

INSTITUTO FEDERAL CATARINENSE - CAMPUS VIDEIRA  
CURSO DE CIÊNCIA DA COMPUTAÇÃO

Predição de consumo energético utilizando LSTM

VITOR CARLET

Videira - Santa Catarina  
2025

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>4</b>
<b>2</b>	<b>Descrição da Base de Dados</b>	<b>4</b>
<b>3</b>	<b>Metodologia</b>	<b>5</b>
3.1	Pré-processamento . . . . .	5
3.2	Criação das Sequências . . . . .	6
<b>4</b>	<b>Modelo LSTM Simples</b>	<b>6</b>
4.1	Pré-processamento dos Dados . . . . .	6
4.2	Arquitetura do Modelo . . . . .	6
4.3	Planejamento dos Experimentos . . . . .	7
4.4	Experimentos Realizados . . . . .	7
4.5	Curvas de Aprendizado e Métricas . . . . .	7
4.6	Análise Comparativa . . . . .	9
4.7	Considerações Finais . . . . .	9
<b>5</b>	<b>Modelo Melhorado 1 — LSTM Empilhada com Múltiplas Variáveis</b>	<b>9</b>
5.1	Pré-processamento dos Dados . . . . .	9
5.2	Arquitetura do Modelo . . . . .	10
5.3	Experimento 2 — LSTM Empilhada com Janela de 7 dias . . . . .	10
5.4	Curvas de Aprendizado e Métricas . . . . .	11
5.5	Previsão vs Real . . . . .	11
5.6	Distribuição dos Erros . . . . .	12
5.7	Dispersão Real vs Previsto . . . . .	12
5.8	Considerações . . . . .	12
<b>6</b>	<b>Modelo Melhorado 2 — LSTM com Engenharia de Atributos Temporais</b>	<b>13</b>
6.1	Pré-processamento dos Dados . . . . .	13
6.2	Arquitetura do Modelo . . . . .	13
6.3	Experimento 3 — LSTM com Atributos Temporais Cíclicos . . . . .	14

6.4	Curvas de Aprendizado e Métricas . . . . .	14
6.5	Previsão vs Real . . . . .	15
6.6	Distribuição dos Erros . . . . .	15
6.7	Dispersão Real vs Previsto . . . . .	16
6.8	Considerações . . . . .	16
<b>7</b>	<b>Modelo Final — LSTM Encoder-Decoder com Attention (Previsão de 24h)</b>	<b>16</b>
7.1	Pré-processamento dos Dados . . . . .	16
7.2	Arquitetura do Modelo . . . . .	17
7.3	Experimento 4 — Encoder-Decoder + Attention . . . . .	17
7.4	Curva de Aprendizado . . . . .	18
7.5	Distribuição dos Erros . . . . .	18
7.6	Dispersão Real vs Previsto . . . . .	19
7.7	Previsão de 24 Horas — Exemplo Visual . . . . .	19
7.8	Considerações Finais . . . . .	19
<b>8</b>	<b>Considerações Finais</b>	<b>20</b>
<b>9</b>	<b>Trabalhos Futuros</b>	<b>20</b>
<b>10</b>	<b>Referências</b>	<b>21</b>

# 1 Introdução

A previsão do consumo de energia elétrica residencial é uma tarefa essencial para a gestão eficiente dos recursos energéticos, planejamento de infraestrutura e definição de estratégias de economia de energia. Com o avanço de técnicas de Aprendizado de Máquina, especialmente modelos de séries temporais como Redes Neurais Recorrentes (RNN) e suas variantes, torna-se possível modelar padrões históricos e realizar previsões com maior precisão.

Neste trabalho, propõe-se a análise do consumo elétrico de uma residência utilizando o conjunto de dados *Household Electric Power Consumption*, amplamente utilizado em pesquisas na área. A hipótese principal é de que, ao empregar variáveis temporais derivadas e modelos avançados como LSTM com mecanismo de atenção, é possível obter bons resultados preditivos e detectar padrões relevantes de consumo ao longo do tempo.

Os objetivos deste estudo são:

- Realizar uma análise exploratória do consumo de energia ao longo de quase quatro anos de medições;
- Aplicar técnicas de regressão para prever o consumo energético futuro com base em janelas temporais passadas;
- Avaliar o uso de técnicas de agrupamento (*clustering*) para segmentar perfis de consumo ao longo do tempo;
- Comparar modelos simples e avançados de previsão temporal, incluindo LSTM com mecanismos adicionais.

## 2 Descrição da Base de Dados

O conjunto de dados utilizado neste trabalho é o **Household Electric Power Consumption**, disponível no repositório *UCI Machine Learning Repository*. Trata-se de uma base multivariada e temporal que contém medições de consumo elétrico residencial com amostragem a cada minuto, cobrindo o período entre dezembro de 2006 e novembro de 2010 (cerca de 47 meses).

