

Modelagem e avaliação de algoritmos

Eficiência Energética de Edifícios Relatório Técnico

Versão 1.0

Londrina, novembro de 2025

Pré-processamento e Transformação dos Dados

O pipeline de pré-processamento foi estruturado para evitar *data leakage* e garantir reprodutibilidade: todas as transformações que aprendem do dado (fit) foram realizadas apenas no conjunto de treino e depois aplicadas ao teste. O processo foi dividido em duas fases sequenciais: **PreSplitProcessor** (pré-divisão) e **PostSplitProcessor** (pós-divisão).

PreSplitProcessor (antes da divisão treino–teste)

- Renomeação das variáveis segundo o padrão adotado na AT1, assegurando consistência semântica.
- Validação da presença das colunas-alvo (*Carga_Aquecimento*, *Carga_Resfriamento*) antes da separação entre X e y.
- Separação formal entre preditores (X) e variáveis-alvo (y).
- Verificação de tipos e coerência do schema, garantindo que variáveis numéricas e discretas mantivessem tipagem adequada.

PostSplitProcessor (após a divisão treino–teste)

Todas as transformações foram ajustadas (fit) exclusivamente no treino e aplicadas (transform) ao teste:

- **StandardScaler**: padronização (média 0, desvio 1) para uso em modelos sensíveis à escala (Regressão Linear, Ridge, Lasso, SVR).
- **MinMaxScaler**: normalização para [0,1], mantendo uma versão adicional dos dados para métodos baseados em distância.
- **Persistência dos scalers treinados**, garantindo a reprodutibilidade e a integração com o ambiente de produção/web.
- **Manutenção da ordem das colunas e mapeamento explícito das features**, assegurando correspondência exata entre os dados transformados e o modelo final.

A sequência adotada previne vazamento de informação entre treino e teste e fornece duas versões padronizadas dos dados (padronizada e normalizada), suportando tanto algoritmos baseados em escala quanto técnicas baseadas em distância.

Metodologia

Após as etapas de pré-processamento, as bases de treino e teste estavam preparadas para alimentar os modelos supervisionados, com garantia de reprodutibilidade, ausência de *data leakage* e compatibilidade com o pipeline de modelagem. O objetivo desta etapa é comparar, de forma controlada, o desempenho de diferentes algoritmos de regressão aplicados ao conjunto de dados de eficiência energética.

O conjunto de dados foi dividido em 80% para treino e 20% para teste utilizando o método *train_test_split* com semente fixa (`random_state = 42`), assegurando consistência entre execuções e permitindo comparação justa entre todos os modelos. Optou-se por uma divisão simples em vez de validação cruzada k-fold, uma vez que o foco deste estudo é avaliar o comportamento dos modelos em um cenário padrão de treino–teste, reduzindo custo computacional e refletindo pipelines reais de produção.

Todos os algoritmos foram avaliados com seus hiperparâmetros padrão (*default*), conforme as implementações oficiais do scikit-learn. Para modelos que não possuem suporte nativo a múltiplas saídas — como Gradient Boosting e SVR — foi utilizado o *MultiOutputRegressor*, permitindo a predição simultânea das variáveis *Carga_Aquecimento* e *Carga_Resfriamento* e garantindo uniformidade na etapa de avaliação.

A performance dos modelos foi mensurada com três métricas complementares:

- **R² (Coeficiente de Determinação):** mede a proporção da variabilidade explicada pelo modelo.
- **MAE (Erro Absoluto Médio):** avalia o desvio médio absoluto das previsões.
- **RMSE (Erro Quadrático Médio):** penaliza erros maiores e fornece sensibilidade a desvios extremos.

Essas métricas foram aplicadas de forma uniforme em todos os modelos, permitindo avaliar precisão, estabilidade e generalização. Os resultados foram consolidados em tabela, sendo o **R²** adotado como critério principal de escolha do modelo com melhor desempenho.

Resultados

Esta seção apresenta o desempenho dos modelos testados, utilizando as métricas **R²**, **RMSE** e **MAE**. Os valores foram calculados a partir das previsões realizadas sobre o conjunto de teste, após o ajuste completo dos modelos no conjunto de treino. Todas as transformações do pipeline foram aplicadas previamente, garantindo padronização dos dados e prevenção de *data leakage*.

