



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

NATHAN FAUSTINO RIBEIRO

**PREVISÃO DE VELOCIDADE DO VENTO COM DADOS REAIS DE UM PARQUE
EÓLICO LOCALIZADO NO CEARÁ UTILIZANDO REDES NEURAIAS
ARTIFICIAIS: UM ESTUDO COMPARATIVO DAS REDES MLP, LSTM, GRU E
CNN**

FORTALEZA
2024

NATHAN FAUSTINO RIBEIRO

PREVISÃO DE VELOCIDADE DO VENTO COM DADOS REAIS DE UM PARQUE
EÓLICO LOCALIZADO NO CEARÁ UTILIZANDO REDES NEURAS ARTIFICIAIS:
UM ESTUDO COMPARATIVO DAS REDES MLP, LSTM, GRU E CNN

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Engenharia Elétrica do
Centro de Tecnologia da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Arthur Plínio de Souza
Braga

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- R37p Ribeiro, Nathan Faustino.
Previsão de velocidade do vento com dados reais de um parque eólico localizado no Ceará utilizando redes neurais artificiais : um estudo comparativo das redes MLP, LSTM, GRU e CNN / Nathan Faustino Ribeiro. – 2024.
70 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia Elétrica, Fortaleza, 2024.
Orientação: Prof. Dr. Arthur Plínio de Souza Braga.
1. Redes Neurais Artificiais; 2. Velocidade do vent. 3. Energia Eólica. I. Título.
- CDD 621.3
-

NATHAN FAUSTINO RIBEIRO

PREVISÃO DE VELOCIDADE DO VENTO COM DADOS REAIS DE UM PARQUE
EÓLICO LOCALIZADO NO CEARÁ UTILIZANDO REDES NEURAIAS ARTIFICIAIS:
UM ESTUDO COMPARATIVO DAS REDES MLP, LSTM, GRU E CNN

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Engenharia Elétrica do
Centro de Tecnologia da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia Elétrica.

Aprovada em: 03/10/2024

BANCA EXAMINADORA

Prof. Dr. Arthur Plínio de Souza Braga (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Lucas Silveira Melo
Universidade Federal do Ceará (UFC)

Prof. Dra. Tatiane Carolyne Carneiro
Universidade Federal do Maranhão (UFMA)

A Deus.

Aos meus pais, Jouseline e Márcio.

AGRADECIMENTOS

Primeiramente a Deus, aquele que é a fonte de tudo o que somos, fomos ou seremos, pelo privilégio da vida e pela força que Ele me concede para superar os obstáculos que surgem no meu caminho.

Aos meus familiares, por sempre me darem suporte, acreditarem no poder dos estudos e dedicação, e ensinarem a ter gratidão pelos desafios e conquistas da vida.

Aos meus amigos de vivência acadêmica, que sempre estiveram me apoiando e instruindo, pelas conversas edificadoras e o companheirismo.

Aos meus colegas de trabalho, por acreditarem e me motivarem no processo de realização deste trabalho.

Ao Professor Arthur, pela orientação prestada e o empenho demonstrado por meio de suas valorosas contribuições.

Ao Professor Lucas e à Professora Tatiane, por terem aceitado participar da banca examinadora desse trabalho.

A todos que, direta ou indiretamente, contribuíram na minha formação pessoal, acadêmica e profissional.

“Quem olha para fora sonha, quem olha para dentro desperta.” (Carl Jung).

RESUMO

O presente trabalho faz um estudo comparativo do desempenho de diferentes Redes Neurais Artificiais para a previsão de velocidade do vento futuro a partir da série histórica de medições lida a partir de uma torre de medição anemométrica localizada na cidade de Aracati-CE, no Parque Eólico Quixaba. Destaca a importância desse tipo de previsão para melhorar o planejamento da operação de usinas eólicas, promover a integração no SIN (Sistema Interligado Nacional) dessa fonte e melhorar a estabilidade da rede elétrica. Menciona os diversos tipos de previsão de velocidade do vento e energia eólica que são utilizados e os trabalhos relacionados com as arquiteturas de redes neurais que foram definidas para a construção dos modelos preditivos (MLP, LSTM, GRU e CNN) mostrando a importância e eficácia desses modelos para o horizonte de tempo avaliado. O trabalho também oferece uma breve introdução às Redes Neurais Artificiais, descrevendo a estrutura de cada uma das arquiteturas exploradas e explicando o processo de aprendizado baseado no algoritmo de retropropagação de erro, que ajusta os pesos das redes. Os dados de velocidade do vento foram tratados e utilizados para otimizar os hiperparâmetros por meio da técnica de validação cruzada aplicada a séries temporais. Com a definição dos hiperparâmetros, os modelos foram construídos e treinados com dados de velocidade do vento com intervalos de 10 minutos dos anos de 2021 a 2022 e, posteriormente, testados com os dados de 2023. Os modelos foram avaliados a partir das métricas MAE, RMSE, MAPE, R^2 e por meio da análise variância (ANOVA), sendo considerados eficazes na previsão da velocidade do vento para a localidade de Quixabá. Dentre os modelos implementados o modelo baseado na rede GRU obteve os melhores resultados, com um MAE de 0,47 m/s, um RMSE de 0,64 m/s, um MAPE de 7,55% (indicando que, em média, a previsão do modelo diverge 7,47% do valor real) e um R^2 de 0,90. Com os resultados analisados, foi possível concluir que considerando o custo dos recursos computacionais envolvido, os modelos baseados nas redes MLP e CNN poderiam ser utilizados em produção para horizontes de previsão de curto período. Contudo, destaca que se o intuito for a busca do máximo desempenho, os modelos baseados nas redes LSTM e GRU são mais recomendados.

