

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
INSTITUTO DE CIÊNCIAS EXATAS  
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

Modelo de rede neural *Long short-term  
memory* para previsão de cargas elétricas  
de sistemas de energia eólica

Diego de Souza Martins

JUIZ DE FORA  
DEZEMBRO, 2023

Modelo de rede neural *Long short-term memory* para previsão de cargas elétricas de sistemas de energia eólica

DIEGO DE SOUZA MARTINS

Universidade Federal de Juiz de Fora

Instituto de Ciências Exatas

Departamento de Ciência da Computação

Bacharelado em Sistemas de Informação

Orientador: Luciana Conceição Dias Campos

Coorientador: Heder Soares Bernardino

JUIZ DE FORA

DEZEMBRO, 2023

MODELO DE REDE NEURAL *Long short-term memory* PARA  
PREVISÃO DE CARGAS ELÉTRICAS DE SISTEMAS DE  
ENERGIA EÓLICA

Diego de Souza Martins

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS  
EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTE-  
GRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
BACHAREL EM SISTEMAS DE INFORMAÇÃO.

Aprovada por:

Luciana Conceição Dias Campos  
Doutora em Engenharia Elétrica (PUC-Rio)

Heder Soares Bernardino  
Doutor em Modelagem Computacional (LNCC, Brasil)

Alex Borges Vieira  
Doutor em Ciências da Computação (UFMG)

Victor Stroele de Andrad Menezes  
Doutor em Engenharia de Sistemas e Computação (UFRJ)

JUIZ DE FORA  
15 DE DEZEMBRO, 2023

## Resumo

Atualmente, a questão energética tem sido debatida mundialmente, em principal, a geração de energia por fontes renováveis. Neste contexto, a energia eólica apresenta-se como uma das principais alternativas que visa atender as demandas sócio-ambientais e econômicas. Segundo o Balanço Energético Nacional (BEN) 2023, a participação da energia eólica na matriz energética brasileira em 2022 foi de 12,6%, ficando atrás apenas da energia hidrelétrica, que teve participação de 58%. Por ser uma fonte de energia limpa e renovável gerada a partir da transformação da energia cinética contida nas massas de ar em energia elétrica, a quantidade de energia produzida pelas turbinas eólicas depende das condições climáticas da região. Tais condições não podem ser controladas, acarretando dificuldade de integralização das usinas eólicas com a rede de matrizes energéticas.

Neste trabalho são adotados modelos baseados em séries temporais recorrentes, especificamente a Rede Neural Recorrente LSTM (*Long Short-Term Memory*), utilizados para identificar os padrões de comportamento da série temporal, gerando previsões futuras de curto prazo. São abordados 2 modelos de previsão: o primeiro modelo utiliza apenas a série de dados de geração de energia eólica, denominado modelo univariado, e o segundo, utiliza a série dados de geração de energia eólica e velocidade prevista do vento como variável exógena, denominado modelo multivariado. Verificou-se que os modelos multivariados apresentaram melhores resultados quando comparados aos modelos univariados. A partir desses resultados, podemos concluir que a utilização de variáveis exógenas contribui de maneira significativa, aumentando a predição de cargas de energia eólicas através Rede Neurais LSTM.

**Palavras-chave:** Series Temporais, Redes Neurais Recorrentes, energia eólica

# Abstract

Currently, the energy issue has been debated worldwide, mainly the generation of energy from renewable sources. In this context, wind energy presents itself as one of the main alternatives that aims to meet socio-environmental and economic demands. According to the National Energy Balance (BEN) 2023, the share of wind energy in the Brazilian energy matrix in 2022 was 12.6%, second only to hydroelectric energy, which had a share of 58%. As it is a source of clean and renewable energy generated from the transformation of kinetic energy contained in air masses into electrical energy, the amount of energy produced by wind turbines depends on the climatic conditions of the region. Such conditions cannot be controlled, causing difficulties in integrating wind power plants into the energy matrix network.

In this work, models based on recurrent time series are adopted, specifically the LSTM Recurrent Neural Network (*Long Short-Term Memory*), used to identify the behavior patterns of the time series, generating short-term future forecasts. Two forecast models are addressed: the first model uses only the wind energy generation data series, called a univalued model, and the second uses the wind energy generation data series and predicted wind speed as an exogenous variable, called the model. multivalued. It was found that multivalued models presented better results when compared to univalued models. From these results, we can conclude that the use of exogenous variables contributes significantly, increasing the prediction of wind energy loads through LSTM Neural Network.

**Keywords:** Time Series, Recurrent Neural Networks, wind energy *Long Short Term Memory*

## Agradecimentos

Agradeço aos meus pais, que, de forma incondicional e constante, foram a base para tudo o que sou e para tudo o que aspiro ser. Ao meu irmão pelo apoio.

Aos meus amigos, que proporcionaram momentos de descontração e amizade genuína, meu sincero obrigado por estarem sempre ao meu lado.

Não posso deixar de expressar gratidão a Luciana C. D. Campos e ao Heder S. Bernardino que guiaram minha jornada acadêmica com suas competências e conhecimentos.

Agradeço também ao Paulo Sergio C. Nascimento e a Ana Carolyne G. França, pela contribuição no trabalho.

# Conteúdo

|   |           |
|---|-----------|
| <b>Lista de Figuras</b>   | <b>6</b>  |
| <b>Lista de Abreviações</b>   | <b>7</b>  |
| <b>1 Introdução</b>   | <b>8</b>  |
| 1.1 Motivações . . . . .  | 9         |
| 1.2 Objetivos . . . . .   | 10        |
| 1.3 Organização do Trabalho . . . . .                               | 10        |
| <b>2 Trabalhos relacionados</b>                                     | <b>12</b> |
| <b>3 Contextualização</b>   | <b>14</b> |
| 3.1 Setor Elétrico Brasileiro . . . . .                             | 14        |
| 3.2 Geração de Energia Eólica . . . . .                             | 15        |
| <b>4 Fundamentação teórica</b>                                      | <b>17</b> |
| 4.1 Redes Neurais Artificiais . . . . .                             | 17        |
| 4.1.1 Função de ativação . . . . .                                  | 19        |
| 4.1.2 Long Short-Term Memory . . . . .                              | 21        |
| 4.1.2.1 <i>Forgot gate</i> . . . . .                                | 23        |
| 4.1.2.2 <i>Input Gate</i> e <i>Cell Gate</i> . . . . .              | 24        |
| 4.1.2.3 <i>Cell Update</i> . . . . .                                | 25        |
| 4.1.2.4 <i>Output Gate</i> e <i>Hidden State</i> . . . . .          | 26        |
| 4.2 Normalização dos dados . . . . .                                | 28        |
| 4.3 Parametrização da Rede Neural . . . . .                         | 29        |
| 4.3.1 Algoritmos de otimização . . . . .                            | 29        |
| 4.3.2 <i>DropOut</i> . . . . .                                      | 30        |
| 4.3.3 <i>EarlyStopping</i> . . . . .                                | 30        |
| 4.3.4 Regularizadores de peso de camada . . . . .                   | 31        |
| 4.4 Ajuste de hiperparâmetros . . . . .                             | 32        |
| <b>5 Conjunto de dados</b>  | <b>34</b> |
| 5.1 Caracterização e análise dos dados . . . . .                    | 36        |
| <b>6 Metodologia</b>  | <b>38</b> |
| 6.1 Séries Temporais das Usinas . . . . .                           | 38        |
| 6.2 Pré-Processamento do Conjunto De Dados . . . . .                | 41        |
| 6.3 Cenário de Teste . . . . .                                      | 43        |
| 6.3.1 LSTM Univalorado . . . . .                                    | 43        |
| 6.3.2 LSTM Multivalorado . . . . .                                  | 44        |
| <b>7 Resultados experimentais</b>                                   | <b>45</b> |
| 7.0.1 Resultados comparativos dos cenários Univalorados . . . . .   | 45        |
| 7.0.2 Resultados comparativos dos cenários Multivalorados . . . . . | 46        |
| 7.0.3 Análise dos resultados dos Modelos . . . . .                  | 47        |

|                                 |           |
|---------------------------------|-----------|
| <b>8 Conclusão</b>              | <b>49</b> |
| 8.1 Trabalhos futuros . . . . . | 49        |
| <b>Bibliografia</b>             | <b>50</b> |



## Lista de Figuras

|      |   |    |
|------|---|----|
| 3.1  | Matriz Elétrica Brasileira de Energia renovável . . . . .                     | 14 |
| 4.1  | Estrutura simplificada de um neurônio biológico . . . . .                     | 17 |
| 4.2  | Estrutura esquemática de um neurônio artificial . . . . .                     | 19 |
| 4.3  | Funções de ativação e seus respectivos gráficos . . . . .                     | 20 |
| 4.4  | Representação de uma célula LSTM. . . . .                                     | 22 |
| 4.5  | Representação do <i>forget gate</i> na célula LSTM. . . . .                   | 24 |
| 4.6  | Representação do <i>input gate</i> e <i>cell gate</i> na célula LSTM. . . . . | 24 |
| 4.7  | Representação da região <i>Cell Update</i> na célula LSTM. . . . .            | 26 |
| 4.8  | Representação do output gate na célula LSTM. . . . .                          | 27 |
| 5.1  | Divisão da base de dados em conjuntos . . . . .                               | 34 |
| 5.2  | Modelo de previsão univariado . . . . .                                       | 35 |
| 5.3  | Modelo de previsão multivariado . . . . .                                     | 36 |
| 6.1  | Base de dados de fontes renováveis eólicas do ONS . . . . .                   | 38 |
| 6.2  | Série histórica da Usina 1 . . . . .  | 39 |
| 6.3  | Série histórica da Usina 2 . . . . .  | 39 |
| 6.4  | Série histórica da Usina 3 . . . . .  | 39 |
| 6.5  | Série histórica da Usina 4 . . . . .  | 40 |
| 6.6  | Série histórica da Usina 5 . . . . .  | 40 |
| 6.7  | Série histórica da Usina 6 . . . . .  | 40 |
| 6.8  | Série histórica da Usina 7 . . . . .  | 41 |
| 6.9  | Quadrículas do modelo GFS para Nordeste. . . . .                              | 42 |
| 6.10 | Quadrículas do modelo GFS para Sul. . . . .                                   | 43 |
| 7.1  | Comparativo entres os cenários do modelo LSTM-UNI . . . . .                   | 46 |
| 7.2  | Comparativo entres os cenários do modelo LSTM-MULT . . . . .                  | 47 |
| 7.3  | Comparativo entres os modelos LSTM-UNI e LSTM-MULT . . . . .                  | 48 |

## Lista de Abreviações

|      |   |
|------|---|
| BEN  | Balanco Energético Nacional                     |
| CCEE | Câmara de Comercialização de Energia Elétrica   |
| DCC  | Departamento de Ciência da Computação           |
| GFS  | Global Forecast System                          |
| LSTM | Long Short Term Memory                          |
| NOAA | National Oceanic and Atmospheric Administration |
| ONS  | Operador Nacional do Sistema Elétrico           |
| RNN  | Rede Neural Recorrente                          |
| SGD  | Gradiente Descendente Estocástico               |
| SIN  | Sistema Interligado Nacional                    |
| UFJF | Universidade Federal de Juiz de Fora            |

# 1 Introdução

A questão energética é um dos temas mais importantes atualmente em discussão no mundo. O consumo global de energia está diretamente ligado ao desenvolvimento econômico e social de uma nação, e há uma preocupação crescente com o suprimento energético e as questões socioambientais. Nesse contexto, a geração de energia eólica surge como uma das principais opções para compor a matriz energética de uma nação, por ser uma fonte de energia renovável e amplamente disponível.

Até o final de julho de 2023, houve uma expansão de 5.673,9 MW na matriz elétrica brasileira. Foi registrada pela ANEEL a iniciação da operação comercial de 176 usinas, com um desdobramento de 79 usinas eólicas (2.713,8 MW), 61 solares fotovoltaicas (2.295,1 MW), 25 termelétricas (531,4 MW), oito pequenas centrais hidrelétricas (122,2 MW) e três centrais geradoras hidrelétricas (11,4 MW). Notavelmente, as usinas solares e eólicas correspondem a 88,3% da total capacidade acrescida no ano (ABEEÓLICA, 2023).

Portanto, é essencial aprimorar os métodos de previsão de energia eólica para desenvolver modelos capazes de gerar previsões de alta qualidade. Isso auxiliará na tomada de decisões pelo Sistema Interligado Nacional - SIN, buscando maior economia, segurança e sustentabilidade (FERREIRA, 2012).

A análise de séries temporais é um recurso importante para resolver problemas que envolvem previsões sobre o comportamento de uma variável ao longo do tempo. Essa análise é frequentemente usada para observar o padrão de variáveis em áreas como finanças, meteorologia, epidemiologia, ciências econômicas, entre outras.

Os métodos estatísticos existentes para análise de séries temporais são bem desenvolvidos na literatura. No entanto, para obter uma boa compreensão e análise, é recomendado ter conhecimento específico para obter uma aplicação eficaz e resultados satisfatórios. O processo de desenvolvimento de modelos estatísticos na área de modelagem de séries temporais teve um crescimento rápido nas últimas décadas, especialmente com a publicação dos trabalhos de Box & Jenkins na década de 70 (GUTIÉRREZ, 2003).

Este trabalho utiliza a Rede Neural Recorrente (RNN), especificamente a arqui-

tutura LSTM (*Long Short Term Memory*), denominada Redes de Memória de longo e curto Prazo, como método de previsão de cargas de energia eólicas. Essa arquitetura é projetada para evitar problemas de dependência de longo prazo, sendo adequada para realizar processos de classificação, processamento e previsão de séries temporais.

## 1.1 Motivações

À medida que a capacidade instalada de energia eólica se expande no Brasil e no mundo, impulsionada por avanços tecnológicos, políticas governamentais e demanda por fontes de energia mais sustentáveis, é fundamental abordar os desafios relacionados à natureza intermitente dessa fonte. A realização de previsões precisas de cargas elétricas eólicas é uma ferramenta essencial para otimizar a integração da energia eólica na rede elétrica. Isso permitiria ao Operador Nacional do Sistema Elétrico - ONS gerenciar de forma eficaz a produção e o despacho de energia eólica, resultando em menor custo e maior confiabilidade na geração de energia (PRADA, 2021). Projetadas para trabalhar com sequências temporais complexas, as LSTM's aprendem dependências temporais de longo prazo, tornando-as ideais para serem utilizadas em modelos de previsão de geração de cargas elétricas de sistemas de energia eólica, que são altamente dependentes das condições climáticas. Estudos recentes demonstram que os modelos LSTMs podem superar outras arquiteturas de redes neurais na previsão de carga elétrica em termos de precisão e eficiência computacional. Além disso, são capazes de lidar com diferentes tipos de dados, como dados meteorológicos e históricos de carga elétrica (HOSSAIN; MAHMOOD, 2020).