Os resultados foram obtidos a partir de uma divisão única de dados (80% treino e 20% teste), utilizando *train_test_split* com semente fixa para assegurar reprodutibilidade. Embora métodos baseados em validação cruzada ofereçam estimativas mais robustas, optou-se pela partição simples, pois o objetivo deste estudo é comparar, de forma controlada e direta, o comportamento dos diferentes algoritmos em um cenário padrão de treino–teste.

A **Tabela 7** apresenta os valores obtidos para R², RMSE e MAE, permitindo uma comparação clara entre modelos lineares, modelos regularizados, métodos baseados em árvores e técnicas avançadas como Gradient Boosting e SVR. A **Figura 7** resume

visualmente essas diferenças, evidenciando o desempenho substancialmente superior dos modelos não lineares em relação às abordagens lineares tradicionais.

	Modelo	RMSE	MAE	R2
0	LinearRegression	3.085987	2.188684	0.902705
1	Ridge	3.094432	2.193333	0.902178
2	Lasso	3.086455	2.188852	0.902677
3	RandomForest	1.409660	0.758669	0.978699
4	GradientBoosting	1.131105	0.721709	0.986351
5	SVR	1.365919	0.849163	0.980188

Figura 7: Métricas de desempenho dos modelos avaliados (RMSE, MAE e R²).

A Figura 8 apresenta o gráfico de barras com os valores de R² obtidos pelos modelos avaliados. Como as predições envolvem duas variáveis-alvo, o valor exibido para cada modelo corresponde ao desempenho agregado na predição conjunta de *Carga_Aquecimento* e *Carga_Resfriamento*. A paleta em tons de azul foi mantida para garantir consistência visual com as demais figuras desta seção.

As Figuras 9 e 10 seguem o mesmo padrão estético, exibindo os resultados referentes às métricas RMSE e MAE, respectivamente. Cada gráfico resume o desempenho obtido com a mesma divisão treino–teste e o mesmo pipeline de pré-processamento, permitindo comparação direta entre os modelos.

A partir do gráfico de R² (Figura 8), observa-se que os modelos lineares — **LinearRegression**, **Ridge** e **Lasso** — apresentam desempenhos muito próximos entre si e claramente inferiores às abordagens não lineares. Entre todos os algoritmos avaliados, o **Gradient Boosting** apresenta o melhor resultado, com **R² ≈ 0,986**, seguido por **SVR (≈ 0,980)** e **Random Forest (≈ 0,978)**. Isso indica que modelos capazes de capturar padrões não lineares conseguem explicar melhor a variabilidade das cargas térmicas.

O mesmo comportamento é observado nas métricas RMSE (Figura 9) e MAE (Figura 10). Novamente, o **Gradient Boosting** apresenta os menores valores de erro — **RMSE ≈ 1,13** e **MAE ≈ 0,72** — reforçando sua superioridade em termos de precisão e estabilidade das previsões. Os modelos lineares, por sua vez, mantêm erros mais elevados, compatíveis com sua limitação em capturar relações mais complexas entre as variáveis.

Em síntese, as Figuras 8, 9 e 10 evidenciam de forma consistente que técnicas não lineares, especialmente o Gradient Boosting, são mais eficazes para este problema, apresentando desempenho superior em todas as métricas avaliadas.

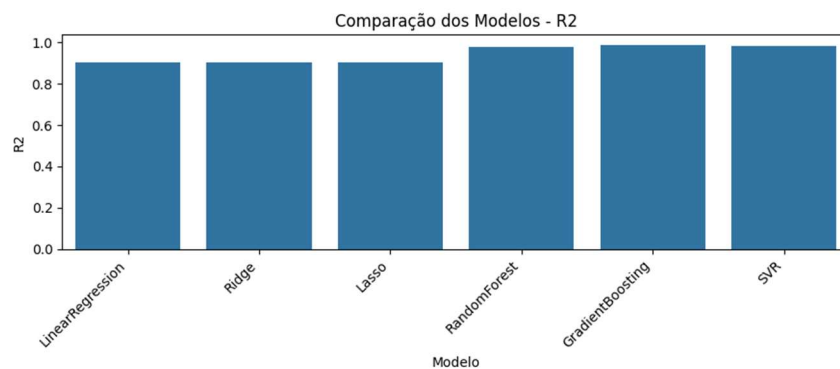


Figura 8- Métricas de Avaliação (R2).