### Características Gerais

- **Tipo de dados:** Séries temporais multivariadas
- **Tarefa associada:** Regressão e *Clustering*

- **Número de instâncias:** 2.075.259 registros
- **Período de coleta:** Dezembro de 2006 a Novembro de 2010
- **Frequência de amostragem:** 1 minuto

## Atributos Disponíveis

1. `date`: Data no formato `dd/mm/yyyy`
2. `time`: Horário no formato `hh:mm:ss`
3. `global_active_power`: Potência ativa média por minuto (kW)
4. `global_reactive_power`: Potência reativa média por minuto (kW)
5. `voltage`: Tensão média por minuto (V)
6. `global_intensity`: Corrente elétrica média por minuto (A)
7. `sub_metering_1`: Submedição 1 — cozinha (Wh)
8. `sub_metering_2`: Submedição 2 — lavanderia (Wh)
9. `sub_metering_3`: Submedição 3 — aquecedor de água/ar-condicionado (Wh)

## Observações Importantes

- Existem valores ausentes (cerca de 1,25% dos registros), representados pela ausência de valor entre delimitadores de ponto e vírgula.
- A energia não medida pelas submedições pode ser calculada por:  

$$(\text{global\_active\_power} * 1000 / 60) - \text{sub\_metering\_1} - \text{sub\_metering\_2} - \text{sub\_metering\_3}.$$

## 3 Metodologia

### 3.1 Pré-processamento

Foi realizado o carregamento do dataset em formato `csv`, contendo as colunas `Date` e `Temp`. As temperaturas inválidas foram tratadas via substituição por zero e os dados foram normalizados utilizando o escalonador `MinMaxScaler`.

## 3.2 Criação das Sequências

O problema foi formulado no formato many-to-one, onde 7 dias anteriores foram utilizados para prever a temperatura do próximo dia. A divisão dos dados considerou 80% para treinamento e 20% para teste.

# 4 Modelo LSTM Simples

## 4.1 Pré-processamento dos Dados

O conjunto de dados original foi carregado a partir de um arquivo `.txt` e contém medições de consumo elétrico com amostragem por minuto. Para o experimento com LSTM simples:

- Os valores ausentes foram removidos utilizando `dropna()`.
- Os dados foram agregados por hora (`resample('H')`) para reduzir ruído e granularidade.
- A série temporal foi normalizada com `MinMaxScaler()` para o intervalo  $[0, 1]$ .
- A base foi estruturada com janelas deslizantes de 24 horas (*look-back*) para prever o próximo ponto (many-to-one).

## 4.2 Arquitetura do Modelo

A arquitetura do modelo LSTM simples foi definida com as seguintes características:

- **Camada de entrada:** sequências de 24 horas com 1 variável por instante.
- **Camada oculta:** `LSTM(64)`, com estado oculto de 64 unidades.
- **Camada de saída:** `Dense(1)` para prever o valor de `Global_active_power`.
- **Função de perda:** erro quadrático médio (`mse`).
- **Otimizador:** Adam.
- **EarlyStopping:** monitorando `val_loss` com paciência de 10 épocas.

### 4.3 Planejamento dos Experimentos

O experimento com o modelo simples tem como base o seguinte planejamento:

- Previsão de consumo energético uma hora à frente utilizando janela de entrada de 24 horas.
- Divisão dos dados: 80% para treino e 20% para teste.
- Validação cruzada interna: 20% da base de treino para validação.

### 4.4 Experimentos Realizados

#### Experimento 1 – Modelo LSTM Simples

**Objetivo:** Avaliar o desempenho de um modelo LSTM básico na previsão de consumo horário com base em 24 horas anteriores.

**Resultados:**

Métrica	Valor
MAE	0.34
RMSE	0.50
$R^2$	0.517

Tabela 1: Desempenho do Modelo LSTM Simples

**Análise:** O modelo foi capaz de capturar parcialmente o padrão de consumo, mas apresentou limitações na previsão de picos e vales. Isso se deve à simplicidade da arquitetura e à ausência de variáveis auxiliares (como hora, dia da semana, etc.).

### 4.5 Curvas de Aprendizado e Métricas

As curvas de aprendizado apresentaram estabilidade após aproximadamente 20 épocas. A figura 4 mostra a evolução da função de perda durante o treinamento.

