

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
BACHARELADO EM ENGENHARIA FÍSICA

VÍTOR DE SOUZA BARBOZA

**PREVISÃO DE DEMANDA DE ENERGIA ELÉTRICA NO
BRASIL UTILIZANDO ARIMA E REDES NEURAIS**

SÃO CARLOS

2025

VÍTOR DE SOUZA BARBOZA

PREVISÃO DE DEMANDA DE ENERGIA ELÉTRICA NO BRASIL
UTILIZANDO ARIMA E REDES NEURAIIS

Trabalho final de curso apresentado ao Programa de Graduação em Engenharia Física da Universidade Federal de São Carlos, para obtenção do título de bacharel em Engenharia Física.
Orientador: Prof. Dr. Pedro Augusto Franco Pinheiro Moreira

SÃO CARLOS

2025

Barboza, Vítor de Souza

Previsão de demanda de energia elétrica no brasil
utilizando Arima e redes neurais / Vítor de Souza
Barboza -- 2025.
48f.

TCC (Graduação) - Universidade Federal de São Carlos,
campus São Carlos, São Carlos
Orientador (a): Pedro Augusto Franco Pinheiro Moreira
Banca Examinadora: Leonardo Kleber Castelano, Paulo
Sérgio da Silva Junior
Bibliografia

1. Previsão. 2. Energia Elétrica. 3. Série Temporal. I.
Barboza, Vítor de Souza. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

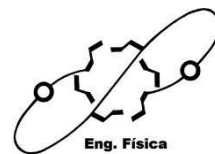
Bibliotecário responsável: Arildo Martins - CRB/8 7180



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA - CCET

Coordenação de Graduação do Curso de
Engenharia Física

Departamento de Física, Rodovia Washington Luiz, Km 235 - Caixa Postal 676 - CEP
13565-905, São Carlos - SP - Brasil - Fone: (16) 3351-8222
e-mail: enfi@df.ufscar.br



FICHA DE AVALIAÇÃO DE TRABALHO FINAL DE CURSO ENGENHARIA FÍSICA

Aluno: Sr. Vitor de Souza Barboza. RA: 791446.

Título: Previsão de Demanda de Energia Elétrica no Brasil utilizando ARIMA e Redes Neurais

Prof. Orientador: Prof. Dr. Pedro Augusto Franco Pinheiro Moreira, DF/UFSCar.


Prof. Examinador 1: Prof. Dr. Leonardo Kleber Castelano, DF/UFSCar.

Prof. Examinador 2: Prof. Dr. Paulo Sérgio da Silva Junior, DF/UFSCar.


Itens Avaliados	Orientador	Examinador 1	Examinador 2
Redação (atribuir notas de 0 a 2)	1,5	1	1,5
Apresentação oral (atribuir notas de 0 a 2)	2	2	2
Conteúdo desenvolvido no trabalho (atribuir notas de 0 a 4)	4	4	4
Arguição (atribuir notas de 0 a 2)	2	2	2

Observações: Não houve observações.


São Carlos, 17 de fevereiro de 2025.

Documento assinado digitalmente
 PEDRO AUGUSTO FRANCO PINHEIRO MOREIRA
Data: 17/02/2025 15:35:23-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Pedro Augusto Franco Pinheiro Moreira

Documento assinado digitalmente
 LEONARDO KLEBER CASTELANO
Data: 17/02/2025 17:20:53-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Leonardo Kleber Castelano

Documento assinado digitalmente
 PAULO SERGIO DA SILVA JUNIOR
Data: 17/02/2025 19:14:28-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Paulo Sérgio da Silva Junior

Resumo

A previsão acurada da demanda de energia elétrica é um elemento importante para o planejamento eficiente e sustentável do setor energético. Este trabalho tem como foco a análise comparativa entre os modelos ARIMA (Autoregressive Integrated Moving Average) e Redes Neurais Recorrentes (RNNs) para a previsão da demanda de energia elétrica no Brasil, utilizando técnicas avançadas de modelagem de séries temporais.

A importância deste estudo se fundamenta na necessidade crescente de otimização do uso dos recursos energéticos, alinhando oferta e demanda de forma eficaz. Uma previsão acurada contribui significativamente para a prevenção de sobrecargas e desperdícios, fatores essenciais para a redução de custos e garantia de um fornecimento estável. Além disso, o uso racional da energia elétrica tem um impacto direto na preservação ambiental, permitindo a redução das emissões de gases de efeito estufa e o cumprimento de metas de sustentabilidade.

No contexto brasileiro, a relevância deste estudo é ainda mais pronunciada, considerando que a matriz energética do país é uma das mais limpas do mundo, com uma participação significativa de fontes renováveis. Segundo dados recentes, as fontes renováveis representarão uma parcela expressiva da capacidade instalada de geração de energia elétrica no Brasil nos próximos anos. A previsão acurada da demanda é fundamental para o sucesso de iniciativas governamentais e para o planejamento a longo prazo do setor, contribuindo para um desenvolvimento energético mais sustentável, econômico e equitativo para a população brasileira.

O objetivo principal deste trabalho foi realizar a previsão da demanda de energia elétrica no Brasil, utilizando e comparando dois métodos robustos: o modelo ARIMA e as RNNs. Os objetivos específicos incluíram a análise exploratória dos dados de demanda energética, a implementação e ajuste dos modelos ARIMA e RNN, a comparação da performance dos modelos utilizando a métrica MAPE (Mean Absolute Percentage Error), e a identificação do método mais eficaz considerando precisão, complexidade e aplicabilidade em dados reais.

Os resultados obtidos neste estudo demonstraram a eficácia de ambos os modelos na previsão da demanda de energia elétrica no Brasil. O modelo ARIMA apresentou uma precisão ligeiramente superior, enquanto o modelo RNN se destacou pela sua eficiência computacional. A análise comparativa revelou que, embora o ARIMA tenha alcançado uma maior acurácia nas previsões, o RNN oferece vantagens significativas em termos de tempo de processamento e escalabilidade. Estes achados têm implicações importantes para o setor energético brasileiro, fornecendo insights para a otimização do planejamento e gestão dos recursos energéticos, alinhando-se com os objetivos de desenvolvimento sustentável e as metas de transição energética do país.

Palavras-chave: ARIMA, Redes Neurais Recorrentes (RNNs), Energia Elétrica, Previsão, Série Temporal

Sumário

1	Introdução	6
1.1	Motivação	6
1.2	Objetivos	7
1.3	Por que utilizar ARIMA e RNNs?	8
2	Fundamentos Teóricos	10
2.1	Séries Temporais	10
2.2	ARIMA	12
2.3	Redes Neurais Recorrentes (RNN)	15
2.4	Métricas de Avaliação	18
3	Metodologia	21
3.1	Base de Dados	21
3.2	Modelagem ARIMA	25
3.3	Redes Neurais Recorrentes (RNNs)	28
4	Resultados	31
4.1	Avaliação do Modelo e Resultados da Previsão da modelagem ARIMA	31
4.2	Avaliação do Modelo e Visualização dos Resultados das RNNs	34
4.3	Avaliação da Variação dos Anos de Validação e Teste	37
5	Conclusão	41
	Referências	43

1 Introdução

1.1 Motivação

A escolha do tema para este Trabalho de Conclusão de Curso, focado na previsão da demanda de energia elétrica no Brasil utilizando modelos ARIMA e Redes Neurais, visa trazer benefícios diretos ao setor energético e à sociedade em geral. A previsão acurada da demanda de energia é essencial para o planejamento eficiente da geração e distribuição de eletricidade, proporcionando uma série de vantagens sociais e ambientais ([HONG; FAN, 2016](#)).

Primeiramente, uma previsão acurada contribui para a otimização do uso dos recursos energéticos, permitindo um melhor alinhamento entre oferta e demanda ([SUGANTHI; SAMUEL, 2012](#)). Isso auxilia na prevenção tanto da sobrecarga quanto do desperdício de energia, fatores essenciais para a redução de custos e para a garantia de um fornecimento estável. Com uma previsão mais acurada, as empresas do setor podem melhorar sua eficiência operacional, resultando em menores custos e potencialmente em tarifas mais acessíveis para os consumidores ([JIANG et al., 2018](#)).

Figura 1 – Imagem panorâmica da Barragem de Itaipu, localizada entre o Brasil e o Paraguai.



Fonte: Wikipedia ([Wikipédia, a enciclopédia livre, 2025](#))

Além disso, o uso racional da energia elétrica tem um impacto direto na preservação ambiental. Com previsões mais apuradas, é possível reduzir a necessidade de acionamento de usinas termelétricas, que apresentam uma emissão de carbono significativa. Assim, contribui-se para a redução das emissões de gases de efeito estufa e para o cumprimento de metas de sustentabilidade e de mitigação do impacto ambiental ([KAYTEZ et al., 2015](#)).

A matriz energética brasileira é uma das mais limpas do mundo, com uma participação significativa de fontes renováveis. Segundo dados da Empresa de Pesquisa Energética (EPE), em 2024 as fontes renováveis representarão 84,62% da capacidade instalada de geração de energia

elétrica no Brasil ([Empresa de Pesquisa Energética, 2024](#)). No primeiro semestre de 2024, o país adicionou 5,6 gigawatts (GW) de potência instalada e 168 novas usinas entraram em operação ([Empresa de Pesquisa Energética, 2024](#)). Além disso, a previsão de crescimento da geração de energia elétrica do país para 2024 é de 10,1 GW, o que será o segundo maior avanço anual desde a criação da ANEEL (Agência Nacional de Energia Elétrica) em 1997 ([Empresa de Pesquisa Energética, 2024](#)).

As ações governamentais neste âmbito são cruciais para garantir a segurança energética e promover a transição para uma matriz ainda mais limpa. O Ministério de Minas e Energia (MME) tem implementado diversas iniciativas, como o Programa de Geração Distribuída (PGDU), que busca fomentar a instalação de sistemas de geração de energia elétrica a partir de fontes renováveis em universidades públicas ([Ministério de Minas e Energia, 2024](#)). Essas ações demonstram a importância da inovação e do planejamento no setor energético.

A previsão acurada da demanda de energia elétrica é fundamental para o sucesso dessas iniciativas e para o planejamento a longo prazo do setor. Os modelos avançados, como as redes neurais e ARIMA, têm demonstrado ser ferramentas valiosas para melhorar a precisão dessas previsões ([AHMAD; ANDERSON; LIE, 2014](#)). Sua aplicação pode contribuir significativamente para a tomada de decisões informadas e para a otimização dos recursos energéticos ([HYNDMAN; ATHANASOPOULOS, 2014](#)).

Em conclusão, este trabalho busca não apenas contribuir para o avanço acadêmico, mas também para um desenvolvimento energético mais sustentável, econômico e equitativo para a população brasileira. A melhoria na previsão da demanda de energia elétrica tem o potencial de impactar positivamente a eficiência do setor, a preservação ambiental e a qualidade de vida dos cidadãos.

1.2 Objetivos

O principal objetivo deste Trabalho de Conclusão de Curso é realizar a previsão da demanda de energia elétrica no Brasil, com o uso de técnicas avançadas de modelagem de séries temporais. A previsão acurada da demanda é essencial para o planejamento eficiente dos recursos energéticos e para a tomada de decisões estratégicas no setor elétrico. Para atingir esse objetivo, serão utilizados dois métodos robustos e amplamente aplicados em previsões de séries temporais: o modelo ARIMA e RNNs. A análise comparativa entre esses métodos visa identificar qual é mais eficaz na previsão da demanda energética, levando em conta a acurácia, a complexidade e a aplicabilidade em dados reais.

Os objetivos específicos deste trabalho são:

1. **Análise e preparação dos dados:** Realizar uma análise exploratória dos dados de demanda energética do Brasil, com foco em identificar tendências, sazonalidades e possíveis outliers.

2. **Implementação do modelo ARIMA:** Desenvolver e ajustar um modelo ARIMA para previsão de séries temporais, validando sua precisão preditiva.
3. **Implementação de RNNs:** Implementar um modelo de RNN, explorando a arquitetura GRU, e ajustar seus hiperparâmetros para a previsão da demanda energética.
4. **Comparação da performance dos modelos:** Avaliar a eficácia dos modelos ARIMA e RNNs utilizando a métrica MAPE para determinar o desempenho de cada abordagem.
5. **Identificação do modelo mais eficiente:** Determinar qual método oferece previsões mais acuradas e robustas, considerando a complexidade e o tempo de execução.

Este trabalho, ao avaliar essas técnicas, busca não apenas identificar a abordagem mais eficaz para previsão de demanda energética, mas também contribuir com a literatura, fornecendo insights práticos e teóricos para a aplicação de técnicas de modelagem em séries temporais no contexto do setor elétrico brasileiro.

1.3 Por que utilizar ARIMA e RNNs?

A previsão de séries temporais é uma área fundamental da estatística e do aprendizado de máquina, com aplicações que abrangem diversos campos, desde economia e finanças até meteorologia e engenharia (BOX et al., 2015). Ao longo dos anos, várias técnicas foram desenvolvidas para abordar esse desafio, cada uma com suas próprias vantagens e limitações. Entre essas técnicas, o modelo Autorregressivo Integrado de Médias Móveis (ARIMA) e as Redes Neurais Recorrentes (RNNs) se destacam como abordagens poderosas e complementares para a previsão de séries temporais (SIAMI-NAMINI; TAVAKOLI; NAMIN, 2019).

O modelo ARIMA, introduzido por Box e Jenkins na década de 1970, tem sido amplamente utilizado devido à sua capacidade de modelar uma ampla gama de padrões de dados (BOX et al., 2015). Este modelo combina três componentes principais: autorregressão (AR), diferenciação (I) e médias móveis (MA). A componente autorregressiva (AR) utiliza a dependência entre uma observação e várias observações anteriores, a diferenciação (I) torna a série estacionária subtraindo uma observação anterior, e a componente de médias móveis (MA) modela o erro em termos de uma combinação linear dos termos de erro anteriores (HYNDMAN; ATHANASSOPOULOS, 2018).

Por outro lado, as RNNs representam uma classe de redes neurais projetadas especificamente para reconhecer padrões em sequências de dados (GOODFELLOW; BENGIO; COURVILLE, 2016). Diferentemente das redes neurais tradicionais, as RNNs possuem conexões recorrentes que formam ciclos direcionados, permitindo que a informação persista ao longo do tempo (ELMAN, 1990). Essa característica é particularmente útil para séries temporais, onde as observações passadas frequentemente influenciam as futuras. Arquiteturas avançadas de RNNs, como Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU), são capazes de cap-

turar dependências de longo prazo nos dados, superando algumas das limitações dos modelos tradicionais como o ARIMA (HOCHREITER; SCHMIDHUBER, 1997) (CHO et al., 2014).

A escolha de utilizar tanto ARIMA quanto RNNs neste estudo se baseia em suas distintas vantagens e complementaridades. O modelo ARIMA é conhecido por sua simplicidade e interpretabilidade (HYNDMAN; ATHANASOPOULOS, 2018). Ele fornece resultados confiáveis para séries temporais lineares e é relativamente fácil de implementar e ajustar. Além disso, ARIMA é um método bem estabelecido com uma sólida base teórica, o que o torna uma escolha segura para muitas aplicações práticas (BOX et al., 2015).