Palavras-chave: Redes Neurais Artificiais; Previsão; Velocidade do vento; Energia Eólica.

ABSTRACT

The present work conducts a comparative study of the performance of different Artificial Neural Networks for predicting future wind speed based on the historical wind speed series read from an anemometric measurement tower located in the city of Quixaba-CE. It highlights the importance of this type of forecasting to improve the operational planning of wind farms, promote the integration of this energy source into the National Interconnected System (SIN), and enhance the stability of the power grid. The work discusses the various types of wind speed and wind energy forecasting methods that are utilized, as well as the related studies on the neural network architectures that were defined for constructing the predictive models (MLP, LSTM, GRU, and CNN), demonstrating the significance and effectiveness of these models for the evaluated time horizon. The work also provides a brief introduction to Artificial Neural Networks, describing the structure of each explored architecture and explaining the learning process based on the backpropagation algorithm, which adjusts the network weights. The wind speed data were processed and used to optimize the hyperparameters through the cross-validation technique applied to time series. With the hyperparameters defined, the models were built and trained with wind speed data at 10-minute intervals from 2021 to 2022 and subsequently tested with data from 2023. The models were evaluated using the metrics MAE, RMSE, MAPE, R^2 , and through variance analysis (ANOVA), and were considered effective in predicting wind speed for the locality of Quixaba-CE. Among the implemented models, the model based on the GRU network obtained the best results, with a MAE of 0.47 m/s, a RMSE of 0.64 m/s, a MAPE of 7.55% (indicating that, on average, the model's prediction diverges 7.47% from the real value) and an R^2 of 0.90. With the results analyzed, it was possible to conclude that, considering the cost of the computational resources involved, the models based on the MLP and CNN networks could be used in production for short period forecast horizons. However, it highlights that if the aim is to seek maximum performance, models based on LSTM and GRU networks are more recommended.

Keywords: Artificial Neural Networks; Prediction; Wind Speed; Wind Power.

LISTA DE FIGURAS

Figura 1 - Matriz de energia elétrica brasileira.....	14
Figura 2 - Diagrama de fluxo das etapas do trabalho.....	16
Figura 3 - Fluxo de vento no rotor do aerogerador.....	20
Figura 4 - Neurônio artificial.....	25
Figura 5 - Rede MLP.....	26
Figura 6 - Camada LSTM.....	28
Figura 7 - Camada GRU.....	29
Figura 8 - Rede CNN.....	31
Figura 9 - Localização da TMA.....	35
Figura 10 - Série histórica das medições de velocidade do vento (m/s).....	37
Figura 11 – Série histórica da média horária da velocidade do vento (m/s).....	37
Figura 12 – Série histórica da média diária da velocidade do vento (m/s).....	37
Figura 13 – Série histórica da média mensal da velocidade do vento (m/s).	38
Figura 14 - Distribuição da série histórica da média horária da velocidade do vento (m/s). ...	38
Figura 15 - Distribuição da série histórica da média diária da velocidade do vento (m/s).	39
Figura 16 - Distribuição da série histórica da média mensal da velocidade do vento (m/s). ...	39
Figura 17 - Entradas e saídas para o treinamento das redes neurais.....	41
Figura 18 - Validação cruzada com janela expansiva.....	43
Figura 19 – (a) Distribuição das amostras por grupo com valor F baixo. (b) Distribuição das amostras por grupo com valor de F alto.	45
Figura 20 - Velocidade do vento Prevista versus Real em janeiro de 2023.....	49
Figura 21 - Distribuição dos erros de previsão em janeiro de 2023.....	50
Figura 22 - Velocidade do vento Prevista versus Real em fevereiro de 2023.	60
Figura 23 Distribuição dos erros de previsão em fevereiro de 2023.	60
Figura 24 - Velocidade do vento Prevista versus Real em março de 2023.....	61
Figura 25 - Distribuição dos erros de previsão em março de 2023.	61
Figura 26 - Velocidade do vento Prevista versus Real em abril de 2023.	62
Figura 27 - Distribuição dos erros de previsão em abril de 2023.....	62
Figura 28 - Velocidade do vento Prevista versus Real em maio de 2023.....	63
Figura 29 - Distribuição dos erros de previsão em maio de 2023.	63
Figura 30 - Velocidade do vento Prevista versus Real em junho de 2023.....	64

