dan *LightGBM Regression*

Muhyiddin Syarif^{1*}, Afiyati²

^{1*,2} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia.

article info

Article history:

Received 15 October 2024

Received in revised form

10 November 2024

Accepted 20 December 2024

Available online April 2025.

Keywords:

Unburned Carbon Prediction;

Machine Learning; Power

Plant Efficiency;

Environmental Sustainability;

LightGBM.

Kata Kunci:

Karbon Tak Terbakar;

Pembelajaran Mesin; Efisiensi

Operasional; Keberlanjutan

Lingkungan; LightGBM.

abstract

This study focuses on predicting unburned carbon levels in coal-fired power plants to enhance operational efficiency. Accurate prediction of unburned carbon is crucial as it directly affects fuel combustion efficiency and environmental sustainability. The research compares three machine learning algorithms: Linear Regression, Random Forest, and LightGBM Regression, using performance metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). The results show that LightGBM Regression performs the best, with MAE of 0.31, MAPE of 1.29, and RMSE of 0.38, outperforming the other two models. This model can be further optimized to improve prediction accuracy, contributing to more efficient and environmentally friendly power plant operations. The application of machine learning in this study supports data-driven decision-making in the energy sector.

abstrak

Penelitian ini berfokus pada prediksi tingkat karbon yang tidak terbakar di pembangkit listrik berbahan bakar batu bara untuk meningkatkan efisiensi operasional. Prediksi yang akurat terhadap karbon yang tidak terbakar sangat penting karena berdampak langsung pada efisiensi pembakaran bahan bakar dan keberlanjutan lingkungan. Penelitian ini membandingkan tiga algoritma machine learning: Linear Regression, Random Forest, dan LightGBM Regression, menggunakan metrik performa seperti Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), dan Root Mean Squared Error (RMSE). Hasil menunjukkan bahwa LightGBM Regression memiliki performa terbaik dengan MAE sebesar 0,31, MAPE sebesar 1,29, dan RMSE sebesar 0,38, mengungguli kedua model lainnya. Model ini dapat dioptimalkan lebih lanjut untuk meningkatkan akurasi prediksi, sehingga berkontribusi pada operasi pembangkit yang lebih efisien dan ramah lingkungan. Penerapan machine learning dalam penelitian ini mendukung pengambilan keputusan berbasis data di sektor energi.

Corresponding Author. Email: syarifmuhyiddin15@gmail.com ^{1}.

Copyright 2025 by the authors of this article. Published by Lembaga Otonom Lembaga Informasi dan Riset Indonesia (KITA INFO dan Riset). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. 

1. Pendahuluan

Pembangkit Listrik Tenaga Uap (PLTU) adalah pembangkit yang mengandalkan energi kinetik dari uap untuk menghasilkan energi listrik. Komponen utama dari pembangkit ini adalah *generator* yang dihubungkan ke *turbin*, yang memerlukan energi kinetik dari uap panas atau kering untuk memutarinya. Dalam PLTU, energi primer yang dikonversikan menjadi energi listrik berasal dari bahan bakar (Lewerissa, 2018). Bahan bakar yang digunakan dapat berupa batubara (padat), minyak (cair), atau gas (Umiyati *et al.*, n.d.). Batubara merupakan bahan bakar yang terbentuk dari endapan batuan organik, terutama terdiri dari karbon, hidrogen, dan oksigen. Dalam beberapa kasus, PLTU menggabungkan beberapa jenis bahan bakar. Konversi energi primer menjadi energi panas (*kalor*) yang terjadi dalam PLTU merupakan tahap pertama dalam proses konversi energi, yang berlangsung dalam ruang bakar *ketel uap*.

Seiring dengan kemajuan teknologi dalam dunia pembangkit listrik, banyak inovasi yang diterapkan untuk meningkatkan efisiensi operasional pembangkit. Pembangkit listrik diharapkan memiliki kinerja yang optimal untuk memastikan operasi yang efisien (Zaman & Sinaga, 2021). Efisiensi pembangkit sangat penting baik untuk peralatan yang digunakan maupun untuk lingkungan. Efisiensi ini dapat dievaluasi melalui beberapa indikator, salah satunya adalah nilai *unburned carbon* yang terkandung dalam batubara. Semakin efisien suatu pembangkit, semakin baik pula performa dan biaya operasionalnya. Dalam *ketel uap* berbahan bakar batubara, salah satu metrik penting dalam mengoptimalkan proses pembakaran adalah proporsi karbon yang tidak terbakar (*UC*) dalam abu (Bajpai *et al.*, 2024).

