

Análise de Modelos para Previsão das Componentes de Tendência e Sazonalidade em Séries Temporais de Preços de Energia Elétrica no Mercado Brasileiro.

José Gabriel Silva Pereira¹, Flávia Pessoa Monteiro¹

¹Instituto de Ciência Exatas e Naturais – Universidade Federal do Pará (UFPA)
Augusto Corrêa, 01 - Guamá, CEP: 66075-110, Belém - PA, Brasil.

jose.pereira@icen.ufpa.br, flaviamonteiro.comp@gmail.com

Abstract. *Electricity price forecasting in brazilian market requires advanced models to handle non-linear high volatility, generated by the need for real-time balance, climate dependence, and demand inelasticity. This article evaluates predictive algorithms—Prophet, Holt-Winters, and LightGBM—in the short-term forecasting of a price time series, comparing them to a baseline model that uses SARIMAX for trend and seasonality, along with XGBoost for residuals. Using the Knowledge Discovery in Databases (KDD) methodology, historical data (2012 to 2025) was subjected to STL decomposition, focusing on the trend and seasonality components for testing. The empirical evaluation, based on MAE, MSE, RMSE, and MAPE, demonstrated that the Holt-Winters and Prophet techniques did not show significant gains. In contrast, the application of LightGBM achieved the best performance, reducing the MAPE error from 17.95% in the baseline model to 9.05%. The study concludes that iterative decision tree-based approaches (Gradient Boosting), with proper temporal treatment, significantly outperform traditional techniques in energy forecasting.*

Resumo. *A previsão de preços de energia elétrica no mercado brasileiro exige modelos avançados para lidar com alta volatilidade não linear, gerada pela necessidade de equilíbrio em tempo real, dependência climática e inelasticidade da demanda. Este artigo avalia algoritmos preditivos — Prophet, Holt-Winters e LightGBM — na previsão de curto prazo de uma série temporal de preços, comparando-os a um modelo base o qual utiliza SARIMAX para tendência e sazonalidade além do XGBoost para resíduo. Utilizando a metodologia Knowledge Discovery in Databases (KDD), dados históricos (2012 a 2025) foram submetidos à decomposição STL, sendo focado os componentes de tendência e sazonalidade para a realização dos testes. A avaliação empírica, baseada em MAE, MSE, RMSE e MAPE, demonstrou que as técnicas HoltWinters e Prophet não apresentaram ganhos expressivos. Em contraste, a aplicação do LightGBM, obteve o melhor desempenho, reduzindo o erro MAPE de 17,95% do modelo base para 9,05%. O trabalho conclui que abordagens baseadas em árvores de decisão iterativas (Gradient Boosting), com tratamento temporal adequado, superam expressivamente as técnicas tradicionais na predição energética.*

1. Introdução

A previsão precisa dos preços no mercado de energia elétrica brasileiro é um desafio analítico complexo. A série temporal de precificação apresenta alta volatilidade não li-

near, caracterizada por picos e quedas bruscas gerados pela necessidade de equilíbrio em tempo real, dependência climática e inelasticidade da demanda. Essa instabilidade exige abordagens de modelagem preditiva avançadas, capazes de capturar tanto as direções de longo prazo quanto as oscilações cíclicas para garantir uma predição com alta eficácia.

Este artigo expande e avança a metodologia proposta por [Monteiro et al. 2025]. Enquanto o estudo anterior utilizou o SARIMAX [Arunraj and Ahrens 2015] para as componentes de tendência e sazonalidade e focou a aplicação de aprendizado de máquina XGBoost [Gono et al. 2023] no componente de resíduo para previsões de médio e longo prazo, o presente trabalho direciona o foco empírico para o curto prazo. A principal contribuição desta pesquisa é investigar o impacto isolado da aplicação de algoritmos avançados (Prophet, Holt-Winters e LightGBM) na modelagem estrita das componentes de tendência e sazonalidade.