LISTA DE TABELAS

Tabela 1 - Funções de ativação.....	26
Tabela 2 - Pseudocódigo do algoritmo de retropropagação de erro.	34
Tabela 3 - Sensores instalados na TMA.	36
Tabela 4 - Resumo estatístico das medições de velocidade do vento (m/s).	40
Tabela 5 – Hiperparâmetros da rede MLP.	47
Tabela 6 - Hiperparâmetros da rede LSTM.	47
Tabela 7 - Hiperparâmetros da rede GRU.	48
Tabela 8 - Hiperparâmetros da rede CNN.	48
Tabela 9 - Erro Médio Absoluto (MAE) em 2023.....	50
Tabela 10 - Raiz do Erro Quadrático Médio (RMSE) em 2023.	51
Tabela 11 - Erro Percentual Médio Absoluto (MAPE) em 2023.....	51
Tabela 12 - Coeficiente de Determinação (R^2) em 2023.....	52
Tabela 13 - Teste ANOVA.	53

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
API	<i>Application Programming Interface</i> (Interface de Programação de Aplicações)
ARIMA	<i>Auto-Regressive Integrated Moving Average</i> (Modelo Auto-Regressivo Integrado de Médias Móveis)
CNN	<i>Convolutional Neural Network</i> (Rede Neural Convolucional)
EVS	<i>Explained Variance Score</i> (Pontuação de Variância Explicada)
GBM	<i>Gradient Boosting Machine</i>
GRU	<i>Gated Recurrent Unit</i> (Unidade Recorrente com Portas)
LSTM	<i>Long Short-Term Memory</i> (Memória de Longo Curto Prazo)
MAE	<i>Mean Absolute Error</i> (Erro Médio Absoluto)
MAPE	<i>Mean Absolute Percentage Error</i> (Erro Percentual Absoluto Médio)
MLP	<i>Multilayer Perceptron</i> (Perceptron Multicamadas)
RMSE	<i>Root Mean Square Error</i> (Raiz do Erro Quadrático Médio)
R^2	<i>R-Squared</i> (Coeficiente de Determinação)
RNA	Rede Neural Artificial
STC-DPN	<i>Spatiotemporal Conversion Deep Predictive Network</i> (Rede Preditiva Profunda de Conversão Espaço-Temporal)
TMA	Torre de Medição Anemométrica

LISTA DE SÍMBOLOS

C_p	Coeficiente de potência do aerogerador
E	Energia gerada pelo aerogerador
K	Filtro ou <i>kernel</i>
E_r	Função perda
I	Matriz de entrada
W_z	Matriz de pesos do portão de atualização associada a entrada
U_z	Matriz de pesos do portão de atualização associada ao estado oculto anterior
W_i	Matriz de pesos do portão de entrada
W_f	Matriz de pesos do portão de esquecimento
W_r	Matriz de pesos do portão de reinicialização associada a entrada
U_r	Matriz de pesos do portão de reinicialização associada ao estado oculto anterior
W_o	Matriz de pesos do portão de saída
ρ	Massa específica do ar
i	Portão de entrada
f	Portão de esquecimento
o	Portão de saída
z	Portão de atualização
r	Portão de reinicialização
P	Potência do aerogerador
w	Peso sináptico
η	Taxa de aprendizado
t	Tempo
$S(i, j)$	Valor da convolução 2D na posição (i, j)
b_z	Vetor de viés do portão de atualização
b_i	Vetor de viés do portão de entrada
b_f	Vetor de viés do portão de esquecimento
b_r	Vetor de viés do portão de reinicialização
b_o	Vetor de viés do portão de saída
b	Viés ou <i>bias</i>