Penurunan nilai *unburned carbon* dapat menjadi indikator utama dalam peningkatan efisiensi PLTU. Nilai *unburned carbon* dipengaruhi oleh berbagai faktor yang kompleks, termasuk karakteristik batubara dan kondisi operasi *boiler*. Gabungan faktor-faktor ini menjadikan prediksi nilai *unburned carbon* sebagai suatu tantangan yang memerlukan pemodelan yang akurat dan analisis yang tepat. Prediksi nilai *unburned carbon* sangat penting bagi pengoperasian pembangkit listrik untuk menilai seberapa efisien sebuah

pembangkit. Selain itu, nilai *unburned carbon* juga berperan dalam menilai kualitas udara di sekitar PLTU. *Unburned carbon* dalam abu terbang dapat mencerminkan proses pembakaran bahan bakar yang tidak efisien (Adamczyk *et al.*, 2021). Dengan mengetahui nilai *unburned carbon*, operator dapat mengatur parameter pembakaran dalam *boiler* untuk mencapai efisiensi yang lebih tinggi dan mengurangi dampak negatif terhadap lingkungan.

Penelitian yang dilakukan oleh Sanaullah, Shailendra Kumar, dan Vikasih Dwivedi pada tahun 2023 mengenai optimalisasi kinerja PLTU melalui pengurangan karbon yang tidak terbakar dengan metode *regresi linier* mengidentifikasi pengaruh pengaturan kemiringan burner dan presentase oksigen terhadap penurunan *unburned carbon*. Penelitian tersebut menggunakan 15 parameter dalam pemodelannya dan menunjukkan bahwa pengaturan ini dapat mengurangi nilai *unburned carbon* melalui penerapan *machine learning* (Kumar & Dwivedi, n.d.). Namun, penelitian ini belum membahas secara spesifik tentang akurasi model *machine learning* yang digunakan serta tahapan pemodelannya. Oleh karena itu, penelitian ini memperluas temuan tersebut dengan membahas pemodelan untuk prediksi nilai *unburned carbon* dan pemilihan algoritma yang optimal.

Penelitian lain yang dilakukan oleh Kuan Lu, Xingsen Yang, Haichao Wang, Ke Liu, Xuhui Zhang, Xiangrong Meng, dan Jun Li pada tahun 2021, yang berfokus pada prediksi karbon yang tidak terbakar dalam debu cerobong berdasarkan metode ansambel Guru-Siswa, menggunakan algoritma *ANN* (Artificial Neural Network) untuk model "guru" dan algoritma integrasi antara *Xgboost* dan *AdaBoost* untuk model "murid". Hasil penelitian menunjukkan error sebesar 7,16% pada model *ANN* dan 5,98% pada algoritma integrasi (Lu *et al.*, 2021).

Penelitian ini mengadopsi tiga algoritma *machine learning* untuk memprediksi nilai *unburned carbon*, yaitu *regresi linier*, *random forest*, dan *LightGBM regression*. Penulis akan membandingkan ketiga algoritma tersebut untuk menentukan algoritma terbaik dalam memprediksi nilai *unburned carbon*, dengan menggunakan parameter evaluasi berupa *MAE* (*Mean Absolute Error*), *MAPE* (*Mean Absolute Percentage Error*), dan *RMSE* (*Root Mean Squared Error*). Penelitian ini

diharapkan dapat berkontribusi pada peningkatan efisiensi operasional PLTU dengan menyediakan prediksi nilai *unburned carbon* yang lebih akurat, sehingga proses pembakaran dapat dioptimalkan dan biaya operasional dapat dikurangi.

2. Metodologi Penelitian

Gambaran *dataset*

Penelitian ini menggunakan *dataset* parameter PLTU sejak bulan April tahun 2016 sampai dengan bulan

Desember tahun 2023. Fitur yang diambil dalam penelitian ini yaitu *time*, *unburned carbon*, nilai kalori batubara, CO, CO₂, outlet generator, *Specific Fuel Consumption (SFC)*, *Eco gas Outlet Temperature*, *primary air flow*, O₂, *fuel temperature*, *secondary air flow*, *carbon*, *hydrogen*, *nitrogen*, *sulfur*, *ash*, *tota moisture*, *oxygen*, *surface moisture*, *inherent moisture*. Semua parameter tersebut digunakan untuk mencari nilai *unburned carbon* yang terkandung dalam batubara. Berikut merupakan contoh sampel *dataset* yang digunakan.