Na literatura é possível encontrar uma grande quantidade de trabalhos que se dedicam ao desenvolvimento de modelos de previsão de séries temporais. Um exemplo importante nesse contexto é a *Makridakis Competition*, organizada por Spyros Makridakis, cujo propósito é avaliar e comparar diferentes métodos de previsão, apresentando novas metodologias e avanços nas discussões em torno dos modelos de previsão. A competição ocorre anualmente e é divulgada no *International Journal of Forecasting* (PRADA, 2021)

## 1.2 Objetivos

O trabalho da (FRANÇA et al., 2020) foi escolhido como inspiração para este trabalho, pois utiliza métodos considerados importantes para a previsão de séries temporais, como o Arima e o Arimax, que são modelos lineares, e também a Rede MLP. O intuito do trabalho é dar continuidade aos experimentos introduzindo modelos mais avançados e explorando novas possibilidades, como as redes neurais recorrentes LSTM, que tem sido muito usada na literatura para previsão de séries temporais, como séries de energia eólica. Contudo, não é possível comparar diretamente os resultados devido a restrições de acesso a todos os dados brutos utilizados no trabalho de inspiração. Os dados de geração eólica e velocidade prevista do vento fornecidos pela ONS, são confidenciais e passaram por normalização, mantendo as características da série temporal original.

O objetivo é analisar o comportamento de algoritmos de redes neurais recorrentes (RNN), modelados com a arquitetura de uma rede *Long short-term memory (LSTM)*, para calcular previsão de geração eólica com base em dois experimentos que utilizam dados de geração de energia de fonte eólica do Sistema Interligado Nacional (SIN) e comparar com os resultados obtidos na literatura e em estudos preliminares.

Como objetivos específicos, podemos destacar:

- Oferecer suporte ao operador SIN em relação à previsão da geração eólica.
- Realizar previsão da série temporal de geração de energia eólica verificada.
- Analisar e comparar o comportamento da rede neural LSTM utilizando a série temporal de geração eólica verificada e da velocidade prevista do vento como variável exógena.

## 1.3 Organização do Trabalho

Este trabalho está organizado em oito capítulos. Neste primeiro capítulo são apresentados a contextualização do problema abordado, com motivações e objetivos.

No capítulo 2, apresenta-se a fundamentação teórica e trabalhos utilizados como base ferramental de pesquisa, no desenvolvimento dos modelos e como ponto de partida

---

para as técnicas e procedimentos propostos.

O capítulo 3, apresenta um panorama do setor elétrico brasileiro, com ênfase na crescente relevância da energia eólica.

No Capítulo 4, apresenta os conceitos fundamentais de séries temporais, redes neurais artificiais e a arquitetura LSTM. Este capítulo serve como base teórica para os métodos utilizados no estudo.

No Capítulo 5, é descrito o conjunto de dados utilizado para o estudo. São detalhados os processos de normalização e preparação dos dados

O Capítulo 6 detalha a metodologia utilizada para realizar os experimentos, incluindo a descrição dos cenários de teste para os modelos univariados e multivariados e as estratégias de validação dos modelos.

No Capítulo 7 são apresentados os resultados dos experimentos. São feitas comparações entre os cenários e modelos, avaliando o desempenho dos modelos LSTM na previsão de cargas eólicas.

No Capítulo 8, o capítulo final resume as principais descobertas, enfatizando a eficácia das redes LSTM na previsão de geração de energia eólica e sugerindo direções para pesquisas futuras.

Além desses capítulos principais, o trabalho inclui uma bibliografia que lista todas as referências utilizadas, bem como uma lista de figuras que auxiliam na visualização dos dados e resultados.

## 2 Trabalhos relacionados

Técnicas de aprendizado de máquina são utilizadas em diversos estudos encontrados na literatura sobre previsões de séries temporais de conjunto muito grandes de dados. No contexto de previsão de geração de energia eólicas os modelos se destacam por apresentar resultados promissores. Esta seção discute alguns trabalhos relevantes sobre previsão de geração eólica usando aprendizado de máquina e arquitetura LSTM e suas particularidades que agregam os modelos de previsão.

(FRANÇA et al., 2020) comparou o desempenho de modelos de previsão de geração de energia eólica baseados em redes neurais MLP e NARX, e métodos tradicionais de séries temporais, como Arima e Arimax. Os resultados indicaram que os modelos baseados em redes neurais MLP e NARX apresentaram melhor desempenho em termos de precisão de previsão, com a rede NARX apresentando resultados ainda mais satisfatórios. A análise dos erros de previsão mostrou que os modelos baseados em redes neurais apresentaram menor erro médio absoluto e menor erro quadrático médio em comparação com os modelos de séries temporais. Além disso, a análise de sensibilidade mostrou que a rede NARX é mais sensível às variações nas variáveis meteorológicas do que a rede MLP. Esses resultados destacam o potencial das redes neurais, especialmente a rede NARX, como ferramentas promissoras para a previsão de geração de energia eólica.

(LIU et al., 2020) utilizaram um modelo de arquitetura híbrida de previsão de curto prazo da energia eólica baseado em uma transformação de onda discreta e redes de memória de curto prazo (WD LSTM). Os autores utilizam 6 variáveis de entradas, sendo elas, Produto Interno Bruto (PIB), Índice de Preços no Consumidor (IPC), Valor Acrescentado Industrial (IVA), Total de Importações e Exportações (TIE), Geração Total de Energia (TPG) e Geração Hidroelétrica (HG), indicadores macroeconômicos e indicadores relacionados de produção de energia do Gabinete Nacional de Estatística da China. Em comparação com outros modelos de arquiteturas distintas, o modelo WD-LSTM alcançou um MAPE de 5.831, prevendo com maior precisão a produção de energia eólica da China.

Em (HUANG; WANG, 2022) é proposto um modelo de previsão da energia eólica que utiliza uma otimização do algoritmo PSO - Particle Swarm Optimization, em português: (Otimização por enxame de partículas), rede de LSTM e fusão classificada. O resultado do conjunto dessas técnicas gerou resultados relevantes, onde foi obtido um RMSE e o MAE no valor de 1,2382 e 0,8210, respectivamente. O modelo se ajusta melhor à curva real da força do vento e imprime uma excelente previsão.

(LIU et al., 2019) projetaram um método de previsão de curto prazo de energia eólica baseado na transformação de onda discreta e LSTM, e relataram um aumento na precisão da previsão em comparação com cinco parâmetros de referência diferentes.



## 3 Contextualização

### 3.1 Setor Elétrico Brasileiro

O Brasil possui uma enorme extensão territorial de clima tropical com um conjunto de bacias hidrográficas em áreas de planalto que cobrem cerca de 80% do território brasileiro. Devido a essas características, a principal fonte da matriz energética do país é a energia hidrelétrica (ALUPAR, 2022).

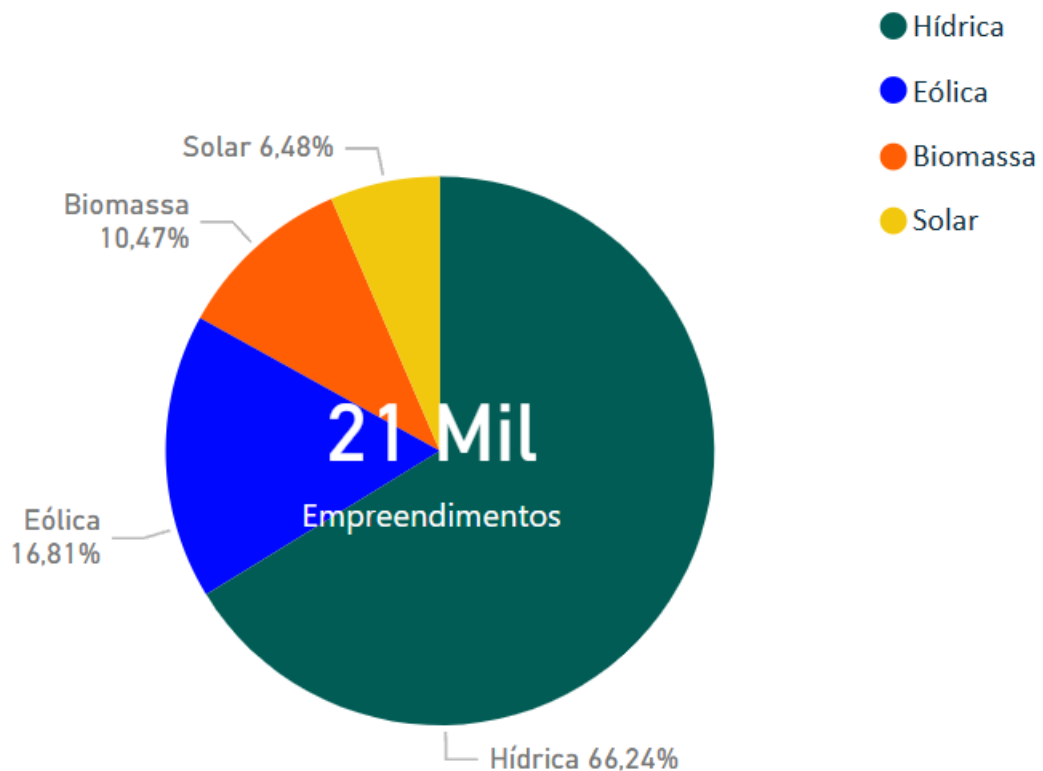


Figura 3.1: Matriz Elétrica Brasileira de Energia renovável

Fonte: Adaptado do SIGA

A Figura 3.1 demonstra a percentual de potência outorgada das principais fontes de energia renováveis na matriz elétrica brasileira no mês de novembro de 2023. A matriz hidroelétrica é responsável por 66,24.1%, em segundo lugar a energia eólica com percentual de participação de 16,81%, seguida pela biomassa e energia solar (ANEEL, 2023).

A energia hidrelétrica, contudo, possui pontos negativos relacionados a sua ins-

talação e manutenção. Entre eles, destacam-se os danos ambientais e sociais causados pelos alagamentos e as oscilações na produção elétrica, que dependem diretamente do regime de chuvas nas regiões onde estão instaladas (ALUPAR, 2022).

Muitos países, inclusive o Brasil, tem adotado ao longo dos últimos anos políticas governamentais de incentivo a diversificação da matriz energética, criando programas de expansão das denominadas fontes de energia renovável, implementando pequenas hidroelétricas, parques eólicos e usinas termoelétricas. Essas iniciativas visam reduzir a emissão de gases que acarretam aumento do efeito estufa e a dependência direta de combustíveis fósseis (ANEEL, 2008).

As fontes de energia renováveis são vistas como a melhor opção para atender à necessidade da sociedade por eletricidade confiável e ecologicamente correta, reduzindo os danos causados ao meio ambiente. Dentre as opções de energia limpa<sup>1</sup>, a energia eólica é uma das fontes com potencial para contribuir significativamente para o atendimento dos requisitos necessários em custos de produção, segurança de abastecimento e sustentabilidade ambiental (MARTINS; GUARNIERI; PEREIRA, 2008).

A energia eólica teve uma participação e contribuição significativa no setor energético brasileiro em 2022. Foram investidos US\$ 6,2 bilhões (R\$ 31,85 bilhões) no setor eólico, representando 42% dos investimentos realizados em energias renováveis no Brasil. A energia eólica encerrou 2022 com 25,63 GW de potência instalada, representando um crescimento de 18,85% em relação a 2021. A energia eólica gerou 9.277,9 MWmed no mesmo ano, totalizando 81,45 TWh, o que representa 13,52% na média de toda a geração injetada no Sistema Interligado Nacional. A geração de energia eólica foi suficiente para abastecer mensalmente, em média, 41,5 milhões de residências, representando cerca de 124 milhões de habitantes (ABEEOLICA, 2022).

## 3.2 Geração de Energia Eólica

A energia eólica é obtida a partir da energia cinética gerada pelo vento, que ao girar as hélices é convertida em energia mecânica rotacional pelo eixo rotativo. O eixo transfere sua rotação a uma caixa de engrenagens acoplada a um gerador elétrico, e este, quando

---

<sup>1</sup>Fontes de energia que acarretam baixa emissão de gases do efeito estufa (MIRANDA, 2012)

---

se encontra em rotação, realiza a conversão eletromecânica produzindo a energia elétrica (IBERDROLA, 2021).

Os fatores que determinam a quantidade de energia produzida estão relacionados à densidade do ar, à área coberta pela rotação das pás e à velocidade do vento. Esses fatores variam de acordo com aspectos da geografia física da região, como topografia, vegetação e interações térmicas entre a superfície da Terra e a atmosfera. Assim como ocorre nas hidroelétricas e em outras fontes de energia renovável, a obtenção da energia eólica existe na presença de condições naturais específicas e favoráveis. Para obter o máximo proveito da energia eólica de uma região, é necessário conhecer os dados sobre a velocidade, direção e regime do vento (ANEEL, 2008).

## 4 Fundamentação teórica

Este capítulo apresenta uma visão geral dos conceitos principais para a realização deste trabalho, apresentando características sobre a organização do setor elétrico brasileiro, geração de energia eólica, definição de Rede Neural Recorrente (RNN) e a arquitetura de redes LSTM.

Segundo (BROCKWELL et al., 2016), Uma série temporal é um conjunto de observações,  $x_t$ , cada uma registrada num tempo específico  $t$ . O objetivo de se trabalhar com séries temporais é utilizar a correlação temporal presente em sua estrutura para obter previsões de eventos futuros.

### 4.1 Redes Neurais Artificiais

O neurônio biológico é a unidade básica do sistema nervoso, responsável por receber, processar e transmitir informações. Ele é composto por um corpo celular, dendritos e um axônio. Os dendritos recebem sinais de outros neurônios e o corpo celular processa esses sinais. Se o sinal for forte o suficiente, o neurônio dispara um impulso elétrico pelo axônio, que transmite a informação para outros neurônios (ANOCHI, 2010).

A Figura 4.1 apresenta uma representação da estrutura de um neurônio biológico.

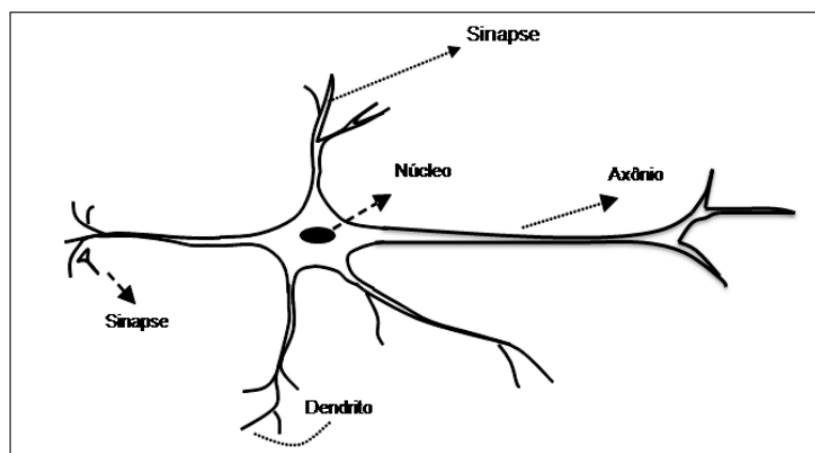


Figura 4.1: Estrutura simplificada de um neurônio biológico

Fonte: (ANOCHI, 2010)

As redes neurais artificiais são modelos matemáticos inspirados no funcionamento do sistema nervoso biológico. Elas são compostas por unidades de processamento (neurônios artificiais) interconectadas por conexões sinápticas (pesos). Cada neurônio artificial recebe entradas de outros neurônios ou do ambiente externo, processa essas entradas e produz uma saída que pode ser enviada para outros neurônios ou para fora da rede (ANOCHI, 2010).

Assim como os dendritos dos neurônios biológicos recebem sinais de outros neurônios, os neurônios artificiais recebem entradas de outros neurônios ou do ambiente externo. Da mesma forma que o corpo celular dos neurônios biológicos processa esses sinais e decide se deve disparar um impulso elétrico pelo axônio, os neurônios artificiais processam as entradas recebidas e produzem uma saída que pode ser enviada para outros neurônios ou para fora da rede (ANOCHI, 2010).