No caso dos gráficos de erro, o RMSE apresentado na Figura 9 mostra que o melhor desempenho foi obtido pelo modelo **Gradient Boosting**, que apresentou o menor erro entre todas as abordagens avaliadas (RMSE $\approx 1,13$). Em seguida, destacam-se os modelos **SVR** (RMSE $\approx 1,36$) e **Random Forest** (RMSE $\approx 1,40$), ambos com desempenho significativamente superior aos métodos lineares. Já os modelos **LinearRegression**, **Ridge** e **Lasso** apresentaram erros substancialmente mais elevados (RMSE $\approx 3,08$), evidenciando sua limitação na captura de relações não lineares presentes no conjunto de dados.

Para a métrica MAE, representada na Figura 10, observa-se um padrão semelhante. O **Gradient Boosting** novamente se destaca como o modelo com menor erro absoluto médio (MAE $\approx 0,72$), seguido pelo **Random Forest** (MAE $\approx 0,75$) e pelo **SVR** (MAE $\approx 0,85$). Os modelos lineares mantiveram desempenhos próximos entre si, porém com erros superiores aos dos métodos não lineares, reforçando sua menor capacidade de generalização nesse contexto.

De forma geral, os resultados demonstram que, embora os modelos lineares apresentem comportamento consistente entre si, eles permanecem inferiores aos métodos não lineares em todas as métricas avaliadas. Entre todas as abordagens testadas, o **Gradient Boosting** obteve simultaneamente o menor erro e a maior capacidade explicativa, justificando sua escolha como o modelo final para este problema.

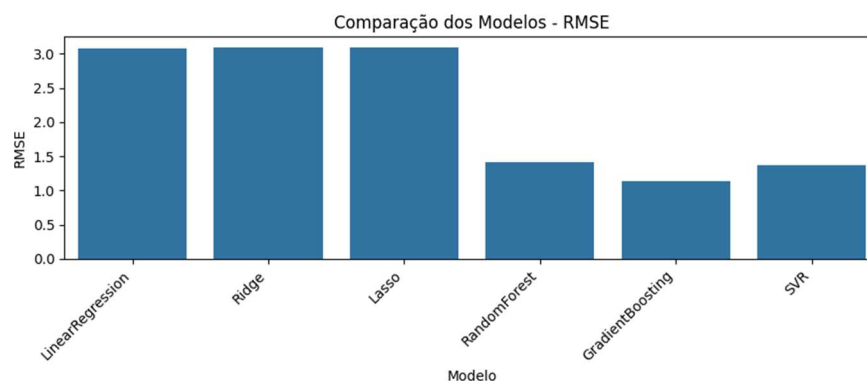


Figura 9 – Métricas de Avaliação (RMSE)

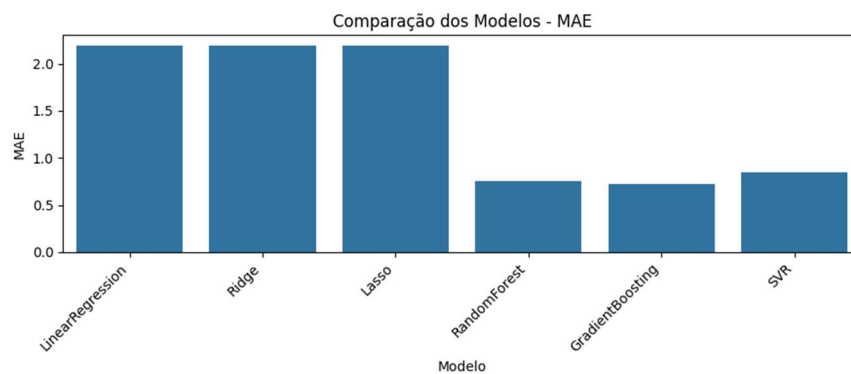


Figura 10 – métricas de Avaliação (MAE)

Para aprofundar a interpretação dos resultados, foram incluídas visualizações complementares que permitem observar tendências globais e a consistência entre as métricas de avaliação. Essas representações adicionais fornecem uma visão integrada do comportamento dos modelos, facilitando a comparação direta do desempenho sob múltiplas perspectivas.