Tabel 1. *Dataset Unburned Carbon*

<i>Time</i>	<i>Unburned carbon</i>	Nilai kalori	...	<i>Surface moisture</i>	<i>Inherent moisture</i>
13-04-16	3.6590	4559.68	...	11.59	17.39
09-05-16	1.5045	4739.94	...	10.93	16.39
07-06-16	2.4490	4832.12	...	10.62	15.93
22-07-16	0.9680	4599.99	...	11.54	17.31
...	...	...	...	...	...
12-2023	0.085	3982.95	...	15.01	22.52

Dataset parameter PLTU yang diambil yaitu data historis *performance test* pembangkit. kegiatan *performance test* dilakukan setiap satu bulan sekali. *Dataset* tersebut terdiri dari 21 fitur dengan 89 data. *Dataset* terdiri dari 1 data objek dan 20 data numeric.

	unburn_carbon	nilai_kalor	co	...	pa_flow	carbon
count	89.0	89.0	89.0	...	89.0	89.0
mean	0.814172	4214.927525	0.013522	...	973.944449	44.307748
std	1.09878	344.497532	0.042256	...	75.06852	3.527327
min	0.026959	3375.90205	0.0	...	730.391084	35.864958
25%	0.211	4002.608435	0.00029	...	931.945168	41.945534
50%	0.499	4144.0	0.00125	...	977.58978	43.904449
75%	0.968	4398.0	0.003	...	1022.365298	46.587114
max	7.5885	5067.0	0.2869	...	1129.129841	52.024112
...	...	...	...	...	...	...

Gambar 1. Deskripsi *Dataset*

Dataset ini menunjukkan variasi signifikan pada beberapa fitur penting pembakaran. *Unburned carbon* memiliki rata-rata 0.81 dengan rentang luas hingga 7.59, mengindikasikan perbedaan besar dalam efisiensi pembakaran. Nilai kalor (rata-rata 4214.93) dan PA Flow (rata-rata 973.94) juga bervariasi, mencerminkan fluktuasi dalam energi dan distribusi udara. CO memiliki rata-rata rendah (0.0135) tetapi maksimum tinggi (0.2869), menunjukkan beberapa pembakaran kurang sempurna. Dengan variasi data

ini, analisis korelasi dapat membantu mengidentifikasi faktor-faktor yang memengaruhi performa pembakaran dan emisi.

Tahapan penelitian

Tahapan pertama dimulai dengan studi pustaka, yang dilakukan untuk menelaah penelitian terdahulu yang relevan dengan *prediksi unburned carbon* serta metode regresi yang digunakan dalam penelitian ini, seperti terlihat pada Gambar 2.



Gambar 2. Tahapan Penelitian

Tahapan kedua adalah pengumpulan data, di mana *dataset* yang digunakan dalam penelitian ini dikumpulkan dari hasil pengukuran historis dari boiler, yang mencakup berbagai parameter. *Dataset* tersebut akan menjadi studi kasus untuk penelitian. Setelah data dikumpulkan, langkah berikutnya adalah pra-pemrosesan data, yang mencakup penanganan *missing data*, *outlier*, dan normalisasi fitur-fitur numerik yang akan digunakan dalam model prediksi. Proses ini dilakukan untuk memastikan kualitas data yang optimal sebelum dilanjutkan ke tahap berikutnya. Tahapan selanjutnya adalah penerapan algoritma prediksi, di mana *dataset* yang sudah diproses digunakan untuk melatih tiga model regresi, yaitu *Linear Regression*, *Random Forest Regression*, dan *LightGBM Regression*. Setiap algoritma diuji secara terpisah untuk memprediksi nilai *unburned carbon* berdasarkan fitur-fitur yang tersedia dalam *dataset*. Setelah pengujian ketiga algoritma selesai, dilakukan evaluasi model dengan menggunakan metrik seperti *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE) dan *Root Mean Squared Error* (RMSE) untuk membandingkan performa ketiga algoritma tersebut. Hasil evaluasi ini digunakan untuk menilai model mana yang memberikan prediksi *unburned carbon* yang paling akurat. Tahapan terakhir adalah analisis hasil, di mana hasil evaluasi dibandingkan, dan didiskusikan mengenai kelebihan dan kekurangan dari setiap model yang telah diuji.