Diante disso, a principal questão de pesquisa que norteia este trabalho é: Modelos baseados em árvores de decisão iterativas (Gradient Boosting) conseguem superar métodos estatísticos clássicos e frameworks automatizados na captura das variações não lineares do Preço de Liquidação das Diferenças? Para responder a esta questão, o estudo adota a metodologia Knowledge Discovery in Databases (KDD). O artigo está organizado apresentando os trabalhos relacionados (Seção 2), a metodologia e pré-processamento (Seção 3), a discussão dos resultados empíricos (Seção 4) e as conclusões (Seção 5).

2. Trabalhos Relacionados

A previsão de séries temporais no mercado de energia elétrica demanda modelos capazes de lidar com seu comportamento não linear. Dentre os métodos estatísticos clássicos, o modelo de suavização exponencial de Holt-Winters permanece como uma linha de base comum na literatura devido ao seu baixo custo computacional e adequação para previsões de curto prazo com sazonalidades bem comportadas [Veríssimo et al. 2013]. Contudo, sua principal limitação reside na dificuldade de mapear padrões não lineares complexos.

Abordagens baseadas em decomposição automatizada, como o Prophet, ganharam visibilidade por permitirem a modelagem explícita de tendência e sazonalidade [Taylor and Letham 2018]. Apesar de sua facilidade de implementação em escala, a literatura indica que seu desempenho pode ser subótimo em séries com elevada variabilidade climática e econômica [Curiël et al. 2025].

Com o avanço das técnicas de aprendizado de máquina, modelos baseados em árvores de decisão iterativas, particularmente o LightGBM [Ke et al. 2017], têm demonstrado resultados superiores. O LightGBM captura relações não lineares complexas através da descida do gradiente, apresentando alta eficácia quando associado a uma rigorosa engenharia de características temporais. Diferenciando-se da literatura existente, este trabalho isola empiricamente os componentes da série via *Seasonal and Trend decomposition using Loess* (STL) para comparar o desempenho exato destas três técnicas frente aos desafios específicos do setor elétrico brasileiro.

3. Metodologia

O desenvolvimento da pesquisa seguiu as etapas do processo KDD, detalhadas a seguir.

3.1. Pré-Processamento e Análise Exploratória de Dados

Uma série temporal é um conjunto de observações ordenadas cronologicamente, com um índice temporal exclusivo mantido constante representado matematicamente por $T_s = TS[k] - TS[k - 1]$ [Kramar and Alchakov 2023]. A série pode ser decomposta através da equação aditiva $Y_t = T_t + S_t + R_t$, onde T_t denota a tendência direcional, S_t representa a sazonalidade e R_t os resíduos.

Os dados utilizados contêm a precificação semanal de energia do período de 2012 a 2025, fornecidos pela base de dados da empresa DCIDE [Dcide 2026]. Para garantir a clareza sobre os dados tratados, a Tabela 1 esquematiza as variáveis utilizadas no estudo.

Tabela 1. Dicionário e Descrição das Variáveis do Dataset

Variável	Descrição	Unidade/Tipo	Frequência
Data	Índice temporal cronológico	Data (<i>Datetime</i>)	Semanal
Preço MPC_SECO	Valor do preço de energia	Númerico (<i>Float</i>)	Semanal

O pré-processamento envolveu a definição das datas como índice e a adequação da tipagem decimal (substituição de vírgula por ponto). Em seguida, aplicou-se a normalização *MinMaxScaler* para converter os valores numéricos para o intervalo entre 0 e 1, utilizando a fórmula $v = new - min_A + R(\frac{v - min_A}{max_A - min_A})$ [Thakker and Buch 2024]. Esta etapa garante eficiência ao gradiente, evitando que grandezas numéricas díspares afetem negativamente os cálculos.

Foi aplicado o método de decomposição de séries temporais denominado de *Seasonal and Trend decomposition using Loess* (STL), em que faz a determinação das relações não lineares dentro de diferentes componentes, sendo utilizado dados de regressão local para ajustar a curva. O primeiro passo da decomposição é aproximar a linha de tendência da série temporal, através de um modelo polinomial. Após isso, o componente sazonal é definido a partir do cálculo sobre a totalidade dos dados. Sendo subtraído os dados sazonais dos originais, o componente de sazonalidade é removido, logo, a série temporal é dessazonalizada. Os resíduos são determinados a partir da sobra após a remoção dos componentes de tendência e sazonalidade [Apaydin et al. 2021].