As Redes Neurais Artificiais (RNAs) são algoritmos que seguem modelos matemáticos baseados na estrutura do sistema nervoso central de organismos inteligentes, possibilitando simular o funcionamento do cérebro humano em computadores. Dessa forma, as Redes Neurais conseguem aprender e tomar decisões com base em seu próprio processo de aprendizado, armazenando e fornecendo conhecimento para a aplicação a qual são expostas (FLECK et al., 2016).

A Figura 4.2 apresenta uma representação esquemática da estrutura de um neurônio artificial. Seu funcionamento pode ser descrito da seguinte maneira: cada entrada  $x_j$  é multiplicada por um peso sináptico  $w_{kj}$  onde  $j$  é o total de entradas e  $k$  representa o índice do neurônio ao qual essa entrada está conectada. O resultado da multiplicação de todas as entradas pelos seus respectivos pesos sinápticos, é somado e aplicado sobre uma função de ativação que gera um sinal de saída do neurônio  $k$ , que representamos por  $y_k$ .

A Equação 4.1 mostra como um neurônio pode ser descrito.

$$V_k = \sum_{i=1}^j (w_{ki}x_i) + b_k \quad (4.1)$$

Os componentes do vetor de entrada são  $x_i$  para  $i = 1, 2, \dots, i$  e o peso sináptico associado é representado por  $w_{ki}$ , que conecta a entrada  $i$  ao neurônio  $k$ . A atividade do neurônio  $k$  é dada por:  $y_k = \varphi(v_k)$  onde  $\varphi$  é a função de ativação e  $y_k$  é o sinal de saída

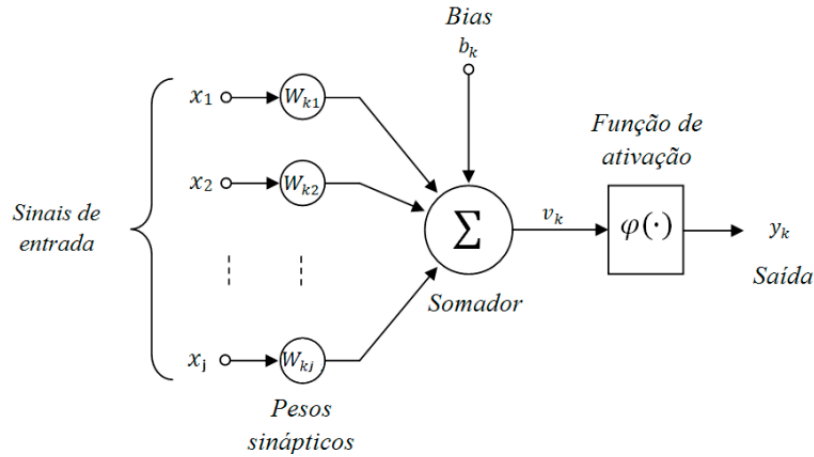


Figura 4.2: Estrutura esquemática de um neurônio artificial

Fonte: (SILVA, 2015)

do neurônio (SILVA, 2015).

Os neurônios de uma RNA devem estar conectados entre si e são dispostos em camadas, sendo que os neurônios de uma mesma camada normalmente se comportam da mesma maneira. A disposição dos neurônios nas camadas e o padrão de conexão entre estas definem a arquitetura da RNA. A camada da qual se extraem os sinais de saída são denominadas camadas visíveis, de entrada e de saída, respectivamente, e as demais camadas intermediárias, caso existentes, são denominadas de camadas escondidas. Cada conexão entre os neurônios é chamada de peso sináptico e tem um valor numérico associado. (SILVA, 2010).

Durante o processo de treinamento da rede, esses pesos são atualizados e otimizados para minimizar a diferença entre a saída prevista pela rede neural e a saída real desejada, maximizando a capacidade de generalização, produzindo saídas precisas e confiáveis para entradas que não foram incluídas no conjunto de treinamento. Em outras palavras, é a capacidade da rede de replicar esse conhecimento adquirido durante o treinamento para novos casos (NELSON, 2017).

#### 4.1.1 Função de ativação

A função de ativação é uma função matemática aplicada à saída de cada neurônio de uma rede neural artificial, transforma a saída do neurônio em um valor não linear determinando se o neurônio deve ser ativado, ou não, com base nas entradas recebidas. A

não-linearidade permite que as redes neurais aprendam relações mais complexas entre as entradas e produzam saídas mais precisas (MEDEIROS, 2019).

Existem vários tipos de funções de ativação, como a função sigmoide, a função ReLU e a função tangente hiperbólica. Cada tipo de função de ativação tem suas próprias características e é adequada para diferentes tipos de problemas.

A Figura 4.3 mostra as principais funções de ativação com seus respectivos gráficos.

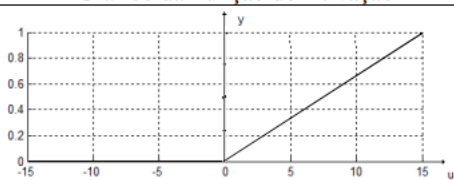
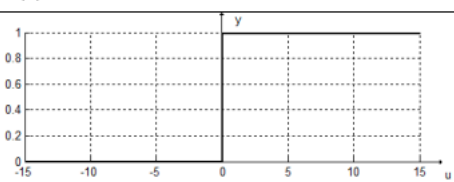
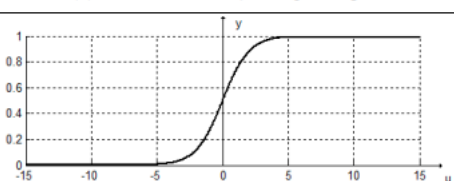
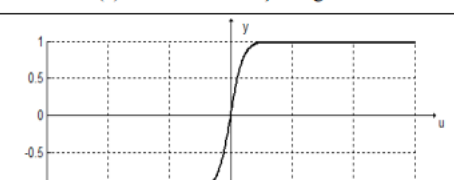
| Função de Ativação  | Gráfico da Função de Ativação   |
|---|---|
| $\text{ReLU} \varphi(v) \begin{cases} 0, & \text{se } v_k < 0 \\ v_k, & \text{se } v_k \geq 0 \end{cases}$    |  <p>(a) Gráfico de ativação unidade linear retificada</p> |
| $\text{Limiar } \varphi(v) = \begin{cases} 1, & \text{se } v_k \geq 0 \\ 0, & \text{se } v_k < 0 \end{cases}$ |  <p>(b) Gráfico de ativação degrau bipolar.</p>          |
| $\text{Sigmoide } \varphi(v) = \frac{1}{1 + \exp(-av)}$   |  <p>(c) Gráfico de ativação logística.</p>              |
| $\text{Tangente hiperbólica } \varphi(v) = \frac{1 - \exp(-av)}{1 + \exp(-av)}$                               |  <p>(d) Gráfico de ativação tangente hiperbólica.</p>   |

Figura 4.3: Funções de ativação e seus respectivos gráficos

Fonte: Adaptado de (SILVA, 2015)

Após experimentos preliminares analisando as seguintes funções de ativação: Tangente Hiperbólica (TanH), Função Sigmoide, Unidade Linear Retificada (ReLU), Unidade Linear Exponencial (ELU) e Função Softmax a ReLU foi a escolhida, pois apresentou melhores resultados em relação à eficiência computacional, capacidade de evitar o problema de gradientes desaparecendo e facilidade de otimização.

### 4.1.2 Long Short-Term Memory

A Rede de memória LSTM, do inglês *Long Short-Term Memory* é uma das muitas variações de Redes Neurais Recorrentes (RNN) utilizadas para tratar sequências longas. As LSTMs são explicitamente projetadas para evitar o problema de dependência de longo prazo, tendo como padrão de comportamento armazenar informações de longos períodos. Sua arquitetura possui portões (*gates*) que regulam a informação que entra e sai das células de memória, essas modificações são projetadas para contornar o problema da dissipação do gradiente, do inglês *Vanishing Gradient Problem*, que é particularmente perceptível em sequências longas, onde os gradientes precisam ser retropropagados por muitas camadas de neurônios (HOCHREITER, 1998). Os portões compostos por camadas de redes neurais que decidem quais informações devem ser lembradas ou esquecidas, e ajudam a manter a estabilidade do gradiente durante o treinamento, o que não ocorre nas RNNs tradicionais (BROWNLEE, 2017).

O problema da dissipação do gradiente (PDV) é um fenômeno que pode ocorrer ao treinar redes neurais profundas usando retropropagação, onde os gradientes dos pesos em relação à função de perda tornam-se muito pequenos. Isso pode dificultar o aprendizado eficaz da rede, porque o otimizador tem dificuldade em fazer atualizações significativas nos pesos com base nesses gradientes com valores muito pequenos.

Os gradientes tendem a ficar menores à medida que passam pelas camadas da rede. Isso pode ser um problema porque os pesos nas camadas anteriores são responsáveis por aprender os recursos mais importantes dos dados de entrada e, se os gradientes forem muito pequenos, esses pesos podem não ser atualizados com eficiência.

Os PVGs surgem durante o ajuste dos pesos realizado a partir do cálculo das derivadas dos erros, esses ajustes podem resultar em valores muito grandes ou muito pequenos, esses erros são denominados *blow up* e *vanish* dos gradientes, respectivamente. A subida do gradiente *blow up* (ou explosão) pode resultar em oscilações dos valores dos pesos e o decaimento do gradiente (*vanish*) pode afetar a capacidade de assimilação das dependências temporais, ocasionando uma demora no tempo e fazendo com que a rede não funcione como previsto (HOCHREITER; SCHMIDHUBER, 1997).

As LSTMs conseguem reter informações de um maior número de entradas pas-



sadas (*long-term memory*) enquanto mantém uma maior relevância dos estados recentes. Esses estados são chamados *cell state* e *hidden state*, responsáveis pelo transporte de informações através dos neurônios da rede. A rede possui portões que funcionam como mecanismo de controle de acesso às células de memória para permitir ou impedir a modificação do conteúdo da célula de memória para as várias etapas de tempo (HOCHREITER; SCHMIDHUBER, 1997).

Essas propriedades possibilitam as redes LSTM processar dados de forma separada e com interdependências complexas. Normalmente, são utilizadas para aprendizado de reconhecimento da fala, tradução automática e geração de imagem para captação (KALCHBRENNER; DANIHELKA; GRAVES, 2015).

A Figura 4.4 demonstra uma visão detalhada dos componentes internos de uma célula LSTM. É possível observar que a estrutura interna possui linhas que transportam as informações dos vetores de inteiros de um nó a outros, concatenando e bifurcando esses valores. As estruturas circulares  $\oplus$  e  $\otimes$  são operações pontuais e representam a soma dos vetores e o produto *Hadamard*<sup>2</sup>, respectivamente. Os retângulos são unidades neurais composta pelas funções de ativação  $\varphi$  (função sigmoide) e  $\tanh$  (tangente hiperbólica).

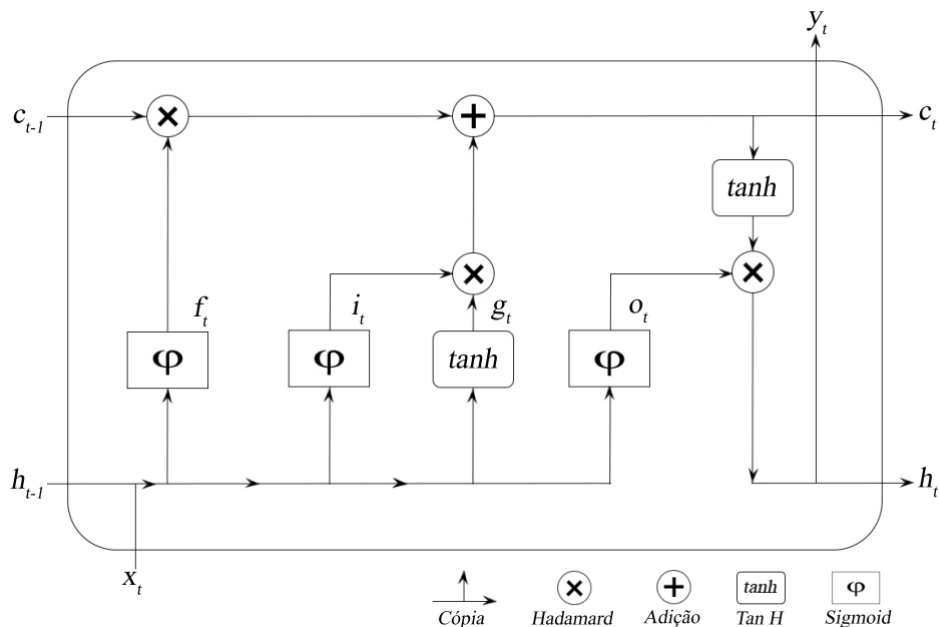


Figura 4.4: Representação de uma célula LSTM.

Fonte: (OLIVEIRA, 2020)

<sup>2</sup>O produto Hadamard é uma operação matemática que combina dois vetores ou matrizes de mesma dimensão, multiplicando cada elemento correspondente.

Para o instante de tempo  $t$ , a célula LSTM é alimentada com informações provenientes do estado atual, representada pela entrada  $x_t$  e por informações recorrentes vinda do instante de tempo anterior  $t - 1$ , sendo,  $c_{t-1}$  o estado da célula (*cell state*) no instante anterior,  $h_{t-1}$  o estado oculto (*hidden state*) no instante anterior e a saída da célula é composta pelo  $c_t$  (*cell state*) e o  $h_t$  (*hidden state*) no instante atual.

A combinação de operações realizadas internamente pela célula LSTM são denominadas como portões de controle (*gates*), sendo estes, portão de esquecimento (*Forget Gate*), portão de entrada (*Input Gate*) e portão de saída (*Output Gate*) (BROWNLEE, 2017).

Ao realiza o processo de inserir ou remover informações das células de memória de forma independente, os portões utilizam o resultado das operações que envolvem as funções de ativação, bias e matriz de pesos associadas a cada portão para decidir se as entradas são relevantes ou desprezíveis, produzindo o  $ct$ , sendo uma das saídas da célula. As seções seguintes apresentam cada uma dessas estruturas.

#### 4.1.2.1 *Forgot gate*

Este portão é o componente da célula de neurônio da LSTM responsável por esquecer uma informação mantida pelos instantes de tempos passados  $t-n, \dots, t-1$ , considerando a entrada no instante atual  $t$ . A Figura 4.5 destaca o *forgot gate* na célula LSTM. Observe as duas das entradas da célula,  $h_{t-1}$  e  $x_t$ , *hidden state* no instante anterior e a entrada no instante atual  $t$ , respectivamente. Elas se concatenam nesse ponto passando por uma função de ativação  $\varphi$  (função sigmoide) gerando  $f_t$  como saída do *forgot gate*.

A Equação 4.2 detalha o processo realizado pelo *forgot gate*.

$$f_t = \varphi(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}), \quad (4.2)$$

Os termos  $W_{if}$ ,  $W_{hf}$ ,  $b_{if}$  e  $b_{hf}$ , são as matrizes de pesos e *bias* associados a entrada do *forgot gate*, respectivamente, onde  $*_{if}$  (*input forgot*) se relaciona com a entrada  $x_t$  e  $*_{hf}$  (*hidden forgot*) à entrada  $h_{t-1}$ .