Algoritma linear regression

Regresi linear sederhana adalah teknik yang digunakan untuk memodelkan hubungan antara satu variabel yang bergantung pada variabel lain yang independen (Muhartini *et al.*, 2021). Tujuan utama dari regresi linear adalah untuk memahami dan memodelkan hubungan ini sehingga kita dapat menggunakan informasi dari variabel independen untuk memprediksi nilai variabel dependen. Model regresi linear sederhana dinyatakan dalam persamaan matematis berikut:

$$Y = a + bX$$

Dimana :

Y = Variabel dependen

X = Variabel independen

a = Konstanta / Intercept

b = Koefisien regresi / Slope

Analisis regresi linear berguna dalam berbagai bidang, termasuk ilmu sosial, ekonomi, ilmu alam, dan banyak lagi. Ini membantu memahami dan memprediksi hubungan antara variabel-variabel yang penting dalam penelitian dan pengambilan Keputusan.

Algoritma random forest

Random Forest (RF) adalah algoritma yang memanfaatkan teknik pemisahan biner secara berulang untuk mencapai simpul akhir dalam struktur pohon (Yoga Religia *et al.*, 2021). *Random Forest* dapat dikatakan sebagai sebuah algoritma *Machine Learning* yang menggunakan konsep *Ensemble Learning*. *Ensemble Learning* mengacu pada pendekatan menggabungkan beberapa model untuk meningkatkan kinerja prediksi secara keseluruhan. *Random Forest* secara khusus menggabungkan pohon keputusan (*Decision Trees*) dalam ensemble-nya. Keunggulan *Random Forest* meliputi stabilitas (karena menggunakan banyak pohon), kemampuan menangani data besar, dan ketahanannya terhadap nilai ekstrem (*outliers*) karena mengambil hasil yang paling umum dari prediksi. Dalam penelitian, *Random Forest* sering digunakan dalam berbagai bidang, seperti prediksi kesehatan, analisis keuangan, dan klasifikasi citra, karena keunggulannya dalam mengatasi *overfitting* dan meningkatkan akurasi prediksi.

Algoritma Lightgbm regression

Algoritma *LightGBM Regression* merupakan pendekatan *Machine Learning* yang menggabungkan kekuatan pohon keputusan (*Decision Trees*) dengan teknik *boosting* untuk membangun model regresi yang kuat. Pohon keputusan digunakan untuk memetakan hubungan antara variabel dependen dan independen dalam data. Teknik *boosting* yang digunakan dalam *LightGBM* memungkinkan kombinasi model-model lemah untuk membentuk model yang lebih kuat melalui iterasi berulang, fokus pada pengurangan kesalahan prediksi dari iterasi sebelumnya (Liao *et al.*, 2022). *LightGBM* juga menggunakan *leaf-wise splitting*, di mana setiap *split* dipilih berdasarkan penurunan kesalahan yang paling signifikan, menghasilkan pohon yang lebih dalam dan kompleks. Metode ini didukung oleh optimasi berbasis *gradien*, yang meminimalkan fungsi kerugian selama pelatihan model. Keunggulan *LightGBM* meliputi kemampuan menangani data besar, efisiensi komputasi yang tinggi, penanganan nilai yang hilang, dan kemampuan untuk menangani

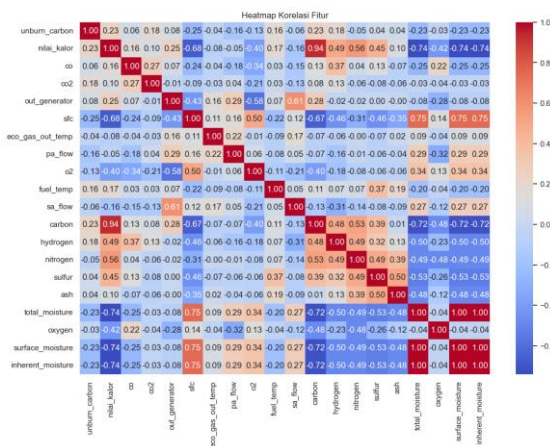
masalah regresi dengan akurat dan efisien. *LightGBM Regression* menjadi pilihan yang populer dalam memprediksi nilai variabel target berdasarkan fitur-fitur yang ada dalam *dataset*.