Com isso, a série é fragmentada em três componentes, a tendência, a qual mostra a direção geral do crescimento ou queda ao longo do tempo, a sazonalidade, em que mostra os padrões que se repetem em intervalos regulares, neste trabalho foi definido um período de 52, sendo representado as 52 semanas que um ano possui, aproximadamente. E os resíduos o qual é a variação aleatória que não foi explicada nem pela tendência nem pela sazonalidade.

O particionamento cronológico foi definido com o período de 2012 a 2024 para o conjunto de treinamento, reservando o ano de 2025 inteiramente para validação. Para garantir rigor metodológico e evitar vazamento de dados, o método STL, foi ajustado estritamente sobre os dados de treino, isolando a tendência polinomial, definindo a componente sazonal com um período de 52 semanas e separando os resíduos aleatórios.

3.2. Mineração de Dados e Implementação

As técnicas Prophet, Holt-Winters e LightGBM foram treinadas isoladamente para cada componente (Tendência e Sazonalidade) e contrapostas ao cenário *baseline* estabelecido com SARIMAX. O processo de testes seguiu a estrutura definida na Tabela 2.

Tabela 2. Metodologia de testes e combinação de modelos por componente.

Etapa	Componente de tendência	Componente de sazonalidade	Componente de Resíduo	Objetivo do Teste
Cenário Base	Modelo Base-line (SARIMAX)	Modelo Base-line (SARIMAX)	Modelo Base-line (XGBoost)	Estabelecer as métricas de erro de referência para comparação.
Passo 1	Avaliação	Modelo Base-line (SARIMAX)	Modelo Base-line (XGBoost)	Avaliar o impacto isolado da substituição do modelo de tendência.
Passo 2	Modelo Base-line (SARIMAX)	Avaliação	Modelo Base-line (XGBoost)	Avaliar o impacto isolado da substituição do modelo de sazonalidade.
Passo Final	Avaliação	Avaliação	Modelo Base-line (XGBoost)	Validar a implementação completa combinando as novas técnicas.

3.2.1. Prophet

O PROPHET é uma técnica para a previsão de dados de séries temporais criado pela equipe de Core Data Science do Facebook. O seu objetivo é ser capaz de realizar previsões em escala, o que significa que o PROPHET procura ser uma ferramenta de previsão de natureza automatizada, proporcionando uma maior facilidade de uso no ajuste de métodos de séries temporais[Aditya Satrio et al. 2021].

3.2.2. Holt-Winters

“Holt-Winters que é uma expansão do método Holt (1957) desenvolvida por Winter (1960) para aplicar em séries temporais que exibem tendência e sazonalidade. É um

dos métodos mais utilizados para previsão de curto prazo da demanda, devido a sua simplicidade, baixo custo de operação, boa precisão e capacidade de ajustamento automático e rápido a mudanças na série em análise. Partindo do princípio que determinadas séries possuem um fator sazonal, além do nível e tendência que capta características da série que se repetem a intervalos regulares de tempo, Winter propõe métodos de projeção para essas séries, considerando dois tipos de efeitos sazonais: multiplicativo e aditivo.”[Veríssimo et al. 2013]

Como neste trabalho as componentes foram isoladas pelo método STL para o teste de diferentes técnicas. Quanto a sazonalidade, foi aplicada a suavização exponencial de Holt-Winters com sazonalidade aditiva e sem componente de tendência, configurada para capturar ciclos semanais, adequando-se à natureza da componente sazonal previamente.

Já para o componente de tendência foi aplicada a Suavização Exponencial Dupla, especificamente o Modelo Linear de Holt. Esta abordagem foi configurada com uma componente de tendência aditiva e com a supressão total da componente sazonal, uma vez que a série temporal de entrada consistia puramente no comportamento direcional dos dados, livre de oscilações cíclicas.