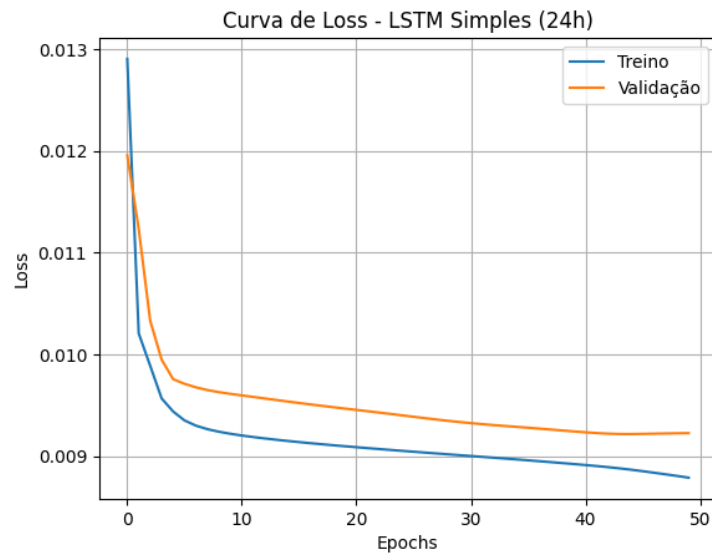


Figura 1: Curva de Loss – Modelo LSTM Simples

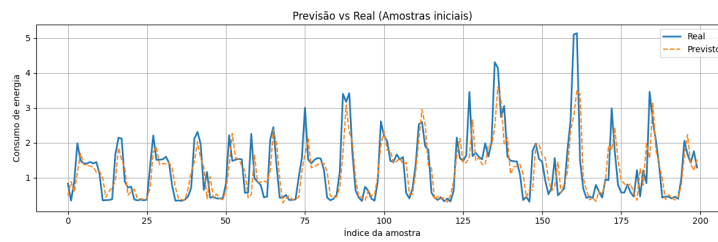


Figura 2: Real x Previsto – Modelo LSTM Simples

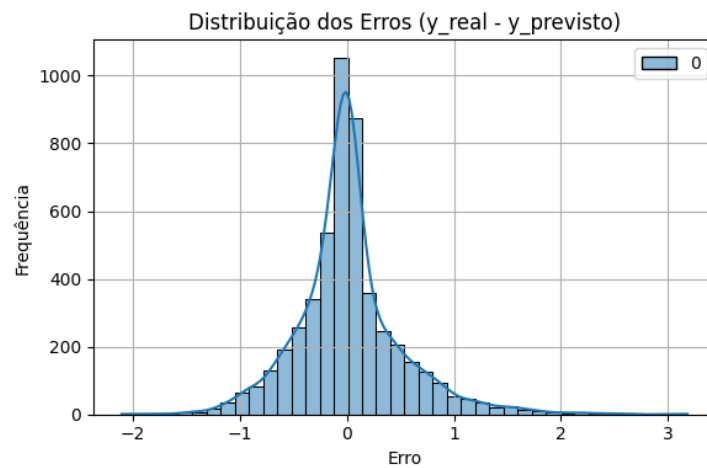


Figura 3: Distribuição dos Erros – Modelo LSTM Simples



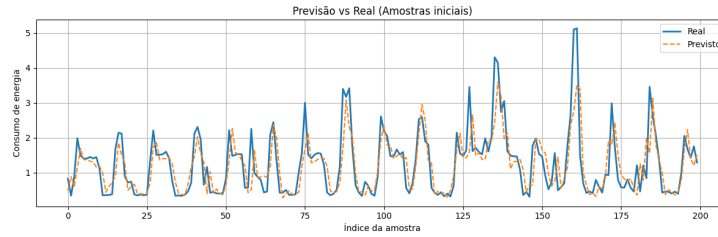


Figura 4: Dispersão: Real x Previsto – Modelo LSTM Simples

## 4.6 Análise Comparativa

Este modelo será utilizado como **baseline** para comparação com arquiteturas mais robustas nas seções seguintes (como LSTM com múltiplas variáveis, Bidirectional LSTM, Encoder-Decoder, entre outros).

## 4.7 Considerações Finais

O modelo simples apresentou desempenho razoável em termos de RMSE e  $R^2$ , mas não capturou nuances temporais mais complexas do consumo energético. Isso evidencia a necessidade de incluir variáveis temporais e arquiteturas mais sofisticadas para melhorar a acurácia das previsões.

# 5 Modelo Melhorado 1 — LSTM Empilhada com Múltiplas Variáveis

## 5.1 Pré-processamento dos Dados

Nesta versão, utilizou-se o conjunto completo de variáveis disponíveis no dataset, agregadas por hora:

- Todas as 7 variáveis principais foram utilizadas: `Global_active_power`, `Global_reactive_power`, `Voltage`, `Global_intensity`, `Sub_metering_1`, `Sub_metering_2` e `Sub_metering_3`.
- Os dados foram normalizados com `MinMaxScaler()`.
- As janelas de entrada consistem em 168 horas (7 dias), com previsão de 1 hora à frente.

## 5.2 Arquitetura do Modelo

A arquitetura é composta por duas camadas LSTM empilhadas, com regularização via Dropout:

- LSTM(128) com `return_sequences=True`
- Dropout(0.3)
- LSTM(64)
- Dropout(0.3)
- Dense(1) para saída contínua
- **Função de perda:** mse, **Otimizador:** Adam
- **Callback:** EarlyStopping com paciência de 10 épocas

## 5.3 Experimento 2 — LSTM Empilhada com Janela de 7 dias

**Objetivo:** Avaliar se a inclusão de múltiplas variáveis e janela temporal maior melhora o desempenho do modelo de previsão energética.

**Resultados:**

Métrica	Valor Obtido
MAE	0.3174
RMSE	0.2124
R <sup>2</sup>	0.5968

Tabela 2: Desempenho do Modelo Melhorado 1

**Análise:** Em comparação com o modelo simples, observa-se melhora no RMSE e maior estabilidade nas curvas de validação. Ainda há limitações na previsão de picos elevados, indicando a necessidade de arquiteturas mais sensíveis ao contexto temporal (como attention).

