



Latin American Journal of Energy Research – Lajer (2026) v. 13, n. 1, p. 71–80
<https://doi.org/10.21712/lajer.2026.v13.n1.p71-80>

Comparação do desempenho dos algoritmos XGBoost e LightGBM na previsão de potência de sistemas fotovoltaicos *Comparison of the performance of XGBoost and LightGBM algorithms in power prediction of photovoltaic systems*

José Airton Azevedo dos Santos^{1,*}, Dalila Natalia Costa da Luz²

¹ Professor do Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná – UTFPR, campus Medianeira, PR, Brasil

² Aluna do Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná – UTFPR, campus Medianeira, PR, Brasil

*Autor para correspondência, E-mail: airton@utfpr.edu.br

Received: 16 February 2026 | Accepted: 21 March 2026 | Published online: 29 April 2026

Resumo: Este trabalho tem como objetivo comparar o desempenho, na previsão da potência gerada por dois sistemas fotovoltaicos, dos algoritmos, baseados em *boosting*, XGBoost (*eXtreme Gradient Boosting*) e LightGBM (*Light Gradient Boosting Machine*). A base de dados foi obtida de uma Estação de Pesquisa em Energia Solar (EPESOL). A EPESOL possui dois sistemas distintos de energia solar fotovoltaica: um utiliza silício monocristalino (m-Si) e o outro silício policristalino (p-Si). A base de dados compreende 25235 observações, para cada tecnologia, e inclui as seguintes variáveis: Potência elétrica DC, Temperatura dos módulos, Irradiação solar e Hora do dia. Os hiperparâmetros dos modelos foram otimizados por meio da biblioteca Optuna. Para avaliar o desempenho, dos modelos LightGBM e XGBoost, foram utilizadas as métricas R^2 , MAE e RMSE. Os resultados obtidos indicaram, para ambos os sistemas fotovoltaicos, uma forte correlação entre valores reais e previstos ($R^2 > 0,96$). Este desempenho evidencia a eficácia, dos modelos XGBoost e LightGBM, na previsão da geração de potência em sistemas fotovoltaicos.

Palavras-chave: aprendizado de máquina; energia solar; modelos preditivos; optuna.

Abstract: This work aims to compare the performance of two photovoltaic systems, using boosting-based algorithms XGBoost (*eXtreme Gradient Boosting*) and LightGBM (*Light Gradient Boosting Machine*). The database was obtained from a Solar Energy Research Station (EPESOL). EPESOL has two distinct photovoltaic solar energy systems: one uses monocrystalline silicon (m-Si) and the other polycrystalline silicon (p-Si). The database comprises 25235 observations for each technology and includes the following variables: DC power, module temperature, solar irradiance, and time of day. The model hyperparameters were optimized using the Optuna library. To evaluate the performance of the LightGBM and XGBoost models, the R^2 , MAE, and RMSE metrics were used. The results obtained indicated, for both photovoltaics systems, a strong correlation between actual and predicted values ($R^2 > 0.96$). This performance demonstrates the effectiveness of the XGBoost and LightGBM models in predicting power generation in photovoltaic systems.

Keywords: machine learning; solar energy; predictive models; optuna.

1 Introdução

O aumento no consumo de energia, impulsionado pelo progresso tecnológico e pelo crescimento socioeconômico, revela dois grandes desafios: a poluição causada pelos combustíveis fósseis e o risco de seu esgotamento. Embora o carvão, petróleo e gás natural ainda dominem a matriz energética mundial, seu uso intensivo emite grandes quantidades de CO_2 , agravando o aquecimento global e a degradação ambiental. Além disso, a disponibilidade limitada desses recursos ameaça o abastecimento futuro, com algumas regiões já enfrentando escassez. É essencial, para evitar crises energéticas, investir em fontes

renováveis, como solar, eólica e hidrelétrica, garantindo sustentabilidade e segurança energética a longo prazo (Abdellatif et al., 2022; Kanwal et al., 2022; Zahoor et al. 2022; Couto et al., 2021; Mlilo et al., 2021).

A energia solar, dentre as fontes renováveis, se destaca como uma alternativa promissora, apresentando, em comparação com as fontes de energia tradicionais, vantagens como: menor necessidade de manutenção, inexistência de emissões de Gases de Efeito Estufa (GEE), maior vida útil e retorno financeiro mais previsível. Contudo, a geração de energia solar, devido à variabilidade natural da irradiação, está sujeita a flutuações. Estas flutuações de energia podem provocar desequilíbrios entre oferta e demanda na rede elétrica. Esses desequilíbrios comprometem a estabilidade e a confiabilidade do sistema. Nestas circunstâncias, uma previsão precisa da produção fotovoltaica torna-se essencial para garantir a operação segura e contínua do sistema elétrico (Akhter et al., 2022; Antonanzas et al., 2016; Ramsami e Oree, 2015; Ye et al., 2022; Bouziane et al., 2024).