A Figura 11 apresenta um **radar plot** que reúne simultaneamente as três métricas analisadas — **RMSE (invertido)**, **MAE (invertido)** e **R²**. Essa forma de visualização evidencia de maneira clara a dominância do modelo **Gradient Boosting**, que ocupa a maior área no gráfico, refletindo seu desempenho superior tanto em erro quanto em capacidade explicativa. Os modelos **Random Forest** e **SVR** aparecem em posições intermediárias, apresentando desempenho consistente, embora inferior ao do Gradient Boosting.

As abordagens lineares, por sua vez, exibem áreas significativamente menores e praticamente sobrepostas, indicando limitações estruturais na captura de padrões não lineares. Esse comportamento reforça as conclusões obtidas nas análises anteriores, confirmando que esses modelos não conseguem se aproximar da precisão alcançada pelas técnicas mais avançadas.

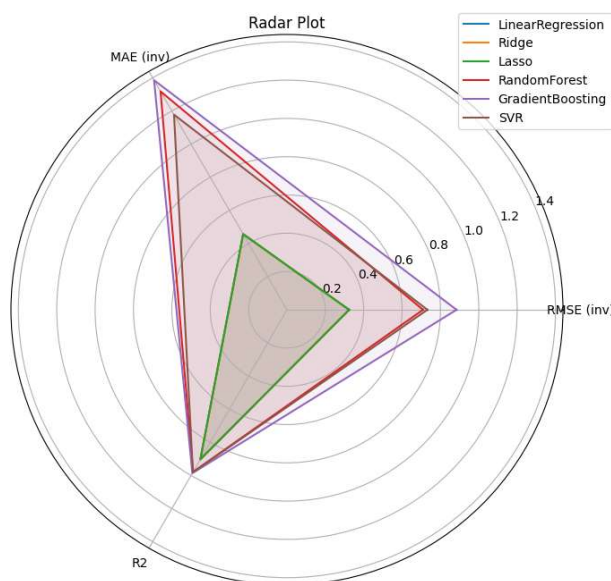


Figura 11 – Radar Plot – Comparação Complementar dos Modelos

A Figura 12 apresenta um heatmap contendo os valores originais das métricas **RMSE**, **MAE** e **R²** para cada modelo avaliado. Essa visualização facilita a identificação de padrões de desempenho, permitindo observar de forma rápida quais técnicas apresentam menores erros e maior capacidade explicativa.

Os resultados mostram que o **Gradient Boosting** mantém o desempenho mais equilibrado entre as métricas, combinando os menores valores de RMSE e MAE com o maior valor de R². Os modelos **Random Forest** e **SVR** aparecem em posição intermediária, apresentando desempenho consistente, mas inferior ao do Gradient Boosting. Por outro lado, os modelos lineares concentram-se nas regiões de maior erro e menor R², evidenciando uma limitação na captura de relações não lineares presentes no conjunto de dados.

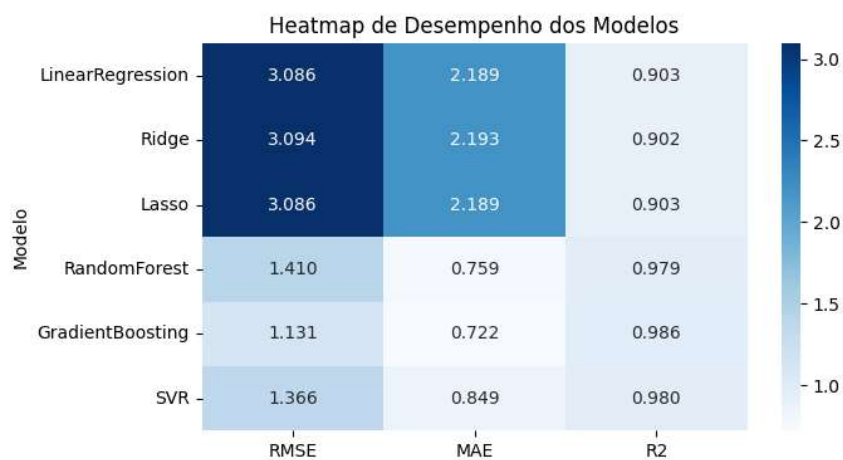


Figura 12 – Heatmap de Desempenho dos Modelo

A Figura 13 apresenta o ranking relativo dos modelos para cada métrica de avaliação, em que a posição **1** representa o melhor desempenho. Essa visualização permite comparar diretamente a posição de cada abordagem em termos de erro e capacidade explicativa, destacando padrões de consistência entre as técnicas.