3.2.3. LightGBM

O LightGBM, é um framework desenvolvido pela Microsoft para a implementação do algoritmo Gradient Boosting Decision Tree (GBDT). O GBDT é um algoritmo de árvores de decisão de natureza iterativa, que utiliza modelos de árvores de classificação e regressão como seus aprendizes fracos. Neste processo, cada novo aprendiz é construído na direção de descida do gradiente da função de perda do aprendiz anterior, sendo gerado de forma iterativa para o treinamento do modelo. [Cao et al. 2023]

Apesar de sua alta capacidade de aprendizado iterativo, modelos baseados em árvores não interpretam séries temporais nativamente, por isso, aplicou-se intensa engenharia de atributos. Foram extraídas variáveis de calendário (semana, mês, ano) e construídos atrasos temporais (*lags* 1, 2 e 3) para prover inércia direcional e contexto histórico, com taxa de aprendizado de 0.01 e 1000 estimadores para prevenir *overfitting* (Tabela 5).

4. Resultados e Discussão

A partir da seleção e pré-processamento realizado, a Figura 1 ilustra o comportamento dos dados iniciais, evidenciando picos e quedas bruscas nos preços. O resultado da decomposição STL controlada nos dados de treino é visualizado na Figura 2.

Os resultados da mineração de dados podem ser avaliados através dos parâmetros de erros na Tabela 6 (sazonalidade) e Tabela 7 (tendência). O comportamento visual das previsões para sazonalidade e tendência encontram-se nas Figuras 3 e 4, respectivamente.

A análise empírica baseou-se predominantemente no Erro Percentual Absoluto Médio (MAPE). O modelo *baseline* obteve um erro geral de 17,95%. A introdução do Prophet demonstrou uma degradação de eficácia, registrando o pior desempenho com um MAPE de 24,11%, como pode ser visto na Tabela 8. Este resultado evidencia que a decomposição automatizada do Prophet, eficiente em cenários comerciais estáveis, sofre