Vários trabalhos, apresentados na literatura, utilizaram modelos, de aprendizado de máquina, para preverem variáveis relacionadas com sistemas fotovoltaicos. Dentre eles podem-se citar: Akhter et al. (2022) que realizaram uma previsão horária, da produção de energia fotovoltaica, analisando três sistemas distintos: policristalino, monocristalino e de filme fino. Utilizaram, como parâmetros de entrada, a irradiação solar, a temperatura ambiente, a temperatura dos módulos e a velocidade do vento. Konstantinou et al. (2021) utilizaram, para previsão da geração de energia solar fotovoltaica, a rede neural recorrente LSTM. Dados históricos de uma usina fotovoltaica em Nicósia, Chipre, foram usados na previsão da geração de energia fotovoltaica. Observaram, neste trabalho, que o modelo LSTM apresentou um bom desempenho de previsão. Abdellatif et al. (2022) compararam o desempenho, dos modelos *Random Forest*, *Stack-ETR*, *XGBoost* e *AdaBoost*, na previsão da geração de energia fotovoltaica na Malásia. O estudo comparou o desempenho, dos quatro algoritmos, em diferentes tecnologias de painéis solares: filme fino, monocristalinos e policristalinos. Concluíram que o modelo *Stack-ETR* apresentou, com relação aos outros modelos, melhor desempenho. Já Oliveira e Santos (2022) utilizaram modelos, baseados em redes neurais artificiais e regressão linear múltipla, para predição da corrente elétrica gerada por um sistema solar fotovoltaico, localizado na região oeste paranaense. Concluíram, neste trabalho, que o modelo de rede neural artificial é mais adequado, para prever a corrente elétrica do sistema fotovoltaico, do que o modelo de regressão linear.

Embora diversos estudos tenham utilizado técnicas de aprendizado de máquina, para previsão de variáveis relacionadas a sistemas fotovoltaicos, ainda são escassas as pesquisas que comparam de forma direta, em condições reais de operação, o desempenho dos algoritmos *XGBoost* e *LightGBM* na previsão de potência de sistemas fotovoltaicos com diferentes tecnologias de módulos fotovoltaicos (monocristalinos e policristalinos). Neste contexto, este trabalho tem como objetivo comparar, na previsão da potência gerada por dois sistemas fotovoltaicos, o desempenho dos algoritmos, baseados em *boosting*, *XGBoost* e *LightGBM*.

O artigo é organizado da seguinte maneira. Na Seção 2 é apresentada a metodologia utilizada na previsão da potência gerada pelos sistemas fotovoltaicos. Na seção 3 são discutidos os resultados obtidos e suas respectivas análises. Na seção 4 conclui-se o trabalho com observações finais e conclusões.

2 Materiais e métodos

Nesta seção apresentam-se os algoritmos de previsão, a estação EPESOL, a base de dados, bem como as métricas utilizadas.

2.1 Algoritmos

eXtreme Gradient Boosting (XGBoost): O XGBoost é um dos algoritmos, de aprendizado de máquina, mais populares e eficazes para tarefas de regressão e classificação. Este algoritmo utiliza a técnica de *gradient boosting*, que consiste em construir um conjunto de árvores de decisão de forma iterativa. Em cada iteração, uma nova árvore é treinada para corrigir os erros das árvores anteriores. O XGBoost, devido a sua excelente capacidade de unir velocidade e precisão nas previsões, tem sido amplamente aplicado em muitas áreas do conhecimento (Solano et al., 2022).

Light Gradient Boosting Machine (LightGBM): O LightGBM é um algoritmo de aprendizado de máquina, desenvolvido pela Microsoft em 2016, que pode ser aplicado para problemas de regressão e classificação. Este algoritmo aplica, com foco específico em eficiência e velocidade, técnicas de *gradient boosting* para construir um conjunto de árvores de decisão. O LightGBM se destaca como um dos

algoritmos mais eficientes e precisos para modelagem preditiva, sendo especialmente eficaz em cenários com grandes volumes de dados e múltiplas variáveis. Sua velocidade, capacidade de lidar com *overfitting* e flexibilidade tornam-no uma excelente escolha para problemas de previsão em energia solar (Tang et al., 2020; Guo et al., 2023).

2.2 Estação EPESOL

A estação de pesquisa EPESOL (Figura 1), localizada no campus Medianeira, foi criada a partir de uma parceria entre a Companhia Paranaense de Energia (COPEL) e o Laboratório de Energia Solar (LABENS) da Universidade Tecnológica Federal do Paraná (UTFPR). A EPESOL possui um sistema fotovoltaico de silício monocristalino (m-Si) de tecnologia PERC com 5,11 kWp (14 módulos de 365Wp cada) e outro sistema de silício policristalino (p-Si) com 4,69 kWp (14 módulos de 335Wp cada). Os módulos foram instalados nas condições ótimas, ou seja, com o ângulo de inclinação igual à latitude do local e orientados para o norte geográfico. A EPESOL possui uma estação solarimétrica de padrão SONDA (Nakahata, 2023; Santos e Polo, 2025)



Figura 1. Estação de pesquisa EPESOL.

2.3 Bases de dados e pré-processamento

Os dados gerados pela EPESOL são armazenados, por meio de um sistema de aquisição da Campbell Scientific, e transmitidos, em arquivos no formato CSV, via internet para um servidor FTP. O sistema de aquisição foi programado para realizar a leitura e armazenar, por minuto, os dados das componentes de irradiação global, temperatura de operação, tensão e corrente gerada pelos sistemas monocristalino e policristalino. As bases de dados, com 25235 observações cada, são formadas pelas variáveis: Potência elétrica DC (*Pot*), Temperatura dos módulos (*TP*), Irradiação solar (*RS*) e *Hora*. Na Figura 2 apresentam-se os cinco primeiros registros das bases de dados dos sistemas monocristalino e policristalino.