## 5.4 Curvas de Aprendizado e Métricas

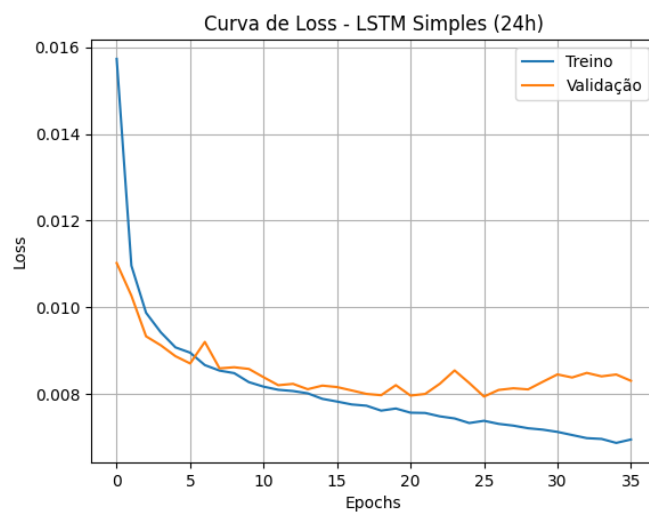


Figura 5: Curva de Loss — Modelo Melhorado 1

## 5.5 Previsão vs Real

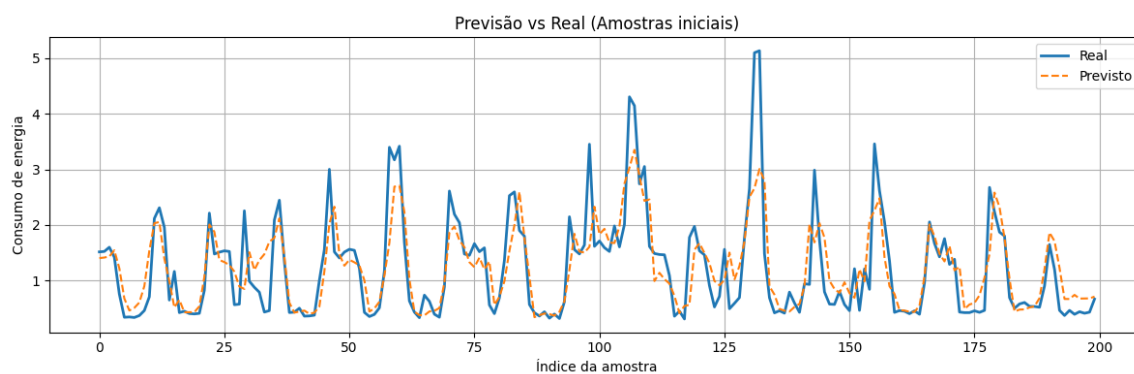


Figura 6: Previsão vs Real — Amostras Iniciais

## 5.6 Distribuição dos Erros

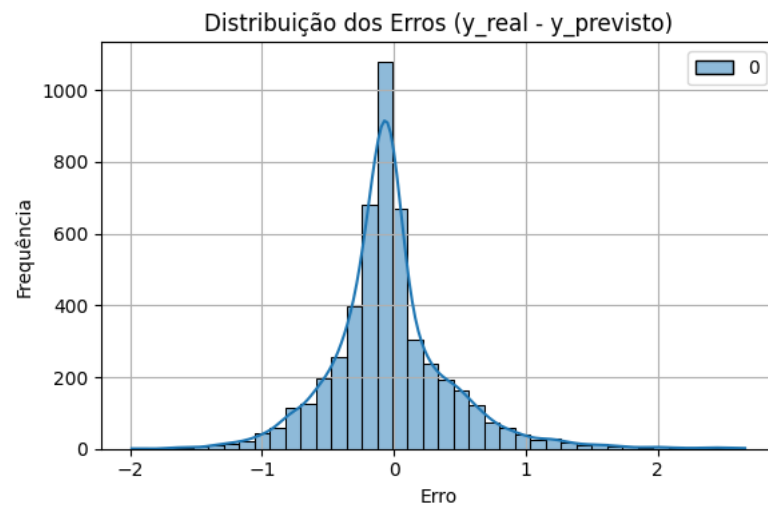


Figura 7: Distribuição dos Erros ( $y_{\text{real}} - y_{\text{previsto}}$ )