3. Hasil dan Pembahasan

Hasil

Eksplorasi Data dan Analisa Visual

Menampilkan gambar grafik pola atau anomali yang ditampilkan pada model.



Gambar 3. *Heatmap Dataset Unburned Carbon*

Berdasarkan *heatmap* korelasi di atas, dapat dilihat beberapa temuan penting terkait hubungan antar variabel dalam pembakaran bahan bakar. Korelasi positif yang kuat antara kandungan karbon dan nilai kalor ($r = 0.94$) menunjukkan bahwa semakin tinggi kandungan karbon dalam bahan bakar, semakin besar energi yang dihasilkan. Di sisi lain, korelasi negatif yang signifikan antara kadar kelembapan total dan nilai kalor ($r = -0.74$) menyoroti bahwa kelembapan yang lebih tinggi dalam bahan bakar dapat mengurangi efisiensi pembakaran karena energi terbuang untuk menguapkan air sebelum menghasilkan panas. Hubungan yang kuat juga terlihat antara kelembapan total, kelembapan permukaan, dan kelembapan bawaan ($r = 1.00$), yang mengindikasikan bahwa ketiga variabel ini secara substansial mewakili aspek yang sama dari kandungan air dalam bahan bakar. Selain itu, beberapa variabel menunjukkan korelasi yang rendah, seperti hubungan yang relatif rendah antara nilai kalor dan output generator ($r = 0.25$), yang

mungkin dipengaruhi oleh faktor operasional lain di luar bahan bakar. Korelasi antara sulfur dan oksigen ($r = -0.40$) juga lebih lemah dari yang diharapkan, menunjukkan adanya faktor lain yang mungkin memengaruhi distribusi sulfur dan oksigen selama pembakaran.

Dalam penelitian ini, *preprocessing* data dilakukan menggunakan *pipeline* untuk memastikan setiap fitur diproses sesuai karakteristiknya, sehingga meminimalkan potensi bias dan meningkatkan akurasi model. *Pipeline* ini terdiri dari dua jalur: satu untuk fitur numerik dan satu lagi untuk fitur kategori. Pada jalur numerik, dilakukan imputasi nilai hilang menggunakan *SimpleImputer* dengan strategi '*median*', yang bertujuan untuk menggantikan nilai yang hilang dengan nilai tengah dari distribusi data, sehingga hasilnya tidak terlalu dipengaruhi oleh nilai ekstrem. Setelah itu, fitur numerik dinormalisasi menggunakan *MinMaxScaler*, yang mengubah nilai ke rentang $[0, 1]$, menjaga proporsi data dan memudahkan konvergensi dalam model yang sensitif terhadap skala, seperti algoritma berbasis jarak atau gradien. Selanjutnya akan dilakukan pemodelan menggunakan algoritma *linear regression*, *random forest* dan *lightgbm regression*.

Hasil evaluasi model

Hasil evaluasi performa model dalam memprediksi nilai *unburned carbon* disajikan menggunakan tiga metrik utama: *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), dan *Root Mean Squared Error* (RMSE). Ketiga metrik ini digunakan untuk mengevaluasi akurasi prediksi dan perbedaan hasil antara model yang diterapkan. Melalui analisis ini, akan terlihat seberapa baik masing-masing model dalam memprediksi *unburned carbon* berdasarkan data yang tersedia, sehingga dapat diidentifikasi model mana yang memiliki performa terbaik. Nilai MAE menunjukkan rata-rata kesalahan (*error*) absolut antara hasil peramalan/prediksi dengan nilai riil (Suryanto, 2019). *Mean Absolute Percentage Error* (MAPE) digunakan apabila ukuran variable pada peramalan yang dilakukan adalah faktor yang berpengaruh dalam melakukan evaluasi akurasi peramalan yang dilakukan (Nabillah & Ranggadara, 2020). *Root mean square error* (RMSE) dan *mean absolute error* (MAE) adalah dua ukuran kesalahan absolut yang digunakan dalam berbagai macam disiplin ilmu; misalnya pembelajaran mesin dan penggalian data (Karunasingha, 2022).

	model	Train_Mean	std	Test_Score
0	LightGBM	0.32	0.03	0.31
1	Random Forest	0.36	0.03	0.31
2	Linear Regression	0.42	0.14	0.35