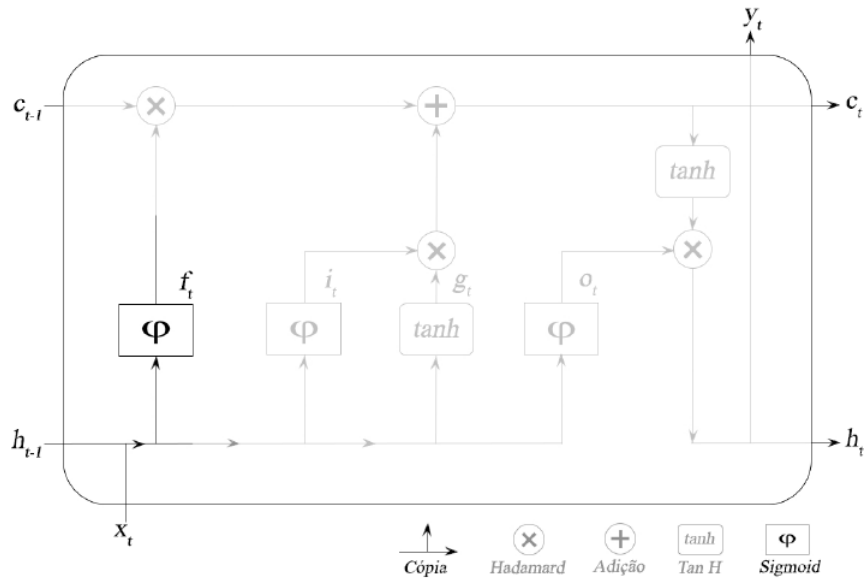


Figura 4.5: Representação do *forget gate* na célula LSTM.

Fonte: (OLIVEIRA, 2020)

#### 4.1.2.2 Input Gate e Cell Gate

O portão *input gate* é responsável por decidir o quanto de informação obtida da nova entrada será inserida na célula.

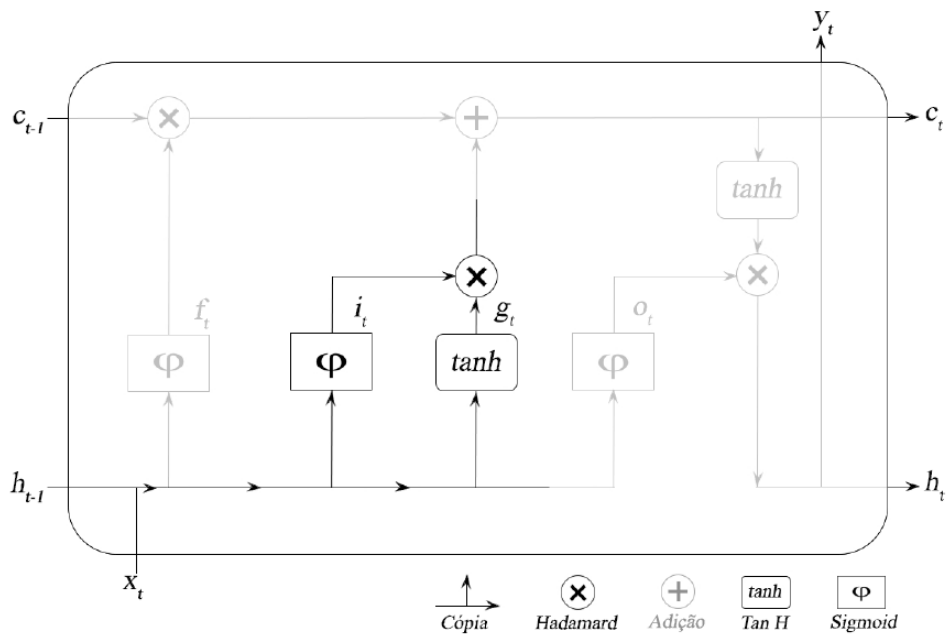


Figura 4.6: Representação do *input gate* e *cell gate* na célula LSTM.

Fonte: (OLIVEIRA, 2020)

A partir da Figura 4.6 observa-se que as duas entradas da célula,  $h_{t-1}$  *hidden state*, no instante anterior  $t - 1$ , e  $x_t$ , no instante atual  $t$ , passam por duas funções de

ativação, representadas na figura pelas funções  $\varphi$  e  $\tanh$ . No primeiro momento, o  $x_t$  e o  $h_{t-1}$  passam pela função sigmoide, quantificando entre 0 e 1, sendo 0 (menos relevante) e 1 (mais relevante) a importância da nova informação transportada por essas entradas. No segundo momento as mesmas entradas  $x_t$  e o  $h_{t-1}$  passam pela função  $\tanh$  resultando em valores entre -1 e 1. Se o valor for negativo, a informação é retirada do estado da célula e se o valor for positivo a informação é adicionada ao estado da célula (*cell state*). As equações 4.3 e 4.4 detalham o processo realizado pelo *input gate* e *cell gate* respectivamente.

$$i_t = \varphi(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}), \quad (4.3)$$

$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}), \quad (4.4)$$

Na primeira Equação 4.3 os termos  $W_{ii}$ ,  $W_{hi}$ ,  $b_{ii}$  e  $b_{hi}$ , são as matrizes de pesos e *bias* associados a entrada do *input gate*, respectivamente, onde  $*_{ii}$  se relaciona com a entrada  $x_t$  e  $*_{hi}$  à entrada  $h_{t-1}$ . Esses termos são usados como parâmetros da função sigmoide  $\varphi$  resultando a saída do *input gate*  $i_t$ .

Na segunda Equação 4.4 os termos  $W_{ig}$ ,  $W_{hg}$ ,  $b_{ig}$  e  $b_{hg}$ , são as matrizes de pesos e *bias* associados a entrada do *cell state*, respectivamente, onde  $*_{ig}$  se relaciona com a entrada  $x_t$  e  $*_{hg}$  à entrada  $h_{t-1}$ . Esses termos são usados como parâmetros da função tangente hiperbólica  $\tanh$  resultando a saída do *cell gate*  $g_t$ .

#### 4.1.2.3 Cell Update

O *Cell Update* é o estado da célula de memória responsável por atualizar o estado da célula  $c_t$  no instante de tempo atual  $t$ . A Figura 4.7 mostra onde ocorre o *cell uptade*. Destaca-se a multiplicação entre o *cell state* no instante de tempo anterior  $c_{t-1}$  e *forgot gate*  $f_t$ . O valor resultante dessa operação é somado com os valores obtidos através do produto realizado entre o *input gate*  $i_t$  e o *cell gate*  $g_t$ . O valor resultante se torna o novo *cell state*  $c_t$ , que irá servir como entrada do *cell gate* para a próxima unidade de neurônio da LSTM.

As operações envolvendo o processo realizado pelo *cell update* pode ser observada

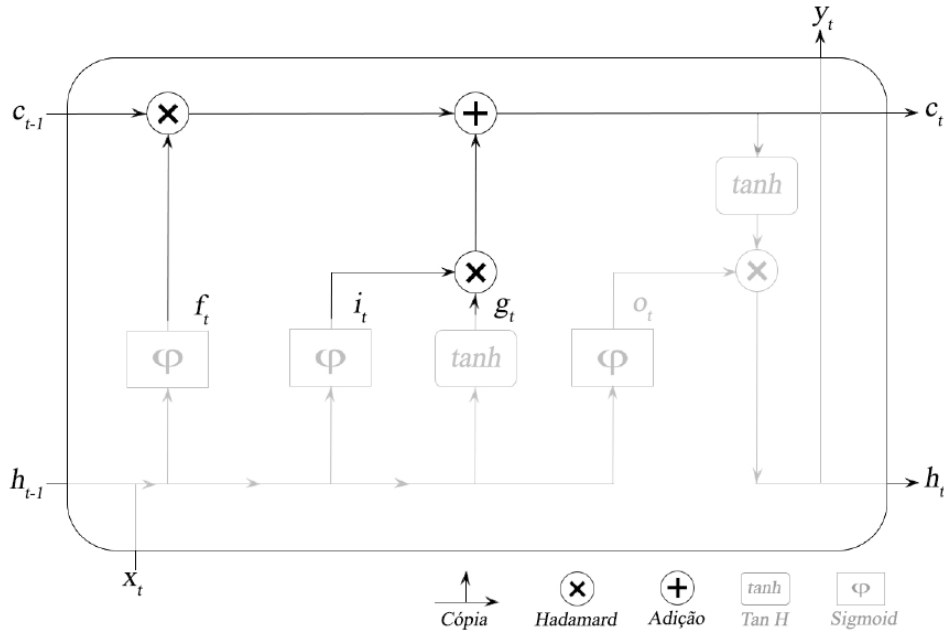


Figura 4.7: Representação da região *Cell Update* na célula LSTM.

Fonte: (OLIVEIRA, 2020)

na Equação 4.5

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (4.5)$$

O Símbolo  $\odot$  representa o produto de Hadamard, denominado como multiplicação de elementos a elementos entre matrizes de mesma dimensão.

#### 4.1.2.4 *Output Gate e Hidden State*

O portão *output gate* processa as informações que geram a saída da célula  $h_t$  (*hidden state*), responsável por propagar as informações para os instantes de tempo subsequentes  $t_{+1}$ . No último instante de tempo de execução da rede a saída da célula pode ser denominada também como  $y_t$  como mostra a Figura 4.8.

Observa-se que as entradas  $x_t$  e  $h_{t-1}$  são usadas como entrada da função de ativação sigmoide  $\varphi$  gerando assim os valores de saída do  $o_t$  (*output gate*). O  $h_t$  (*hidden state*), no instante atual  $t$ , é gerado pelo produto do  $o_t$  com o resultado da aplicação da função de ativação tangente hiperbólica  $\tanh$ , utilizando como entrada  $c_t$  (*cell state*). As Equações 4.6 e 4.7 demonstram como são calculados o  $o_t$  e o  $h_t$ .

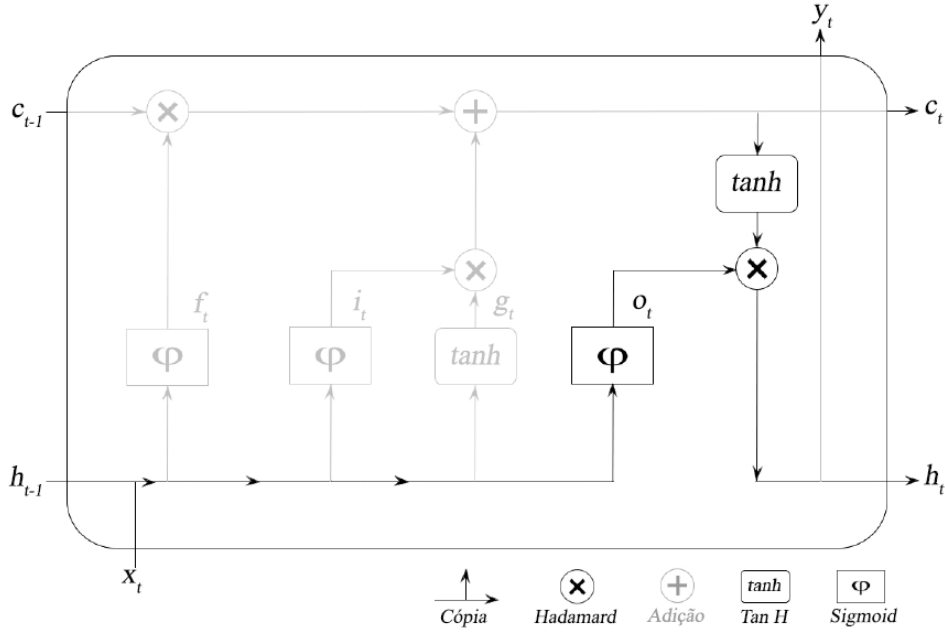


Figura 4.8: Representação do output gate na célula LSTM.

Fonte: (OLIVEIRA, 2020)

$$o_t = \varphi(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \quad (4.6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (4.7)$$

Na Equação 4.6 os termos  $W_{io}$ ,  $W_{ho}$ ,  $b_{io}$  e  $b_{ho}$ , são as matrizes de pesos e *bias* associados a entrada do *output gate*  $o_t$ , respectivamente, onde  $*_{io}$  (*input output*) se relaciona com a entrada  $x_t$  e  $*_{ho}$  (*hidden output*) à entrada  $h_{t-1}$ . A função de ativação sigmoide é representada pela letra  $\varphi$ . Na Equação 4.7 o símbolo  $\odot$  representa o produto *Hadamard* e  $h_t$  representa a saída *hidden state*.

Em geral, é desta forma que pode se descrever o funcionamento de uma célula LSTM. A composição interna dessas redes faz com que os estados internos das células sejam responsáveis pela propagação de informações de estado passado  $t_{-1}$  através da rede de uma forma que não necessite uma retroalimentação explícita, o que as distingue estruturalmente das RNNs comuns. Essa estrutura de células independentes as aproxima das multilayer perceptron <sup>3</sup> na maneira como as informações fluem em um caminho direto

<sup>3</sup>São redes neurais artificiais do tipo feedforward, compostas de várias camadas de neurônios artificiais, ou 'perceptrons'. Os Multilayer perceptrons são usados para tarefas de aprendizado supervisionado, como classificação e regressão.

através da entrada da rede até a saída (SILVA, 2015).

## 4.2 Normalização dos dados

A normalização de um conjunto de dados é um requisito comum para muitos algoritmos de aprendizado de máquina. Isso ocorre porque a normalização ajuda a garantir que os dados estejam em uma escala semelhante, possibilitando que cada característica de entrada (*input*) contribua igualmente para a capacidade de aprendizado do modelo.

Segundo (SHAH, 2017), os conjuntos de dados são divididos em três categorias: treinamento, validação e teste. O conjunto de dados de treinamento é usado para ajustar o modelo, o conjunto de dados de validação é usado para avaliar o ajuste do modelo e o conjunto de dados de teste é usado para fornecer uma avaliação imparcial do modelo.

A normalização de um conjunto de dados é um requisito comum para muitos algoritmos de aprendizado de máquina. A diferença entre as escalas das variáveis pode dificultar o aprendizado do modelo caso os valores de entrada diferem do padrão de dados normalmente distribuídos (BROWNLEE, 2010).

A base de dados utilizada é normalizada para que os dados de entrada se mantenham no intervalo entre 0 e 1. Para realizar esse procedimento é adotado o estimador *MinMaxScaler* presente na biblioteca Scikit-learn da linguagem Python, desenvolvida especificamente para aplicação prática de *MachineLearning*.

A Equação para a transformação *MinMaxScaler* é representada pela equação 4.8. Onde  $X$  é o valor original da variável,  $X_{min}$  é o valor mínimo da variável no conjunto de dados,  $X_{max}$  é o valor máximo da variável no conjunto de dados, e  $X_{scaled}$  é o valor transformado ou normalizado da variável, que fica no intervalo  $[0, 1]$ .

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4.8)$$

## 4.3 Parametrização da Rede Neural

A configuração de redes LSTM para previsão de séries temporais é um processo complexo que envolve a experimentação com diferentes configurações. O modelo de treinamento de redes neurais envolve a definição de vários hiperparâmetros que afetam diretamente o desempenho da rede como: quantidade de dados de entrada, números de camadas ocultas, números de neurônios por camada, função de ativação dos neurônios em cada camada, algoritmos de treinamento e números de épocas de treinamento. Portanto, é necessário ajustar esses valores para que o modelo seja eficaz em termos de tempo e capacidade de generalização, evitando os problemas de *overfitting*<sup>4</sup> ou *underfitting*.<sup>5</sup>

### 4.3.1 Algoritmos de otimização

Em redes neurais são utilizados algoritmos de otimização para ajustar os pesos e os *bias* das conexões entre os neurônios durante o treinamento da rede. Eles são responsáveis por encontrar os valores ótimos dos parâmetros da rede neural, de forma a minimizar a métrica de erro analisada.