	Monocristalino				Policristalino				
	TP	RS	Pot	hora	TP	RS	Pot	hora	
0	18.48	-5.864128	0.0	0.000000	0	18.40	-5.864128	0.0	0.000000
1	18.52	-5.912101	0.0	0.016667	1	18.39	-5.912101	0.0	0.016667
2	18.54	-5.952342	0.0	0.033333	2	18.39	-5.952342	0.0	0.033333
3	18.56	-6.001678	0.0	0.050000	3	18.33	-6.001678	0.0	0.050000
4	18.56	-6.009477	0.0	0.066667	4	18.32	-6.009477	0.0	0.066667

Figura 2. Cinco primeiros registros das bases de dados.

Inicialmente, na etapa de pré-processamento, observou-se que as bases de dados continham valores ausentes. Optou-se, para garantir a integridade dos dados, por remover todas as instâncias com valores

ausentes. Em seguida, dividiram-se as bases de dados nos conjuntos de treinamento (20188 observações (80%)) e conjuntos de teste (5047 observações (20%)). Realizou-se também, nesta etapa, a normalização dos dados de treinamento, de ambos os sistemas fotovoltaicos, por meio da função `MinMaxScaler`.

Na sequência realizaram-se, para os dois sistemas fotovoltaicos, uma análise da correlação das variáveis *hora*, *TP* e *RS* com a variável *POT*. Na Figura 3 apresentam-se as matrizes de correlação para os dois sistemas (monocristalino e policristalino).

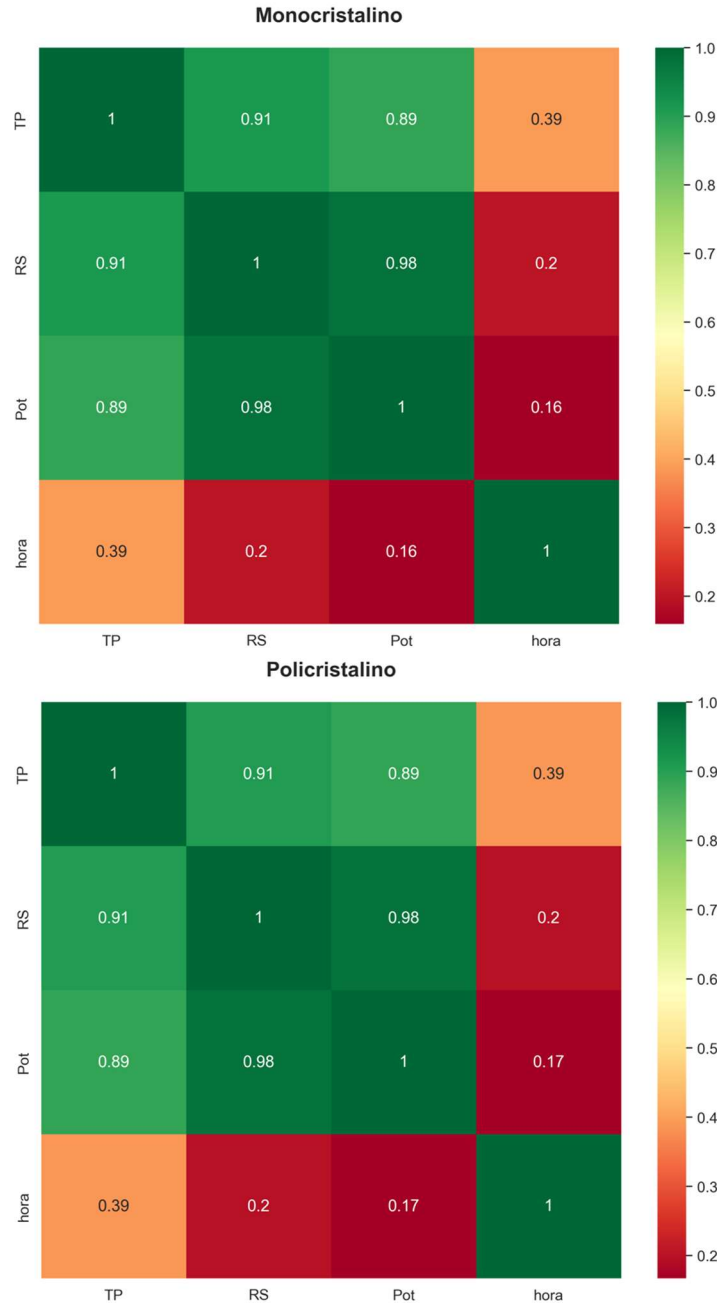


Figura 3. Matriz de correlação – monocristalino e policristalino.

Observou-se, para ambos os sistemas fotovoltaicos, uma estrutura de correlação praticamente idêntica, em que a potência (*Pot*) apresenta correlação linear muito forte com a irradiação solar (*RS*), com coeficientes de 0,98 em ambos os casos. Isto indica que a irradiação é o principal fator explicativo da geração de potência. A temperatura dos módulos (*TP*) também exibe, para as duas tecnologias, correlação elevada com a Potência ($Pot=0,89$). Contudo, a hora do dia apresenta correlação fraca com *Pot* (0,16 no sistema monocristalino e 0,17 no policristalino), o que evidencia que o horário atua mais como um marcador do ciclo diário de geração do que como preditor linear direto.