As RNNs, por sua vez, oferecem uma abordagem mais flexível e poderosa para lidar com séries temporais não lineares e complexas (GOODFELLOW; BENGIO; COURVILLE, 2016). Com a capacidade de aprender padrões diretamente dos dados sem a necessidade de especificar explicitamente a estrutura do modelo, as RNNs podem capturar interações complexas e não lineares (SCHMIDHUBER, 2015). Esta capacidade é particularmente importante em contextos onde os dados apresentam comportamentos dinâmicos e variáveis que não são facilmente capturados por modelos lineares (ZHANG, 2003).

Estudos comparativos entre ARIMA e RNNs têm demonstrado que cada método pode superar o outro em diferentes contextos. Por exemplo, descobriram que as LSTMs superaram significativamente o ARIMA na previsão de séries temporais financeiras, reduzindo o erro de previsão em até 85% (SIAMI-NAMINI; TAVAKOLI; NAMIN, 2018). No entanto, outros estudos mostraram que a combinação de ARIMA e redes neurais pode produzir resultados superiores a qualquer um dos métodos individualmente (ZHANG, 2003).

Utilizar ambas as técnicas permite explorar e comparar suas capacidades preditivas em diferentes contextos. O ARIMA, com sua robustez em modelar padrões lineares, e as RNNs, com sua flexibilidade e poder para capturar dependências não lineares, oferecem um panorama abrangente das possibilidades de previsão (MAKRIDAKIS; SPILOTIS; ASSIMAKOPOULOS, 2018). A comparação entre esses métodos visa identificar não apenas qual modelo fornece previsões mais acuradas, mas também entender as condições sob as quais cada método é mais eficaz (AHMED et al., 2010).

Além disso, a crescente disponibilidade de dados e o aumento da capacidade computacional têm impulsionado o desenvolvimento e a aplicação de técnicas de aprendizado profundo, como as RNNs, em problemas de previsão de séries temporais (SCHMIDHUBER, 2015). Isso torna ainda mais relevante a comparação entre métodos tradicionais e técnicas de aprendizado profundo.

Portanto, a motivação para utilizar ARIMA e RNNs neste trabalho reside em suas reconhecidas vantagens e na oportunidade de avaliar suas performances de maneira comparativa. Esta abordagem permite uma análise abrangente e fundamentada das melhores práticas para a previsão de demanda energética, contribuindo para o desenvolvimento de metodologias mais eficientes e robustas para o setor energético (HONG; FAN, 2016).

2 Fundamentos Teóricos

2.1 Séries Temporais

Uma série temporal é definida como uma sequência de observações de uma variável coletadas em intervalos regulares ao longo do tempo (BOX et al., 2015). Matematicamente, podemos representar uma série temporal como:

$$X = X_t : t \in T \quad (2.1)$$

onde X_t é o valor observado no tempo t , e T é o conjunto de índices temporais. As séries temporais são frequentemente decompostas em componentes fundamentais: tendência, sazonalidade e ruído (HYNDMAN; ATHANASOPOULOS, 2018). Esta decomposição pode ser aditiva ou multiplicativa, dependendo da natureza dos dados:

$$X_t = T_t + S_t + R_t \quad (\text{Aditiva}) \quad (2.2)$$

$$X_t = T_t \times S_t \times R_t \quad (\text{Multiplicativa}) \quad (2.3)$$

onde T_t é a componente de tendência, S_t é a componente sazonal, e R_t é o componente de ruído.

A tendência (T_t) representa o comportamento de longo prazo da série. Pode ser modelada de diversas formas, incluindo:

- Linear: $T_t = \alpha + \beta t$
- Quadrática: $T_t = \alpha + \beta_1 t + \beta_2 t^2$
- Exponencial: $T_t = \alpha e^{\beta t}$

onde α , β , β_1 , e β_2 são parâmetros a serem estimados (CHATFIELD, 2016).

A sazonalidade (S_t) captura padrões cíclicos que se repetem em intervalos regulares. Pode ser modelada usando séries trigonométricas:

$$S_t = \sum_{k=1}^K \left[\gamma_k \sin\left(\frac{2\pi kt}{P}\right) + \delta_k \cos\left(\frac{2\pi kt}{P}\right) \right] \quad (2.4)$$

onde K é o número de harmônicos, P é o período sazonal, e γ_k e δ_k são coeficientes a serem estimados (BROCKWELL; DAVIS, 2016).

O ruído (R_t) representa as flutuações aleatórias não explicadas pelos outros componentes. Geralmente, assume-se que o ruído segue uma distribuição normal com média zero e variância constante:

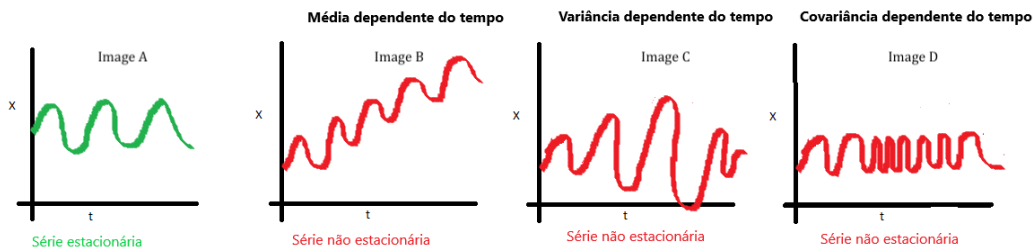
$$R_t \sim \mathcal{N}(0, \sigma^2) \quad (2.5)$$

onde σ^2 é a variância do ruído (WEI, 2006).

A estacionariedade é um conceito fundamental em séries temporais. Uma série é considerada estacionária se suas propriedades estatísticas (média, variância, autocorrelação) permanecem constantes ao longo do tempo (HAMILTON, 2020). Formalmente, para uma série estacionária:

$$\begin{aligned} E[X_t] &= \mu \quad (\text{constante}) \\ \text{Var}(X_t) &= \sigma^2 \quad (\text{constante}) \\ \text{Cov}(X_t, X_{t+k}) &= \gamma_k \quad (\text{depende apenas de } k) \end{aligned} \tag{2.6}$$

Figura 2 – Representação ilustrativa de diferentes tipos de séries temporais indicando a estacionariedade para cada caso.



Fonte: Médium (Medium, 2025)

Para verificar a estacionariedade, utiliza-se frequentemente o teste de Dickey-Fuller Aumentado (ADF) (DICKEY; FULLER, 1979). A hipótese nula do teste ADF é que a série tem uma raiz unitária (não é estacionária). O teste é baseado na seguinte regressão:

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \sum_{i=1}^p \delta_i \Delta X_{t-i} + \epsilon_t \tag{2.7}$$

onde $\Delta X_t = X_t - X_{t-1}$, α é uma constante, β é o coeficiente de uma tendência temporal, γ e δ_i são parâmetros a serem estimados, e ϵ_t é o termo de erro.

Vale introduzir o conceito de "lag" em séries temporais. Este conceito refere-se ao intervalo de tempo entre observações. Por exemplo, um lag de 1 compara cada observação com a observação imediatamente anterior, enquanto um lag de 2 compara com a observação dois períodos antes.

A função de autocorrelação (ACF) mede a correlação entre observações em diferentes lags (BOX et al., 2015):

$$\rho_k = \frac{\text{Cov}(X_t, X_{t+k})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+k})}} = \frac{\gamma_k}{\gamma_0} \tag{2.8}$$

onde γ_k é a autocovariância no lag k . A função de autocorrelação parcial (PACF) mede a correlação entre X_t e X_{t+k} após remover os efeitos das observações intermediárias (WEI, 2006):

$$\phi_{kk} = \text{Corr}(X_t, X_{t+k} | X_{t+1}, \dots, X_{t+k-1}) \tag{2.9}$$

A PACF pode ser calculada recursivamente usando as equações de Yule-Walker:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j} \quad (2.10)$$

2.2 ARIMA

O modelo ARIMA (AutoRegressive Integrated Moving Average) é uma técnica estatística amplamente utilizada para a análise e previsão de séries temporais (BOX et al., 2015). Este modelo combina três componentes principais: Autoregressão (AR), Média Móvel (MA) e Integração (I), proporcionando uma abordagem robusta para capturar padrões complexos em dados temporais (HYNDMAN; ATHANASOPOULOS, 2018). O componente autoregressivo do modelo ARIMA captura a dependência linear entre uma observação e suas observações passadas. Um processo AR de ordem p é definido pela seguinte equação:

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad (2.11)$$

onde X_t é o valor da série no tempo t , c é uma constante, ϕ_i são os parâmetros autoregressivos e ϵ_t é o termo de erro (ruído branco). Para facilitar a representação de modelos em séries temporais, utiliza-se o operador de defasagem (ou operador lag), denotado por B . Esse operador é definido como $BX_t = X_{t-1}$, ou seja, ele desloca a série em um período de tempo. Com isso, o operador de defasagem pode ser utilizado para expressar o modelo AR de forma mais compacta:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = c + \epsilon_t \quad (2.12)$$

ou, de forma ainda mais concisa:

$$\phi(B) X_t = c + \epsilon_t \quad (2.13)$$

onde $\phi(B)$ é o polinômio autoregressivo de ordem p (BROCKWELL; DAVIS, 2016). O componente de média móvel modela a dependência entre uma observação e os erros de previsão passados. Um processo MA de ordem q é definido como:

$$X_t = \mu + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad (2.14)$$

onde μ é a média do processo, θ_i são os parâmetros de média móvel e ϵ_t e ϵ_{t-i} são termos de erro. Utilizando o operador de defasagem, o modelo MA pode ser expresso como:

$$X_t = \mu + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \epsilon_t \quad (2.15)$$

ou, de forma mais compacta:

$$X_t = \mu + \theta(B)\epsilon_t \quad (2.16)$$

onde $\theta(B)$ é o polinômio de média móvel de ordem q (WEI, 2006).

O componente de integração lida com a não-estacionariedade da série temporal através da diferenciação. A diferenciação de ordem d é definida como:

$$\Delta^d X_t = (1 - B)^d X_t \quad (2.17)$$

Para $d = 1$, temos a diferenciação de primeira ordem:

$$\Delta X_t = X_t - X_{t-1} \quad (2.18)$$

Para $d = 2$, a diferenciação de segunda ordem:

$$\Delta^2 X_t = \Delta(\Delta X_t) = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} \quad (2.19)$$

A diferenciação transforma séries não-estacionárias em estacionárias, permitindo a aplicação dos modelos AR e MA (TSAY, 2005).

Combinando os componentes AR, I e MA, obtemos o modelo ARIMA(p, d, q), que pode ser expresso como:

$$\phi(B)(1 - B)^d X_t = c + \theta(B)\epsilon_t \quad (2.20)$$

onde p é a ordem do componente autoregressivo, d é a ordem de diferenciação e q é a ordem do componente de média móvel. Expandindo esta equação, obtemos:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d X_t = c + (1 + \theta_1 B + \dots + \theta_q B^q)\epsilon_t \quad (2.21)$$

Para séries com componentes sazonais, utiliza-se o modelo SARIMA (Seasonal ARIMA), que incorpora termos sazonais:

$$\Phi(B^s)\phi(B)(1 - B^s)^D(1 - B)^d X_t = c + \Theta(B^s)\theta(B)\epsilon_t \quad (2.22)$$

onde $\Phi(B^s)$ e $\Theta(B^s)$ são os polinômios sazonais AR e MA, respectivamente, s é o período sazonal e D é a ordem de diferenciação sazonal.

Variáveis exógenas são aquelas que não pertencem à série temporal que se deseja modelar, mas que influenciam diretamente o comportamento da série. Essas variáveis são fornecidas como entradas adicionais ao modelo e podem representar fatores externos, como temperatura, preços de commodities ou eventos sazonais, dependendo do contexto. No modelo SARIMAX

(Seasonal ARIMA with eXogenous variables), essas variáveis são representadas por Z_t , um vetor de variáveis exógenas, e sua contribuição é ponderada pelos coeficientes β , como descrito na equação:

$$\Phi(B^s)\phi(B)(1 - B^s)^D(1 - B)^d(X_t - \beta^T Z_t) = c + \Theta(B^s)\theta(B)\epsilon_t \quad (2.23)$$

permitindo que o modelo SARIMAX capture efeitos externos relevantes que podem melhorar a acurácia das previsões (HYNDMAN; ATHANASOPOULOS, 2018).

A estimação dos parâmetros do modelo ARIMA é geralmente realizada através do método de máxima verossimilhança, que busca encontrar os valores dos parâmetros que maximizam a função de log-verossimilhança, avaliando a adequação do modelo aos dados observados com base nos resíduos. A função de log-verossimilhança para um modelo ARIMA gaussiano é dada por:

$$\ell(\phi, \theta, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^n \epsilon_t^2 \quad (2.24)$$

onde n é o número de observações e σ^2 é a variância dos resíduos (BROCKWELL; DAVIS, 2016). Nesse contexto, os parâmetros ϕ , θ e σ^2 devem ser ajustados de forma a maximizar o valor de $\ell(\phi, \theta, \sigma^2)$.

O diagnóstico do modelo inclui a análise dos resíduos, que devem apresentar características de ruído branco. Testes como o de Ljung-Box, que avalia se as autocorrelações dos resíduos são estatisticamente significativas para detectar a presença de dependência temporal, são utilizados para verificar a autocorrelação dos resíduos:

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (2.25)$$

onde $\hat{\rho}_k$ é a autocorrelação amostral no lag k e h é o número de lags testados (LJUNG; BOX, 1978). A seleção do modelo pode ser realizada utilizando critérios de informação como o AIC (Akaike Information Criterion) ou o BIC (Bayesian Information Criterion):

$$AIC = -2 \log(L) + 2k \quad (2.26)$$

$$BIC = -2 \log(L) + k \log(n) \quad (2.27)$$

onde L é a verossimilhança maximizada do modelo e k é o número de parâmetros (AKAIKE, 1974).

Esta formulação matemática e conceitual dos modelos ARIMA, SARIMA e SARIMAX proporciona uma base sólida para a análise e previsão de séries temporais em diversos campos, incluindo a previsão de demanda de energia elétrica (TAYLOR; LETHAM, 2018).