O **Gradient Boosting** mantém-se como o modelo mais robusto, ocupando a primeira posição tanto em RMSE quanto em MAE e apresentando desempenho competitivo em R². Os modelos **Random Forest** e **SVR** aparecem em posições intermediárias, refletindo boa capacidade de modelagem, embora com resultados inferiores aos do Gradient Boosting. Já as abordagens lineares permanecem nas últimas colocações em todas as métricas, reforçando sua limitação estrutural na captura de relações não lineares presentes no conjunto de dados.

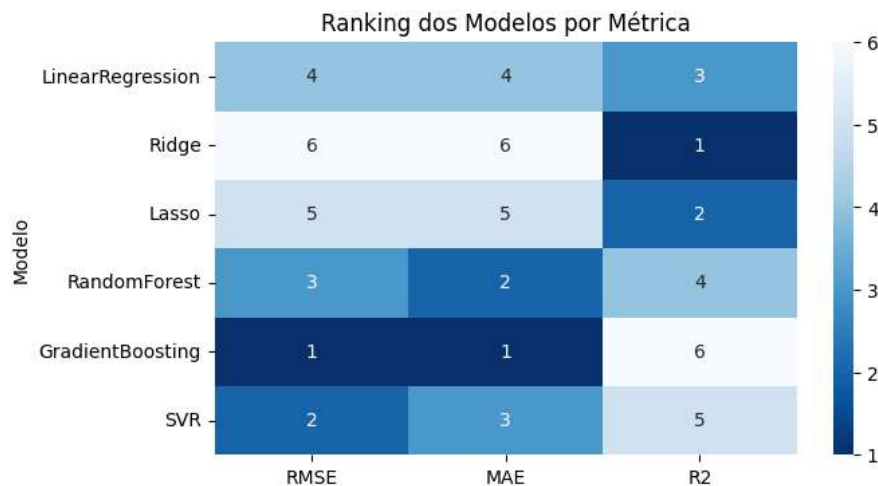


Figura 13 – Ranking dos Modelos por Métrica

Análise Complementar

Os resultados apresentados nas Figuras 7 a 13 evidenciam que os modelos não lineares — em especial **Gradient Boosting**, **SVR** e **Random Forest** — apresentaram desempenho significativamente superior aos modelos lineares em todas as métricas avaliadas. Tanto o RMSE quanto o MAE confirmam que essas técnicas são mais adequadas para capturar as relações complexas presentes no dataset de eficiência energética.

O coeficiente de determinação **R²** reforça a elevada capacidade explicativa do Gradient Boosting, que se destacou como o melhor modelo, combinando os menores erros com o maior poder preditivo entre todas as abordagens testadas. Esse comportamento consistente justifica sua seleção como modelo final desta etapa.

As análises complementares apresentadas nas Figuras 11, 12 e 13 reforçam os achados anteriores, oferecendo diferentes perspectivas sobre o desempenho dos algoritmos. O **radar plot** evidencia visualmente a dominância do Gradient Boosting, enquanto o **heatmap** sintetiza os valores numéricos de forma compacta, facilitando a comparação entre os modelos. O **ranking por métrica** confirma a estabilidade do desempenho, com o Gradient Boosting ocupando posições superiores na maior parte das avaliações.

Por fim, embora os resultados obtidos sejam satisfatórios, é importante destacar que todos os modelos foram avaliados utilizando **hiperparâmetros padrão**. Dessa forma, existe amplo espaço para aprimoramento por meio de técnicas como ajuste fino (tuning), validação cruzada, busca em grade (*Grid Search*) ou algoritmos de otimização. Essas melhorias serão exploradas nas próximas fases do trabalho, com o objetivo de identificar configurações ainda mais eficientes para aplicação prática no sistema de eficiência energética.