2.4 Métricas

Os modelos, implementados neste trabalho, foram avaliados pelas seguintes métricas (Santos, 2021; Santos e Chaukoski, 2020).

- Coeficiente de Determinação (R^2): Indica a qualidade do ajuste de um modelo a variável que se pretende explicar (Eq. 1). O coeficiente de determinação varia entre 0 e 1. Quanto mais próximo de 1 melhor o ajuste.

$$R^2 = \frac{\sum_{i=1}^n ((y_i - \hat{y})^2)}{\sum_{i=1}^n ((y_i - \bar{y})^2)} \quad 1$$

- Erro Médio Absoluto (MAE): Média dos valores absolutos dos erros (Eq. 2). Valor perfeito – MAE=0.

$$MAE = \frac{1}{n} \sum_1^n |(y_i - \hat{y}_i)| \quad 2$$

- Raiz do Erro Quadrático Médio (RMSE): Raiz quadrada da média dos quadrados dos erros (Eq. 3). Valor perfeito – RMSE=0.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i)^2)} \quad 3$$

Onde: y_i é o valor real do período i , \bar{y}_i é a média, \hat{y}_i a previsão para o período i e n é o número de observações.

3. Resultados e discussões

Inicialmente, neste trabalho, realizou-se uma análise descritiva dos dados de Temperatura (TP - °C), Irradiação (RS - W/m²) e Potência (Pot - W) (Tabela 1).

Tabela 1. Análise descritiva de dados.

	TP (°C)	RS (W/m ²)	Pot (W)
Monocristalino			
Média	37,01	437,78	1247,85
Mínimo	6,58	0	0
Máximo	63,25	1348,1	4260,68
Desvio padrão	13,59	348,85	1055,58
Policristalino			
Média	36,37	437,77	1168,41
Mínimo	6,67	0	0
Máximo	62,33	1348,1	4395,6
Desvio padrão	13,31	336,22	992,82

A temperatura média dos módulos monocristalinos (37,01 °C) é levemente superior à dos policristalinos (36,37 °C), com desvios padrão próximos (13,59 e 13,31 °C, respectivamente), indicando condições ambientais de operação semelhantes para ambos os sistemas. As temperaturas mínimas e máximas mostram variação significativa ao longo do dia, com valores entre aproximadamente 6,6 °C e superior a 62 °C, o que é esperado em aplicações fotovoltaicas reais.

A irradiância média foi praticamente idêntica para ambas as tecnologias (aproximadamente 437,78 W/m²), o que é consistente, pois os módulos estavam submetidos às mesmas condições ambientais. Os valores máximos de 1348,1 W/m² confirmam períodos de alta incidência solar, enquanto o valor mínimo

de 0 W/m² sugere baixa intensidade no início do período (por volta das 6h) ou ao final (por volta das 18h). Os desvios padrão elevados (348,85 W/m² e 336,22 W/m²) evidenciam a grande variação da irradiância ao longo do dia, com oscilações influenciadas por nuvens, inclinação solar e condições atmosféricas.

A potência média gerada pelos módulos monocristalinos (1247,85 W) foi superior à dos policristalinos (1168,41 W), o que confirma sua maior eficiência de conversão. Ainda assim, os policristalinos apresentaram um pico de geração ligeiramente superior (4395,6 W contra 4260,68 W), sugerindo que, sob condições ideais, o desempenho pode se equiparar ou até superar momentaneamente os monocristalinos. O desvio padrão da potência também foi elevado (1055,58 W para monocristalinos e 992,82 W para policristalinos), o que é típico em medições horárias de um sistema solar ao longo do dia, devido à variabilidade da irradiação e da temperatura.

3.1 Treinamento

A capacidade de generalização de um algoritmo de aprendizado de máquina está diretamente relacionada à escolha adequada de seus hiperparâmetros. Estes hiperparâmetros controlam aspectos essenciais do processo de aprendizado, como regularização, profundidade de árvores, taxa de aprendizado, entre outros. Hiperparâmetros mal ajustados podem levar tanto ao *overfitting* (quando o modelo aprende demais os dados de treinamento, perdendo desempenho em dados novos) quanto ao *underfitting* (quando o modelo não aprende padrões suficientes, resultando em baixa acurácia).

Nesse contexto, o *framework* Optuna se destaca como uma poderosa ferramenta para otimização de hiperparâmetros. Trata-se de uma biblioteca de otimização automática e eficiente, baseada em métodos de Otimização Bayesiana, com técnicas de busca adaptativas como o *Tree-structured Parzen Estimator* (TPE). O Optuna oferece uma abordagem inteligente e automatizada para encontrar combinações de hiperparâmetros, em tarefas específicas, que maximizem o desempenho de modelos de aprendizado de máquina (Srinivas e Katarya, 2022).

Os melhores modelos XGBoost e LighGBM, encontrados pelo Optuna, apresentam os seguintes hiperparâmetros (Tabela 2).

Tabela 2. Parâmetros dos modelos.