Gambar 4. Evaluasi model menggunakan MAE

Gambar evaluasi *Mean Absolute Error* menunjukkan bahwa model terbaik dalam memprediksi nilai *unburned carbon* yaitu *lightGBM regression* dengan nilai eror sebesar 0,31 pada *test set* dan 0,32 pada *training set*. Sedangkan pada posisi kedua yaitu terdapat algoritma *random forest* yang memiliki nilai eror 0,31 pada *test set* dan 0,36 pada *training set*. Sedangkan nilai eror yang paling besar yaitu algoritma *linear regression* dengan nilai eror 0,35 pada *test set* dan 0,42 pada *training set*.

	model	Train_Mean	std	Test_Score
0	LightGBM	1.44	0.62	1.29
1	Random Forest	1.80	0.96	1.35
2	Linear Regression	2.29	1.73	1.38

Gambar 5. Evaluasi model menggunakan MAPE

Gambar Evaluasi MAPE menunjukkan hal yang sama yaitu algoritma terbaik *lightGBM regression* dengan nilai eror 1,44 pada *training set* dan 1,29 pada *test set*. Kemudian algoritma *random forest* memiliki nilai eror 1,80 pada *training set* dan 1,35 pada *test set*. *Linear regression* masih memiliki nilai eror yang terbesar yaitu 2,29 pada *training set* dan 1,38 pada *test set*.

	model	Train_Mean	std	Test_Score
0	LightGBM	0.42	0.04	0.38
1	Random Forest	0.44	0.05	0.41
2	Linear Regression	0.55	0.20	0.43

Gambar 6. Evaluasi model menggunakan RMSE

Pada matriks evaluasi RMSE hasil evaluasi menunjukkan hal yang sama yaitu algoritma terbaik dalam memprediksi nilai *unburned carbon* adalah *lightGBM regression* dengan nilai eror 0,42 pada *training set* dan 0,38 pada *test set*. Kemudian dilanjutkan dengan algoritma *Random Forest* dengan nilai eror 0,44 pada *training set* dan 0,41 pada *test set*. *Linear regression* memiliki peringkat akhir dalam evaluasi yaitu dengan nilai eror 0,55 pada *training set* dan 0,43 pada *test set*.

Analisis hasil dan implikasi

Pada bagian ini, akan dianalisis hasil dari pemodelan prediksi nilai *unburned carbon* yang telah dilakukan dengan menggunakan tiga algoritma *machine learning* yang berbeda. Hasil yang diperoleh akan dibandingkan berdasarkan metrik evaluasi yang telah ditentukan, yaitu MAE (*Mean Absolute Error*), MAPE (*Mean Absolute Percentage Error*), dan RMSE (*Root Mean Squared Error*). Selanjutnya, implikasi dari temuan tersebut akan dibahas, dengan fokus pada pengaruhnya terhadap efisiensi operasional PLTU dan kontribusinya terhadap pengurangan dampak lingkungan yang dihasilkan oleh pembangkit listrik tenaga uap.

Tabel 2. Evaluasi Model Pada Test Set

Algoritma	MAE	MAPE	RMSE
<i>LightGBM Regression</i>	0,31	1,29	0,38
<i>Random Forest</i>	0,31	1,35	0,41
<i>Linear Regression</i>	0,35	1,38	0,43

Berdasarkan hasil evaluasi prediksi *unburned carbon* pada *test set*, *LightGBM Regression* terbukti memiliki performa terbaik dibandingkan dengan *Random Forest* dan *Linear Regression*. Dengan nilai MAE sebesar 0,31, MAPE 1,29, dan RMSE 0,38, *LightGBM* menunjukkan kemampuan prediksi yang lebih akurat dan stabil. Algoritma ini mengungguli dua algoritma lainnya dalam hal kesalahan prediksi yang lebih rendah pada ketiga metrik evaluasi. *Random Forest* juga menunjukkan performa yang cukup baik, namun dengan MAPE yang sedikit lebih tinggi (1,35) dan RMSE sebesar 0,41. Sementara itu, *Linear Regression* memiliki kesalahan prediksi yang lebih tinggi, menunjukkan bahwa model linear ini kurang sesuai untuk menangani kompleksitas data prediksi *unburned carbon*. Tinjauan dari hasil evaluasi menunjukkan bahwa *LightGBM Regression* adalah model dengan performa terbaik untuk prediksi *unburned carbon* dan memiliki potensi untuk dikembangkan lebih lanjut. Potensi pengembangan ini dapat dilakukan melalui optimisasi lebih dalam terhadap *hyperparameter* atau dengan mempertimbangkan penambahan fitur relevan lainnya untuk lebih meningkatkan akurasi prediksi. Selain itu, pemahaman yang lebih baik terhadap data dan hubungan antar fitur dapat membantu memperbaiki hasil prediksi di masa mendatang.