O Gradiente Descendente Estocástico (*Stochastic Gradient Descent, SGD*) é um algoritmo de otimização já conhecido pela comunidade de aprendizado de máquina, de abordagem simples, e muito eficiente para ajustar classificadores lineares e regressores sob funções de perda. Ele é particularmente útil para trabalhar com grandes conjuntos de dados e pode ser aprimorado com técnicas como a taxa de aprendizado adaptativa. (SRINIVASAN, 2019).

Após a realização de experimentos preliminares, o SGD foi o otimizador escolhido por ter o processo de treinamento mais rápido e apresentar melhores resultados de previsão.

---

<sup>4</sup>Sobreajuste do modelo, onde prever bons resultados em relação ao conjunto de dados observado, mas é ineficaz para prever novos resultados.

<sup>5</sup>Subajuste do modelo, tornando incapaz de capturar a relação entre os dados de entrada e saída com precisão, gerando uma alta taxa de erro no conjunto observado e nos dados não vistos.



### 4.3.2 *DropOut*

O *DropOut* é uma técnica de regularização usada em redes neurais profundas para evitar o *overfitting* (ajuste excessivo) durante o treinamento. O *DropOut* consiste em desativar aleatoriamente alguns neurônios durante cada época de treinamento, de modo que eles não contribuam para a saída final da rede. Isso é feito de forma a reduzir a dependência de cada neurônio em relação aos outros e a evitar que a rede fique muito ajustada aos dados de treinamento.

Após realizar experimentos preliminares, o desempenho mais promissor foi obtido ao aplicar a técnica *DropOut* entre a camada de entrada e a camada oculta, atribuindo o valor de 0.1. Isso significa que, a cada época, um novo conjunto aleatório contendo 10% dos neurônios é desativado durante o processo de treinamento.

### 4.3.3 *EarlyStopping*

O *EarlyStopping* é uma técnica usada em aprendizado de máquina que possibilita determinar o número ideal de épocas de treinamento de uma rede neural. A cada época do treinamento da rede neural, o desempenho da rede no conjunto de validação é monitorado de forma que, caso o desempenho piore ou se mantenha estagnado por um número consecutivo de épocas, isso se torna um indicativo que a rede neural está sofrendo *overfitting* (ajuste excessivo), ou seja, está se ajustando aos dados de treinamento, mas não consegue generalizar o conhecimento adquirido durante o treinamento para novos dados. Neste caso o treinamento é interrompido precocemente, ou seja, antes de atingir o total de épocas para treinamento preestabelecido e evitando que a rede neural continue a se ajustar excessivamente aos dados de treinamento.

Com base nos experimentos preliminares, o *EarlyStopping* foi definido para 50 épocas, o que significa que o treinamento da rede neural será interrompido se o desempenho no conjunto de validação piorar ou se manter estagnado por 50 épocas consecutivas.

### 4.3.4 Regularizadores de peso de camada

A regularização de peso é uma técnica complementar utilizada para reduzir o ajuste excessivo nos dados de treinamento evitando o *overfitting* no modelo de rede neural de aprendizado profundo. É aplicada nos dados de treinamento para melhorar o desempenho do modelo em relação aos conjunto de dados de validação, ainda não conhecidos pela rede.

Para otimizar o cálculo do valor do peso que será aplicado nas camadas da rede é utilizado o método de regularização L2, disponíveis na API da biblioteca de programação Keras<sup>6</sup>.

Essa técnica possibilita equilibrar a complexidade do modelo para que ele possa fazer boas previsões sem ficar muito complicado. Isso é feito adicionando uma penalidade ao modelo com base nos valores dos pesos das conexões entre os neurônios da rede. Quanto maior o peso, maior a penalidade. Isso ajuda a evitar que o modelo fique muito complexo e instável.

A regularização L2 é representada pela Equação 4.9, onde

$$Loss_{regularizada} = Loss_{original} + \lambda \sum_{i=1}^n w_i^2 \quad (4.9)$$

- $Loss_{regularizada}$  é a função de perda regularizada.
- $Loss_{original}$  é a função de perda original do modelo (sem regularização).
- $\lambda$  é o parâmetro de regularização que controla a força da penalização dos pesos.
- $n$  é o número total de pesos na rede neural.
- $w_i$  são os pesos individuais da rede neural.

A penalidade é calculada ao somar o quadrado de cada coeficiente, multiplicado por um parâmetro de regularização. Esse parâmetro controla o equilíbrio entre o ajuste aos dados de treinamento e a complexidade do modelo.

---

<sup>6</sup><https://keras.io/api/layers/regularizers/l2-class>

## 4.4 Ajuste de hiperparâmetros

Para ajustar os hiperparâmetros é utilizado o Keras Tuner<sup>7</sup>, uma biblioteca de otimização de hiperparâmetros para modelos de aprendizado de máquina treinados com o framework de *deep learning* Keras. É definido um conjunto de valores para cada um dos hiperparâmetros do modelo analisados, em seguida, é utilizada a técnica de otimização Bayesiana<sup>8</sup> para encontrar o conjunto ótimo de hiperparâmetros para o problema de aprendizado de máquina (INVERNIZZI JAMES LONG, 2021).

O conjunto de valores descritos na Tabela 4.2 foi escolhido com base em experimentos preliminares, realizados visando encontrar uma configuração que proporcionasse um bom desempenho do modelo. Os experimentos foram realizados utilizando o mesmo conjunto de dados de treinamento e validação e teste, porém com o número total de épocas reduzido, dessa forma, é possível realizar uma avaliação rápida para os diferentes valores de hiperparâmetros.

| Hiperparâmetros                   | Intervalo de Valores  |
|-----------------------------------|-----------------------|
| N.º Neurônios (Camada de Entrada) | 512, 712 ou 1024      |
| N.º Neurônios (Camada Oculta)     | 512, 712 ou 1024      |
| N.º de Camadas Ocultas            | 1 ou 2                |
| <i>DropOut</i>                    | 0.1, 0.2, 0.3 ou 0.4  |
| Taxa de Aprendizado do SGD        | 0.01, 0.001 ou 0.0001 |

Tabela 4.1: Conjunto de hiperparâmetros de entrada do KerasTuner

Os resultados demonstraram que o número de neurônios na camada de entrada e na camada oculta é um dos hiperparâmetros mais importantes para o desempenho do modelo. Modelos com mais neurônios nas respectivas camadas apresentam um desempenho melhor, mas são propensos ao *overfitting* e possuem alto custo computacional, enquanto modelos com menos neurônios nas camadas tem um custo computacional reduzido, porém não apresentava bons resultados.

Para os demais hiperparâmetros, a faixa de valores foi determinada com base em uma análise dos resultados de experimentos preliminares. Os resultados mostraram que os valores escolhidos proporcionam um bom desempenho do modelo, com um equilíbrio entre

<sup>7</sup>Keras é uma API de aprendizado profundo multiestruturada

<sup>8</sup>A otimização Bayesiana é uma técnica de otimização que combina uma distribuição de probabilidade sobre os parâmetros de um modelo com informações empíricas para encontrar o valor ótimo dos parâmetros

tempo de execução, custo computacional e menor diferença percentual entre os valores previstos e os valores reais.

Para o ajuste dos hiperparâmetros utilizando Keras Tuner, foi definido com base em testes preliminares, o conjunto 20 combinações distintas, repetindo o treinamento 3 vezes para cada combinação. Isso significa que o algoritmo irá explorar  $20 \times 3 = 60$  combinações de hiperparâmetros. A cada passo, o algoritmo de Otimização Bayesiana ajusta a distribuição de probabilidade dos hiperparâmetros com base nos resultados obtidos. Esse procedimento é repetido até que o algoritmo descubra os melhores hiperparâmetros, ou seja, aqueles que maximizam a precisão em relação ao conjunto de validação.

A Tabela 4.2 demonstra a relação de todos os hiperparâmetros definidos com base no processo de otimização e nos experimentos preliminares.

Tabela 4.2: Conjunto de hiperparâmetros de entrada da LSTM

| <b>Hiperparâmetro</b>                         | <b>Valor</b> |
|---|--------------|
| N. <sup>o</sup> Neurônios (Camada de Entrada) | 1024         |
| N. <sup>o</sup> Neurônios (Camada Oculta)     | 512          |
| N. <sup>o</sup> de Camadas Ocultas            | 1            |
| Número máximo de épocas                       | 300          |
| Função de Atização                            | ReLU         |
| <i>DropOut</i>                                | 0.1          |
| Algoritmo de Otimização                       | SGD          |
| <i>EarlingStopping</i>                        | 50 épocas    |
| Taxa de Aprendizado do otimizador             | 0.01         |

## 5 Conjunto de dados

O conjunto de dados utilizado para o treinamento das redes LSTM contém séries temporais de geração verificada e velocidade prevista do vento, de 1º de janeiro de 2017 a 31 de dezembro de 2019, com discretização de 30 minutos. Para cada dia, há um conjunto de 48 registros, correspondendo aos dados de 0h a 23h30.

Deste modo, a base de dados é constituída por 52.560 amostras que compõem as entradas dos modelos abordados, contendo os seguintes campos:

- Data e hora: data e hora referente ao dia e intervalo da medição;
- Geração Verificada: valor da geração registrada no horário da medição;
- Velocidade prevista do vento: estimativa prevista da velocidade do vento para o horário da medição.

A série temporal é dividida em três conjuntos distintos:

- Conjunto de dados de treinamento: contém informações do período de 01/01/2017 até 30/06/2018 (26.040 amostras).
- Conjunto de dados de validação: contém informações do período de 01/07/2018 até 31/12/2018 (8.688 amostras)
- Conjunto de dados de teste: contém informações do período de 01/01/2019 até 31/12/2019 (17.376 amostra).

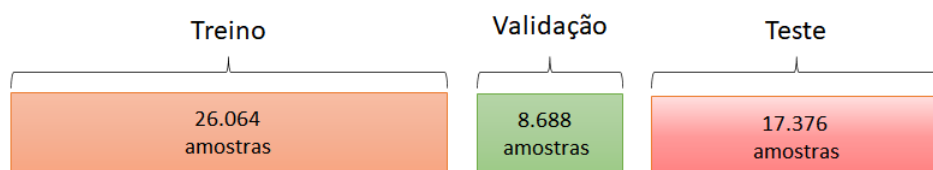


Figura 5.1: Divisão da base de dados em conjuntos

Fonte: Autor

Para realizar o treinamento é necessário determinar o intervalo de tempo a ser utilizado como memória, ou seja, a quantidade de dados passados (*lookback*), a quantidade de dados previsto (*lookforward*) e o conjunto de dados composto pela quantidade variáveis de entrada do modelo.

Neste trabalho foram realizados experimentos em dois modelos de previsão, com conjunto de variáveis de entradas distintos, demonstrados nas Figuras 5.2 e 5.3.

- Univalorado: Considera apenas dados de geração verificada, onde  $D$  representa o dia de previsão,  $D_{-n}$ , onde,  $n$  representa o número de dias de dados passados utilizados para prever o dia seguinte  $D_{+1}$ . A entrada de dados é representada por  $(D_{-n}, \dots, D_{-2}, D_{-1})$ .

Por exemplo, se  $n=2$ , estamos usando os dados dos dois dias anteriores (ontem e anteontem) para prever o dia de amanhã,  $D_{+1}$ .

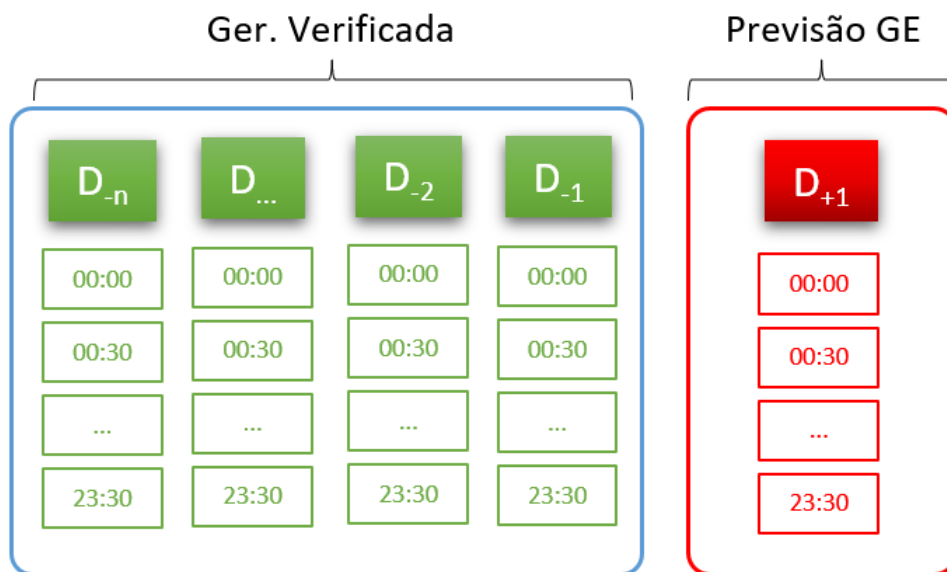


Figura 5.2: Modelo de previsão univalorado

Fonte: Autor

- Multivalorado: Considera os dados de geração verificada e velocidade de vento previsto, onde  $D$  representa o dia de previsão,  $D_{-n}$ , onde,  $n$  representa o número de dias de dados passados utilizados para prever o dia seguinte  $D_{+1}$ . A entrada de dados é composta por duas variáveis:

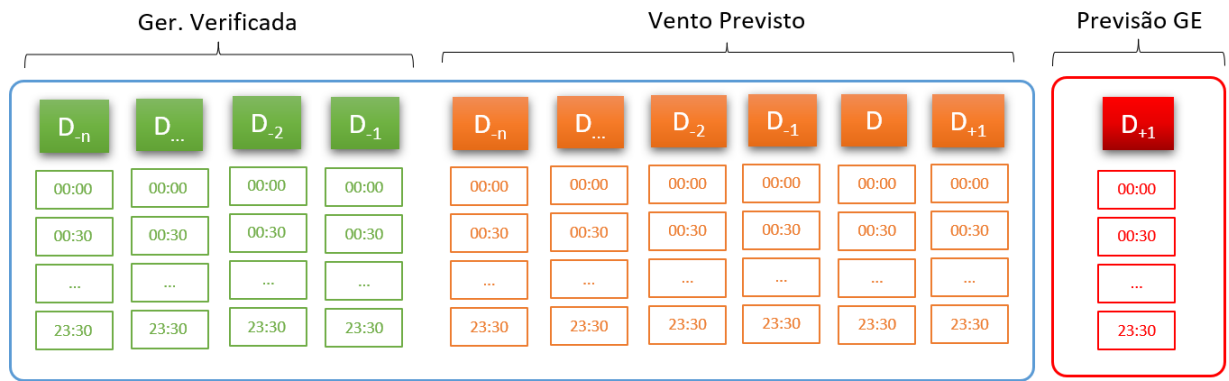


Figura 5.3: Modelo de previsão multivalorado

Fonte: Autor

- Geração verificada:  $(D_{-n}, \dots, D_{-2}, D_{-1})$
- Velocidade prevista do vento:  $(D_{-n}, \dots, D_{-2}, D_{-1}, D, D_{+1})$ .