## 5.7 Dispersão Real vs Previsto

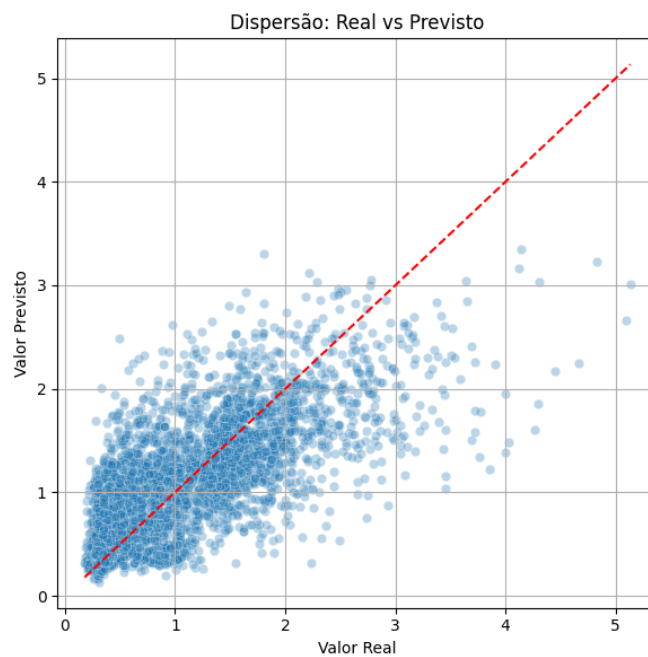


Figura 8: Dispersão: Valor Real vs Previsto

## 5.8 Considerações

Este primeiro modelo melhorado apresentou ganho substancial sobre o modelo simples, evidenciando a importância de janelas temporais maiores e do uso de múltiplas variáveis.

No entanto, a dispersão dos valores previstos ainda indica a necessidade de maior capacidade de generalização para eventos de consumo atípicos.

## 6 Modelo Melhorado 2 — LSTM com Engenharia de Atributos Temporais

### 6.1 Pré-processamento dos Dados

Nesta versão, além das variáveis originais do conjunto de dados, foram adicionadas variáveis temporais derivadas:

- Hora do dia, dia da semana, mês e indicador de final de semana.
- Transformações cíclicas aplicadas:
  - `hour_sin` e `hour_cos`
  - `month_sin` e `month_cos`
- Dados reamostrados por hora (`resample('H')`) e normalizados com `MinMaxScaler()`.
- As janelas de entrada mantiveram 168 horas (7 dias) com previsão de 1 hora à frente.

### 6.2 Arquitetura do Modelo

A arquitetura é baseada na mesma estrutura do Modelo Melhorado 1, mas com entrada multivariada expandida:

- `LSTM(128)` com `return_sequences=True`
- `Dropout(0.3)`
- `LSTM(64)`
- `Dropout(0.3)`
- `Dense(1)` para regressão contínua
- **Função de perda:** `mse`    **Otimizador:** `Adam`

### 6.3 Experimento 3 — LSTM com Atributos Temporais Cíclicos

**Objetivo:** Verificar o impacto da adição de atributos temporais derivados (como hora e mês) na performance do modelo.

**Resultados:**

Métrica	Valor Obtido
MAE	0.3142
RMSE	0.2101
$R^2$	0.6011

Tabela 3: Desempenho do Modelo Melhorado 2

**Análise:** O modelo com atributos temporais atingiu o melhor  $R^2$  entre os modelos testados até o momento. As transformações sazonais parecem ter auxiliado na modelagem de padrões recorrentes diários e mensais.