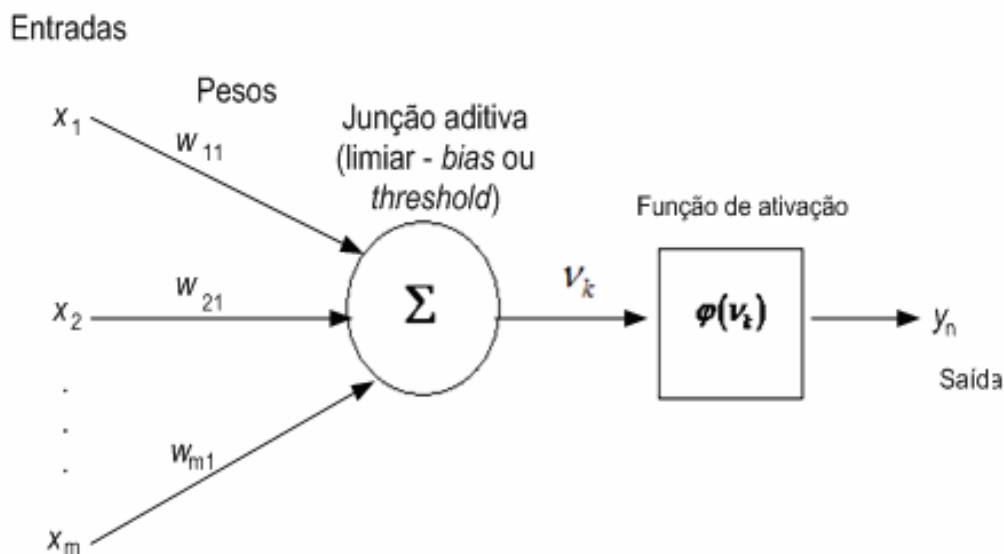
2.3 Redes Neurais Recorrentes (RNN)

As redes neurais artificiais são modelos computacionais inspirados no funcionamento do cérebro humano, compostos por camadas de neurônios artificiais interconectados que aprendem a realizar tarefas específicas a partir de exemplos (HAYKIN, 2009). A estrutura básica de uma rede neural artificial consiste em uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada camada é composta por um conjunto de nós (neurônios) que realizam operações matemáticas, sendo a saída de cada nó da camada anterior usada como entrada para os nós da próxima camada (GOODFELLOW; BENGIO; COURVILLE, 2016). A função que define a saída de um neurônio y pode ser representada pela Equação 2.28:

$$y = \phi(Wx + b) \quad (2.28)$$

Nesta equação, W é o vetor de pesos associado às entradas x , b é o *bias*, que é um termo adicional utilizado para deslocar a função de ativação, permitindo que o neurônio se ative mesmo quando as entradas x sejam iguais a zero, e ϕ é a função de ativação, que pode ser, por exemplo, a função sigmoide ou ReLU, dependendo do tipo de rede (LECUN; BENGIO; HINTON, 2015).

Figura 3 – Representação esquemática de uma rede neural.



Fonte: Wikipedia ([Wikipédia, a enciclopédia livre, 2025](#))

As Redes Neurais Recorrentes (RNNs) surgiram como uma evolução das redes neurais convencionais para lidar com dados sequenciais ou temporais (RUMELHART; HINTON; WILLIAMS, 1986). Enquanto as redes neurais tradicionais consideram apenas as entradas atuais, as RNNs possuem conexões cíclicas que permitem reter informações ao longo do tempo (ELMAN, 1990). A principal vantagem das RNNs é a habilidade de modelar sequências e capturar padrões temporais complexos, o que as torna ideais para aplicações como previsão de séries temporais, reconhecimento de voz e processamento de texto (GRAVES; MOHAMED; HINTON, 2013).

As RNNs conseguem processar informações sequenciais de forma mais eficiente que as redes neurais convencionais, justamente por manter um estado oculto que é atualizado a cada nova entrada (MIKOLOV et al., 2010).

A base de funcionamento das RNNs está na propagação de um estado oculto ao longo do tempo, o que pode ser descrito pela equação 2.29:

$$h_t = \phi(W_h h_{t-1} + W_x x_t + b_h) \quad (2.29)$$

Aqui, h_t representa o estado oculto no instante t , enquanto W_h e W_x são as matrizes de pesos responsáveis por ajustar o impacto do estado oculto anterior h_{t-1} e da entrada x_t , respectivamente. O termo b_h é o vetor de *bias*, que adiciona flexibilidade ao modelo, e ϕ é a função de ativação (PASCANU; MIKOLOV; BENGIO, 2013).

Entre as funções de ativação mais comuns estão a tangente hiperbólica (\tanh) e a ReLU (Unidade Linear Retificada). A tangente hiperbólica é dada pela Equação 2.30:

$$\phi(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.30)$$

Já a ReLU, que é amplamente utilizada em redes neurais devido à sua simplicidade e desempenho, pode ser expressa matematicamente pela Equação 2.31:

$$\phi(x) = \text{ReLU}(x) = \max(0, x) \quad (2.31)$$

Enquanto a tangente hiperbólica mapeia os valores de entrada para um intervalo entre -1 e 1, a ReLU retorna 0 para entradas negativas e o valor da própria entrada para valores positivos, introduzindo uma não-linearidade que ajuda a rede a aprender representações complexas dos dados (NAIR; HINTON, 2010).

Embora as RNNs convencionais sejam eficazes em diversas aplicações, elas apresentam dificuldades quando se trata de capturar dependências de longo prazo devido ao problema de desvanecimento do gradiente, que ocorre quando os gradientes das funções de perda se tornam extremamente pequenos durante a retropropagação, dificultando o ajuste eficaz dos pesos (HOCHREITER; SCHMIDHUBER, 1997). Para mitigar esse problema, uma variação poderosa chamada Long Short-Term Memory (LSTM) foi desenvolvida (HOCHREITER; SCHMIDHUBER, 1997). As LSTMs introduzem o conceito de células de memória, que mantêm informações por longos períodos de tempo. O comportamento de uma célula LSTM pode ser descrito por um conjunto de equações, começando com o cálculo da ativação dos vetores de entrada e esquecimento:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.32)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.33)$$

A função sigmoide σ , aplicada aqui, garante que os valores resultantes estejam entre 0 e 1, determinando a fração de informação que deve ser preservada ou esquecida (GERS; SCHMIDHUBER;

CUMMINS, 2000). Em seguida, a célula de memória é atualizada com uma nova estimativa, \tilde{c}_t , calculada pela função \tanh :

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.34)$$

O estado da célula c_t é então ajustado pela combinação dos vetores de esquecimento f_t e entrada i_t , multiplicados pelos estados anteriores e pelas novas informações, respectivamente:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (2.35)$$

onde \odot representa a multiplicação elemento a elemento (também conhecida como multiplicação Hadamard), realizada entre os valores correspondentes de dois vetores, preservando sua dimensionalidade.

A saída final da LSTM, h_t , é calculada com base no estado atualizado da célula e no vetor de saída o_t :

$$h_t = o_t \odot \tanh(c_t) \quad (2.36)$$

Embora as LSTMs sejam poderosas, outra variação mais simples, chamada Gated Recurrent Unit (GRU), também é amplamente utilizada(CHO et al., 2014). O GRU simplifica as operações ao combinar as funções das portas de *reset* e atualização. Suas equações incluem o cálculo do vetor de *reset* r_t e do vetor de atualização z_t :

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \quad (2.37)$$

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \quad (2.38)$$

Com o vetor de *reset* ajustado, o estado oculto proposto \tilde{h}_t é atualizado usando a função \tanh , que pondera o estado anterior e a nova entrada:

$$\tilde{h}_t = \tanh(W_h[r_t \odot h_{t-1}, x_t] + b_h) \quad (2.39)$$

Finalmente, o estado oculto atual h_t é calculado como uma interpolação entre o estado anterior h_{t-1} e o novo estado \tilde{h}_t , controlado pelo vetor de atualização z_t , vide Equação 2.40:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (2.40)$$

A principal diferença entre as LSTM e as GRU está na estrutura das operações e no número de portas. Na LSTM, há três portas: a porta de entrada i_t , que controla o quanto da nova informação será incorporada ao estado da célula; a porta de esquecimento f_t , que controla o quanto do estado anterior será mantido; e a porta de saída o_t , que determina a saída da célula. Além disso, a LSTM mantém uma célula de memória c_t , que armazena o estado da célula ao longo do tempo.

Já na GRU, as operações são simplificadas com apenas duas portas: a porta de *reset* r_t e a porta de atualização z_t . A porta de *reset* controla quanto do estado oculto anterior deve ser esquecido (semelhante à porta de esquecimento na LSTM), e a porta de atualização controla

a mistura entre o estado anterior e o novo estado proposto (\tilde{h}_t), semelhante à função da porta de entrada e saída combinadas da LSTM. Como a GRU não possui uma célula de memória separada como a LSTM, o estado oculto h_t diretamente reflete tanto a informação anterior quanto a nova informação ponderada pelas portas de atualização e *reset*. Isso torna o GRU mais simples e computacionalmente mais eficiente, com um número menor de parâmetros.

O treinamento de uma rede neural envolve conceitos importantes como épocas, dimensão latente e tamanho do lote (BENGIO, 2012). Uma época corresponde a uma varredura completa do conjunto de dados durante o treinamento. Se o conjunto de dados possui N exemplos e o lote processado contém B exemplos, o número de iterações por época será $\frac{N}{B}$. A dimensão latente refere-se ao número de unidades nas camadas ocultas, sendo um indicativo da capacidade de representação da rede (GOODFELLOW; BENGIO; COURVILLE, 2016). O tamanho do lote, por sua vez, é a quantidade de exemplos processados antes de cada atualização dos pesos. Assim, para um lote de tamanho B , a rede atualiza seus pesos $\frac{N}{B}$ vezes por época (BENGIO, 2012).

No treinamento de redes neurais, a função perda é um componente central, pois mede o erro entre as previsões do modelo e os valores reais (GOODFELLOW; BENGIO; COURVILLE, 2016). Para problemas de regressão, uma função amplamente utilizada é o Erro Quadrático Médio (MSE), definido pela Equação 2.41:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.41)$$

Nesta equação, y_i representa o valor real, enquanto \hat{y}_i é o valor previsto pelo modelo (BISHOP, 2006). O objetivo do treinamento é minimizar essa perda, ajustando os pesos da rede neural com o auxílio de otimizadores como o Gradiente Descendente (RUDER, 2016). Neste método, a atualização dos pesos é feita de acordo com a Equação 2.42:

$$\theta := \theta - \eta \nabla_{\theta} J(\theta) \quad (2.42)$$

Aqui, θ denota os pesos da rede, η é a taxa de aprendizado e $\nabla_{\theta} J(\theta)$ representa o gradiente da função perda em relação aos pesos. A operação $:=$ simboliza a atualização dos pesos, orientada pelo gradiente, de forma a reduzir a função de custo ao longo do processo de treinamento (RUDER, 2016). Esses conceitos e equações formam a base teórica para o funcionamento de redes neurais recorrentes, sendo essenciais para modelar e prever padrões temporais em diversas aplicações, incluindo a previsão de demanda de energia elétrica, tema central deste trabalho (ZHANG; WANG; ZHANG, 2018).

2.4 Métricas de Avaliação

A avaliação precisa de modelos de previsão de séries temporais é fundamental para garantir a confiabilidade e a aplicabilidade das previsões em diversos campos, como economia, meteorologia e engenharia (HYNDMAN; ATHANASOPOULOS, 2018). Esta seção apresenta

uma análise aprofundada das principais métricas utilizadas para avaliar o desempenho desses modelos, com ênfase no MAPE (Mean Absolute Percentage Error) e suas implicações.

O desenvolvimento de métricas de avaliação para modelos de séries temporais tem sido um processo contínuo ao longo das últimas décadas. Inicialmente, medidas simples como o erro médio eram utilizadas, mas logo se percebeu a necessidade de métricas mais sofisticadas que pudessem capturar nuances específicas dos dados temporais (MAKRIDAKIS, 1993). Na década de 1970, houve um avanço significativo com a introdução de métricas como o MAPE e o RMSE, que se tornaram padrões na indústria e na academia (ARMSTRONG; COLLOPY, 1992). Desde então, pesquisadores têm proposto várias outras métricas, cada uma com suas vantagens e limitações específicas.

O Erro Quadrático Médio (MSE) é uma das métricas mais básicas e amplamente utilizadas. É calculado como:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.43)$$

onde y_i é o valor observado e \hat{y}_i é o valor previsto. O MSE penaliza erros maiores mais severamente devido à sua natureza quadrática (CHAI; DRAXLER, 2014).

O RMSE é a raiz quadrada do MSE:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.44)$$

Esta métrica é preferida por muitos pesquisadores porque está na mesma unidade que os dados originais, facilitando a interpretação (WILLMOTT; MATSUURA, 2005).

O Erro Absoluto Médio (MAE) é calculado como:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.45)$$

O MAE é menos sensível a outliers comparado ao RMSE, fornecendo uma medida mais robusta do erro médio (CHAI; DRAXLER, 2014).

O MAPE é uma métrica que merece atenção especial devido à sua ampla utilização e características únicas. É calculado como:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (2.46)$$

O MAPE (Erro Percentual Absoluto Médio) apresenta diversas vantagens e desvantagens que merecem atenção. Entre os pontos positivos, destaca-se sua fácil interpretação, pois expressa o erro em termos percentuais, tornando-o intuitivo para a compreensão. Além disso, o MAPE permite comparações entre séries de diferentes escalas, o que é particularmente útil em análises comparativas. Outra vantagem é a normalização do erro em relação à magnitude dos valores observados, proporcionando uma perspectiva mais equilibrada (MYTTENAERE et al., 2016).

Por outro lado, o MAPE também possui algumas limitações. Uma desvantagem significativa é que ele se torna indefinido para valores observados iguais a zero, podendo levar a uma

divisão por zero. Além disso, o MAPE apresenta uma assimetria, tendendo a penalizar previsões acima do valor real mais do que previsões abaixo, o que pode distorcer a interpretação dos resultados. Por fim, o MAPE é sensível a outliers, podendo ser distorcido por valores extremos, o que requer cuidado na análise de dados com variações significativas (MYTTENAERE et al., 2016).

A avaliação de modelos de previsão de séries temporais é um campo em constante evolução, com novas métricas e metodologias sendo desenvolvidas para abordar desafios específicos. O MAPE, apesar de suas limitações, continua sendo uma métrica valiosa devido à sua interpretabilidade e ampla aceitação. No caso específico deste trabalho, a escolha do MAPE como métrica de avaliação se deve à ausência de valores nulos, valores próximos de zero e outliers na série temporal utilizada. Essas características tornam o MAPE especialmente adequado, uma vez que eliminam os principais fatores que poderiam distorcer seus resultados, garantindo uma análise confiável e intuitiva da precisão do modelo (MYTTENAERE et al., 2016).

3 Metodologia

3.1 Base de Dados

A previsão acurada da demanda de energia elétrica é fundamental para o desenvolvimento econômico e social, permitindo a otimização dos recursos energéticos e garantindo a confiabilidade no fornecimento(BHATTACHARYYA, 2011). Este estudo utiliza dados históricos de demanda energética para construir modelos preditivos que auxiliem na gestão eficiente do sistema elétrico brasileiro(SUGANTHI; SAMUEL, 2012). A importância deste tema se estende além da simples previsão, abrangendo aspectos cruciais como planejamento de infraestrutura, políticas energéticas e sustentabilidade ambiental(KAYTEZ et al., 2015). Uma previsão acurada pode resultar em economias significativas, redução de emissões de carbono e melhor alocação de recursos(HONG; FAN, 2016). Todo o código e procedimento realizado está disponível no github: <https://github.com/vitorsbarboza/Energy-Demand—Time-Series-Forecasting>

Os dados utilizados nesta pesquisa foram obtidos do Operador Nacional do Sistema Elétrico (ONS), entidade responsável pela operação do sistema elétrico no Brasil(Operador Nacional do Sistema Elétrico (ONS), 2024). Para facilitar o acesso e manipulação, optou-se por uma versão organizada disponível na plataforma Kaggle, intitulada *23 Years of Hourly Electric Energy Demand (Brazil)*(KAGGLE, 2023). A escolha desta fonte se justifica por sua conveniência no acesso e integração com ferramentas de análise, pela organização dos dados em formato adequado para análises estatísticas e pela abrangência temporal, cobrindo um período significativo de 23 anos. A utilização de dados públicos e bem estruturados como estes promove a reprodutibilidade da pesquisa e facilita comparações com outros estudos no campo da previsão de demanda energética(HYNDMAN; ATHANASOPOULOS, 2018).