Hiperparâmetro	Intervalo de busca	Monocristalino		Policristalino	
		XGBoost	LighGBM	XGBoost	LighGBM
<i>n_estimators</i>	[50,500]	180	220	177	486
<i>learning_rate</i>	[1e ⁻⁵ ,1e ⁻¹]	0,083	0,078	0,076	0,017
<i>max_depth</i>	[3,15]	5	4	5	5
<i>min_child_weight</i>	[1,10]	3	3	9	9
<i>subsample</i>	[0.0,1.0]	0,27	0,27	0,88	0,49
<i>colsample_bytree</i>	[0.0,1.0]	0,88	0,88	0,8	0,84

Aplicou-se também, para garantir a robustez e a confiabilidade do desempenho dos modelos, a validação *TimeSeriesSplit* da biblioteca *scikit-learn*. É uma técnica de avaliação de modelos que respeita estritamente a ordem cronológica dos dados, impedindo o vazamento de informações futuras para o passado (*data leakage*). Nessa abordagem, utiliza-se o conceito de "janelas expansivas", onde a cada iteração o conjunto de treinamento cresce progressivamente, englobando observações passadas, enquanto o conjunto de validação é sempre formado por um período imediatamente subsequente ao treino (Kashefi et al., 2026).

3.2 Testes

Para testar os modelos LightGBM e XGBoost, compararam-se, ao longo do Conjunto de Teste, os valores de potência prevista e observada (Figura 4). Observa-se que os valores do conjunto de teste, 5047 observações, não participaram da etapa de treinamento dos modelos.

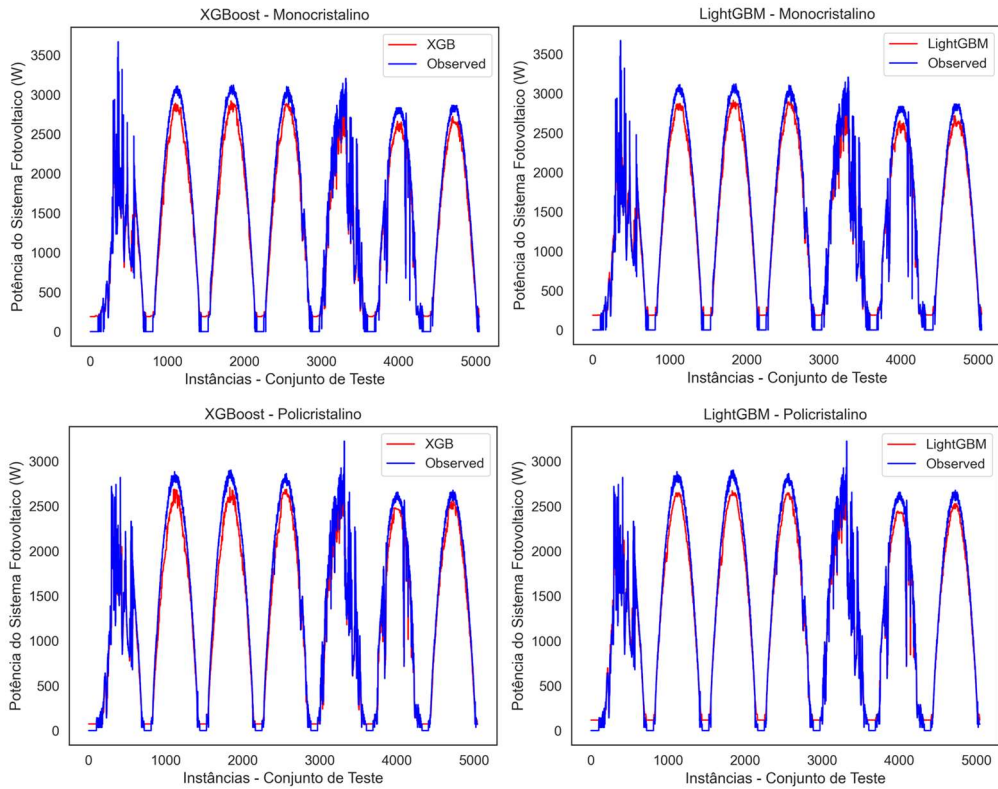


Figura 4. Resultados de previsão – XGBoost e LightGBM.

Pode-se observar, por meio desta figura, que as potência previstas pelos modelos, LightGBM e XGBoost nos dois sistemas, tem um bom ajuste com os valores de potências observados. Observa-se também, nestas figuras, a presença de evidências claras de dias com nebulosidade. Em vários trechos a potência observada apresenta quedas e recuperações abruptas, padrão característico da passagem de nuvens.

A análise visual por si só não permite avaliar com precisão a eficiência dos resultados. Para maior confiabilidade, foram calculados os valores das métricas R^2 , MAE e RMSE. Estas métricas avaliam a qualidade do ajuste dos modelos (Figura 5).