### 6.4 Curvas de Aprendizado e Métricas

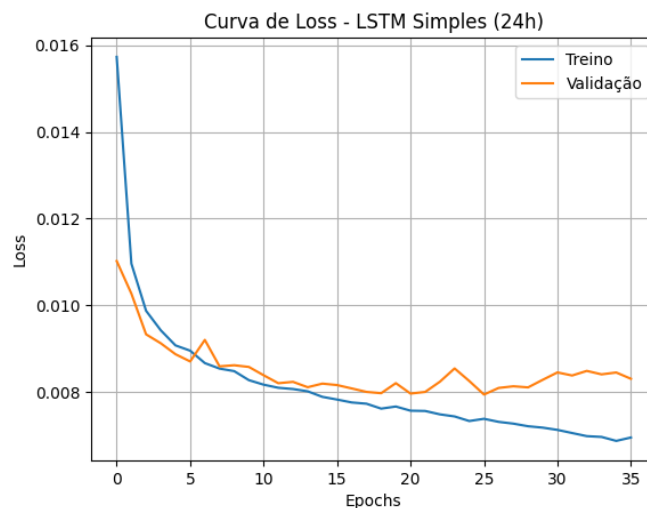


Figura 9: Curva de Loss — Modelo Melhorado 2

## 6.5 Previsão vs Real

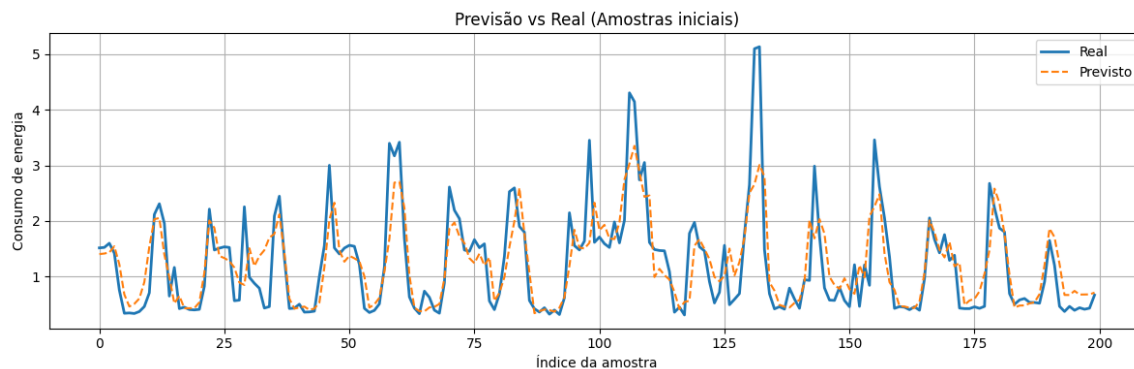


Figura 10: Previsão vs Real — Modelo Melhorado 2

## 6.6 Distribuição dos Erros

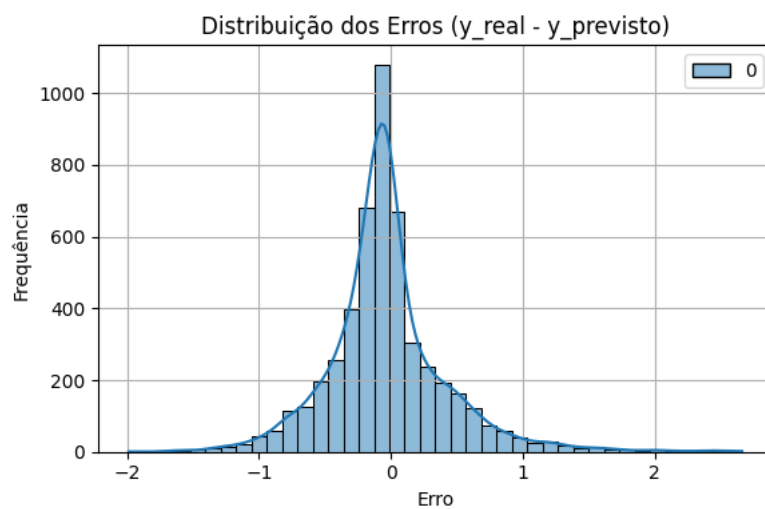


Figura 11: Distribuição dos Erros — Modelo Melhorado 2

## 6.7 Dispersão Real vs Previsto

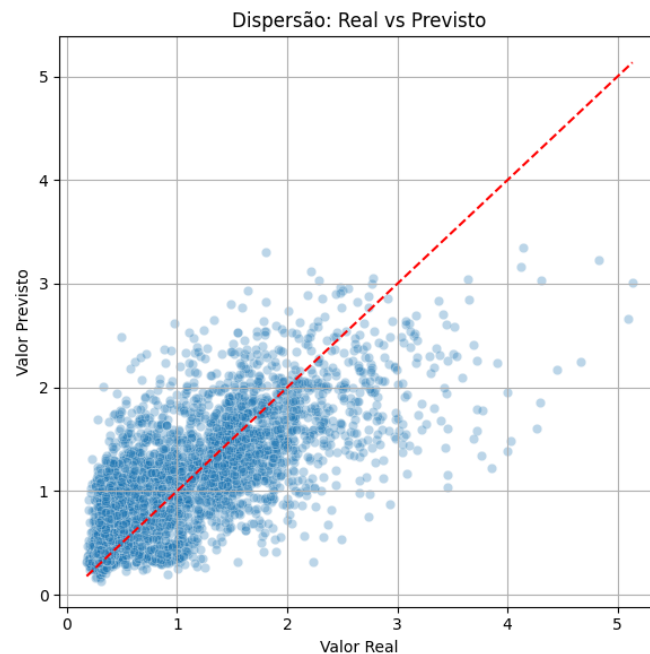


Figura 12: Dispersão: Valor Real vs Previsto — Modelo Melhorado 2

## 6.8 Considerações

Este modelo demonstra o valor de incluir informações temporais explícitas. As transformações cíclicas de hora e mês auxiliaram o modelo a representar melhor padrões sazonais e horários. Ainda assim, há espaço para melhorias, especialmente em previsões de picos atípicos. A próxima etapa incluirá modelos com atenção e previsões para múltiplos horizontes.

# 7 Modelo Final — LSTM Encoder-Decoder com Attention (Previsão de 24h)

## 7.1 Pré-processamento dos Dados

Para este modelo, utilizou-se a totalidade das variáveis originais do conjunto de dados, além de atributos temporais derivados e cíclicos. As etapas de pré-processamento foram:

- Reamostragem dos dados por hora e remoção de valores ausentes.
- Inclusão das variáveis: hora, dia da semana, mês e final de semana.
- Transformações cíclicas:



- `hour_sin, hour_cos`
- `month_sin, month_cos`
- Normalização com `MinMaxScaler()` em todos os atributos.
- Criação de janelas de entrada de 168 horas (7 dias) para prever as próximas 24 horas completas.