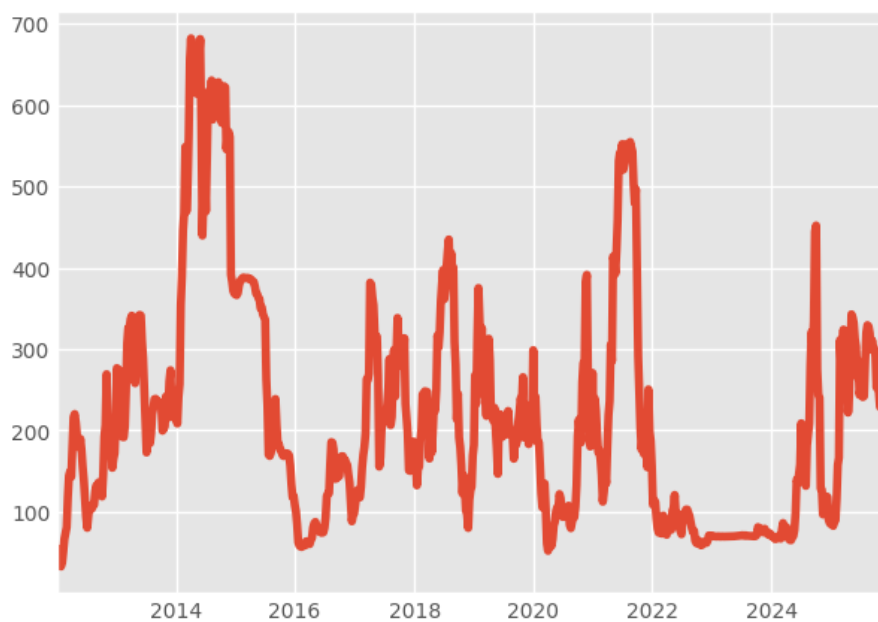


Figura 1. Comportamento da Base de Dados

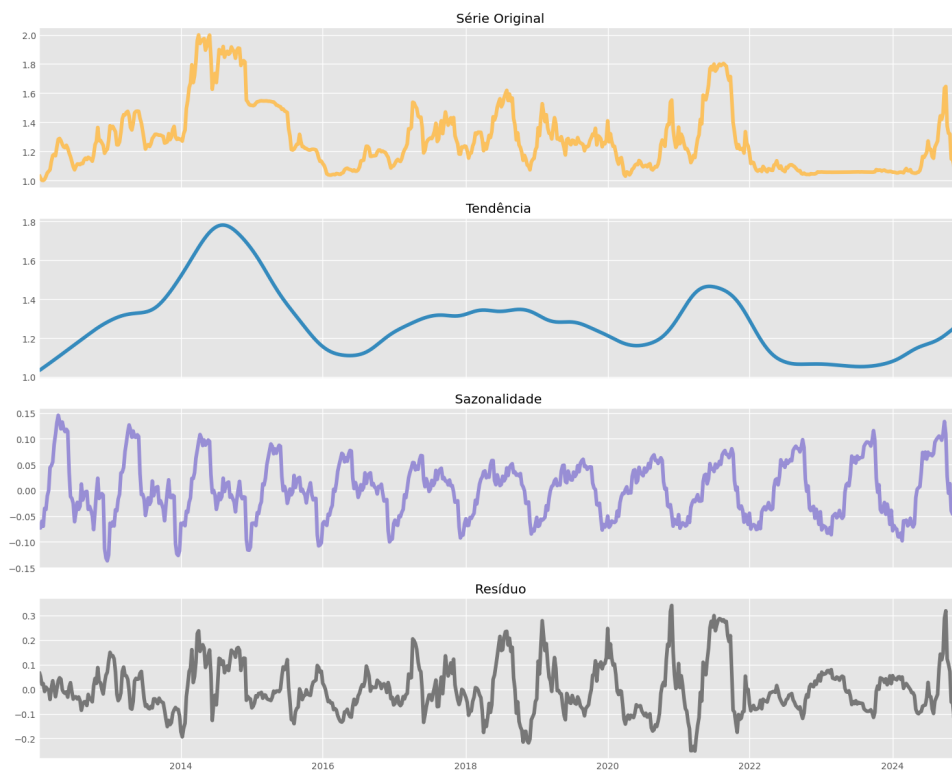


Figura 2. Dados normalizados usando MinMax Scaler e decompostos em Tendência+Sazonalidade+Resíduo.

severas limitações diante da alta volatilidade e choques abruptos presentes no mercado de energia elétrica brasileiro.

O método estatístico Holt-Winters apresentou um resultado de 18,34%, situando-se em um patamar de precisão quase idêntico ao *baseline*, falhando em prover ganhos substanciais na modelagem de variações complexas.

Em contraste absoluto, a aplicação do LightGBM obteve o melhor desempenho consolidado, reduzindo o erro MAPE geral para 9,05%. A superioridade analítica deste algoritmo justifica-se pela sua arquitetura iterativa. Diferente do SARIMAX ou do Holt-Winters, o LightGBM, quando alimentado de forma correta pela engenharia de atributos (especialmente os *lags* de até 3 semanas), conseguiu mapear as interações não lineares. Essa característica permitiu que o algoritmo se ajustasse mais rapidamente às inflexões direcionais (Figura 4) e capturasse os microciclos sazonais (Figura 3) com notável precisão. O reflexo desta assertividade é consolidado na reconstrução final da série, ilustrada na Figura 5.