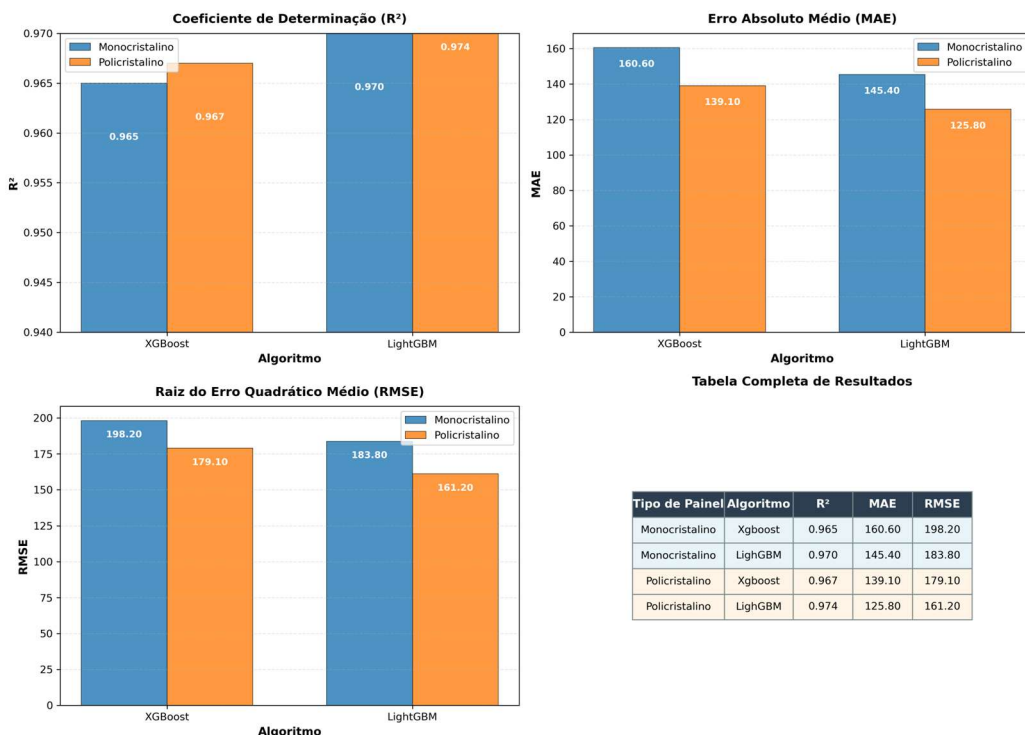


Figura 5. Resultados das métricas.

Pode-se observar, por meio dos resultados apresentados na Figura 5, que modelo LightGBM apresentou, em ambos tipos de sistemas fotovoltaicos, melhor desempenho. O LightGBM, para o sistema monocristalino, alcançou $R^2 = 0,970$, superando o XGBoost ($R^2 = 0,965$), além de reduzir o RMSE em 14,40 W (7,27%) e o MAE em 15,20 W (9,46%). As melhorias, no sistema policristalino, foram ainda mais expressivas: o R^2 subiu para 0,974 (ante o 0,967 do XGBoost), acompanhado de uma redução de 17,90 W (9,99%) no RMSE e de 13,30 W (9,56%) no MAE. Esses resultados indicam que, ambos os modelos apresentem muito boa precisão preditiva ($R^2 > 0,96$). Contudo, o LightGBM demonstra, nas condições analisadas destes sistemas fotovoltaicos, melhor capacidade de generalização e menor erro de previsão.

Na Figura 6 apresentam-se, para as tecnologias monocristalina e policristalina, os gráficos de dispersão para o modelo LightGBM. A análise dos gráficos de dispersão, entre os valores reais e preditos, evidencia uma forte correlação para ambas as tecnologias. Os pontos estão fortemente concentrados ao longo da linha de referência (linha ideal), indicando alta precisão dos modelos. O coeficiente de determinação (R^2) reforça essa observação: para o sistema monocristalino, obteve-se $R^2 = 0,970$, enquanto o policristalino apresentou $R^2 = 0,974$, demonstrando baixa dispersão dos erros e elevada capacidade preditiva em ambos os casos.

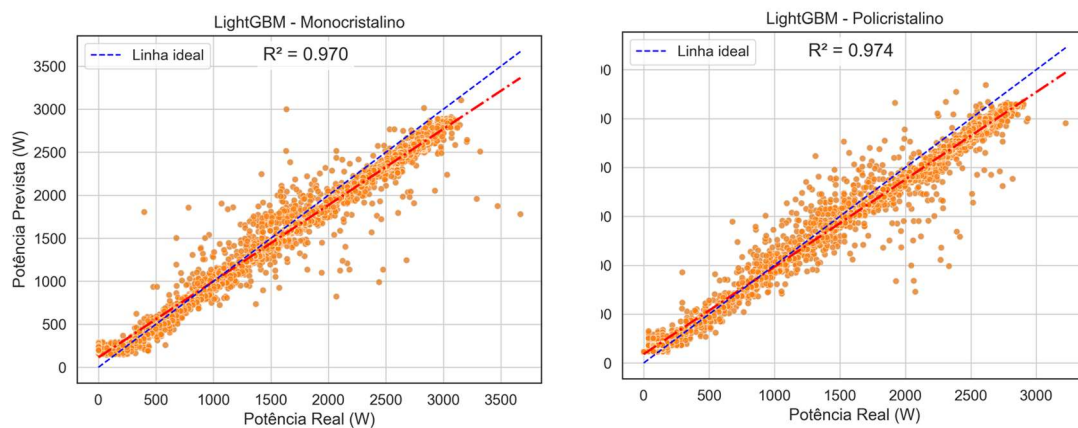


Figura 6. Dispersão entre valores reais e previstos (LightGBM) – Conjunto de Teste.