O conjunto de dados apresenta características importantes, abrangendo o período de 01/01/2000 a 31/12/2022 (23 anos), com resolução horária e unidade de medida em Megawatts (MW), disponibilizado em formato CSV (Comma-Separated Values). Esta granularidade temporal permite análises detalhadas de padrões diários, semanais e sazonais, essenciais para a compreensão do comportamento da demanda energética(TAYLOR; LETHAM, 2018).

Para o processamento dos dados, utilizou-se a linguagem Python, amplamente adotada na ciência de dados(MCKINNEY, 2012). A biblioteca *pandas* foi empregada para carregar e manipular os dados(MCKINNEY et al., 2010).

```
import pandas as pd
Carregar os dados
df = pd.read_csv('demanda_energia.csv')
```

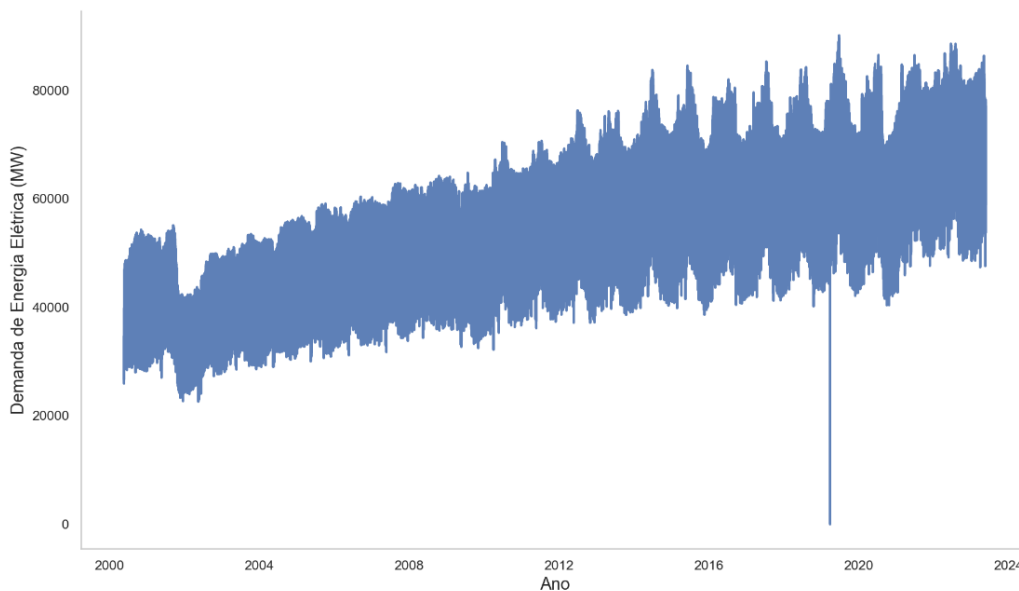
A escolha do Python e do *pandas* se justifica pela eficiência no tratamento de grandes volumes

de dados e pela vasta comunidade de desenvolvedores, o que facilita a resolução de problemas e a implementação de técnicas avançadas de análise(VANDERPLAS, 2016).

A análise inicial focou em compreender as principais características da série temporal(BOX et al., 2015). Observou-se um aumento consistente no consumo de energia ao longo dos anos (veja Fig.4), refletindo o crescimento econômico e populacional do país, caracterizando assim a tendência da série (veja Fig.6a). Além disso, identificaram-se variações cíclicas anuais, possivelmente relacionadas a fatores climáticos e padrões de consumo sazonais, evidenciando a sazonalidade presente nos dados (veja Fig.6b). Por fim, notaram-se flutuações imprevisíveis no consumo de energia, que podem ser atribuídas a eventos pontuais ou variações aleatórias, representando o componente de ruído da série temporal (veja Fig.6c).

A compreensão destes componentes é importante para a seleção e ajuste de modelos preditivos adequados(HYNDMAN; ATHANASOPOULOS, 2018).

Figura 4 – Série temporal da demanda de energia no Brasil, em MW (hora em hora).



Fonte: Próprio Autor

Identificaram-se valores zero no registro de consumo, o que é irrealista para demanda de eletricidade(HYNDMAN; ATHANASOPOULOS, 2018). Para corrigir:

1. Valores zero foram convertidos em nulos.
2. Utilizou-se interpolação linear para preencher pontos faltantes (até oito valores consecutivos).

Converter zeros para NaN

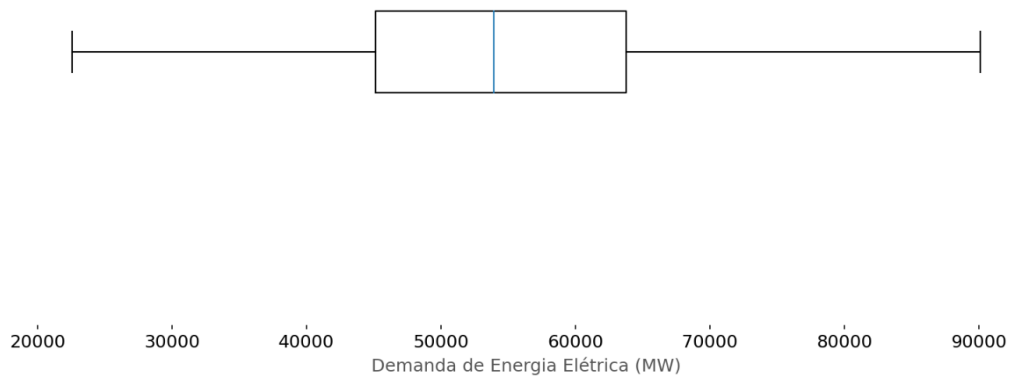
```
df['demanda'] = df['demanda'].replace(0, np.nan)
```

Interpolação linear

```
df['demanda'] = df['demanda'].interpolate(method='linear', limit=8)
```

Este processo de limpeza é fundamental para garantir a integridade dos dados e evitar distorções nas análises subsequentes(GUO et al., 2016). A distribuição dos valores após o tratamento é ilustrada na Figura 5.

Figura 5 – Boxplot da distribuição da demanda de energia elétrica, em MW (hora em hora).



Fonte: Próprio Autor

Para reduzir a demanda computacional e facilitar análises de longo prazo, os dados horários foram agregados em dados diários(TAYLOR; LETHAM, 2018):

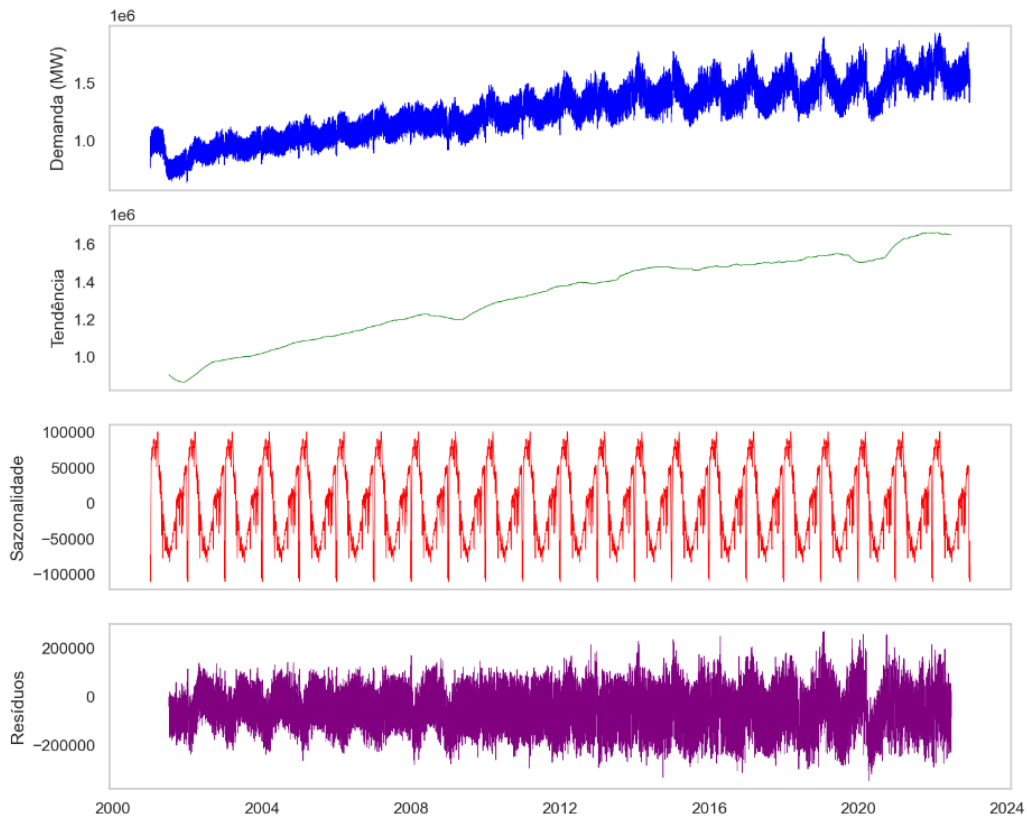
Agregação diária

```
df_daily = df.resample('D').sum()
```

Esta agregação permite uma visão mais ampla das tendências de consumo, mantendo informações suficientes para análises significativas(CLEVELAND et al., 1990).

A decomposição da série temporal diária em suas componentes (tendência, sazonalidade e resíduos) foi realizada para uma compreensão mais profunda do comportamento dos dados(CLEVELAND et al., 1990), como a Fig.6 mostra.

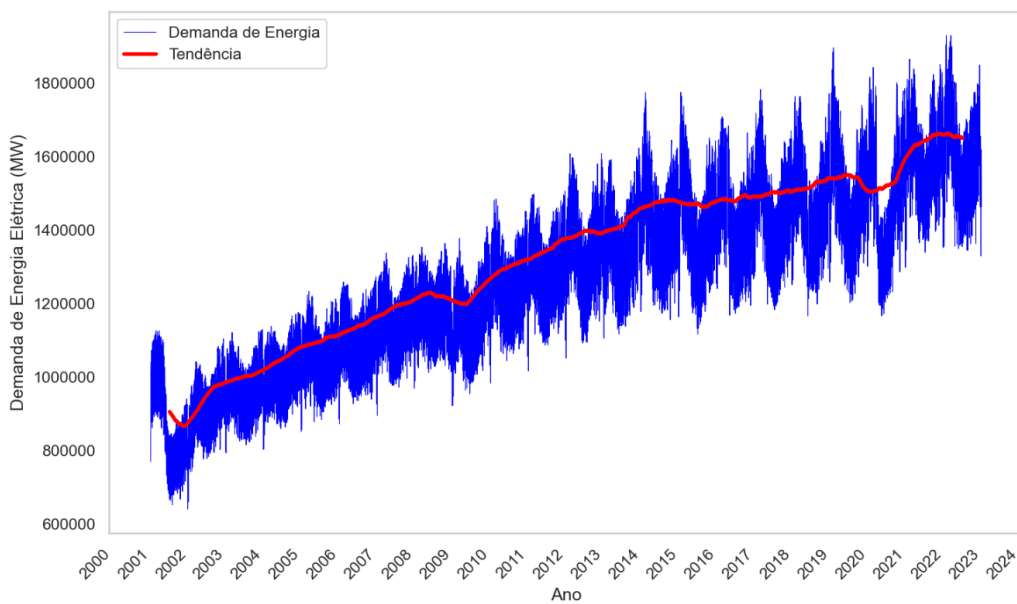
Figura 6 – Decomposição da série temporal de demanda diária de energia elétrica (MW) em tendência, sazonalidade e resíduos.



Fonte: Próprio Autor

Para uma análise mais clara do comportamento ao longo do tempo, gerou-se um gráfico da série temporal diária com a curva de tendência (Figura 7).

Figura 7 – Série temporal diária juntamente com a tendência.



Fonte: Próprio Autor

O gráfico evidencia o crescimento acentuado do consumo de energia elétrica no Brasil, especialmente na última década do período analisado. Este aumento pode ser atribuído a diversos fatores, como crescimento populacional, aumento da industrialização e desenvolvimento econômico do país(WOLDE-RUFAEL, 2015).

A base de dados utilizada neste estudo oferece uma visão abrangente da demanda de energia elétrica no Brasil ao longo de 23 anos. O tratamento e a análise realizados proporcionam uma base para a construção de modelos preditivos, capazes de capturar as nuances da demanda energética e contribuir para uma gestão mais eficiente do sistema elétrico brasileiro(HONG; FAN, 2016).

3.2 Modelagem ARIMA

A modelagem ARIMA (Autoregressive Integrated Moving Average) é uma técnica estatística amplamente utilizada para análise e previsão de séries temporais (BOX et al., 2015). Esta seção detalha a metodologia empregada na aplicação do modelo ARIMA para a previsão da demanda de energia elétrica, abordando desde a preparação dos dados até a validação do modelo.

A etapa inicial de qualquer análise de séries temporais é a preparação adequada dos dados. No contexto deste estudo, os dados de demanda de energia elétrica abrangem o período de 2000 a 2022, totalizando 8401 registros diários(HYNDMAN; ATHANASOPOULOS, 2018).

Neste estudo, adotou-se uma estratégia de separação temporal, que é particularmente adequada para séries temporais. O conjunto de dados foi dividido da seguinte forma: o conjunto de treino abrange os dados de 2000 a 2021, totalizando 8036 registros. Este conjunto mais extenso permite que o modelo aprenda os padrões e tendências de longo prazo presentes nos dados. Já o conjunto de teste é composto pelos dados de 2022, com 365 registros, representando um ano de observações.

Esta divisão temporal permite que o modelo aprenda com uma extensa base histórica e seja testado com dados recentes, simulando um cenário real de previsão. A escolha de um ano completo para o conjunto de teste é particularmente vantajosa, pois permite avaliar o desempenho do modelo ao longo de todas as estações do ano, capturando possíveis variações sazonais na demanda de energia (WERON, 2014).