Pembahasan

Pada penelitian ini, tiga algoritma *machine learning linear regression*, *random forest*, dan *LightGBM regression* digunakan untuk memprediksi nilai *unburned carbon* pada PLTU. Berdasarkan hasil yang diperoleh, setiap algoritma menunjukkan hasil yang bervariasi dalam hal akurasi prediksi, dengan *LightGBM regression* menghasilkan performa terbaik. Dari evaluasi yang dilakukan menggunakan metrik MAE (*Mean Absolute Error*), MAPE (*Mean Absolute Percentage Error*), dan RMSE (*Root Mean Squared Error*), model *LightGBM regression* menunjukkan hasil yang lebih unggul dibandingkan dengan model *linear regression* dan *random forest*.

Nilai MAE yang diperoleh adalah 0.31, MAPE sebesar 1.29%, dan RMSE sebesar 0.38, yang menunjukkan tingkat kesalahan yang lebih rendah jika dibandingkan dengan hasil dua algoritma lainnya (Liao *et al.*, 2022). Hasil ini sesuai dengan temuan dari penelitian sebelumnya yang juga menemukan bahwa *LightGBM* memberikan hasil yang lebih akurat dalam aplikasi prediksi berbasis *machine learning* (Liao *et al.*, 2022; Lu *et al.*, 2021). Implikasi dari temuan ini sangat signifikan untuk efisiensi operasional PLTU. Penggunaan model prediktif seperti *LightGBM regression* memungkinkan prediksi nilai *unburned carbon* yang lebih akurat, yang dapat digunakan untuk mengatur parameter pembakaran dalam *boiler* dengan lebih optimal. Hal ini akan berkontribusi dalam meningkatkan efisiensi bahan bakar dan menurunkan konsumsi energi, yang pada gilirannya dapat mengurangi biaya operasional pembangkit listrik (Bajpai *et al.*, 2024).

Selain itu, pengurangan nilai *unburned carbon* secara langsung dapat menurunkan emisi yang dihasilkan, berkontribusi pada keberlanjutan lingkungan dan mengurangi polusi udara, yang merupakan tantangan penting bagi industri energi (Adamczyk *et al.*, 2021). Penurunan *unburned carbon* juga berperan dalam meningkatkan efisiensi pembakaran dalam pembangkit listrik berbahan bakar batubara, seperti yang dijelaskan dalam penelitian oleh Kumar dan Dwivedi (n.d.), yang menunjukkan bahwa pengaturan kemiringan burner dan presentase oksigen dapat menurunkan kadar *unburned carbon*. Selain itu, berdasarkan hasil penelitian oleh Zaman dan Sinaga (2021), efisiensi pembangkit listrik juga

dapat dievaluasi melalui pengurangan nilai *unburned carbon* sebagai salah satu faktor kunci yang mempengaruhi performa keseluruhan pembangkit. Penerapan algoritma *machine learning* untuk memprediksi nilai *unburned carbon* menawarkan berbagai keuntungan, termasuk pengurangan emisi, penghematan biaya, dan peningkatan efisiensi operasional, yang sejalan dengan tren global menuju pembangkit listrik yang lebih ramah lingkungan dan efisien. Implementasi lebih lanjut dari model ini dapat mendorong inovasi teknologi dalam sektor energi, seperti yang tercermin dalam penelitian-penelitian terbaru (Lewerissa, 2018; Yoga Religia *et al.*, 2021). Hasil penelitian ini menunjukkan bahwa penerapan *machine learning* dalam prediksi *unburned carbon* dapat memberikan kontribusi signifikan terhadap pengelolaan pembangkit listrik yang lebih efisien dan berkelanjutan. Model prediktif yang lebih akurat, seperti *LightGBM regression*, menawarkan potensi untuk mengurangi emisi dan biaya operasional, serta memperbaiki kinerja PLTU dalam jangka panjang.

4. Kesimpulan

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa model *LightGBM Regression* menunjukkan performa terbaik dalam memprediksi nilai *unburned carbon* pada pembangkit listrik tenaga uap, dengan evaluasi yang sangat baik, yakni MAE sebesar 0,31, MAPE sebesar 1,29%, dan RMSE sebesar 0,38. Jika dibandingkan dengan algoritma lainnya, *LightGBM Regression* mengungguli model *Random Forest* yang memiliki performa cukup baik, serta *Linear Regression* yang menunjukkan tingkat kesalahan prediksi paling tinggi.