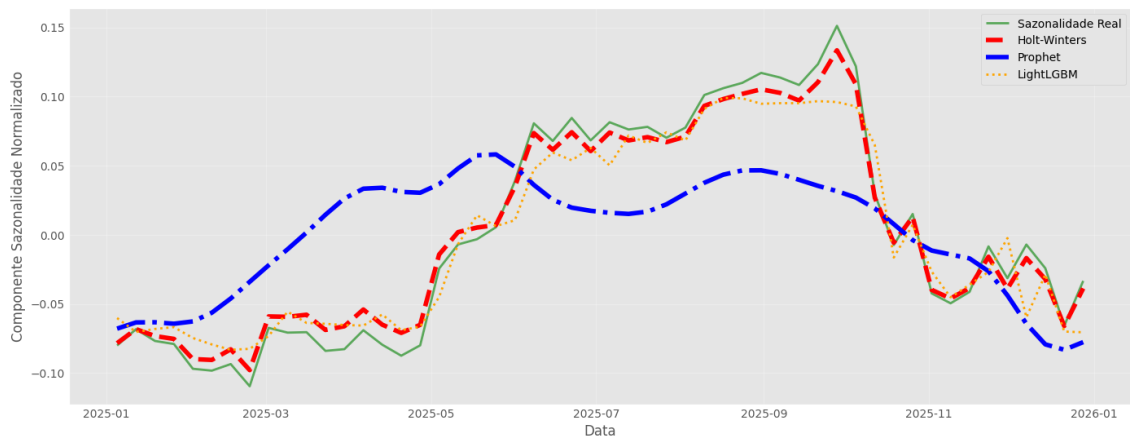


Figura 3. Resultado de sazonalidade com algoritmos aplicados.

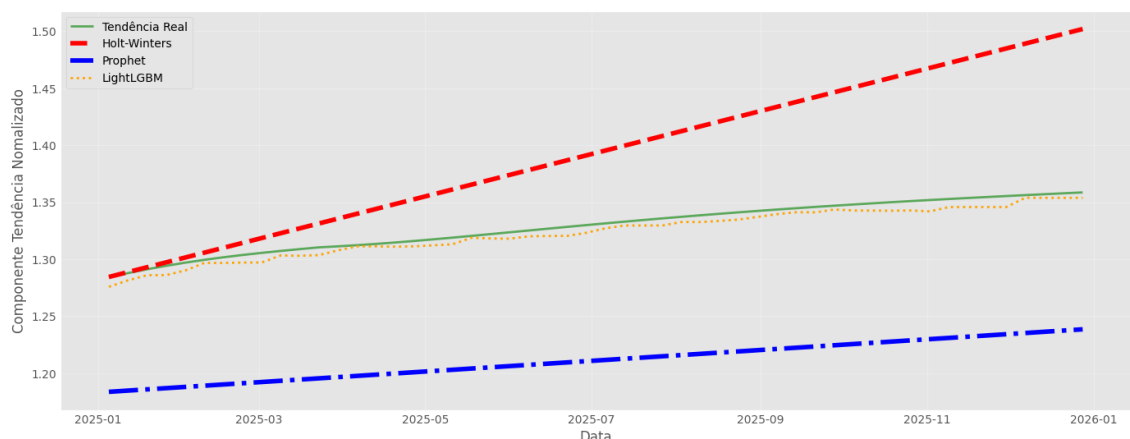


Figura 4. Resultado de tendência com algoritmos aplicados.

5. Conclusão

Este estudo avaliou a eficácia de abordagens preditivas clássicas e de aprendizado de máquina isoladas nos componentes de tendência e sazonalidade para a série temporal

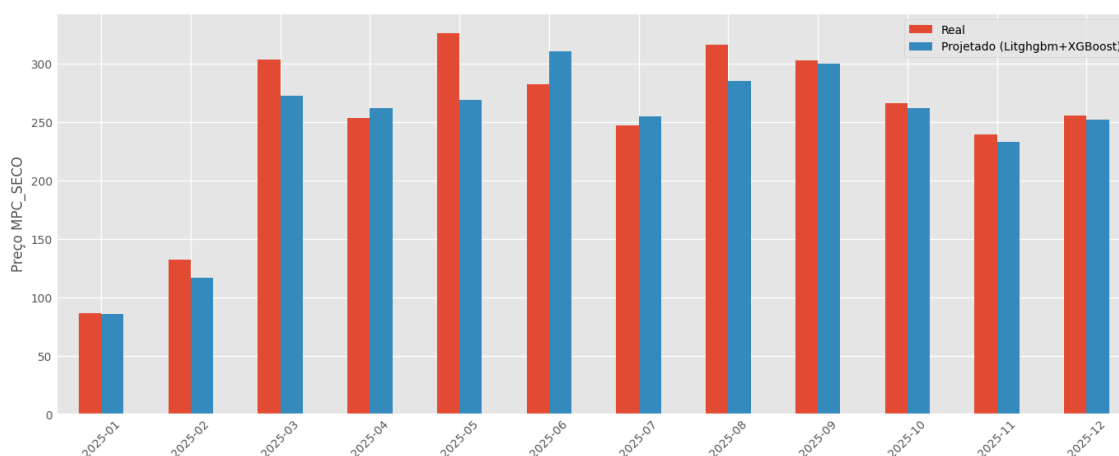


Figura 5. Previsão realizada com LightGBM + XGBOOST.