A normalização é um passo essencial no pré-processamento de dados para modelos de aprendizado de máquina, pois evita que variáveis com magnitudes diferentes influenciem desproporcionalmente os cálculos de aprendizado (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). No contexto deste estudo, optou-se pela técnica MinMaxScaler, que transforma os dados para um intervalo entre 0 e 1. Esta escolha é justificada por preservar a distribuição original dos dados, manter as relações entre os valores extremos e facilitar a interpretação dos resultados. A utilização dessa técnica permite que os dados sejam ajustados de forma uniforme, garantindo que todas as variáveis tenham a mesma escala e não haja distorções na análise devido a diferen-

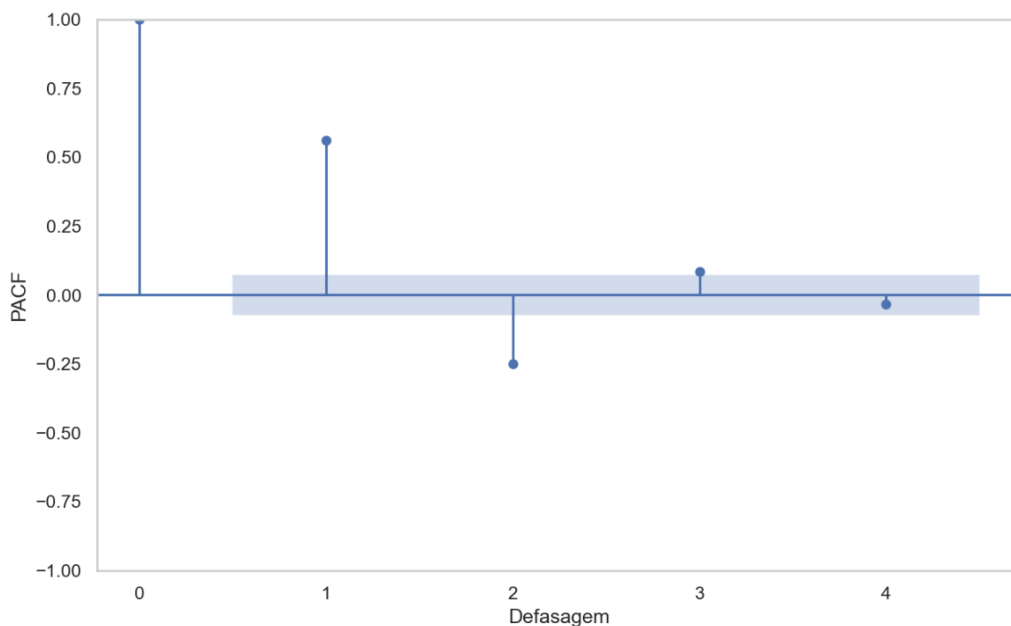
ças de magnitude entre as características. Além disso, ao manter a estrutura relativa dos dados, o MinMaxScaler possibilita uma comparação mais precisa entre as variáveis, contribuindo para uma modelagem mais eficiente e resultados mais confiáveis.

A normalização foi aplicada primeiro ao conjunto de treino e, posteriormente, a mesma transformação foi utilizada no conjunto de teste, garantindo consistência entre os conjuntos (KUHN; JOHNSON, 2013).

O modelo ARIMA é caracterizado por três parâmetros principais: p , d e q (BOX et al., 2015). A determinação destes parâmetros é essencial para o desempenho do modelo.

O parâmetro p representa a ordem da parte autorregressiva (AR) do modelo. Para sua determinação, utilizou-se a Função de Autocorrelação Parcial (PACF) (TSAY, 2005). A PACF revela a correlação entre um ponto da série e seus valores defasados, eliminando o efeito das defasagens intermediárias. A análise da PACF (Figura 8) indicou que, a partir da terceira defasagem, os valores não são mais estatisticamente significativos. Portanto, definiu-se $p = 3$.

Figura 8 – Função de Autocorrelação Parcial da Série Temporal da Demanda de Energia.



Fonte: Próprio Autor

Segue a reescrita do texto:

O parâmetro d representa o número de diferenciações aplicadas à série temporal para torná-la estacionária, uma condição essencial para a aplicação do modelo ARIMA (BROCKWELL; DAVIS, 2016). No presente trabalho, foi adotado $d = 1$, o que corresponde à aplicação de uma diferenciação de primeira ordem. Essa escolha se fundamenta na premissa de que, em muitos casos práticos, uma única diferenciação é suficiente para eliminar tendências ou flutuações sazonais, garantindo a estacionariedade da série. Embora não tenham sido realizados testes específicos para validar essa decisão, ela se justifica pela observação da série e pela busca de

uma configuração simples e eficaz para atender às exigências do modelo.

O parâmetro q refere-se à ordem da parte de médias móveis (MA) do modelo. Para sua determinação, testaram-se valores de 0 a 5, avaliando o desempenho do modelo através do erro percentual médio absoluto (MAPE). O menor MAPE foi obtido com $q = 0$.

Considerando a natureza sazonal da demanda de energia elétrica (TAYLOR; LETHAM, 2018), estendeu-se o modelo ARIMA para SARIMAX (Seasonal ARIMA with eXogenous factors). Esta extensão incorpora componentes sazonais ao modelo, permitindo capturar padrões cíclicos na série temporal. Os parâmetros sazonais P , D , Q , e m foram determinados através de experimentações, resultando na configuração `seasonal_order = (1,1,0,28)`. Esta configuração indica uma sazonalidade mensal com período de 28 dias, alinhando-se com os ciclos típicos de consumo de energia (HONG; FAN, 2016).

A implementação do modelo foi realizada utilizando a biblioteca `statsmodels` em Python (SEABOLD; PERKTOLD, 2010). Especificamente, utilizou-se o método SARIMAX, que permite modelar tanto componentes ARIMA quanto sazonais. A configuração final dos parâmetros foi:

- `order = (3, 1, 0) # p, d, q`
- `seasonal_order = (1, 1, 0, 28) # P, D, Q, m`

Para validação do modelo, adotou-se a técnica de "*walk forward validation*" (HYNDMAN; ATHANASOPOULOS, 2018). Esta abordagem é particularmente adequada para séries temporais, pois simula um cenário de previsão em tempo real. O processo envolve o treinamento do modelo com um subconjunto inicial de dados, seguido pela previsão para um ponto futuro. Após essa etapa, o ponto previsto é incorporado aos dados de treinamento, e o processo é repetido para o próximo ponto de previsão. Essa metodologia permite uma avaliação mais robusta do desempenho do modelo em condições que se aproximam da realidade, onde novas informações são continuamente adicionadas à base de dados. Além disso, a "*walk forward validation*" ajuda a identificar possíveis problemas de overfitting e a avaliar a capacidade do modelo de se adaptar a mudanças nas tendências dos dados ao longo do tempo.

Embora computacionalmente intensiva, esta técnica proporciona uma avaliação robusta da capacidade preditiva do modelo ao longo do tempo, considerando a evolução dinâmica da série temporal (BERGMEIR; HYNDMAN; KOO, 2018).

A metodologia ARIMA/SARIMAX, ao capturar padrões complexos e sazonalidades na demanda de energia, oferece uma ferramenta poderosa para enfrentar estes desafios. Sua aplicação não se limita apenas ao setor energético, mas estende-se a diversos campos que lidam com séries temporais, como economia, meteorologia e epidemiologia (BOX et al., 2015). Em suma, esta abordagem metodológica rigorosa, combinando técnicas estatísticas avançadas com validação robusta, proporciona uma base sólida para análises e previsões confiáveis no complexo cenário da demanda de energia elétrica.

3.3 Redes Neurais Recorrentes (RNNs)

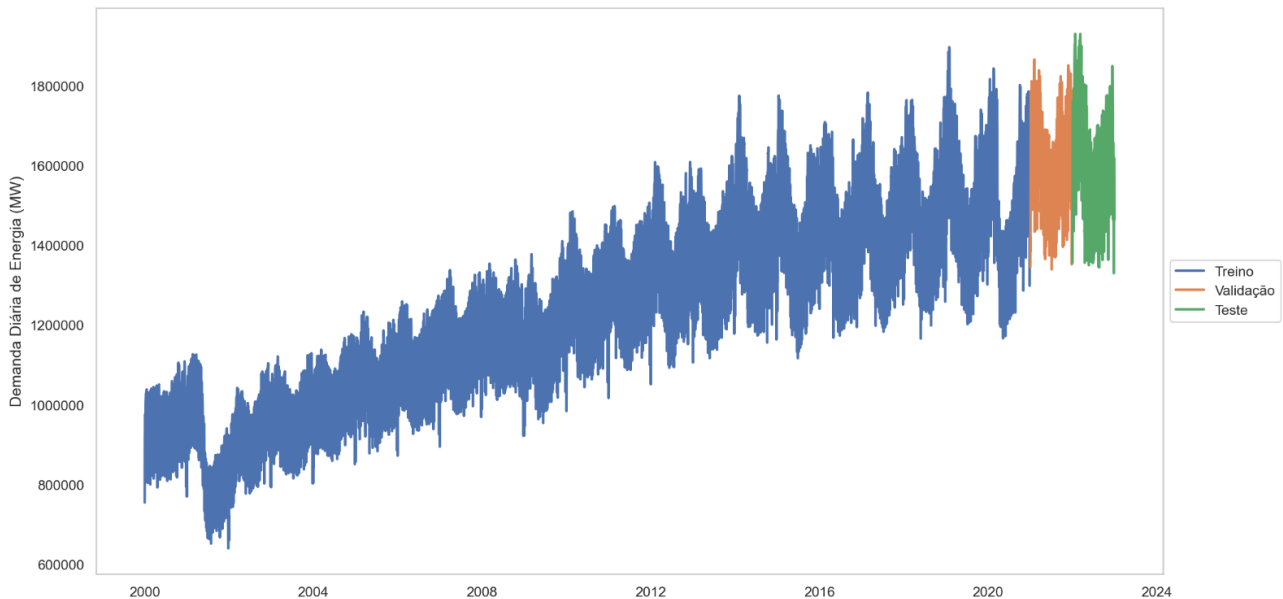
As Redes Neurais Recorrentes (RNNs) têm se destacado como uma ferramenta poderosa no campo da aprendizagem de máquina, especialmente quando se trata de lidar com dados sequenciais e temporais (GOODFELLOW; BENGIO; COURVILLE, 2016). Sua capacidade de manter uma "memória" de informações passadas as torna particularmente adequadas para tarefas como previsão de séries temporais, processamento de linguagem natural e análise de dados sequenciais (LIPTON; BERKOWITZ; ELKAN, 2015).

A preparação dos dados é uma etapa fundamental no desenvolvimento de modelos de aprendizagem de máquina, e no caso das RNNs, isso não é diferente. O processo de preparação dos dados para este estudo foi meticulosamente planejado para garantir a robustez e confiabilidade do modelo final.

Os dados foram estrategicamente divididos em três subconjuntos: treino, validação e teste. Esta abordagem é fundamental para avaliar adequadamente o desempenho do modelo e evitar problemas como overfitting (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). O Conjunto de Treino abrangeu o período de 2000 a 2020, fornecendo uma base sólida de informações históricas. Este extenso período foi escolhido para permitir que o modelo capture padrões e tendências de longo prazo nos dados (HYNDMAN; ATHANASOPOULOS, 2018). O Conjunto de Validação correspondeu ao ano de 2021. Este conjunto é decisivo para o ajuste fino dos hiperparâmetros do modelo, permitindo otimizar seu desempenho sem incorrer em overfitting (BISHOP, 2006). Por fim, definiu-se o conjunto de teste como sendo o ano de 2022, representando o período mais recente disponível na base de dados. Este conjunto serve para avaliar o desempenho final do modelo em dados completamente novos, simulando um cenário de aplicação real (JAMES et al., 2013). Esta divisão estratégica dos dados permite uma avaliação robusta e confiável do modelo, garantindo sua capacidade de generalização e eficácia em situações reais de previsão.

A representação visual desta divisão pode ser observada na Figura 9, que ilustra claramente a separação temporal dos conjuntos de dados.

Figura 9 – Representação da separação do conjunto de dados em treino, validação e teste.



Fonte: Próprio Autor

Assim como descrito na seção anterior, foi necessária a realização da normalização dos dados para este processo. O mesmo método de normalização Min-Max, implementado com a função `MinMaxScaler`, foi empregado, garantindo a uniformidade do pré-processamento ao longo do estudo.

Para capturar efetivamente a natureza temporal dos dados de demanda de energia, foi implementada uma abordagem de features lag. Especificamente, foram utilizadas 20 features lag, o que significa que o modelo considera as 20 observações anteriores para fazer uma previsão. Esta escolha foi baseada em um equilíbrio entre fornecer informações suficientes sobre o passado recente e evitar a inclusão de dados excessivos que poderiam prejudicar o desempenho do modelo (BOX et al., 2015).

A remodelagem dos dados para um formato tridimensional é um passo para adequá-los à entrada da rede neural recorrente. Utilizou-se a função `reshape` para criar uma estrutura de dados com as dimensões: (número de amostras, 20 features lag, 1 dimensão de saída). As features lag, também conhecidas como variáveis defasadas, são valores de uma série temporal em momentos anteriores, utilizados como preditores para o valor futuro. Esta organização permite que a RNN processe eficientemente as sequências temporais (CHOLLET, 2018).

A configuração e o treinamento da rede neural são etapas críticas que determinam a capacidade do modelo de aprender e generalizar a partir dos dados temporais. Neste estudo, foi implementada uma arquitetura cuidadosamente projetada para lidar com séries temporais de demanda de energia.

A rede neural consiste em uma camada GRU (Gated Recurrent Unit) com 30 unidades latentes, seguida por uma camada densa final. As GRUs são uma variante eficiente das RNNs,

capazes de capturar dependências de longo prazo nos dados(CHO et al., 2014). Esta estrutura permite que o modelo aprenda padrões complexos e relações temporais nos dados de demanda de energia.

O processo de treinamento foi configurado com 50 épocas, um número escolhido para permitir um aprendizado suficiente, mas evitando o risco de overfitting(GOODFELLOW; BENGIO; COURVILLE, 2016). O tamanho do batch foi definido como 200 amostras, buscando um equilíbrio entre a estabilidade do treinamento e a velocidade de atualização dos parâmetros(MASTERS; LUSCHI, 2018). Além disso, foram utilizadas 20 features lag como entradas, representando os valores passados mais relevantes para prever a demanda de energia.

Foram testadas algumas configurações distintas de parâmetros para determinar os valores que melhor se adequassem ao problema. Cada configuração foi avaliada com base na métrica MAPE (Mean Absolute Percentage Error), e os valores finais foram escolhidos por terem proporcionado o menor MAPE, indicando maior precisão preditiva do modelo. A escolha inicial dos intervalos testados para os parâmetros foi orientada por recomendações da literatura, como (LAZZERI, 2020).

Para otimização, foi escolhido o algoritmo RMSprop, conhecido por sua eficácia em problemas de séries temporais devido à sua capacidade de adaptar as taxas de aprendizado(TIELEMAN; HINTON, 2012). A função de perda utilizada foi o Erro Quadrático Médio (MSE), que mede a diferença quadrática entre as previsões e os valores reais(BISHOP, 2006).

Uma técnica importante implementada para prevenir o overfitting foi o early stopping. Este método monitora o desempenho no conjunto de validação e interrompe o treinamento se não houver melhoria após um número especificado de épocas (neste caso, 5 épocas)(PRECHELT, 1998).

O treinamento do modelo segue um processo iterativo onde o modelo é alimentado com os dados de treinamento, ajustando seus parâmetros para minimizar a função de perda. O desempenho é avaliado no conjunto de validação após cada época, e o early stopping é aplicado se não houver melhoria no desempenho de validação. Após o treinamento, o modelo é testado no conjunto de teste para avaliar sua capacidade de generalização.

Esta metodologia rigorosa garante que o modelo RNN seja capaz de capturar eficientemente as dependências temporais nos dados de demanda de energia, fornecendo previsões acuradas e confiáveis. A combinação de uma arquitetura adequada, técnicas de otimização eficientes e métodos de regularização robustos resulta em um modelo capaz de aprender padrões complexos nos dados e fazer previsões acuradas sobre a demanda futura de energia.