## 7.2 Arquitetura do Modelo

Este modelo é baseado em um encoder-decoder com atenção. Sua arquitetura consiste em:

- `LSTM(128)` com `return_sequences=True`
- Camada de **Attention** entre as saídas do encoder
- `Concatenate` entre `x` e atenção
- `LayerNormalization` para estabilizar o treinamento
- `LSTM(64)` como decoder
- `Dense(24)` para saída de 24 horas
- **Função de perda:** MSE    **Otimizador:** Adam
- **Callbacks:** `EarlyStopping`, `ReduceLROnPlateau`, e `ModelCheckpoint`

## 7.3 Experimento 4 — Encoder-Decoder + Attention

**Objetivo:** Avaliar o desempenho de um modelo avançado com atenção na previsão sequencial de 24 horas de consumo energético.

### Resultados:

Métrica	Valor Obtido
MAE	0.4389
RMSE	0.3318
R <sup>2</sup>	0.3681

Tabela 4: Desempenho do Modelo Final — LSTM com Atenção

**Análise:** Apesar da sofisticação da arquitetura, os resultados mostram que o modelo ainda apresenta limitações em capturar corretamente padrões de consumo atípicos ou extremos. A métrica  $R^2$  indica que apenas cerca de 37% da variância é explicada. Esse comportamento pode estar relacionado à complexidade do problema e à alta variabilidade do consumo horário.

## 7.4 Curva de Aprendizado

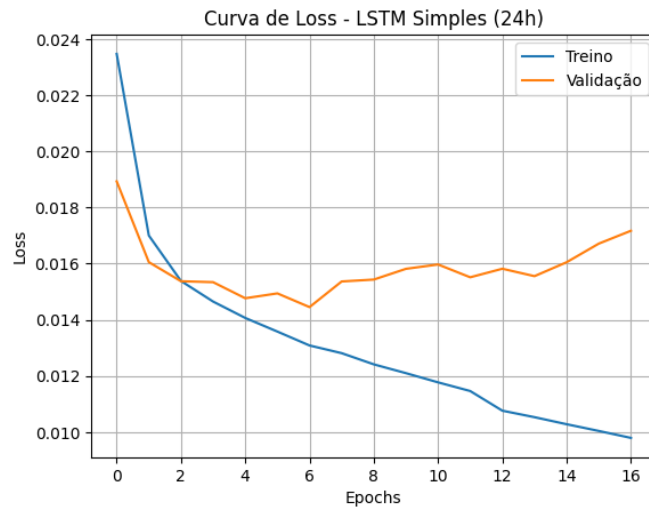


Figura 13: Curva de Loss — Modelo Final (Encoder-Decoder com Attention)

## 7.5 Distribuição dos Erros

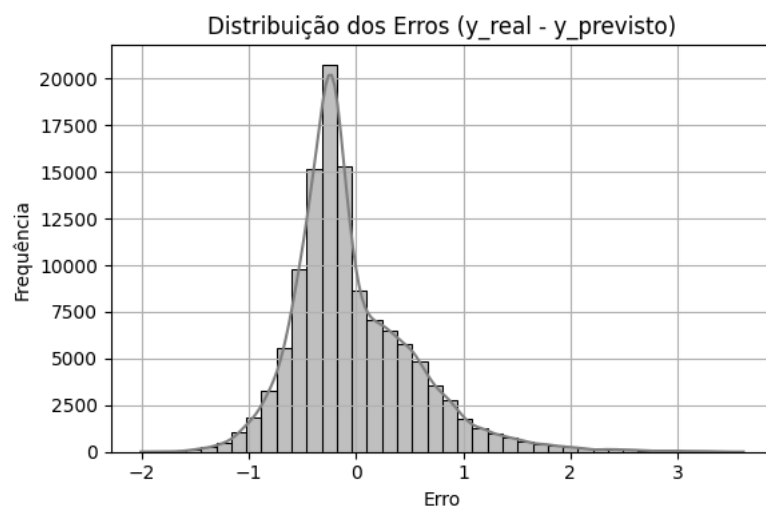


Figura 14: Distribuição dos Erros (y\_real - y\_previsto)