Por exemplo, se  $n=2$ , os dados de geração verificada inclui os dados de 2 dias passados, (ontem e anteontem), enquanto para os dados de velocidade de vento previsto, são considerados os dias (anteontem, ontem, hoje e amanhã). O tamanho da entrada de dados é determinado pela concatenação dos dados de geração verificada e velocidade prevista do vento.

## 5.1 Caracterização e análise dos dados

O processo de treinamento dos modelos adotados neste trabalho segue a lógica de um treinamento supervisionado. Esse método implica alimentar o modelo com dados prévios das séries temporais, permitindo que a rede neural aprenda os padrões e relações entre os dados. Durante o treinamento, o modelo ajusta seus parâmetros com base na diferença entre o que é previsto e os valores reais, esses ajustes permitem que a rede aprenda as relações e seja capaz de generalizar de forma a realizar previsões precisas em conjuntos de dados desconhecidos. Essa abordagem é essencial para a caracterização e análise dos dados, fornecendo uma base sólida para compreender as nuances das séries temporais provenientes de sete usinas eólicas distintas.

Para avaliar o desempenho do modelo ao longo das épocas, é utilizada a combinação das métricas de perda. A perda (*loss*) é calculada nos dados de treinamento,

enquanto a validação de perda (*valloss*) é calculada nos dados de validação.

As métricas são usadas para avaliar o desempenho do modelo e para ajustar seus parâmetros de modo que a perda seja minimizada. Quanto menor o valor da perda, melhor o modelo se adapta aos dados de treinamento.

A perda é calculada pela função de perda, erro quadrado médio (MSE), que mensura a média dos erros quadrados entre as previsões do modelo e os valores observados, descrita pela Equação 5.1.

$$MSE = \frac{1}{n} \sum_{i=1}^n (V_i - P_i)^2 \quad (5.1)$$

Onde  $n$  é número de observações,  $V_i$  são os valores verificados e  $P_i$  são os valores previstos.

Para avaliar o desempenho do modelo em relação ao conjunto de teste é utilizado a métrica *NMAPE*, erro percentual médio normalizado pela potência instalada de cada usina. Conforme descrito na equação 5.2.

$$NMAPE(\%) = \frac{1}{N \cdot P_{ins}^{tot}} \sum_{i=1}^N |(P_{verif}^i - P_{prev}^i)| \times 100\% \quad (5.2)$$

Onde  $N$  é número de amostras,  $P_{ins}^{tot}$  é a potência total instalada da usina,  $P_{verif}^i$  são os valores verificados e  $P_{prev}^i$  são os valores previstos.

Cada usina possui um valor da potência instalada, observados na tabela 5.1.

| Usina | Potência Instalada |
|-------|--------------------|
| U1    | 12.2657            |
| U2    | 11.8856            |
| U3    | 11.5592            |
| U4    | 9.5177             |
| U5    | 9.7434             |
| U6    | 10.3301            |
| U7    | 12.3335            |

Tabela 5.1: Potência Instalada do conjunto de Usinas



## 6 Metodologia

### 6.1 Séries Temporais das Usinas

A base de dados utilizada neste trabalho é obtida por meio de uma parceria com Operador Nacional do Sistema Elétrico (ONS)<sup>9</sup>. São utilizados dados de 7 usinas eólicas, uma de cada estado brasileiro com capacidade instalada de geração de energia eólica significativa. Por se tratar de dados privados, os conjunto de dados foram modificados com a finalidade de manter a sua confiabilidade, porém, preservando as características das séries temporais, não impactando nos resultados obtidos pelos modelos.

O histórico de geração verificada não transformado é coletado pelo Sistema de Supervisão do ONS em conjunto com o Sistema de Medição de Faturamento da Câmara de Comercialização de Energia Elétrica (CCEE) (discretização de hora e 5 minutos). As séries históricas não transformadas de velocidade prevista do vento, são provenientes do modelo *Global Forecast System (GFS)* e do *National Oceanic and Atmospheric Administration (NOAA)*, com resolução espacial de 25 km e temporal de 1h, com previsão estendida para o dia corrente e alguns dias para frente. (ONS, 2020).

A Figura 6.1 ilustra a origem da base de dados construída no ONS para dar suporte as previsões eólicas desenvolvidas nesse trabalho.



Figura 6.1: Base de dados de fontes renováveis eólicas do ONS

Fonte: Adaptado de: (ONS, 2020)

<sup>9</sup><https://www.ons.org.br/>

As Figuras 4.2 até 4.8 ilustram as séries históricas das usinas U1, U2, U3, U4, U5, U6 e U7, contendo dados de Geração verificada (MW) e Velocidade prevista do vento (m/s) compreendidos entre os anos de 2017 a 2019, totalizando 52.560 registros.

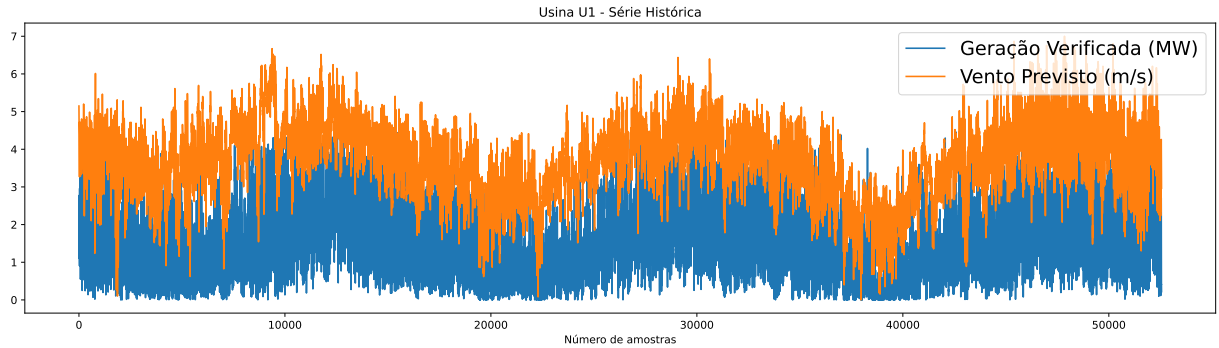


Figura 6.2: Série histórica da Usina 1

A série temporal da Usina 1 mostra um padrão de geração de energia relativamente sazonal com flutuações previsíveis que pode sugerir uma dependência forte e direta da velocidade do vento.

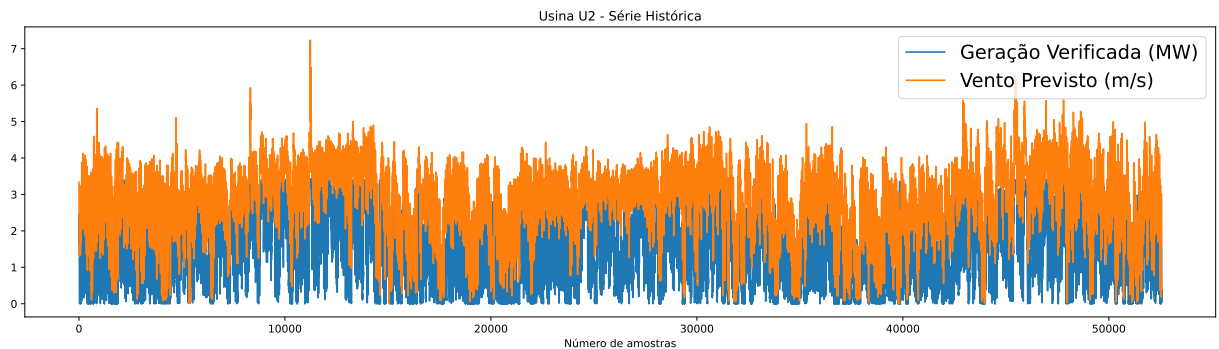


Figura 6.3: Série histórica da Usina 2

A série temporal da Usina 2 pode revelar uma variabilidade maior na geração de energia, com picos abruptos e quedas acentuadas.

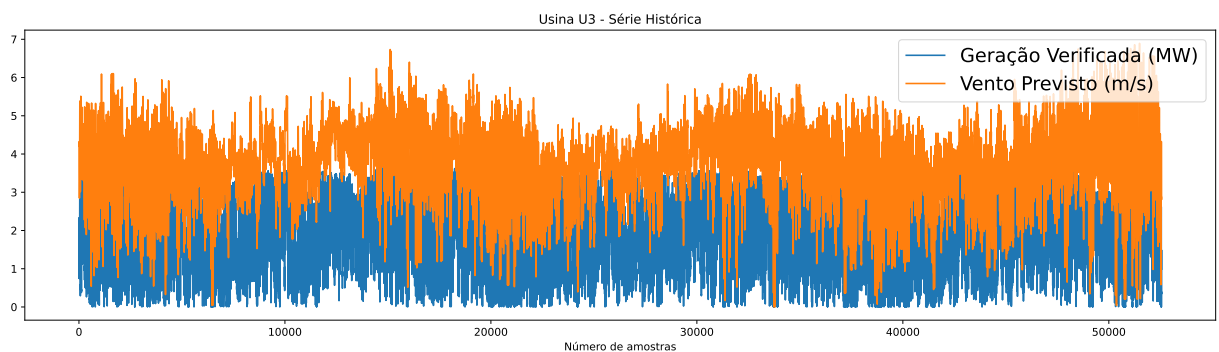


Figura 6.4: Série histórica da Usina 3

A série temporal da Usina 3 apresenta um variabilidade baixa da geração de energia em relação a intensidade do vento.

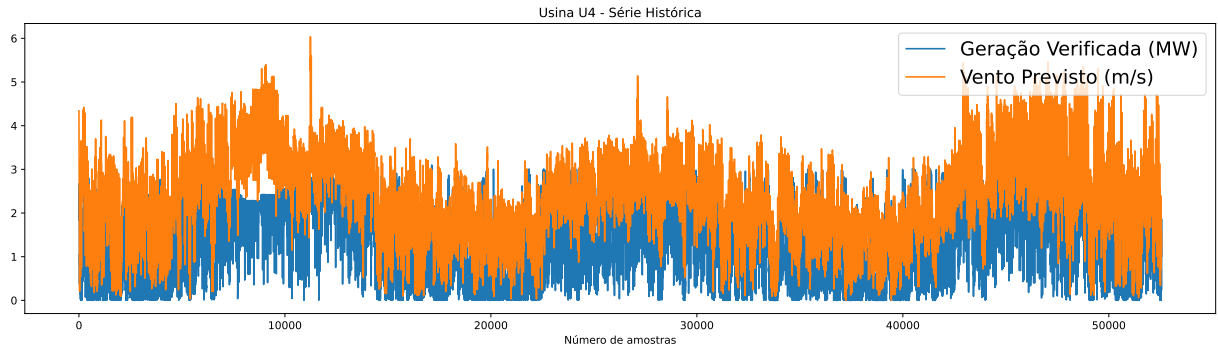


Figura 6.5: Série histórica da Usina 4

A série temporal da Usina 4 pode indicar períodos alternados de alta e baixa produção de energia, que podem ser atribuídos a um ciclo específico de variação da intensidade do vento na região.

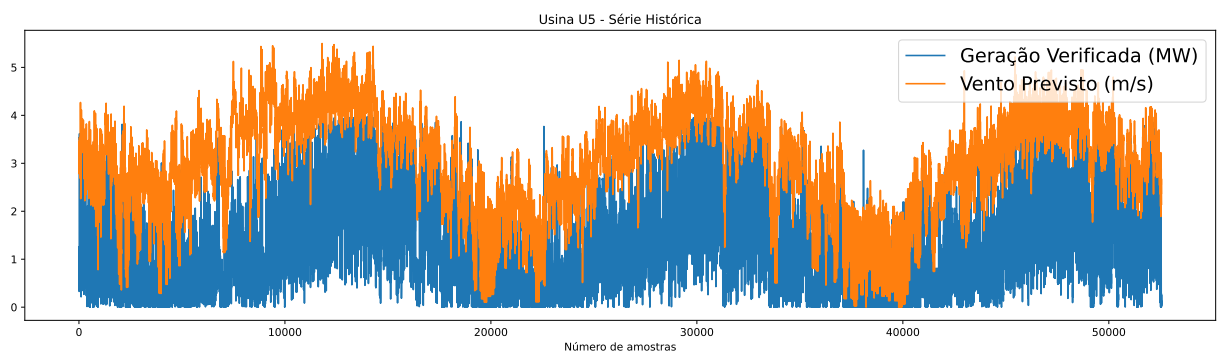


Figura 6.6: Série histórica da Usina 5

A série temporal da Usina 5 apresenta um padrão sazonal claro, com períodos de alta produção de energia que pode indicar a correspondência direta com as estações específicas do ano.

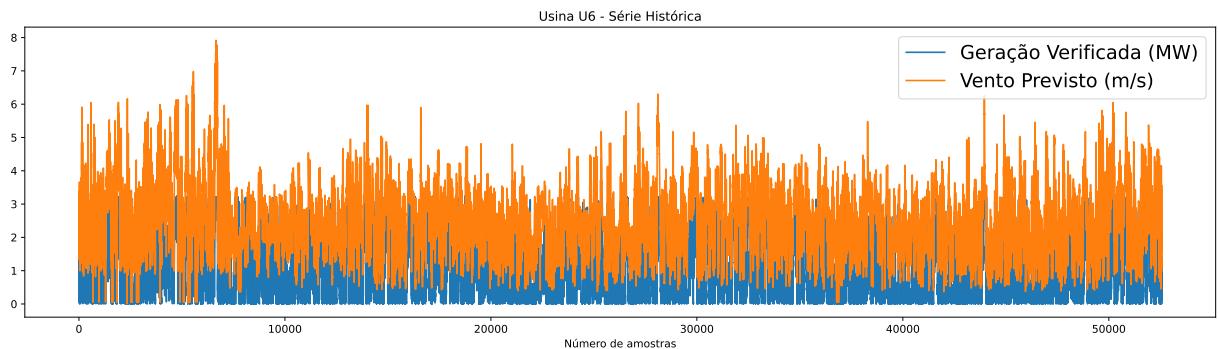


Figura 6.7: Série histórica da Usina 6

A série temporal da Usina 6 pode indicar uma produção de energia muito consistente com variações mínimas. embora apresente pico abruptos em relação a velocidade do vento.

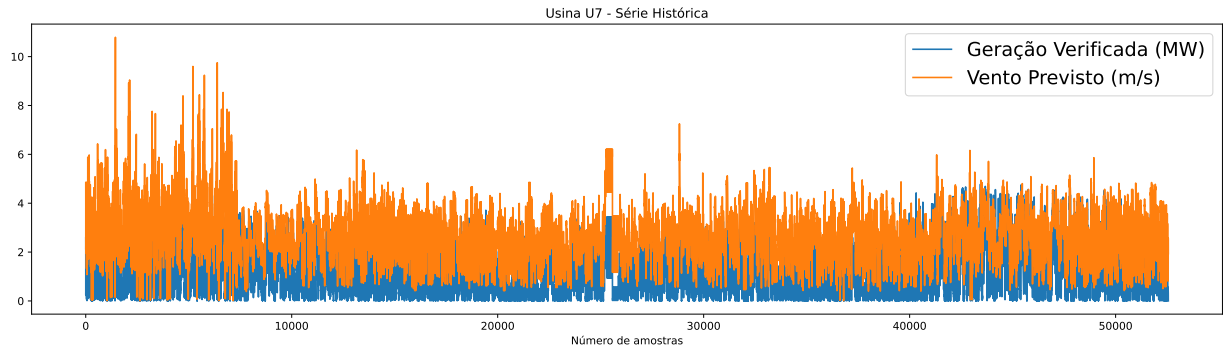


Figura 6.8: Série histórica da Usina 7

A série temporal da Usina 7 pode indicar uma produção de energia muito consistente com variações mínimas.