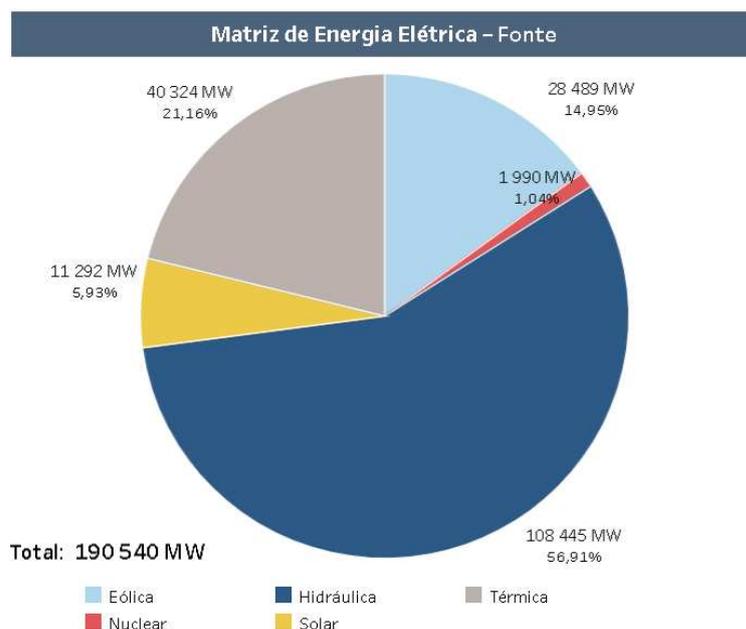
SUMÁRIO

1. INTRODUÇÃO	14
1.1 Objetivos.....	15
1.2 Metodologia.....	16
1.3 Organização do trabalho	17
2. FORMALIZAÇÃO DO PROBLEMA.....	19
2.1 Energia eólica.....	19
2.2 Métodos de previsão de energia eólica.....	20
2.3 Trabalhos relacionados	21
2.4 Conclusão parcial	23
3. FUNDAMENTAÇÃO TEÓRICA	24
3.1 Redes Neurais Artificiais	24
3.2 Rede de Perceptron Multicamadas (MLP)	26
3.3 Rede de Memória de Longo e Curto Prazo (LSTM).....	27
3.4 Rede de Unidade Recorrente com Portas (GRU)	29
3.5 Rede Neural Convolutacional (CNN).....	30
3.6 Treinamento	32
3.7 Algoritmo de retropropagação de erro.....	33
4. METODOLOGIA	35
4.1 Conjunto de dados utilizado	35
4.2 Preparação dos dados.....	40
4.3 Otimização de hiperparâmetros.....	41
4.3.1 <i>Busca aleatória de parâmetros</i>	<i>41</i>
4.3.2 <i>Validação cruzada com janela expansiva</i>	<i>42</i>
4.4 Métricas de avaliação	43
5. RESULTADOS.....	47
5.1 Otimização dos hiperparâmetros	47
5.2 Previsão de velocidade do vento	48
5.3 Observações.....	50
6. CONCLUSÃO	55
6.1 Sugestões para trabalhos futuros	55

1. INTRODUÇÃO

A demanda por fontes de energia renovável tem crescido significativamente nas últimas décadas, impulsionada pela necessidade de reduzir as emissões de gases de efeito estufa e mitigar os impactos das mudanças climáticas. Em um país tropical com abundância de fonte solar e eólica, como é caso do Brasil, o suprimento energético tende cada vez mais de ser atendido por fontes alternativas de energia. Nesse sentido, como pode ser visto na Figura 1, que mostra a participação das fontes de energia elétrica em termos de carga instalada no Sistema Interligado Nacional até o ano de 2023, a matriz de energia elétrica brasileira tem caminhado cada vez mais na direção de uma maior participação das fontes renováveis de energia. Como pode ser visto, a fonte eólica compreendia 14,95% e a fonte solar 5,93%, totalizando quando ambas 39.781MW de potência instalada até 2023.

Figura 1 - Matriz de energia elétrica brasileira.



Fonte: <https://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>

Por mais que a energia eólica seja uma fonte renovável e limpa, sua eficiência na geração de energia é fortemente influenciada por fatores meteorológicos imprevisíveis, como temperatura, umidade, altitude e velocidade do vento. Realizar a previsão dessas variáveis, em especial, da velocidade do vento que possui grande influência na geração da energia eólica, se torna crucial para melhorar o planejamento da operação de usinas

eólicas, promover a integração no SIN (Sistema Interligado Nacional) desse tipo de fonte e melhorar a estabilidade da rede elétrica (Tsai et. al., 2023). O próprio ONS (Operador Nacional do Sistema Elétrico) realiza diariamente a programação da geração eólica nacional por meio de métodos de previsão que são capazes de informar a geração que se espera do sistema como um todo e dos subsistemas em diferentes horizontes de tempo.

Existem muitos métodos para prever a velocidade do vento, incluindo métodos físicos, métodos estatísticos, métodos de aprendizado profundo (*deep learning*) e métodos de modelos híbridos. Entre esses o método de aprendizado profundo, que dispõe da utilização de redes neurais artificiais, tem mostrado grande potencial em melhorar a acurácia das previsões de velocidade do vento (Tsai et. al., 2023).