4 Resultados

4.1 Avaliação do Modelo e Resultados da Previsão da modelagem ARIMA

A análise e previsão de séries temporais desempenham um papel em diversos campos, desde a economia até a engenharia, e são particularmente relevantes no setor energético (BOX *et al.*, 2015). No contexto deste estudo, focamos na aplicação do modelo ARIMA (Autoregressive Integrated Moving Average) para prever a demanda de energia, um desafio complexo e de grande importância para o planejamento e gestão eficiente dos recursos energéticos (HYNDMAN; ATHANASOPOULOS, 2018).

A avaliação da performance do modelo é realizada primariamente através do cálculo do MAPE (Mean Absolute Percentage Error), uma métrica amplamente reconhecida e utilizada na análise de séries temporais (MYTTENAERE *et al.*, 2016). O MAPE oferece diversas vantagens que o tornam particularmente adequado para este tipo de análise:

No caso do modelo ARIMA desenvolvido neste estudo, o MAPE obtido foi de 1,97%. Este resultado é notável e merece uma análise mais detalhada. Um MAPE de 1,97% indica que, em média, as previsões do modelo desviaram apenas 1,97% dos valores reais. Isto se traduz em uma precisão de 98,03%, um desempenho excepcionalmente alto para previsões de demanda de energia (HONG; FAN, 2016). Na literatura, modelos considerados bem-sucedidos para previsão de demanda de energia frequentemente apresentam MAPEs na faixa de 2-5% (SUGANTHI; SAMUEL, 2012). Portanto, nosso modelo se destaca positivamente neste aspecto. Uma precisão tão elevada pode resultar em significativas economias e melhorias na eficiência operacional para as empresas do setor energético (WERON, 2014).

Este nível de precisão permite que as empresas de energia façam planejamentos mais acurados, otimizem a distribuição de recursos e minimizem desperdícios, o que pode levar a uma redução de custos operacionais e a um melhor atendimento à demanda dos consumidores. Além disso, a alta precisão do modelo pode contribuir para uma gestão mais eficiente da rede elétrica, auxiliando na prevenção de sobrecargas e na manutenção da estabilidade do sistema.

É importante notar que este nível de precisão foi alcançado apesar da complexidade inerente às séries temporais de demanda de energia, que são influenciadas por múltiplos fatores como condições climáticas, padrões de consumo e eventos sazonais (TAYLOR; LETHAM, 2018).

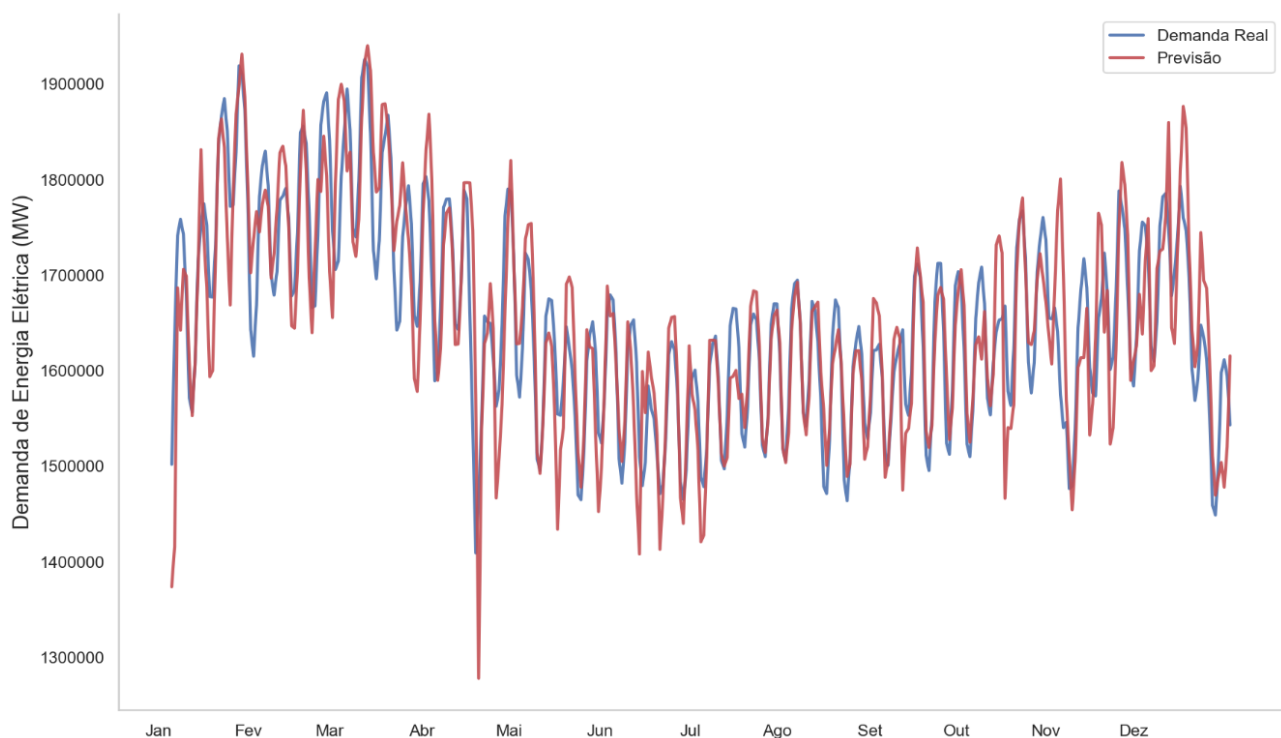
Um aspecto relevante a ser considerado na aplicação de modelos de séries temporais é o custo computacional associado. No presente estudo, o processo completo de treinamento, validação e realização das previsões demandou entre 1h30 e 2 horas. Este tempo de processamento relativamente longo reflete a complexidade do modelo e a extensão da série temporal

analisada (BOX et al., 2015). Os experimentos foram realizados em um computador com processador Intel(R) Core(TM) i5-4210U CPU @ 1.70GHz (frequência máxima de 2.40 GHz), 8 GB de memória RAM e sistema operacional Windows 10 Pro. O custo computacional elevado ressalta alguns pontos importantes. Primeiramente, a complexidade do modelo ARIMA, especialmente quando incorpora componentes sazonais (SARIMA), pode ser computacionalmente intensiva devido à necessidade de estimar múltiplos parâmetros (HYNDMAN; ATHANASOPOULOS, 2018). Além disso, as séries temporais de demanda de energia tipicamente envolvem um grande volume de dados, especialmente quando se trabalha com granularidade horária ou sub-horária (HONG; FAN, 2016). Por fim, o alto custo computacional é frequentemente o preço a se pagar por uma maior precisão nas previsões (PETROPOULOS et al., 2020). Este trade-off entre precisão e tempo de processamento é uma consideração importante na implementação prática de modelos de previsão de demanda de energia. Embora o tempo de processamento possa parecer significativo, é importante contextualizar esse custo em relação aos benefícios obtidos em termos de precisão das previsões e ao valor que essas previsões acuradas podem proporcionar na gestão eficiente de sistemas de energia.

Este aspecto do estudo sublinha a importância de considerar não apenas a precisão, mas também a eficiência computacional ao selecionar e implementar modelos de previsão em contextos práticos (HYNDMAN; ATHANASOPOULOS, 2018).

A Figura 10 apresenta uma comparação visual entre os valores reais da demanda de energia e as previsões geradas pelo modelo ARIMA. Esta representação gráfica é fundamental para uma compreensão intuitiva do desempenho do modelo (TUFTE, 2001).

Figura 10 – Previsão do modelo ARIMA vs. valores reais da demanda diária de energia elétrica nos dados de teste (ano de 2022).



Fonte: Próprio Autor

Analisando o gráfico, podemos observar vários aspectos importantes que demonstram o desempenho do modelo. O ajuste geral é notável, com a proximidade entre as curvas de valores reais e previstos indicando um excelente ajuste do modelo à série temporal (CHATFIELD, 2000). Esta correspondência estreita sugere que o modelo foi capaz de capturar com precisão as características fundamentais da demanda de energia ao longo do tempo.

Além disso, o modelo parece capturar eficazmente tanto as tendências gerais quanto os padrões sazonais presentes na série (BOX et al., 2015). Isso é particularmente importante no contexto da demanda de energia, onde fatores sazonais como variações climáticas e padrões de consumo podem ter um impacto significativo. A capacidade do modelo de refletir essas nuances demonstra sua sofisticação e adequação para a tarefa em questão.

A consistência do ajuste ao longo de todo o período de teste indica que o modelo é robusto e não está sofrendo de overfitting (HYNDMAN; ATHANASOPOULOS, 2018). Em um modelo que estivesse sofrendo overfitting, o que se observaria seria uma performance muito boa nos dados de treinamento, mas uma queda significativa na acurácia nos dados de teste. Isso geraria um gráfico onde a linha de erro para o conjunto de treinamento seria muito baixa, enquanto a linha de erro para o conjunto de teste apresentaria um aumento considerável. A robustez observada, portanto, indica que o modelo não apenas se ajusta bem aos dados de treinamento, mas também consegue generalizar adequadamente para novos dados. Essa característica garante que o modelo seja eficaz e útil na previsão de demanda de energia em cenários

reais.

A capacidade do modelo em acompanhar de perto as variações na demanda de energia é particularmente notável, considerando a complexidade típica dessas séries temporais (WERON, 2014). Isso pode ser atribuído, em parte, à inclusão do componente sazonal no modelo SARIMA, que permite capturar padrões cíclicos na demanda (TAYLOR; LETHAM, 2018).

Os resultados obtidos com o modelo ARIMA têm implicações significativas para o setor energético. Previsões acuradas de curto prazo, geralmente compreendendo períodos de até um ano, podem melhorar significativamente o planejamento da geração e distribuição de energia (HONG; FAN, 2016). Uma previsão precisa da demanda no curto prazo permite ajustes rápidos, garantindo uma alocação mais eficiente de recursos e, potencialmente, reduzindo custos operacionais (WERON, 2014). Modelos precisos de previsão de demanda são também fundamentais para a integração de fontes de energia renováveis, que podem apresentar variações imprevisíveis, exigindo ajustes rápidos e eficazes (KAUR; NONNENMACHER; COIMBRA, 2016). Além disso, previsões acuradas de curto prazo podem otimizar a implementação de programas de resposta à demanda, contribuindo para a estabilidade da rede (HONG; FAN, 2016). Já o longo prazo, geralmente considerado acima de um ano, abrange projeções para planejamento estratégico e infraestrutura, com um enfoque mais abrangente e menos suscetível a flutuações imediatas, mas igualmente relevante para garantir a continuidade e crescimento do setor energético.

Apesar do excelente desempenho do modelo ARIMA neste estudo, é importante reconhecer suas limitações e considerar possíveis direções para pesquisas futuras. O modelo ARIMA atual não incorpora variáveis exógenas como condições climáticas ou eventos especiais. A inclusão dessas variáveis em um modelo ARIMAX poderia potencialmente melhorar ainda mais as previsões (TAYLOR; LETHAM, 2018). Este estudo focou em previsões de curto prazo. Investigações futuras poderiam explorar o desempenho do modelo para horizontes de previsão mais longos (HYNDMAN; ATHANASOPOULOS, 2018).

Em conclusão, o modelo ARIMA desenvolvido neste estudo demonstrou um desempenho excepcional na previsão de demanda de energia, com implicações significativas para a eficiência e planejamento no setor energético. Os resultados obtidos não apenas validam a eficácia do modelo ARIMA para este tipo de série temporal, mas também abrem caminhos para futuras investigações e refinamentos na área de previsão de demanda de energia.

4.2 Avaliação do Modelo e Visualização dos Resultados das RNNs

As Redes Neurais Recorrentes (RNNs) têm se destacado como uma ferramenta poderosa na análise e previsão de séries temporais, especialmente em campos complexos como a previsão de demanda de energia (GOODFELLOW; BENGIO; COURVILLE, 2016). Esta seção apresenta uma análise detalhada dos resultados obtidos com o modelo RNN desenvolvido, discutindo sua precisão, interpretação e implicações práticas.

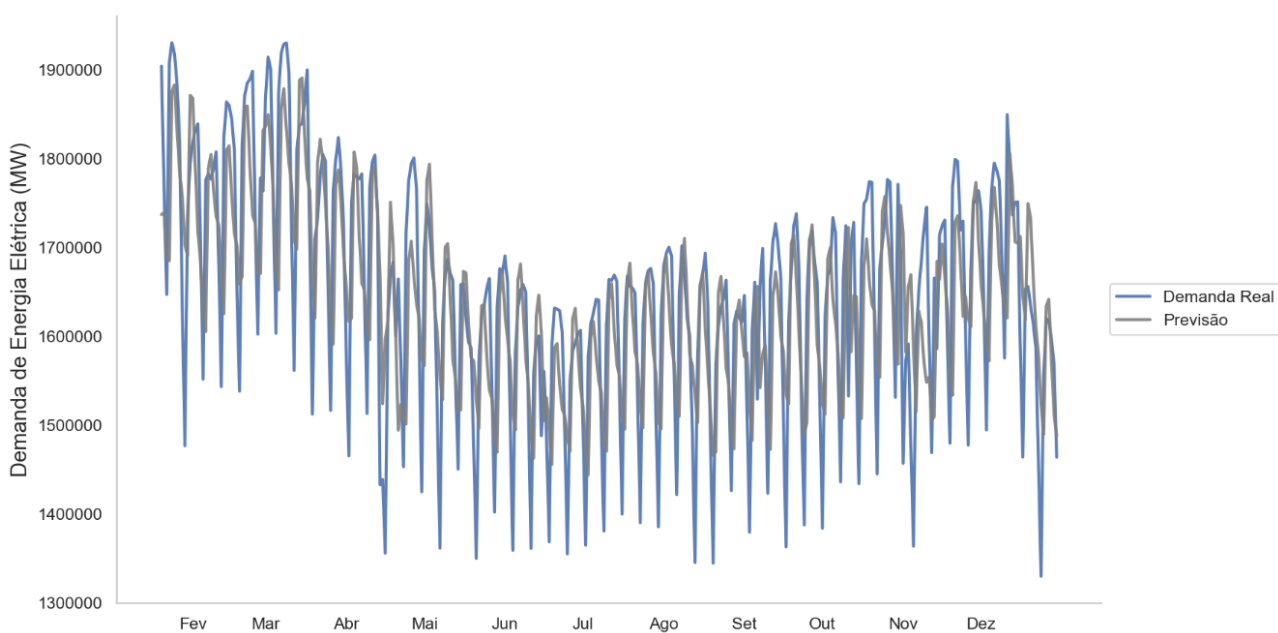
A avaliação da precisão do modelo é realizada utilizando o Erro Percentual Absoluto Médio (MAPE), uma métrica amplamente adotada na literatura de previsão de séries temporais (HYNDMAN; ATHANASOPOULOS, 2018). O MAPE oferece várias vantagens que o tornam particularmente adequado para este tipo de análise. Ele fornece uma medida percentual do erro, facilitando a compreensão intuitiva da precisão do modelo (MYTTENAERE et al., 2016). Por ser uma medida relativa, o MAPE permite comparações entre diferentes séries temporais, independentemente da escala dos dados (KIM; WON, 2018). Em contextos práticos, como o planejamento energético, uma medida percentual do erro é frequentemente mais útil para os tomadores de decisão (HONG; FAN, 2016).