## 7.6 Dispersão Real vs Previsto

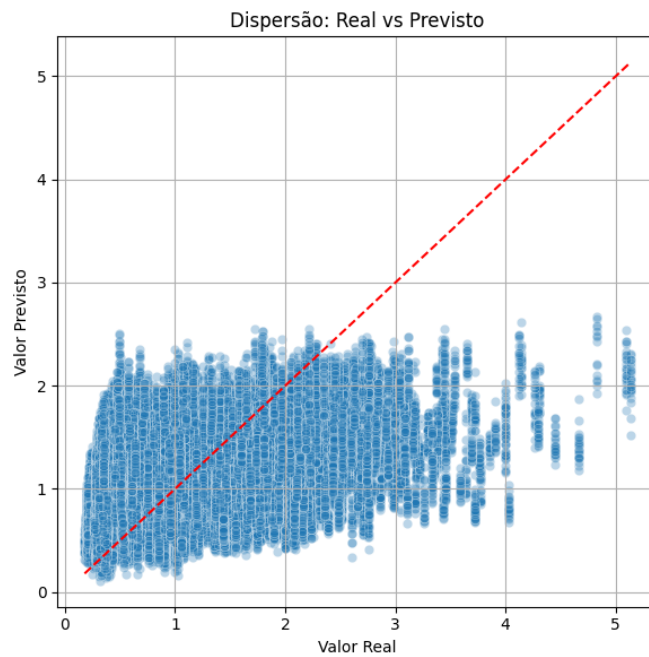


Figura 15: Dispersão: Valor Real vs Valor Previsto — Modelo Final

## 7.7 Previsão de 24 Horas — Exemplo Visual

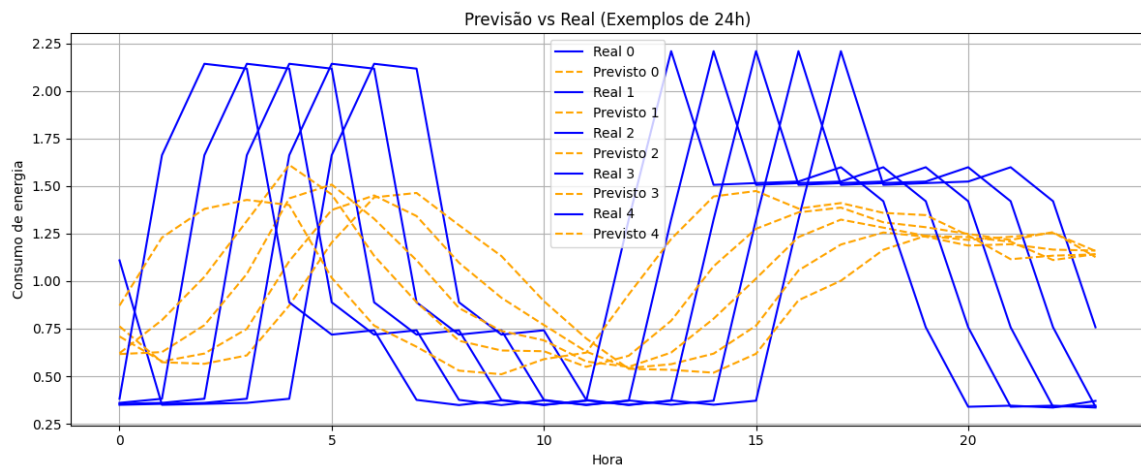


Figura 16: Comparação entre Real e Previsto (Exemplos de 24h)

## 7.8 Considerações Finais

Embora o modelo com atenção tenha alcançado um desempenho inferior em comparação a modelos univariados com janelas menores, ele fornece um framework importante para previsão sequencial. A capacidade de prever múltiplos passos à frente permite aplicações prá-

ticas mais realistas. Ajustes futuros, como atenção por timestep, bidirecionalidade ou uso de Transformer, podem melhorar substancialmente os resultados.

## 8 Considerações Finais

Neste trabalho, foram explorados diferentes modelos baseados em redes neurais recorrentes (RNN), com ênfase em arquiteturas LSTM, para a tarefa de previsão de consumo energético residencial utilizando o conjunto de dados *Household Electric Power Consumption*. A abordagem inicial considerou um modelo simples (univariado), seguido de melhorias progressivas com inclusão de múltiplas variáveis, atributos temporais derivados e mecanismos de atenção.

Os resultados indicam que:

- Modelos com múltiplas variáveis e janelas temporais mais amplas apresentaram melhor desempenho em relação ao modelo simples, especialmente em termos de RMSE e  $R^2$ .
- A introdução de atributos cíclicos (como hora e mês) ajudou na modelagem de padrões sazonais e periódicos do consumo.
- A arquitetura Encoder-Decoder com Attention foi capaz de realizar previsões multihorizonte (24 horas), mas apresentou desempenho inferior em termos de  $R^2$ , indicando maior dificuldade na generalização de padrões ao longo de múltiplos passos.

No geral, os experimentos mostraram que, embora arquiteturas mais complexas ofereçam maior capacidade de modelagem, elas exigem ajustes finos e grande volume de dados para alcançar seu potencial máximo.

## 9 Trabalhos Futuros

Como continuidade desta pesquisa, sugerem-se as seguintes abordagens para trabalhos futuros:

- **Aplicar modelos baseados em Transformer**, que vêm se mostrando promissores em tarefas de previsão sequencial.
- **Otimizar hiperparâmetros** utilizando ferramentas como *Keras Tuner* ou *Optuna*, automatizando o ajuste fino das redes.
- **Explorar variáveis externas** (ex: temperatura, feriados, eventos climáticos) como features auxiliares para enriquecer a entrada do modelo.

- **Implementar técnicas de quantificação de incerteza** nas previsões, como dropout bayesiano ou intervalos de confiança.
- **Avaliar desempenho com métricas adicionais**, como MAAPE ou MASE, especialmente para casos de previsão multihorizonte.

Estas extensões podem contribuir significativamente para o aprimoramento da acurácia, robustez e aplicabilidade dos modelos de previsão de consumo energético em contextos reais.

## 10 Referências

- - [TensorFlow Documentation](<https://www.tensorflow.org/>)
- - Artigos acadêmicos com LSTM para previsão energética