1.1 Objetivos

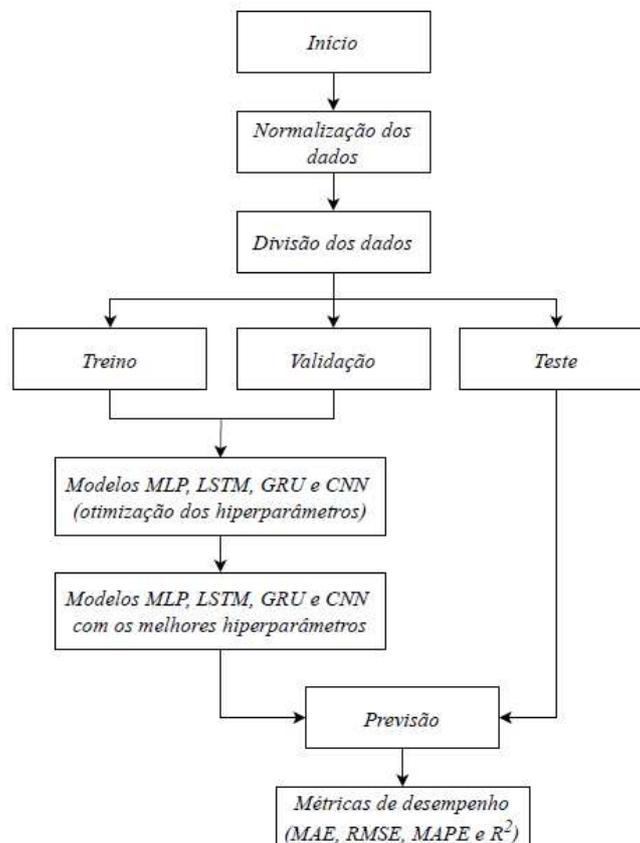
O objetivo geral deste trabalho consiste em realizar um estudo comparativo do desempenho de diferentes modelos de redes neurais artificiais na previsão de velocidade do vento. Para atingir este objetivo, os objetivos específicos são:

- Investigar os conceitos e fundamentos teóricos das redes neurais artificiais (RNAs);
- Realizar a revisão da literatura do uso de RNAs na previsão de energia eólica e velocidade do vento;
- Explorar o conjunto de dados de uma TMA (Torre de Medição Anemométrica) instalada no Parque Eólico Quixaba, na cidade de Aracati-CE, e realizar a preparação desses dados para o treinamento de modelos baseado em RNAs;
- Realizar o treinamento e avaliação de modelos de redes neurais MLP (rede perceptron multicamadas), LSTM (rede de memória de longo e curto prazo), GRU (unidade recorrente com portas) e CNN (rede neural convolucional), utilizando a *API* (Application Programming Interface) *Keras 2.15.0* do Python;
- Analisar os resultados obtidos via indicadores de desempenho de modelos preditivos como MAE, RMSE, MAPE e R^2 e discutir as contribuições e limitações dos modelos implementados com o intuito de propor melhorias para trabalhos futuros com base nas observações verificadas.

1.2 Metodologia

Para que os objetivos expostos anteriormente fossem alcançados algumas etapas necessárias foram definidas. Primeiramente, na etapa de início é feita a procura da base de dados a ser utilizada no trabalho e o pré-processamento das informações contidas nessa base para o adequado treinamento das redes. Posteriormente, a normalização dos dados é feita com os dados já definidos e pré-processados, transformando as magnitudes da variável velocidade do vento em valores num intervalo de 0 a 1. Com os dados normalizados, eles são enfim divididos em dados de treino, validação e teste. Os dados de treino e validação são utilizados para treinar as redes neurais e otimizar os parâmetros que as definem. Enquanto isso, os dados de teste servem para realizar a previsão com os modelos treinados e comparar o desempenho através de indicadores de desempenho. Com isso, é possível realizar o estudo comparativo dos modelos e inferir os que possuem bom e mal desempenho para o conjunto de dados em análise. A Figura 2, mostra na forma de diagrama as etapas descritas.

Figura 2 - Diagrama de fluxo das etapas do trabalho.