O LightGBM, no que tange à eficiência computacional, demonstrou melhor desempenho nos cenários avaliados (monocristalino e policristalino) (Figura 7). Para a tecnologia monocristalina, o algoritmo reduziu o tempo de execução em aproximadamente 13,6% (77,25 s contra 89,39 s), vantagem que foi ampliada na análise dos dados policristalinos, onde o ganho de eficiência atingiu cerca de 34,0% (62,20 s contra 94,25 s). Essa diferença de desempenho corrobora a literatura especializada, atribuindo-se a maior velocidade do LightGBM à sua estratégia de crescimento de árvores baseada em folhas (*leaf-wise*) e ao uso de técnicas de otimização como o GOSS (*Gradient-based One-Side Sampling*).

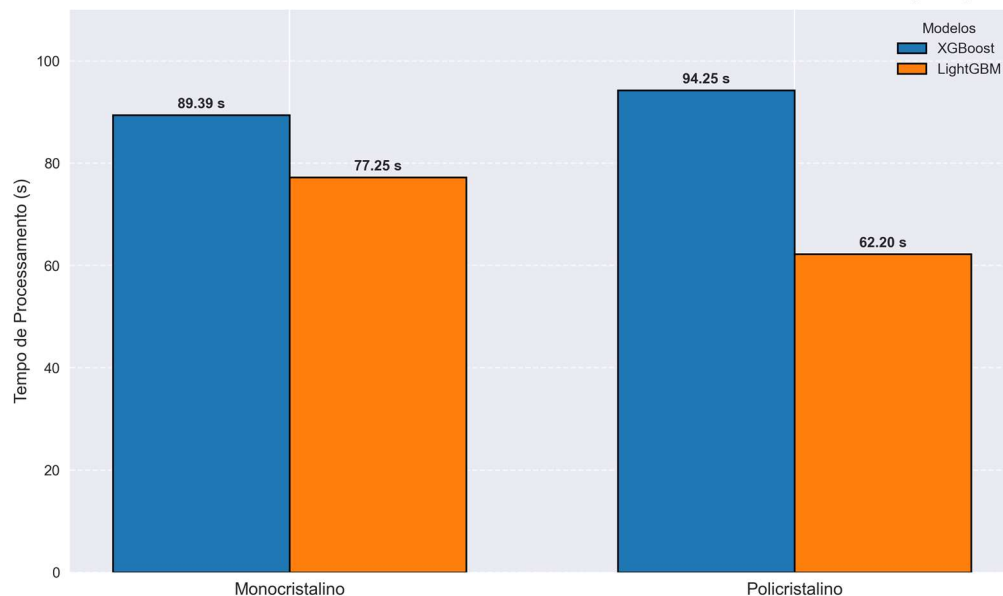


Figura 7. Tempos de processamento (LightGBM e XGBoost).

4 Conclusões

Este trabalho comparou o desempenho, na previsão da potência gerada por dois sistemas fotovoltaicos, dos algoritmos, baseados em *boosting*, XGBoost (*Extreme Gradient Boosting*) e LightGBM (*Light Gradient Boosting Machine*). Utilizou-se, para realizar a comparação de desempenho, duas base de dados obtidas da estação de pesquisa EPESOL, localizada no campus Medianeira da UTFPR. Os modelos, após o pré-processamento, foram treinados e tiveram seus hiperparâmetros otimizados com o auxílio da biblioteca Optuna.

Dentre os modelos avaliados, o LightGBM destacou-se como o mais eficiente, alcançando os melhores índices de desempenho (R^2 de 0,970 para monocristalino e 0,974 para policristalino) e os menores erros (MAE e RMSE). Esse melhor desempenho também se refletiu na eficiência computacional, onde o LightGBM apresentou, para os dois sistemas fotovoltaicos, os menores tempos de execução. Contudo, o XGBoost também apresentou resultados sólidos (R^2 de 0,965 para monocristalino e 0,967 para policristalino). Estes resultados confirmam uma forte correlação entre os valores reais e preditos, validando a aplicabilidade dessas técnicas de *machine learning* em sistemas de energia solar.

Portanto, este estudo demonstrou a eficácia, dos algoritmos XGBoost e LightGBM, na previsão da potência gerada por sistemas fotovoltaicos monocristalinos e policristalinos, com ambos os modelos apresentando alta precisão ($R^2 > 0,96$). Os resultados reforçam a viabilidade de integrar essas técnicas em sistemas de energia renovável, contribuindo para maior eficiência operacional e confiabilidade na gestão de recursos energéticos. Futuros estudos podem explorar a aplicação desses modelos em outras tecnologias fotovoltaicas ou incorporar variáveis climáticas adicionais para aprimorar ainda mais a precisão preditiva.

Agradecimentos

Os autores agradecem a UTFPR pelo apoio e infraestrutura disponibilizada para o desenvolvimento desta pesquisa e ao LABENS pelos dados fornecidos no âmbito do projeto de P&D ANEEL / COPEL-DIS / UTFPR - PD 2866-0464/2017 - Metodologia para Análise, Monitoramento e Gerenciamento da GD por Fontes Incentivadas.