Selain itu, model *LightGBM Regression* memiliki potensi untuk dikembangkan lebih lanjut, terutama melalui optimasi parameter dan penambahan fitur relevan yang dapat meningkatkan akurasi prediksi. Dengan prediksi nilai *unburned carbon* yang lebih akurat, diharapkan dapat tercapai peningkatan efisiensi pembakaran, penurunan biaya operasional, serta pengurangan dampak lingkungan, seperti emisi yang dihasilkan oleh pembangkit listrik tenaga uap.

5. Daftar Pustaka

- Adamczyk, Z., Komorek, J., Bialecka, B., Calus-Moszek, J., & Klupa, A. (2021). Possibilities of graphitization of unburned carbon from coal fly ash. *Minerals*, 11(9), 1027. <https://doi.org/10.3390/min11091027>.
- Bajpai, P., Sheorey, T., Kumar, D. S., Singh, A. K., & Singh, J. P. (2024). Modeling of Unburned Carbon in Pulverized Coal-Fired Supercritical Steam Generators. *Combustion Science and Technology*, 1-16. <https://doi.org/10.1080/00102202.2024.2326861>.
- Galih, M., & Atika, P. D. (2022). Prediksi Penjualan Menggunakan Algoritma Regresi Linear pada Koperasi Karyawan Usaha Bersama. *Journal of Informatic and Information Security*, 3(2), 193-202.
- Hetharia, M., & Lewerissa, Y. J. (2018). analisis energi pada perencanaan pembangkit listrik tenaga uap (PLTU) dengan cycle tempo. *Jurnal Voering*, 3(1).
- Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, 585, 609-629. <https://doi.org/10.1016/j.ins.2021.11.036>.
- Kumar, S., & Dwivedi, V. (2023). Performance Optimization of Steam Generator by Reduction of Unburned Carbon. *International Journal of Scientific Research in Mechanical and Materials Engineering*, 7(3), 51-65.
- Liao, S., Tian, X., Liu, B., Liu, T., Su, H., & Zhou, B. (2022). Short-Term Wind Power Prediction Based on LightGBM and Meteorological Reanalysis. *Energies*, 15(17), 6287. <https://doi.org/10.3390/en15176287>.
- Lu, K., Yang, X., Wang, H., Liu, K., Zhang, X., Meng, X., & Li, J. (2021, February). Prediction of Unburned Carbon in Flue Dust based on Teacher-student Ensemble Method. In *IOP Conference Series: Earth and Environmental Science* (Vol. 634, No. 1, p. 012030). IOP Publishing. <https://doi.org/10.1088/1755-1315/634/1/012030>.
- Muhartini, A. A., Sahroni, O., Rahmawati, S. D., Febrianti, T., & Mahuda, I. (2021). Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana. *Jurnal Bayesian: Jurnal Ilmiah Statistika dan Ekonometrika*, 1(1), 17-23. <https://doi.org/10.46306/bay.v1i1.2>.
- Nabillah, I., & Rangadara, I. (2020). Mean absolute percentage error untuk evaluasi hasil prediksi komoditas laut. *Journal of Information System*, 5(2), 250-255.
- Religia, Y., Nugroho, A., & Hadikristanto, W. (2021). Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), 187-192. <https://doi.org/10.29207/resti.v5i1.2813>.
- Suryanto, A. A., & Muqtadir, A. (2019). Penerapan metode mean absolute error (MEA) dalam algoritma regresi linear untuk prediksi produksi padi. *Saintekbu*, 11(1), 78-83. <https://doi.org/10.32764/saintekbu.v11i1.298>.
- Umiyati, A., Dasari, D., & Agustina, F. (2021). Peramalan Harga Batubara Acuan Menggunakan Metode PSOSVR dan IPSOSVR. *Jurnal EurekaMatika*, 9(1), 69-94.
- Zaman, M. R., & Sinaga, N. (2021). THERMAL LOSSES EVALUATION IN 660 MW COAL-FIRED POWER PLANT USING INDIRECT EFFICIENCY METHOD. *Jurnal Sains dan Teknologi Reaksi*, 19(01). <https://doi.org/10.30811/jstr.v19i01.2263>.