Fonte: Adaptado de Zhang et.al. (2021)

Todas as tarefas necessárias para atender ao fluxo desejado foram possíveis de serem realizadas a partir da linguagem *Python v. 3.10.12* no ambiente de codificação em nuvem *Colaboratory Google*. O Python é uma linguagem de alto nível, fácil interpretação, gratuita, de grande robustez e, além disso, de código aberto, fazendo com que sua utilização permita um rápido desenvolvimento com uma ampla comunidade ativa de usuários que garantem suporte e acesso a recursos continuamente. As seguintes bibliotecas do Python foram utilizadas:

- *Pandas v.2.1.4*: Biblioteca para manipulação e análise de dados, oferecendo estruturas de dados flexíveis e eficientes como DataFrames;
- *Numpy v.1.26.4*: Biblioteca para computação numérica, que fornece suporte a arrays multidimensionais e funções matemáticas de alto desempenho;
- *Matplotlib v.3.7.1*: Biblioteca para criação de gráficos estáticos, animados e interativos em Python, ideal para visualização de dados;
- *Seaborn v.0.13.1*: Biblioteca de visualização de dados baseada em *Matplotlib*, que oferece gráficos estatísticos de alto nível e estética aprimorada;
- *Scipy v. 1.13.1*: Biblioteca que oferece ferramentas para matemática, ciência e engenharia. Ela estende as capacidades do NumPy, incluindo funcionalidades avançadas para otimização, estatísticas e processamento de sinais;
- *Scikit-learn v.1.3.2*: Biblioteca de aprendizado de máquina que fornece ferramentas simples e eficientes para mineração e análise de dados;
- *TensorFlow v.2.17.0*: Biblioteca de código aberto para computação numérica e criação de modelos de aprendizado de máquina;
- *Keras v.3.4.1*: Biblioteca de alto nível do TensorFlow para criar e treinar modelos de aprendizado profundo. Ela é usada para prototipagem rápida, pesquisa de ponta e produção.

1.3 Organização do trabalho

O presente trabalho está organizado em cinco capítulos principais. No Capítulo 2, é apresentada a formalização do problema, abordando a importância e necessidade da tarefa de previsão de velocidade do vento. Também apresenta uma revisão

dos trabalhos relacionados ao assunto, destacando aqueles que utilizam modelos baseados em redes neurais na previsão de velocidade do vento e energia eólica. A fundamentação teórica, apresentando os conceitos básicos de redes neurais artificiais e a teoria envolvida em cada uma das arquiteturas de redes neurais em estudo no presente trabalho é descrita no Capítulo 3. A exploração do conjunto de dados é apresentada no Capítulo 4, bem como a metodologia empregada na aquisição dos resultados, mostrando de que forma os modelos foram validados e comparados. No Capítulo 5, as principais observações dos resultados obtidos são mostradas com a conclusão de quais modelos tiveram melhor desempenho. Por fim, no Capítulo 6, as conclusões são feitas destacando as contribuições do trabalho, limitações e possíveis direções para pesquisas futuras.

2. FORMALIZAÇÃO DO PROBLEMA

Devido a incessante busca por fontes alternativas de energia nos últimos anos, é possível notar uma redução cada vez maior de fontes que não sejam limpas, renováveis e de baixo custo (Zhang, et. al., 2021). Com intuito de tornar a matriz de energia elétrica mais renovável, a energia eólica passou então a ser largamente incentivada. Entretanto, por mais que a energia eólica promova vantagens econômicas e sustentáveis, ela também possui desvantagens.

Sua flutuação e intermitência devido a variação dos ventos, traz problemas para a rede de energia elétrica, principalmente no que diz respeito ao gerenciamento energético (Zhang, et. al., 2021). Ademais, pelo lado dos agentes, a controlabilidade da fonte é altamente dificultada pelas características já citadas. Nesse sentido, a previsão da velocidade do vento pode auxiliar em um maior aproveitamento do recurso eólico e no planejamento energético de outras fontes, como as térmicas (Carneiro et. al., 2014).

Na Seção 1 do presente capítulo será feita a formalização do problema de previsão de velocidade do vento, e por consequência de energia eólica, abordando de que forma essa energia é produzida. Na Seção 2, os diferentes métodos existentes de previsão de velocidade do vento serão expostos juntamente com a escolha do método utilizado para o presente trabalho. Na Seção 3, será feita uma revisão bibliográfica de trabalhos relacionados ao problema destacando o desempenho dos modelos implementados pelos autores. Por fim, na Seção 4, a partir revisão realizada será possível ter o sentimento dos modelos que têm tido grande visibilidade na literatura recente e que servem de alicerce para a construção dos modelos escolhidos.

2.1 Energia eólica

A energia eólica é produzida pelos aerogeradores (também chamados de turbinas eólicas), e advém da energia cinética dos ventos. Nos aerogeradores existem três elementos principais, que são o rotor, o eixo e o gerador, e vários elementos secundários que variam de acordo com o tipo e projeto do aerogerador. O rotor é o conjunto das pás e cubo do aerogerador responsável por capturar a energia no vento, o eixo é o elo que transfere a energia captada no rotor para o gerador, e o gerador é o responsável pela conversão de energia mecânica em elétrica (Tolmasquim, 2016).