Tabela 3. Parâmetros utilizados na modelagem com Prophet (tendência e sazonalidade).

Componente	Parâmetro (Valor)	Descrição
Tendência	changepoint_prior_scale (0.05)	Controla a flexibilidade. O valor gera uma curva equilibrada, evitando overfitting.
	yearly_seasonality (False)	Desabilitado, pois o objetivo é modelar exclusivamente a tendência.
	weekly_seasonality (False)	Desabilitado, pois o foco é apenas na tendência em dados semanais.
	daily_seasonality (False)	Desabilitado, pois não há captura de sazonalidade intradiária.
Sazonalidade	yearly_seasonality (True)	Habilita a captura e modelagem matemática de padrões anuais.
	weekly_seasonality (False)	Desabilitado, uma vez que a granularidade já é semanal.
	daily_seasonality (False)	Desabilitado por ausência de dados intradiários.
Ambos	periods (52)	Horizonte de previsão exato para o período de teste (2025).
	freq ('W')	Frequência temporal configurada como semanal.

Tabela 4. Parâmetros utilizados na modelagem com Holt-Winters para tendência e sazonalidade.

Componente	Parâmetro (Valor)	Descrição
Tendência	trend ('add')	Define a modelagem da tendência como aditiva e linear ao longo do tempo.
	seasonal (None)	Desabilitado, uma vez que a entrada consiste apenas na tendência isolada pelo STL.
	steps (<i>len(test)</i>)	Estabelece o horizonte temporal de projeção em semanas futuras para validação.
Sazonalidade	trend (None)	Desabilitado, pois o modelo recebe exclusivamente a componente sazonal isolada.
	seasonal ('add')	Define a modelagem da sazonalidade como aditiva, alinhada com a decomposição STL.
	seasonal_periods (52)	Estabelece a periodicidade do ciclo sazonal de um ano completo (52 semanas).
	steps (52)	Define o horizonte temporal da previsão correspondente ao conjunto de teste.

Tabela 5. Parâmetros e variáveis utilizadas na modelagem com LightGBM para tendência e sazonalidade.

Componente	Parâmetro/Variável (Valor)	Descrição
Ambos	n_estimators (1000)	Número máximo de árvores de decisão. Permite ao modelo aprender padrões complexos.
	learning_rate (0.01)	Taxa de aprendizado regularizadora, prevenindo o sobreajuste (<i>overfitting</i>).
	random_state (42)	Semente fixa que garante a total reprodutibilidade do treinamento e dos resultados.
	Atraso (<i>lag_1, lag_2, lag_3</i>)	Valores das 3 semanas anteriores, fornecendo inércia direcional e contexto histórico.
Tendência	Temporais (<i>ordinal_date, year</i>)	Variáveis contínuas e numéricas fundamentais para o modelo mapear o avanço direcional.
Sazonalidade	Temporais (<i>week, month</i>)	Variáveis de calendário extraídas para fornecer informações sobre o ciclo anual repetitivo.

Tabela 6. Resultados das métricas de erro para sazonalidade.

Modelo (sazonalidade)	MAE sazonalidade	MSE sazonalidade	RMSE sazonalidade
SARIMAX	0.0001746	4.1726748	0.0002042
PROPHET	0.0546087	0.0039184	0.0625972
HOLTWINTERS	0.0081377	8.6599061	0.0093058
LIGHTGBM	0.0166401	0.0004265	0.0206520

Tabela 7. Resultados das métricas de erro para tendência.