## 6.2 Pré-Processamento do Conjunto De Dados

Os dados relativos às séries temporais de geração de fonte eólica e velocidade do vento verificada e prevista são resultantes das diversas fontes de informação colhidas pelo ONS. Como a discretização temporal difere entre as fontes, os dados são transformados em discretização de 30 minutos. Embora alguns estimadores apresentem falhas, os indicadores das medidas coletadas garantem a qualidade dos dados.

Segundo (ONS, 2020), o tratamento dos dados de geração considerados espúrios segue, de forma simplificada, as seguintes etapas:

- Correlacionar dados do ONS com dados da CCEE de 1h;
- Eliminar resultados ruins do Estimador de Estados do ONS ;
- Eliminar dados onde:  $Geração < 0$  (Megawatt) ou  $Geração > 1.1 \times P_{inst}$  (Megawatt);
- Eliminar dados congelados (ao longo dos dias e das horas);
- Eliminar dados dos dias com menos de 2,5 horas válidas.

As previsões de vento coletadas pelo ONS mostram desvios em relação aos dados reais de vento. Há uma tendência observada para que as previsões de vento se desviem dos dados de vento verificados em quase todos os parques eólicos, oferecendo vantagens no tratamento de dados de previsão de geração de fontes eólicas.

Os arquivos oriundos desse modelo numérico são recebidos por quadrícula. Sendo assim é necessário associar cada usina à sua respectiva quadrícula<sup>10</sup>. Há um algoritmo que gera arquivos associando cada usina à sua correspondente quadrícula em relação aos arquivos de velocidade prevista do vento.

A Figura. 6.9 e a Figura. 6.10 ilustram as quadrículas para o Nordeste e Sul, respectivamente, do modelo GFS. Nestas figuras são marcados pontos em vermelho representando exemplos de usinas eólicas consideradas pelo ONS.

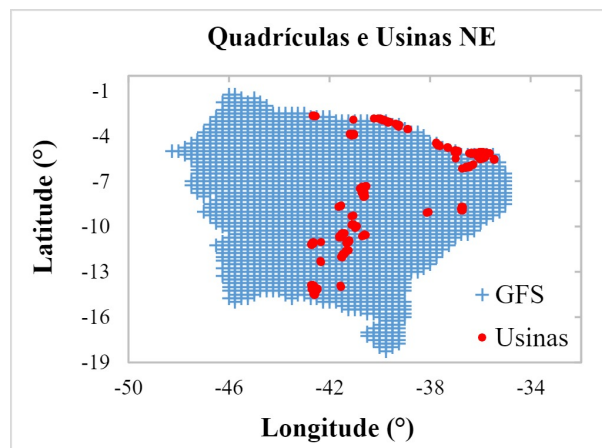


Figura 6.9: Quadrículas do modelo GFS para Nordeste.

Fonte:(ONS, 2020)

<sup>10</sup>As quadrículas do sistema GFS são uma divisão do espaço em retângulos de tamanhos iguais, usadas para representar os dados meteorológicos gerados pelo sistema GFS (NOAA, 2023).

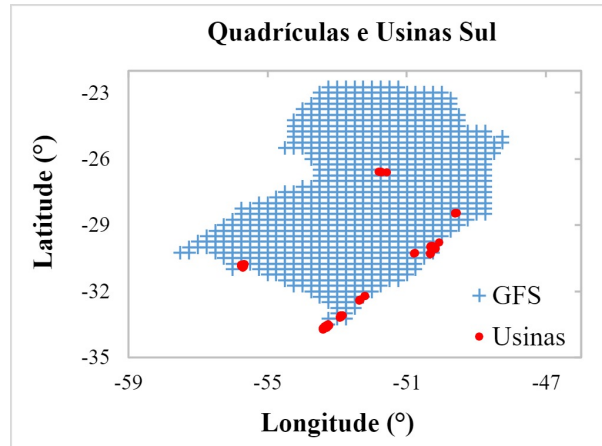


Figura 6.10: Quadrículas do modelo GFS para Sul.

Fonte:(ONS, 2020)

## 6.3 Cenário de Teste

### 6.3.1 LSTM Univalorado

O modelo univalorado denominado LSTM-UNI, utiliza apenas a série histórica de geração de energia eólica como variável endógena. Para este modelo, são realizados dois cenários:

- Cenário 1: considera 3 dias passados (144 dados de entrada) para prever 1 dia (48 dados).
- Cenário 2: considera 7 dias passados (336 dados de entrada) para prever 1 dia (48 dados).

A Tabela 6.1 mostra os hiperparâmetros para o modelo LSTM-UNI para os dois cenários.

Tabela 6.1: Conjunto de hiperparâmetros utilizados no modelo univalorado

| Hiperparâmetros                   | Cenário 1           | Cenário 2           |
|-----------------------------------|---------------------|---------------------|
| N.º Neurônios (Camada de Entrada) | 1024                | 1024                |
| N.º Neurônios (Camada Oculta)     | 512                 | 512                 |
| N.º de Camadas Ocultas            | 1                   | 1                   |
| Número máximo de épocas           | 300                 | 300                 |
| <i>DropOut</i>                    | 0.1                 | 0.1                 |
| Taxa de Aprendizagem do SGD       | 0.01                | 0.01                |
| Variáveis                         | Geração Eólica (MW) | Geração Eólica (MW) |
| Dias Anteriores                   | 3 dias              | 7 dias              |

Os cenários são organizados conforme a Tabela 6.1. Cada cenário difere pelo

número de dias passados observados. O objetivo é verificar a eficiência do modelo ao aumentar o horizonte de dias observados.

### 6.3.2 LSTM Multivalorado

O modelo multivalorado, denominado LSTM-MULT, utiliza a série histórica de geração de energia eólica e a série histórica de velocidade prevista do vento como entradas. Nesse modelo, também são realizados 2 cenários:

Cenário 1: considera 3 dias de dados passados de geração verificada e 5 dias de dados de velocidade de vento previsto.

Cenário 2: considera 7 dias de dados passados de geração verificada e 9 dias de dados de velocidade de vento previsto.

A Tabela 6.2 mostra os cenários de testes para o modelo LSTM-MULT, acrescentando a série de velocidade prevista do vento como variável exógena.

Tabela 6.2: Conjunto de hiperparâmetros utilizados no modelo multivalorado

| Hiperparâmetros                   | Cenário 1   | Cenário 2   |
|-----------------------------------|---|---|
| N.º Neurônios (Camada de Entrada) | 1024  | 1024  |
| N.º Neurônios (Camada Oculta)     | 512   | 512   |
| Número máximo de épocas           | 300   | 300   |
| N.º de Camadas Ocultas            | 1   | 1   |
| <i>DropOut</i>                    | 0.1   | 0.1   |
| Taxa de Aprendizado do SGD        | 0.1   | 0.1   |
| Variáveis                         | Geração Eólica(MW) e<br>Velocidade do vento<br>previsto (m/s) | Geração Eólica(MW) e<br>Velocidade do vento<br>previsto (m/s) |
| Dias Anteriores                   | 3 dias  | 7 dias  |

Os cenários do modelo LSTM-MULT são organizados de acordo com o número de dias passados observados. O modelo avalia o impacto da inclusão de uma variável exógena na rede neural, bem como o quanto a combinação dos dados (número de variáveis de entrada e quantidade de dados observados) contribui para a eficiência do modelo.

## 7 Resultados experimentais

Para obter as métricas de desempenho, são executados 5 treinamentos para cada cenário de cada modelo e calculando a média dos valores. A comparação entre os modelos e cenários considera apenas o conjunto de teste, que contém dados relativos a todo o ano de 2019.

### 7.0.1 Resultados comparativos dos cenários Univalorados

A Tabela 4.3 contém os valores médios da métrica NMAPE das 7 usinas avaliadas para os 2 cenários. Cenário 1 considera o conjunto 144 dados de entrada, referente a 3 dias de dados passados, e o Cenário 2 com 7 dias de dados passados.

Tabela 7.1: Resultados comparativos entre cenários do modelo LSTM-UNI

| <b>LSTM-UNI NMAPE Pinst(%)</b> |                           |                           |
|--------------------------------|---------------------------|---------------------------|
| <b>Usinas</b>                  | <b>Cenário 1 (3 Dias)</b> | <b>Cenário 2 (7 Dias)</b> |
| U1                             | 12.754                    | 11.821                    |
| U2                             | 20.083                    | 19.150                    |
| U3                             | 18.788                    | 17.855                    |
| U4                             | 18.212                    | 17.279                    |
| U5                             | 11.944                    | 11.012                    |
| U6                             | 24.473                    | 23.540                    |
| U7                             | 17.448                    | 16.515                    |

A 7.1 mostra que o modelo treinado por 7 dias de dados observados obteve um desempenho superior ao modelo treinado por 3 dias. Isso é observado em todas as unidades (U1 a U7), com uma diferença melhora no desempenho que varia de 3.8% a 7.2%. Esse resultado sugere que um treinamento com um número maior de entrada de dados apresenta um bom desempenho com um modelo LSTM-UNI. Isso ocorre porque um treinamento com maior quantidade de dados permite que o modelo aprenda mais sobre os dados de treinamento, o que resulta em um modelo mais preciso.



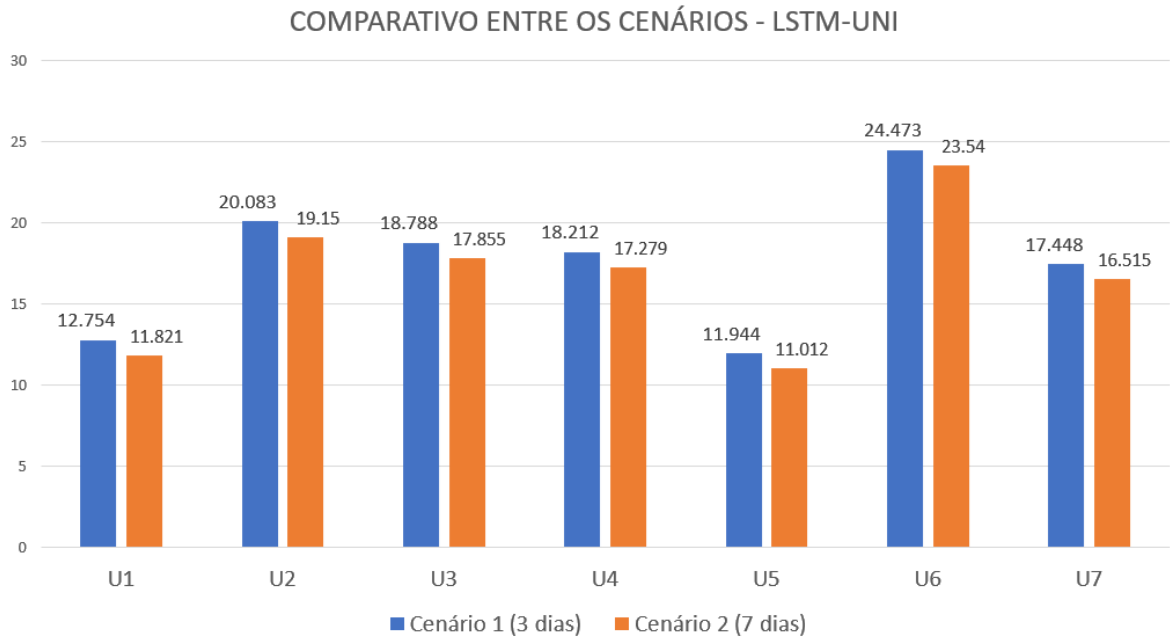


Figura 7.1: Comparativo entres os cenários do modelo LSTM-UNI

## 7.0.2 Resultados comparativos dos cenários Multivalorados

Na Tabela 4.3 está demonstrado os valores médios da métrica NMAPE das 7 usinas avaliadas para os dois cenários. Cenário 1 considera o conjunto 144 dados de entrada, referente a 3 dias de dados passados e o Cenário 2 com 7 dias de dados de entrada, totalizando 336 dados.

Tabela 7.2: Resultados comparativos entre cenários do modelo LSTM-MULT

| <b>LSTM-MULT NMAPE Pinst(%)</b> |                           |                           |
|---------------------------------|---------------------------|---------------------------|
| <b>Usinas</b>                   | <b>Cenário 1 (3 Dias)</b> | <b>Cenário 2 (7 Dias)</b> |
| U1                              | 11.393                    | 11.137                    |
| U2                              | 14.038                    | 13.391                    |
| U3                              | 17.521                    | 16.765                    |
| U4                              | 15.494                    | 14.008                    |
| U5                              | 10.405                    | 9.997                     |
| U6                              | 17.557                    | 16.429                    |
| U7                              | 14.808                    | 14.505                    |

A Figura 7.2 mostra que o modelo no cenário 2 tem um melhor desempenho do que o cenário 1 em todas as colunas. Essa melhora na precisão da previsão da geração pode ser explicada pelo fato de que a geração verificada e a velocidade prevista do vento são duas variáveis que estão correlacionadas. A geração verificada é uma medida da quantidade de energia elétrica que foi gerada por uma turbina eólica em um determinado

período de tempo. A velocidade prevista do vento é uma estimativa da velocidade do vento que será observada em um determinado tempo. Ao considerar as duas variáveis, o modelo de previsão é capaz de aprender a relação entre elas e, assim, prever a carga elétrica eólica com mais precisão.

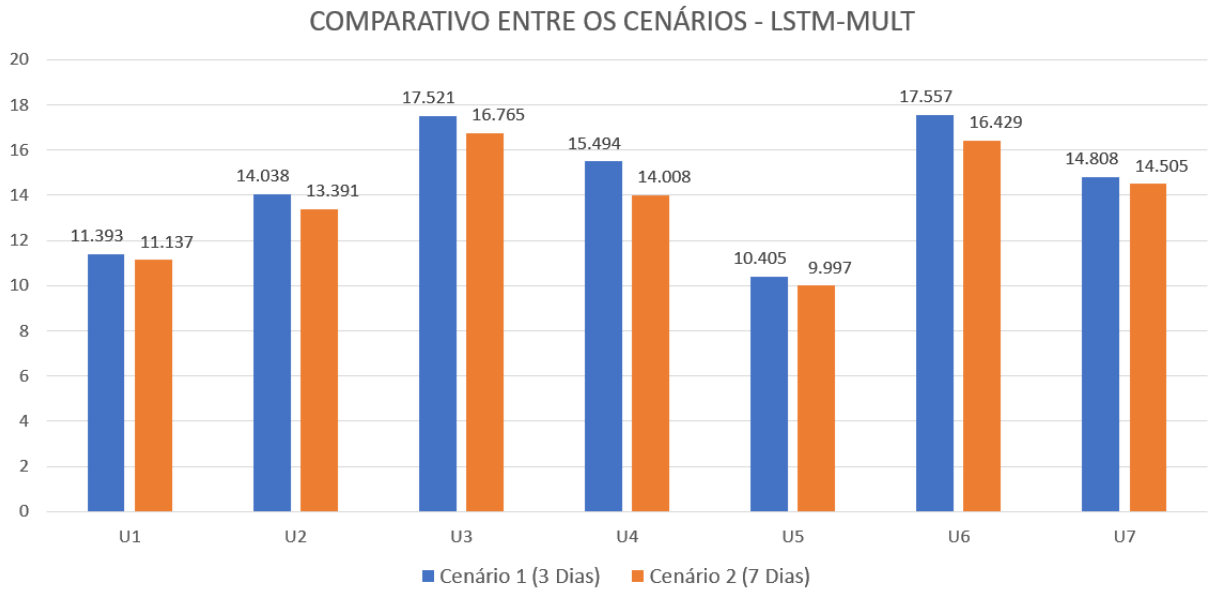


Figura 7.2: Comparativo entres os cenários do modelo LSTM-MULT

### 7.0.3 Análise dos resultados dos Modelos

Os experimentos realizados com os modelos LSTMUNI e LSTM-MULT para previsão de cargas elétricas de sistemas de energia eólica apresentaram resultados interessantes.