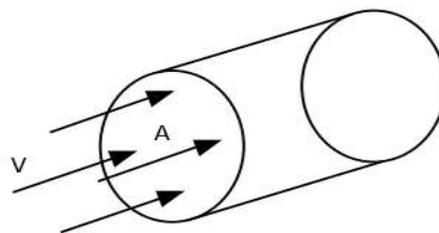
A Equação 2.1 mostra de que forma a potência (P) de um aerogerador pode

ser extraída do vento e, por consequência, a energia gerada ($E = P \cdot t$) quando analisado o tempo (t) em que o ele ficou em funcionamento. Nessa equação, A é a área varrida pelo aerogerador em m^2 , ρ é a massa específica do ar que em condições normais equivale a $1,225 \text{ kg/m}^3$, C_p o coeficiente de potência do aerogerador que representa o rendimento do rotor e v é a velocidade do vento em metros por segundo (m/s), sendo esta variável a principal, já que uma variação dela gera uma variação de ordem cúbica de P (Tolmasquim, 2016). Logo, realizar a previsão da velocidade do vento permite que a energia que será gerada possa ser facilmente mensurada ao saber as características específicas do aerogerador e da área onde este está localizado.

$$P = \frac{1}{2} \cdot A \cdot \rho \cdot C_p \cdot v^3 \quad (2.1)$$

O coeficiente de potência C_p é um fator de eficiência que depende do design do aerogerador e da velocidade do vento. O limite teórico máximo para o C_p é dado pela Lei de Betz, que estabelece que nenhum aerogerador pode converter mais de 59,3% da energia cinética do vento em energia mecânica, ou seja, C_p tem um valor máximo de 0,593. Na prática, os aerogeradores comerciais operam com valores de C_p entre 0,3 e 0,5. A Figura 3 ilustra o fluxo do vento dentro do rotor de área A .

Figura 3 - Fluxo de vento no rotor do aerogerador.



Fonte: Adaptado de Manwell et. al. (2009).

2.2 Métodos de previsão de energia eólica

Existem diversos métodos de previsão de energia eólica, que podem ser classificados em métodos físicos, métodos estatísticos, métodos de aprendizado profundo (*deep learning*) e métodos de modelos híbridos. Nos métodos físicos, informações

geográficas e meteorológicas, como clima e temperatura, são usadas para estimar a energia gerada. Nos métodos estatísticos, os dados históricos de velocidade do vento ou energia eólica são necessários para efetuar a previsão por meio de análise estatística, reconhecendo padrões nos dados. Nos métodos de aprendizado profundo, a relação não linear entre entradas e saídas é utilizada para capturar a característica de alta variabilidade da velocidade do vento, e por consequência da energia eólica gerada. Por fim, nos métodos de modelos híbridos há a combinação de duas ou mais técnicas de aprendizagem profunda ou algoritmos de otimização. Essa combinação evita as deficiências quando se utiliza apenas um único modelo preditivo, possibilitando encontrar características, hiperparâmetros e algoritmos de treinamento ótimos (Tsai et. al., 2023).

Além disso, as previsões podem ser diretas, que estimam a produção de energia com base em séries temporais, ou indiretas, que preveem a velocidade do vento para depois converter em geração de energia. Para o presente trabalho, optou-se por utilizar a abordagem que prevê a velocidade do vento por meio de quatro diferentes tipos de redes neurais e, em seguida, de acordo com as características do aerogerador e do próprio local onde ele se encontra os agentes de despacho podem obter a produção de energia estimada. Essa metodologia permite uma análise mais precisa e fundamentada da geração eólica, contribuindo para a eficiência na integração dessa fonte de energia na matriz elétrica.

2.3 Trabalhos relacionados

Singh & Rizwan (2022) realizam a previsão da energia gerada por uma turbina eólica a partir de dados coletados de um sistema SCADA (*Supervisory Control and Data Acquisition*). Os modelos *Random Forest*, *LSTM (Long Short-Term Memory)* e *Gradient Boosting* são treinados e comparados por vários indicadores de desempenho. Os resultados obtidos com o modelo *XGBoost* sugerido mostram melhor desempenho na média, com um R^2 de cerca de 0,969, MSE de 0,003, RMSE de 0,064, MAPE de 0,282 e MAE de 0,026. O modelo *LSTM* apresentou desempenho com um R^2 de cerca de 0,940, MSE de 0,067, RMSE de 0,078, MAPE de 0,333 e MAE de 0,034.

Solas et. al. (2019) fazem a previsão da energia eólica a partir dos dados horários normalizados de energia gerada, comparando o desempenho obtido entre um modelo baseado em redes neurais do tipo *CNN (Convolutional Neural Network)*, um modelo baseado em dados estatísticos de séries temporais através do modelo *ARIMA*

(*Auto-Regressive Integrated Moving Average*) e outro, baseado em aprendizado de máquina através do GBM (*Gradient Boosting Machine*). Os resultados obtidos mostram que o modelo CNN é superior aos outros dois na tarefa da previsão de energia eólica gerada. Ele apresentou desempenho com um MAE de 0,049, um RMSE de 0,077 e um EVS de 0,938.