Modelo (Tendência)	MAE Tendência	MSE Tendência	RMSE Tendência
SARIMAX	0.0550638	0.0053634	0.0732358
PROPHET	0.1174495	0.0138226	0.1175696
HOLTWINTERS	0.0646779	0.0060615	0.0778560
LIGHTGBM	0.0000343	0.0054925	0.0058537

Tabela 8. Comparativo do melhor resultado de MAPE (Erro Percentual Absoluto Médio) obtido por cada técnica.

Modelo (Sazonalidade)	Modelo (Tendência)	MAPE Geral
SARIMAX	SARIMAX	17.95%
PROPHET	SARIMAX	24.11%
HOLTWINTERS	SARIMAX	18.34%
LIGHTGBM	LIGHTGBM	9.05%

de preços de energia elétrica brasileira. Os testes empíricos evidenciaram que o uso do framework LightGBM resultou em um ganho preditivo expressivo, reduzindo o MAPE de 17,95% para 9,05% no ano de validação (2025).

A diminuição quase pela metade do erro preditivo demonstra que abordagens orientadas a árvores iterativas geram valor direto ao lidar com a instabilidade e não linearidade de mercados energéticos de curto prazo. Como limitação da pesquisa, aponta-se o horizonte restrito de dados históricos utilizados para o treinamento, o que restringe a validação frente a choques estruturais massivos. Para trabalhos futuros, sugere-se a expansão da técnica LightGBM também para a modelagem da componente de resíduos, visando estabelecer um modelo unificado de *Gradient Boosting* de máxima precisão preditiva no setor.

Referências

- Aditya Satrio, C. B., Darmawan, W., Nadia, B. U., and Hanafiah, N. (2021). Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet. *Procedia Computer Science*, 179:524–532. 5th International Conference on Computer Science and Computational Intelligence 2020.
- Apaydin, H., Taghi Sattari, M., Falsafian, K., and Prasad, R. (2021). Artificial intelligence modelling integrated with singular spectral analysis and seasonal-trend decomposition using loess approaches for streamflow predictions. *Journal of Hydrology*, 600:126506.
- Arunraj, N. S. and Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics*, 170:321–335.
- Cao, Q., Wu, Y., Yang, J., and Yin, J. (2023). Greenhouse temperature prediction based on time-series features and lightgbm. *Applied Sciences*, 13(3).
- Curiël, R., Alshag, A. M. M., and Mohammadi Ziabari, S. S. (2025). Integrating climate and economic predictors in hybrid prophet-(q)lstm models for sustainable national energy demand forecasting: Evidence from the netherlands. *Sustainability*, 17(19).
- Dcide (2026). Dcide — preços de energia, curva forward, análise de risco. <https://www.dcide.com.br/>. Acesso em: 16 maio 2026.
- Gono, D. N., Napitupulu, H., and Firdaniza (2023). Silver price forecasting using extreme gradient boosting (xgboost) method. *Mathematics*, 11(18).
- Ke, G., Meng, Q., et al. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30.
- Kramar, V. and Alchakov, V. (2023). Time-series forecasting of seasonal data using machine learning methods. *Algorithms*, 16(5).
- Monteiro, F. P., Monteiro, S., Rodrigues, C., Reis, J., Bezerra, U., Tostes, M. E., and Almeida, F. A. F. (2025). A hybrid methodology using machine learning techniques and feature engineering applied to time series for medium- and long-term energy market price forecasting. *Energies*, 18(6).
- Taylor, S. J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1):37–45.

- Thakker, Z. L. and Buch, S. H. (2024). Effect of feature scaling pre-processing techniques on machine learning algorithms to predict particulate matter concentration for gandhinagar, gujarat, india. *Int. J. Sci. Res. Sci. Technol*, 11(1):410–419.
- Veríssimo, A. J., da Cunha Alves, C., Henning, E., do Amaral, C. E., and da Cruz, A. C. (2013). Métodos estatísticos de suavização exponencial holt-winters para previsão de demanda em uma empresa do setor metal mecânico. *Revista Gestão Industrial*, 8(4).
- Veríssimo, A. J. et al. (2013). Métodos estatísticos de suavização exponencial holt-winters para previsão de demanda. *Revista Gestão Industrial*, 8(4).