No caso do modelo RNN desenvolvido neste estudo, o MAPE obtido foi de 4,11%. Este resultado merece uma análise detalhada. Um MAPE de 4,11% indica que, em média, as previsões do modelo desviaram 4,11% dos valores reais. Isto se traduz em uma precisão de 95,89%, um desempenho notável para previsões de demanda de energia (SUGANTHI; SAMUEL, 2012). Na literatura de previsão de demanda de energia, modelos com MAPE entre 3-10% são geralmente considerados de boa qualidade (WERON, 2014). Portanto, nosso modelo se posiciona na faixa superior deste espectro. Uma precisão tão elevada pode resultar em significativas melhorias no planejamento e operação de sistemas de energia (HONG; FAN, 2016).

É importante notar que este nível de precisão foi alcançado apesar da complexidade inerente às séries temporais de demanda de energia, que são influenciadas por múltiplos fatores como condições climáticas, padrões de consumo e eventos sazonais (TAYLOR; LETHAM, 2018).

A Figura 11 apresenta uma visualização dos resultados, com a comparação entre os valores reais e previstos da demanda de energia.

Figura 11 – Previsão do modelo de RNNs vs. valores reais da demanda diária de energia elétrica nos dados de teste (ano de 2022).



Fonte: Próprio Autor

Analisando o gráfico, podemos observar que a proximidade entre as curvas de valores reais e previstos indica um excelente ajuste do modelo à série temporal (CHATFIELD, 2000). O modelo RNN parece capturar eficazmente tanto as tendências gerais quanto os padrões sazonais presentes na série (GOODFELLOW; BENGIO; COURVILLE, 2016). O desempenho consistente nos conjuntos de treino, validação e teste sugere que o modelo é robusto e generaliza bem para dados não vistos (ZHANG, 2003). As RNNs demonstram capacidade de lidar com a volatilidade típica das séries de demanda de energia (KONG et al., 2017).

A habilidade do modelo em refletir com precisão as flutuações na demanda de energia é impressionante, especialmente devido à natureza complexa das séries temporais envolvidas (WERON, 2014). Essa performance é atribuída à estrutura recorrente da rede neural, que se destaca na captura das relações temporais de longo prazo, essenciais para uma previsão eficaz (GOODFELLOW; BENGIO; COURVILLE, 2016).

Os resultados obtidos com o modelo RNN têm implicações significativas para o setor energético. Previsões acuradas de curto prazo podem melhorar significativamente o planejamento da geração e distribuição de energia (HONG; FAN, 2016). Uma melhor previsão da demanda pode levar a uma alocação mais eficiente de recursos, potencialmente reduzindo custos operacionais (WERON, 2014). Modelos precisos de previsão de demanda são cruciais para a integração eficiente de fontes de energia renováveis, que são frequentemente intermitentes (KAUR; NONNENMACHER; COIMBRA, 2016). Previsões acuradas podem melhorar a implementação de programas de resposta à demanda, contribuindo para a estabilidade da rede (KONG et al., 2017).

O modelo RNN apresentou um bom desempenho ao longo deste estudo, mas é fundamental reconhecer suas limitações. Uma das principais críticas às RNNs é sua característica de "caixa preta", o que dificulta a interpretação dos resultados. Para contornar esse problema, investigações futuras poderiam focar no desenvolvimento de técnicas que aumentem a transparência do modelo (MONTAVON; SAMEK; MÜLLER, 2018). Além disso, a inclusão de variáveis exógenas, como indicadores meteorológicos e econômicos, poderia contribuir para previsões mais precisas (KONG et al., 2017). Outra possibilidade seria explorar arquiteturas mais sofisticadas, como modelos baseados em atenção ou Transformers, que têm demonstrado avanços promissores (VASWANI et al., 2017). Ademais, aprimorar o modelo para gerar previsões probabilísticas, em vez de apenas pontuais, poderia ser relevante para auxiliar na tomada de decisões em cenários incertos (GNEITING; KATZFUSS, 2014).

Dessa forma, os resultados obtidos evidenciam a eficácia das RNNs na previsão da demanda de energia, o que pode trazer benefícios para o planejamento e a eficiência do setor energético. Além de reforçar a viabilidade dessa abordagem para séries temporais desse tipo, os achados também apontam para novas possibilidades de aprimoramento e expansão.

4.3 Avaliação da Variação dos Anos de Validação e Teste

A análise da variação dos anos de validação e teste é um aspecto chave na avaliação de modelos de previsão de séries temporais, pois permite compreender a robustez e a capacidade de generalização dos modelos em diferentes cenários temporais (HYNDMAN; ATHANASOPOULOS, 2018). Esta seção apresenta uma análise detalhada da performance dos modelos ARIMA e RNN (Redes Neurais Recorrentes) em diversos contextos de validação e teste, explorando suas implicações para a precisão das previsões e eficiência computacional.

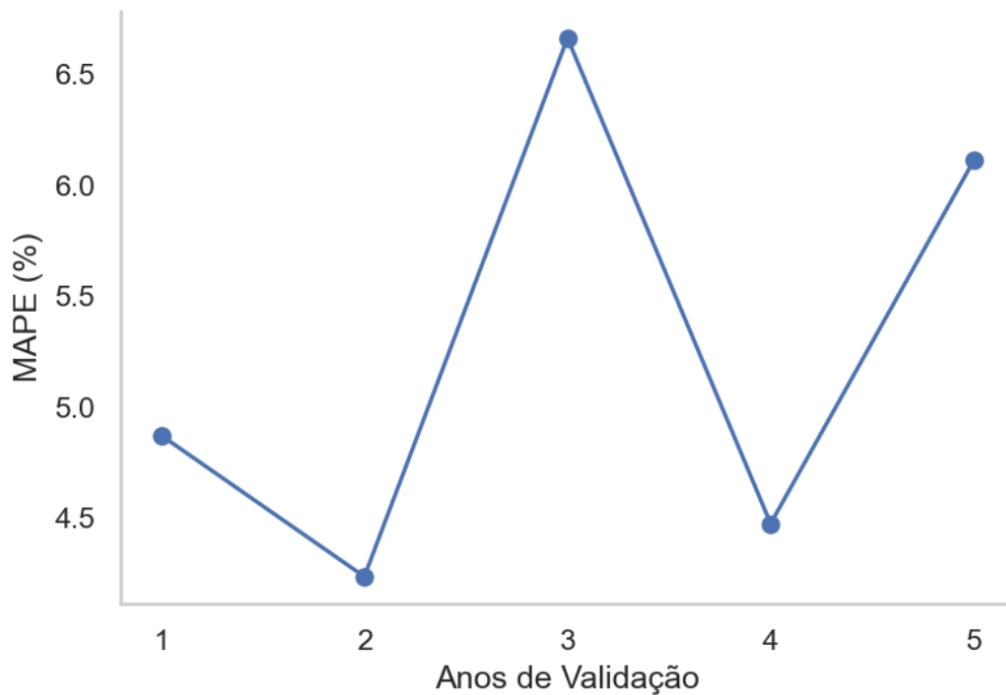
Inicialmente, observou-se que o modelo ARIMA apresentou um desempenho ligeiramente superior em termos da métrica MAPE (Mean Absolute Percentage Error) em comparação com o modelo baseado em RNNs. No entanto, é fundamental considerar não apenas a precisão, mas também a eficiência computacional na avaliação global dos modelos (MAKRIDAKIS; SPILIOTIS; ASSIMAKOPOULOS, 2018).

O modelo ARIMA, apesar de sua precisão marginalmente superior, demandou um tempo de treinamento significativamente maior, aproximadamente 2 horas. Em contraste, o modelo RNN completou seu treinamento em questão de minutos. Esta disparidade no custo computacional é um fator crítico, especialmente em aplicações que requerem atualizações frequentes ou processamento em tempo real (ZHANG, 2003).

Considerando a relação entre precisão e eficiência, o modelo RNN demonstrou uma vantagem substancial. A diferença marginal na métrica MAPE é compensada pela significativa redução no tempo de processamento, tornando o RNN uma escolha mais prática e escalável para muitas aplicações de previsão de séries temporais (HEWAMALAGE; BERGMEIR; BANDARA, 2021).

Para aprofundar a compreensão do comportamento do modelo RNN, realizou-se uma análise sistemática variando a quantidade de anos utilizados para validação, de 1 a 5 anos. É importante destacar que, ao aumentar o período destinado à validação, o intervalo de dados disponível para treinamento foi reduzido enquanto a quantidade de dados de teste foi mantida constante a 1 ano de teste. Esta abordagem tem como objetivo identificar o período ideal de validação que otimiza a performance do modelo (BERGMEIR; HYNDMAN; KOO, 2018). A Figura 12 ilustra a variação da métrica MAPE em função dos diferentes cenários de validação.

Figura 12 – Variação da métrica MAPE em diferentes cenários de quantidade de anos de validação para o modelo preditivo baseado em RNNs.



Fonte: Próprio Autor

A análise revela que o modelo com 2 anos de validação apresentou o menor valor de MAPE, atingindo 4,16%. Este resultado sugere que um período de validação de 2 anos pode ser ideal para capturar padrões relevantes sem incorrer em overfitting (GOODFELLOW; BENGIO; COURVILLE, 2016). No entanto, é importante notar que não se observou uma tendência clara na relação entre o número de anos de validação e o desempenho do modelo.

A ausência de uma tendência definida indica que a performance do modelo RNN não é significativamente afetada pelo tamanho do conjunto de validação. Isso sugere que fatores como a natureza específica dos dados da série temporal e a arquitetura da rede podem ter uma influência mais significativa no desempenho do modelo do que a extensão do período de validação, sem comprometer a acurácia do método, como indicado pela aparente estabilidade do MAPE (LAI et al., 2018).

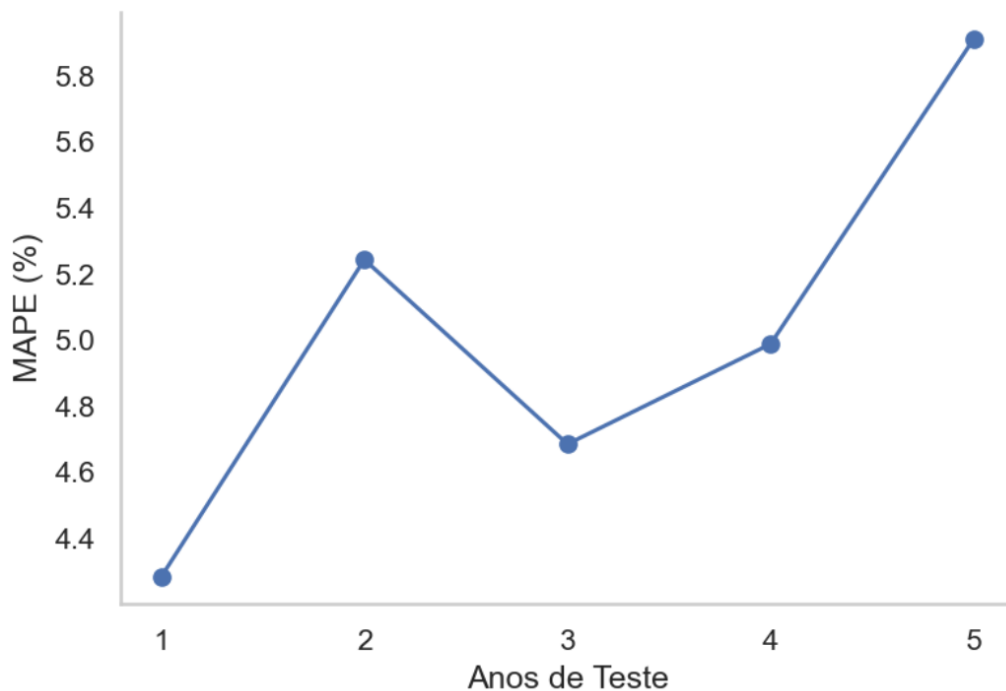
Esta observação tem implicações importantes para a prática de modelagem de séries temporais. Ela destaca a necessidade de uma abordagem cuidadosa e específica para cada conjunto de dados, em vez de seguir regras gerais sobre o tamanho ideal do conjunto de validação, como a divisão 60-20-20 (60% para treinamento, 20% para validação e 20% para teste) ou a regra 70-15-15 (70% para treinamento, 15% para validação e 15% para teste)(HYNDMAN; ATHANASOPOULOS, 2018).

Complementando a análise anterior, investigou-se também o impacto da variação dos anos de teste na performance do modelo RNN. Vale ressaltar que, ao aumentar o período destinado aos dados de teste, o intervalo de dados disponível para treinamento foi reduzido

enquanto a quantidade de dados de validação foi mantida constante a 1 ano. Esta etapa é essencial para avaliar a capacidade do modelo de generalizar para períodos futuros de diferentes extensões (TAIEB et al., 2016).

A Figura 13 apresenta os resultados desta análise, mostrando como a métrica MAPE varia em função do número de anos utilizados para teste.

Figura 13 – Variação da métrica MAPE em diferentes cenários de quantidade de anos de teste para o modelo preditivo baseado em RNNs.



Fonte: Próprio Autor

Os resultados indicam que o modelo apresentou o menor valor de MAPE (4,27%) quando testado com 1 ano de dados. Diferentemente da análise de validação, observou-se uma tendência clara na relação entre o número de anos de teste e o desempenho do modelo.

Esta tendência sugere que o modelo RNN é mais preciso em previsões de curto prazo, com sua performance degradando à medida que o horizonte de previsão se estende (BENIDIS et al., 2020). Tal comportamento é comum em modelos de séries temporais e reflete a dificuldade inerente de prever eventos futuros distantes, especialmente em sistemas complexos e dinâmicos (MAKRIDAKIS; SPILIOTIS; ASSIMAKOPOULOS, 2018).

A observação de que o modelo tem melhor desempenho com horizontes de previsão mais curtos tem implicações práticas significativas. Ela sugere que, para aplicações que requerem previsões de longo prazo, pode ser mais eficaz atualizar o modelo frequentemente com novos dados, em vez de confiar em previsões de longo alcance (TAIEB et al., 2016).

Em conclusão, esta análise detalhada da variação dos anos de validação e teste fornece informações sobre o comportamento e a aplicabilidade de modelos RNN em previsões de séries

temporais. Ela destaca a importância de uma abordagem equilibrada que considere não apenas a precisão das previsões, mas também a eficiência computacional e a capacidade de generalização do modelo para diferentes horizontes temporais.

5 Conclusão

O presente Trabalho de Conclusão de Curso teve como propósito central a previsão da demanda de energia elétrica no Brasil, utilizando métodos de modelagem de séries temporais. A pesquisa concentrou-se na implementação e análise comparativa de duas abordagens: o modelo ARIMA e as Redes Neurais Recorrentes (RNNs). Os resultados evidenciaram a eficiência e aplicabilidade dessas técnicas no setor elétrico brasileiro.

Inicialmente, foi realizada uma análise exploratória dos dados históricos de demanda energética, destacando tendências, padrões sazonais e possíveis anomalias. Essa etapa preliminar foi essencial para compreender a complexidade da série temporal e preparar adequadamente os dados para as etapas subsequentes de modelagem (BOX *et al.*, 2015). A aplicação do modelo ARIMA mostrou-se altamente eficaz, alcançando um MAPE de 1,97%, o que corresponde a uma precisão notável de 98,03% (LEWIS, 2000). Apesar de seu desempenho superior na identificação de padrões sazonais e tendências, o modelo apresentou alto custo computacional, com tempos de execução variando entre 1h30 e 2 horas (HYNDMAN; ATHANASOPOULOS, 2018).