Referências bibliográficas

- Abdellatif, A, Mubarak, H and Ahmad, S (2022) Forecasting Photovoltaic Power Generation with a Stacking Ensemble Model. *Sustainability*, v. 14. <https://doi.org/10.3390/su141711083>
- Akhter, MN, Mekhilef, S and Mokhlis, H (2022) A hybrid deep learning method for an hour ahead power output forecasting of three different photovoltaic systems. *Applied Energy*, v. 307. <https://doi.org/10.1016/j.apenergy.2021.118185>
- Antonanzas, J, Osorio, N, Escobar, R and Urraca, R (2016) Review of photovoltaic power forecasting. *Solar Energy*, v. 136. <https://doi.org/10.1016/j.solener.2016.06.069>
- Bouziane, A, Bouziane, M and Naima, K (2024) Enhancing Photovoltaic Power Forecasting through Hybrid Deep Learning Models: A CNN-RNN Approach for Grid Stability and Renewable Energy Optimization. *J. Ren. Energies*, v. 28. <https://doi.org/10.54966/jreen.v1i3.1294>
- Couto, A, Estanqueiro, A (2021) Assessment of wind and solar PV local complementarity for the hybridization of the wind power plants installed in Portugal. *J. Clean. Prod.* v. 319. <https://doi.org/10.1016/j.jclepro.2021.128728>
- Guo, J (2023) Prediction of heating and cooling loads based on light gradient boosting machine algorithms. *Building and Environment*, v.236. <https://doi.org/10.1016/j.buildenv.2023.110252>.
- Kanwal, S, Mehran, MT, Hassan, M, Anwar, M, Naqvi, SR and Khoja, AH (2022) An integrated future approach for the energy security of Pakistan: Replacement of fossil fuels with syngas for better environment and socio-economic development. *Renewable and Sustainable Energy Reviews*, v. 156. <https://doi.org/10.1016/j.rser.2021.111978>

- Kashefi, M, Karami, H, Niksefat, M and Ghazvinian, H (2026) Prediction of monthly precipitation and maximum 24 h precipitation using Random Forest, Decision Tree and XGBoost models. *Earth Syst. Environ*, v. 12. <https://doi.org/10.1007/s40808-025-02714-3>
- Konstantinou, M, Peratikou, S, Charalambide AG (2021) Solar Photovoltaic Forecasting of Power Output Using LSTM Networks. *Atmosphere*, v. 12. <https://doi.org/10.3390/atmos12010124>
- Mlilo, N, Brown, J and Ahfock, T (2021) Impact of intermittent renewable energy generation penetration on the power system networks—A review. *Technol Econ Smart Grids Sustain Energy*, v. 6. <https://doi.org/10.1007/s40866-021-00123-w>
- Nakahata, B. Y. (2023) Análise de desempenho de quatro tecnologias aplicadas a sistemas fotovoltaicos conectados a rede elétrica instalados em diferentes campus da UTFPR no Paraná. Dissertação de Mestrado. Universidade Tecnológica Federal do Paraná. Medianeira.
- Özbay, H, Özer, I, Dalcali, A, Çetin, O and Temurta, F (2025) Sensorless position estimation in electromagnetic launchers using recurrent neural networks with repeated k-Fold Cross-Validation. *Arabian Journal for Science and Engineering*, v.50. <https://doi.org/10.1007/s13369-024-09905-7>
- Oliveira, L and Santos, JAA (2022) As Redes Neurais Aplicadas à Previsão de Corrente Elétrica de um Sistema Fotovoltaico. *Revista FSA*, v. 19. <http://dx.doi.org/10.12819/2022.19.5.8>
- Ramsami, P and Oree, V (2015) A hybrid method for forecasting the energy output of photovoltaic systems. *Energy Conversion and Management*, v. 135, <https://doi.org/10.1016/j.enconman.2015.02.052>
- Santos, JAA and Chaukoski, Y (2020) Previsão do consumo de energia elétrica na região sudeste: um estudo de caso usando SARIMA e LSTM. *Revista CEREU*, v. 12, <https://10.18605/21757275/cereus.v12n4p93-104>
- Santos, JAA (2021) Aplicação de Redes Neurais na Previsão de Custos de Produção de Frangos de Corte no Estado do Paraná. *Revista FSA*, v. 18, p. 6, 169-185. <http://doi.org/10.12819/2021.18.6.9>
- Santos JAA and Polo ANB (2025) Comparação do desempenho dos algoritmos RF e KNN na previsão da energia gerada por quatro tecnologias fotovoltaicas. *Latin American Journal of Energy Research*, v. 12. <https://doi.org/10.21712/lajer.2025.v12.n2.p77-85>
- Srinivas, P, and Katarya R (2022) hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost. *Biomedical Signal Processing and Control*, v. 73. <https://doi.org/10.1016/j.bspc.2021.103456>
- Solano, ES, Dehghanian P and Affonso CM (2022) Solar radiation forecasting using machine learning and ensemble feature selection. *Energies*, v.15. <https://doi.org/10.3390/en15197049>
- Tang, M, Zhao, Q, Long, W and Huang, B (2020) An Improved LightGBM Algorithm for Online Fault Detection of Wind Turbine Gearboxes. *Energies*, v.13. <https://doi.org/10.3390/en13040807>
- Ye, Z, Yang, B, Han, Y and Chen, N (2022) State-of-the-art solar energy forecasting approaches: critical potentials and challenges. *Front. Energy Res.*, v. 10. <https://doi.org/10.3389/fenrg.2022.875790>
- Zahoor, Z, Khan, I and Hou, F (2022) Clean energy investment and financial development as determinants of environment and sustainable economic growth: Evidence from China. *Environ. Sci. Pollut*, v. 29. <https://doi.org/10.1007/s11356-021-16832-9>