Com base nos resultados apresentados na Tabela 7.3, é possível concluir as redes LSTM são capazes de melhorar a precisão das previsões de séries temporais ao considerar um horizonte de dados maior e a inclusão de variáveis exógenas relevantes.

As redes LSTM são projetadas para capturar padrões e tendências que se estendem por períodos de tempo prolongados. Isso as torna particularmente adequadas para séries temporais, que podem ser caracterizadas por dependências de longo prazo. Ao aumentar o número de dias passados observados, o modelo pode incorporar uma visão mais abrangente da evolução da série temporal, o que é crucial para a previsão precisa.

No contexto da geração de energia eólica, a velocidade do vento é um fator determinante da geração de energia. A inclusão da velocidade do vento como variável exógena no modelo LSTM-MULT permitiu uma modelagem mais precisa da realidade física da

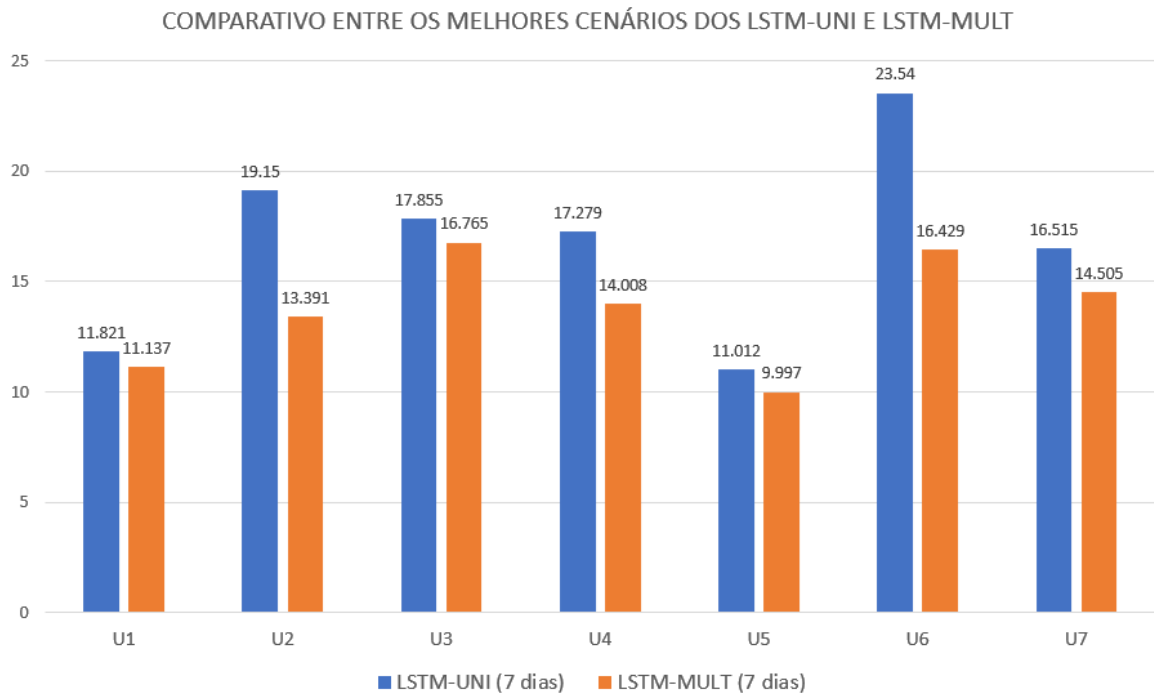


Figura 7.3: Comparativo entres os modelos LSTM-UNI e LSTM-MULT

geração de cargas elétricas eólicas. Isso está alinhado com a literatura e com pesquisas anteriores que demonstram a eficácia de incorporar variáveis exógenas em modelos de previsão para melhorar a precisão.

O modelo LSTM-MULT foi superior ao modelo LSTM-UNI em ambos os cenários testados. Isso sugere que a arquitetura LSTM é capaz de capturar as relações complexas e não-lineares entre múltiplas variáveis de entrada. Essa capacidade é essencial para previsões precisas em sistemas dinâmicos como a geração de energia eólica.

## 8 Conclusão

As redes neurais recorrentes baseadas em LSTM são provavelmente a abordagem mais poderosa para aprender com dados sequenciais e séries temporais são apenas um caso especial. O potencial dos modelos baseados em LSTM é totalmente revelado ao aprender com grandes conjuntos de dados onde podemos detectar padrões complexos. Esses resultados reforçam a importância de modelos avançados de séries temporais no apoio à tomada de decisões no Sistema Interligado Nacional, contribuindo para uma gestão mais eficiente e sustentável da geração de energia eólica. A pesquisa destaca o potencial das redes LSTM em aplicações energéticas e abre caminho para futuros estudos que explorem ainda mais variáveis e modelos híbridos, visando otimizar a previsão de cargas elétricas eólicas.

### 8.1 Trabalhos futuros

Considerando a relevância das redes neurais recorrentes baseadas em LSTM para aprender com dados sequenciais e analisando os resultados apresentados, fica claro a importância de incluir variáveis exógenas de alta correlação com a variável geração eólica, pois essas variáveis podem fornecer informações adicionais sobre os padrões de variação da geração eólica e, conseqüentemente, melhorar a precisão das previsões.

No caso específico de séries de geração eólica, incorporar mais variáveis exógenas além da velocidade do vento, outras variáveis meteorológicas, como temperatura, umidade e pressão atmosférica, podem influenciar a geração de energia eólica. Estudos futuros poderiam explorar o impacto da inclusão dessas variáveis nos modelos LSTM para verificar se há um aumento na precisão das previsões.

A utilização de dados com maior resolução temporal, como medições a cada 15 minutos ou menos, pode proporcionar percepções mais detalhadas sobre as flutuações na geração de energia eólica e, portanto, melhorar as previsões.

## Bibliografia

ABEEOLICA. Boletim anual 2022. ABEEolica, 2022. Disponível em: <https://abeeolica.org.br/wp-content/uploads/2023/06/Boletim-de-Geracao-Eolica-2022.pdf>. Acesso em: 19 de novembro de 2022.

ABEEÓLICA. *Usinas eólicas dominam expansão da oferta de energia elétrica em julho*. 2023. Disponível em: <https://abeeolica.org.br/usinas-eolicas-dominam-expansao-da-oferta-de-energia-eletrica-em-julho/>. Acesso em: 05 de Novembro de 2023.

ALUPAR. Setor elétrico brasileiro. 2022. Disponível em: <http://ri.alupar.com.br/a-companhia/setor-eletrico-brasileiro/>. Acesso em: 22 de Junho de 2022.

ANEEL. Atlas de energia elétrica do brasil. Aneel, p. 233, 2008. Disponível em: <http://www2.aneel.gov.br/arquivos/PDF/atlas3ed.pdf>. Acesso em: 28 de Junho de 2022.

ANEEL. *Sistema de Informações de Geração da ANEEL SIGA*. 2023. Disponível em: <https://lnkd.in/d/mevsbH>. Acesso em: 20 de abril de 2023.

ANOCHI, J. A. *Modelos baseados em redes neurais para o estudo de padrões climáticos sazonais a partir de dados tratados com a teoria dos conjuntos aproximativos*. Tese (Doutorado) — Tese de Doutorado. M. Sc. dissertation on Applied Computing—Intituto Nacional de Pesquisas Espaciais., 2010. Acesso em: 01 de Maio de 2023.

BROCKWELL, P. J. et al. *Introduction to time series and forecasting*. [S.l.]: Springer, 2016.

BROWNLEE, J. *How to Use StandardScaler and MinMaxScaler Transforms in Python*. 2010. Disponível em: <https://machinelearningmastery.com/standardscaler-and-minmaxscaler-transforms-in-python/>. Acesso em: 14 de Julho de 2022.

BROWNLEE, J. *Long short-term memory networks with python: develop sequence prediction models with deep learning*. [S.l.]: Machine Learning Mastery, 2017.

FERREIRA, A. A. B. *Metodologia de previsão do potencial eólico de curto prazo para planejamento da operação de sistemas elétricos*. Dissertação (Mestrado) — Dissertação apresentada ao curso de Pós-Graduação Stricto Sensu da Universidade Federal do Pampa., 2012. Disponível em: <https://cursos.unipampa.edu.br/cursos/ppgee/files/2010/03/ferreira.dissertation.pdf>. Acesso em: 14 de Maio de 2023.

FLECK, L. et al. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, v. 1, n. 13, p. 47–57, 2016.

FRANÇA, C. G. et al. Análise comparativa de modelos de previsão de geração de energia eólica baseados em machine learning. *Revista de Sistemas e Computação-RSC*, v. 9, n. 2, 2020.

GUTIÉRREZ, J. L. *Monitoramento da instrumentação da Barragem de Corumbá-I por redes neurais e modelos de Box & Jenkins*. Dissertação (Mestrado) — Dissertação apresentada ao Programa de Pósgraduação em Engenharia Civil da PUC-Rio, 2003. Disponível em: [https://www.maxwell.vrac.puc-rio.br/4244/4244\\_5.PDF](https://www.maxwell.vrac.puc-rio.br/4244/4244_5.PDF). Acesso em: 14 de Maio de 2023.

HOCHREITER, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, World Scientific, v. 6, n. 02, p. 107–116, 1998.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural computation*, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.

HOSSAIN, M. S.; MAHMOOD, H. Short-term load forecasting using an lstm neural network. In: *2020 IEEE Power and Energy Conference at Illinois (PECI)*. [S.l.: s.n.], 2020. p. 1–6.

HUANG, Q.; WANG, X. A forecasting model of wind power based on ipso-lstm and classified fusion. *Energies*, Multidisciplinary Digital Publishing Institute, v. 15, n. 15, p. 5531, 2022. Disponível em: <http://dx.doi.org/10.3390/en15155531>. Acesso em: 19 de Julho de 2022.

IBERDROLA. *Energia Eólica*. Iberdrola, 2021. Disponível em: <https://www.iberdrola.com/sustentabilidade/energia-eolica>. Acesso em: 04 de Julho de 2022.

INVERNIZZI JAMES LONG, F. C. T. O. H. J. L. *Getting started with KerasTuner*. 2021. Disponível em: [https://keras.io/guides/keras\\_tuner/getting\\_started/](https://keras.io/guides/keras_tuner/getting_started/). Acesso em: 25 de Dezembro de 2022.

KALCHBRENNER, N.; DANIHELKA, I.; GRAVES, A. Grid long short-term memory. 07 2015.

LIU, B. et al. A novel deep learning approach for wind power forecasting based on wd-lstm model. *Energies*, MDPI, v. 13, n. 18, p. 4964, 2020. Disponível em: <https://www.mdpi.com/1996-1073/13/18/4964>. Acesso em: 14 de Julho de 2022.

LIU, Y. et al. Wind power short-term prediction based on lstm and discrete wavelet transform. *Applied Sciences*, v. 9, n. 6, 2019. ISSN 2076-3417. Disponível em: <https://www.mdpi.com/2076-3417/9/6/1108>.

MARTINS, F.; GUARNIERI, R.; PEREIRA, E. O aproveitamento da energia eólica. *Revista Brasileira de Ensino de Física*, v. 30, p. 1304, 01 2008.

MEDEIROS, I. Aplicação de redes neurais artificiais para a previsão de demanda em subestações de distribuição considerando a influência da temperatura. *Trabalho de Conclusão de Curso (Bacharelado em Engenharia Elétrica)*, Universidade Federal da Paraíba, João Pessoa, 73p, 2019. Disponível em: <http://www.cear.ufpb.br/juan/wp-content/uploads/2022/07/2019-Aplica%C3%A7%C3%A3o-de-redes-neurais-artificiais-para-a-previs%C3%A3o-de-demanda-em-subesta%C3%A7%C3%B5es-de-distribui%C3%A7%C3%A3o-considerando-a-influ%C3%Aancia-da-temperatura.pdf>. Acesso em: 26 de novembro 2023.

MIRANDA, M. M. d. Fator de emissão de gases de efeito estufa da geração de energia elétrica no Brasil: implicações da aplicação da avaliação do ciclo de vida. 2012. Disponível em: <https://www.teses.usp.br/teses/disponiveis/18/18139/tde-22012013-112737/publico/DissertacaoMarianaMaiaDeMiranda.pdf>. Acesso em: 25 de novembro de 2023.

NELSON, D. M. Q. Uso de redes neurais recorrentes para previsão de séries temporais financeiras. Universidade Federal de Minas Gerais, 2017.

NOAA. *Global Forecast System (GFS)*. 2023. Disponível em: <https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast>. Acesso em: 26 de novembro 2023.

OLIVEIRA, E. V. d. *Análise de desempenho de método baseado em rede LSTM para classificação de falhas em um processo de controle de nível*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, 2020.

ONS. *O Planejamento da operação energética no sistema : conceitos, modelagem matemática, previsão de geração e carga*. 2020. Disponível em: <http://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Livros-da-Diretoria-de-Planejamento-na-biblioteca-digital-ONS/LIVRO-O-PLANEJAMENTO-DA-OPERA%C3%87%C3%83O-ENERG%C3%89TICA-NO-SISTEMA%20INTERLIGADO-NACIONAL.pdf>. Acesso em: 12 de Julho de 2022.

PRADA, R. B. *Modelo de previsão de carga elétrica considerando a influência da mini e microgeração distribuída no perfil da curva de carga*. Dissertação (Mestrado) — Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio, 2021. Acesso em: 14 de Maio de 2023.

SHAH, T. *Conjuntos de treinamento, validação e teste*. 2017. Disponível em: <https://tarangshah.com/blog/2017-12-03/train-validation-and-test-sets/>. Acesso em: 14 de Julho de 2022.

SILVA, C. A. d. A. *Implementação de uma matriz de neurônios dinamicamente reconfigurável para descrição de topologias de redes neurais artificiais multilayer perceptrons*. Tese (Doutorado) — Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da UFRN., 2015. Disponível em: [http://www.nupeg.ufrn.br/documentos\\_finais/teses\\_de\\_doutorado/teses/carlosalbertodealbuquerque-silva-out.pdf](http://www.nupeg.ufrn.br/documentos_finais/teses_de_doutorado/teses/carlosalbertodealbuquerque-silva-out.pdf). Acesso em: 01 de Maio de 2023.

SILVA, M. *Aplicação de Redes Neurais Artificiais no Diagnóstico de Falhas de Turbinas a Gás*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio, 2010. Disponível em: [https://www.maxwell.vrac.puc-rio.br/16580/16580\\_4.PDF](https://www.maxwell.vrac.puc-rio.br/16580/16580_4.PDF). Acesso em: 20 de agosto 2023.

SRINIVASAN, A. V. *Stochastic Gradient Descent — Clearly Explained !!* [S.l.]: towardsdatascience.com, 2019. Acesso em: 14 de Julho de 2022.