Em paralelo, foi implementado um modelo de RNN com arquitetura baseada em GRU (Gated Recurrent Unit), cujos hiperparâmetros foram ajustados para a previsão da demanda energética. Este modelo apresentou um MAPE de 4,11%, refletindo uma precisão de 95,89% (ZHANG, 2003). Embora ligeiramente inferior em termos de acurácia quando comparado ao ARIMA, a RNN destacou-se pela rapidez no processamento, com tempos de treinamento significativamente reduzidos (HEWAMALAGE; BERGMEIR; BANDARA, 2021).

A análise comparativa entre os modelos revelou aspectos importantes. Enquanto o ARIMA demonstrou maior precisão, o modelo de RNN sobressaiu-se pela eficiência computacional, sendo mais adequado para cenários que demandam atualizações frequentes ou processamento em tempo real (GOODFELLOW; BENGIO; COURVILLE, 2016). Adicionalmente, a análise do impacto de diferentes períodos de validação e teste forneceu informações sobre o desempenho dos modelos. Por exemplo, o modelo RNN apresentou seu melhor resultado com 2 anos de validação, atingindo um MAPE de 4,16

Ao avaliar o modelo mais apropriado, não se considerou apenas a precisão das previsões, mas também a complexidade e o tempo de execução. Embora o ARIMA tenha se destacado em termos de acurácia, a vantagem computacional expressiva da RNN a torna uma alternativa mais prática e escalável para diversas aplicações no setor energético (WERON, 2014).

Os achados desta pesquisa possuem relevância significativa para o setor elétrico brasileiro. Previsões mais assertivas possibilitam um alinhamento mais eficiente entre oferta e demanda, favorecendo a otimização da geração e distribuição de energia (HONG; FAN, 2016). Isso pode levar a uma utilização mais racional dos recursos, com potencial para reduzir custos operacionais e tarifas para os consumidores. Além disso, modelos preditivos confiáveis desem-

penham um papel na integração de fontes renováveis, contribuindo para uma matriz energética mais sustentável (KAUR; NONNENMACHER; COIMBRA, 2016).

Este trabalho não apenas alcançou seus objetivos ao implementar e comparar modelos de previsão de demanda energética. Os resultados abrem espaço para investigações futuras, como a inclusão de variáveis exógenas, a exploração de arquiteturas mais modernas como os Transformers e o desenvolvimento de modelos que ofereçam previsões probabilísticas (LIM; ZOHREN, 2021)(MONTAVON; SAMEK; MÜLLER, 2018).

Em síntese, este estudo destaca o potencial transformador das técnicas de modelagem de séries temporais na previsão da demanda de energia elétrica. A adoção bem-sucedida desses modelos pode contribuir para um setor energético mais eficiente, econômico e sustentável no Brasil, alinhando-se às metas de transição energética e desenvolvimento sustentável do país (SUGANTHI; SAMUEL, 2012).

Referências

- AHMAD, A.; ANDERSON, T.; LIE, T. A review of the applications of neural networks in short-term electricity load forecasting. **International Journal of Forecasting**, Elsevier, v. 30, n. 4, p. 1076–1088, 2014.
- AHMED, N. K. et al. An empirical comparison of machine learning models for time series forecasting. **Econometric Reviews**, Taylor Francis, v. 29, n. 5-6, p. 594–621, 2010.
- AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, IEEE, v. 19, n. 6, p. 716–723, 1974.
- ARMSTRONG, J. S.; COLLOPY, F. Error measures for generalizing about forecasting methods: Empirical comparisons. **International Journal of Forecasting**, Elsevier, v. 8, n. 1, p. 69–80, 1992.
- BENGIO, Y. Practical recommendations for gradient-based training of deep architectures. **Neural networks: Tricks of the trade**, Springer, p. 437–478, 2012.
- BENIDIS, K. et al. Neural forecasting: Introduction and literature overview. **arXiv preprint arXiv:2004.10240**, 2020.
- BERGMEIR, C.; HYNDMAN, R. J.; KOO, B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. **Computational Statistics Data Analysis**, Elsevier, v. 120, p. 70–83, 2018.
- BHATTACHARYYA, S. C. Energy economics: concepts, issues, markets and governance. **Springer Science Business Media**, Springer, 2011.
- BISHOP, C. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006.
- BOX, G. E. et al. **Time series analysis: forecasting and control**. [S.l.]: John Wiley Sons, 2015.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to time series and forecasting**. [S.l.]: Springer, 2016.
- CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature. **Geoscientific Model Development**, Copernicus GmbH, v. 7, n. 3, p. 1247–1250, 2014.
- CHATFIELD, C. **Time-series forecasting**. [S.l.]: CRC press, 2000.
- CHATFIELD, C. **The analysis of time series: an introduction**. [S.l.]: CRC press, 2016.
- CHO, K. et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation. **arXiv preprint arXiv:1406.1078**, 2014.
- CHOLLET, F. **Deep learning with Python**. [S.l.]: Manning Publications Co., 2018.
- CLEVELAND, R. B. et al. Stl: A seasonal-trend decomposition. **Journal of official statistics**, Statistics Sweden (SCB), v. 6, n. 1, p. 3–73, 1990.

- DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American statistical association**, Taylor Francis, v. 74, n. 366a, p. 427–431, 1979.
- ELMAN, J. L. Finding structure in time. **Cognitive science**, Wiley Online Library, v. 14, n. 2, p. 179–211, 1990.
- Empresa de Pesquisa Energética. **Plano Decenal de Expansão de Energia 2024**. Brasília, 2024.
- GERS, F. A.; SCHMIDHUBER, J.; CUMMINS, F. Learning to forget: Continual prediction with lstm. **Neural computation**, MIT Press, v. 12, n. 10, p. 2451–2471, 2000.
- GNEITING, T.; KATZFUSS, M. Probabilistic forecasting. **Annual Review of Statistics and Its Application**, Annual Reviews, v. 1, p. 125–151, 2014.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT press, 2016.
- GRAVES, A.; MOHAMED, A.-r.; HINTON, G. Speech recognition with deep recurrent neural networks. **2013 IEEE international conference on acoustics, speech and signal processing**, p. 6645–6649, 2013.
- GUO, Y. et al. Data mining techniques for complex and big data. **Neurocomputing**, Elsevier, v. 196, p. 1–2, 2016.
- HAMILTON, J. D. **Time series analysis**. [S.l.]: Princeton university press, 2020.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. [S.l.]: Springer Science Business Media, 2009.
- HAYKIN, S. Neural networks and learning machines. **Upper Saddle River, NJ, USA**, Pearson, v. 3, 2009.
- HEWAMALAGE, H.; BERGMEIR, C.; BANDARA, K. Recurrent neural networks for time series forecasting: Current status and future directions. **International Journal of Forecasting**, Elsevier, v. 37, n. 1, p. 388–427, 2021.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.
- HONG, T.; FAN, S. Probabilistic electric load forecasting: A tutorial review. **International Journal of Forecasting**, Elsevier, v. 32, n. 3, p. 914–938, 2016.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. Forecasting: principles and practice. **OTexts**, 2014.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. [S.l.]: OTexts, 2018.
- JAMES, G. et al. **An introduction to statistical learning**. [S.l.]: Springer, 2013. v. 112.
- JIANG, H. et al. Short-term and mid-term load forecasting using a bilevel optimization model. **IEEE Transactions on Power Systems**, IEEE, v. 33, n. 2, p. 1404–1415, 2018.
- KAGGLE. **23 Years of Hourly Electric Energy Demand (Brazil)**. 2023. Disponível em: <https://www.kaggle.com/datasets/your-dataset-url>.

- KAUR, A.; NONNENMACHER, L.; COIMBRA, C. F. Short-term wind power forecasting using empirical mode decomposition based ensemble wavelets. **Applied Energy**, Elsevier, v. 164, p. 778–791, 2016.
- KAYTEZ, F. et al. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. **International Journal of Electrical Power Energy Systems**, Elsevier, v. 67, p. 431–438, 2015.
- KIM, H. Y.; WON, C. H. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. **Expert Systems with Applications**, Elsevier, v. 103, p. 25–37, 2018.
- KONG, W. et al. Short-term residential load forecasting based on lstm recurrent neural network. **IEEE Transactions on Smart Grid**, IEEE, v. 10, n. 1, p. 841–851, 2017.
- KUHN, M.; JOHNSON, K. **Applied predictive modeling**. [S.l.]: Springer, 2013. v. 26.
- LAI, G. et al. Modeling long-and short-term temporal patterns with deep neural networks. **In The 41st International ACM SIGIR Conference on Research Development in Information Retrieval**, ACM, p. 95–104, 2018.
- LAZZERI, F. **Machine Learning for Time Series Forecasting with Python**. [S.l.]: Wiley, 2020. 224 p. ISBN 9781119682363.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015.
- LEWIS, R. J. An introduction to classification and regression tree (cart) analysis. **Annual meeting of the society for academic emergency medicine in San Francisco, California**, v. 14, 2000.
- LIM, B.; ZOHREN, S. Time series forecasting with deep learning: a survey. **Philosophical Transactions of the Royal Society A**, The Royal Society Publishing, v. 379, n. 2194, p. 20200209, 2021.
- LIPTON, Z. C.; BERKOWITZ, J.; ELKAN, C. A critical review of recurrent neural networks for sequence learning. **arXiv preprint arXiv:1506.00019**, 2015.
- LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. **Biometrika**, Oxford University Press, v. 65, n. 2, p. 297–303, 1978.
- MAKRIDAKIS, S. Accuracy measures: theoretical and practical concerns. **International Journal of Forecasting**, Elsevier, v. 9, n. 4, p. 527–529, 1993.
- MAKRIDAKIS, S.; SPILIOTIS, E.; ASSIMAKOPOULOS, V. Statistical and machine learning forecasting methods: Concerns and ways forward. **PloS one**, Public Library of Science San Francisco, CA USA, v. 13, n. 3, p. e0194889, 2018.
- MASTERS, D.; LUSCHI, C. Deep learning with python. **arXiv preprint arXiv:1801.05894**, 2018.
- MCKINNEY, W. **Python for data analysis: Data wrangling with Pandas, NumPy, and IPython**. [S.l.]: O'Reilly Media, Inc., 2012.
- MCKINNEY, W. et al. Data structures for statistical computing in python. **Proceedings of the 9th Python in Science Conference**, Austin, TX, v. 445, p. 51–56, 2010.

- Medium. **Medium**. 2025. Acesso em: 18 fev. 2025. Disponível em: <<https://medium.com/>>.
- MIKOLOV, T. et al. Recurrent neural network based language model. **Interspeech**, v. 2, n. 3, p. 1045–1048, 2010.
- Ministério de Minas e Energia. **Programa de Geração Distribuída em Universidades Públicas**. 2024. <<https://www.gov.br/mme/pt-br/assuntos/noticias/mme-lanca-chamada-publica-de-r-100-milhoes-para-predios-publicos-eficientes>>. Acessado em: 20 de novembro de 2024.
- MONTAVON, G.; SAMEK, W.; MÜLLER, K.-R. Methods for interpreting and understanding deep neural networks. **Digital Signal Processing**, Elsevier, v. 73, p. 1–15, 2018.
- MYTTENAERE, A. D. et al. Mean absolute percentage error (mape) in forecast accuracy. **International Conference on Machine Learning and Data Mining in Pattern Recognition**, Springer, p. 106–117, 2016.
- NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: **Proceedings of the 27th international conference on machine learning (ICML-10)**. [S.l.: s.n.], 2010. p. 807–814.
- Operador Nacional do Sistema Elétrico (ONS). **Histórico da Operação - Curva de Carga Horária**. 2024. Acessado em: 19 fev. 2025. Disponível em: <https://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/curva_carga_horaria.aspx>.
- PASCANU, R.; MIKOLOV, T.; BENGIO, Y. On the difficulty of training recurrent neural networks. **International conference on machine learning**, p. 1310–1318, 2013.
- PETROPOULOS, F. et al. Forecasting: theory and practice. **International Journal of Forecasting**, Elsevier, v. 36, n. 1, p. 1–47, 2020.
- PRECHELT, L. Early stopping-but when? **Neural Networks: Tricks of the trade**, Springer, p. 55–69, 1998.
- RUDER, S. An overview of gradient descent optimization algorithms. **arXiv preprint arXiv:1609.04747**, 2016.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural networks**, Elsevier, v. 61, p. 85–117, 2015.
- SEABOLD, S.; PERKTOLD, J. statsmodels: Econometric and statistical modeling with python. **9th Python in Science Conference**, 2010.
- SIAMI-NAMINI, S.; TAVAKOLI, N.; NAMIN, A. S. A comparison of arima and lstm in forecasting time series. **2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)**, p. 1394–1401, 2018.
- SIAMI-NAMINI, S.; TAVAKOLI, N.; NAMIN, A. S. A comparison of arima and lstm in forecasting time series. **2019 17th IEEE International Conference on Machine Learning and Applications (ICMLA)**, p. 1394–1401, 2019.
- SUGANTHI, L.; SAMUEL, A. A. Energy models for demand forecasting—a review. **Renewable and sustainable energy reviews**, Elsevier, v. 16, n. 2, p. 1223–1240, 2012.

- TAIEB, S. B. et al. Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression. **IEEE Transactions on Smart Grid**, IEEE, v. 7, n. 5, p. 2448–2455, 2016.
- TAYLOR, S. J.; LETHAM, B. Forecasting at scale. **The American Statistician**, Taylor Francis, v. 72, n. 1, p. 37–45, 2018.
- TIELEMAN, T.; HINTON, G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. **COURSERA: Neural networks for machine learning**, v. 4, n. 2, p. 26–31, 2012.
- TSAY, R. S. **Analysis of financial time series**. [S.l.]: John Wiley Sons, 2005.
- TUFTE, E. R. **The visual display of quantitative information**. [S.l.]: Graphics press Cheshire, CT, 2001. v. 2.
- VANDERPLAS, J. **Python data science handbook: Essential tools for working with data**. [S.l.]: O'Reilly Media, Inc., 2016.
- VASWANI, A. et al. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017.
- WEI, W. W. **Time series analysis: univariate and multivariate methods**. [S.l.]: Pearson Addison Wesley, 2006.
- WERON, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. **International journal of forecasting**, Elsevier, v. 30, n. 4, p. 1030–1081, 2014.
- Wikipédia, a enciclopédia livre. **Wikipédia**. 2025. Acesso em: 18 fev. 2025. Disponível em: <<https://www.wikipedia.org/>>.
- WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. **Climate Research**, Inter-Research Science Center, v. 30, n. 1, p. 79–82, 2005.
- WOLDE-RUFAEL, Y. Energy demand forecasting using neural network: a case study of turkey with projections to 2025. **Energy Policy**, Elsevier, v. 87, p. 656–666, 2015.
- ZHANG, G. P. Time series forecasting using a hybrid arima and neural network model. **Neurocomputing**, Elsevier, v. 50, p. 159–175, 2003.
- ZHANG, X.; WANG, J.; ZHANG, K. Short-term load forecasting based on deep learning: A survey. **Energies**, MDPI, v. 11, n. 11, p. 3039, 2018.