Souza et. al. (2020) utilizam dados de energia gerada de uma turbina eólica, localizada no litoral norte de Holguín em Cuba, no intervalo de 01 de fevereiro de 2018 a 31 de março de 2018, para treinar modelos de previsão de energia eólica baseado em modelos de aprendizado de máquina. Os modelos baseados em árvore de regressão, máquinas de vetores de suporte, redes neurais artificiais, regressão linear, lasso e redes LSTM são treinados e comparados com as métricas RMSE e MSE. Os resultados que foram obtidos mostraram que o modelo baseado em redes LSTM foi o que apresentou melhor desempenho, com um RMSE de 35,68 kW e um MSE de 23,84 kW.

Chandran et. al. (2021) fazem a coleta de dados por meio de um sistema SCADA de um parque eólico localizado na Bélgica e utilizam eles para prever a energia eólica gerada em uma turbina usando métodos baseados em redes neurais profundas. Os modelos baseados em redes RNN (*Recurrent Neural Network*), GRU (*Gated Recurrent Unit*) e LSTM são treinados e comparados utilizando a métrica MSE. Os resultados obtidos mostraram que o modelo RNN apresentou um MSE de 0,143, o modelo LSTM um MSE de 0,136, e com melhor desempenho, o modelo GRU apresentou um MSE de 0,130.

Zhang et.al. (2021) utilizam um modelo híbrido CNN-LSTM que consiste na combinação da rede LSTM com a CNN para ser treinado em um conjunto de dados de um parque eólico localizado na China, com informações de velocidade do vento, direção do vento, temperatura, pressão do ar, umidade e energia gerada no decorrer de um ano. Esse conjunto é subdividido de acordo com as estações do ano de modo a mitigar os efeitos da sazonalidade inerentes dos ventos. O modelo híbrido é treinado juntamente com outros dois: um utilizando apenas redes LSTM e outro apenas redes CNN. A comparação feita a partir das métricas MAE, RMSE e R^2 sugerem um melhor desempenho do modelo híbrido frente aos outros modelos treinados.

Yu et. al. (2022) utilizam um modelo híbrido espaço-temporal combinando uma rede neural CNN e uma rede neural bidirecional LSTM, conhecida como BI-LSTM para ser treinado em um conjunto de dados composto pela potência real, por dados meteorológicos e por dados de previsão numérica no tempo (NWP) disponibilizados no

WIND (*Wind Integration National Dataset*) Toolkit. A CNN é utilizada para extrair as características espaciais de múltiplos parques eólicos com distribuição em nuvem de pontos e a BI-LSTM para aprender as características temporais. Os resultados do modelo proposto mostraram um excelente desempenho em termos de execução e precisão com um MAPE de 4,86%.

Liu et. al. (2023) propõem um novo modelo híbrido que combina métodos de previsão unificada e correção de erros para previsão da velocidade do vento. Ele utiliza dados de previsão numérica no tempo (NWP) em conjunto com um modelo de rede neural convolucional temporal (TCN) para extrair características espaciais e temporais e foi denominado de rede preditiva profunda de conversão espaço-temporal (STC-DPN). Para comprovar o bom desempenho do modelo proposto frente a outros, foram utilizados dados de um parque eólico localizado em Jining, na China. Os resultados mostraram que para um horizonte de tempo de previsão de 1 hora o modelo STC-DPN apresentou um MAE de 0,85 m/s, um RMSE de 1,14 m/s e uma correlação de 0,84.

2.4 Conclusão parcial

Em vista dos materiais relacionados ao assunto que foram apresentados, o presente trabalho busca realizar um estudo comparativo de modelos baseados nas redes neurais MLP (Multilayer Perceptron), LSTM (Long-Short Term Memory), GRU (Gated Recurrent Unit) e CNN (Convolutional Neural Network) para a tarefa de previsão de velocidade do vento, usando para isso dados reais de velocidade do vento coletados de uma TMA (Torre de Medição Anemométrica) instalada no Parque Eólico Quixaba, localizado na cidade de Aracati-CE. A escolha dessas redes foi feita após a constatação que elas são as que mais tem sido objeto de estudo na literatura recente, como pôde ser visto no subtópico anterior. Sendo assim, realizar um estudo comparativo de modelos que são baseados nessas redes permite identificar vantagens e desvantagens no treinamento e teste desses modelos e verificar o desempenho em um conjunto de dados real de um parque eólico do nordeste brasileiro.

3. FUNDAMENTAÇÃO TEÓRICA

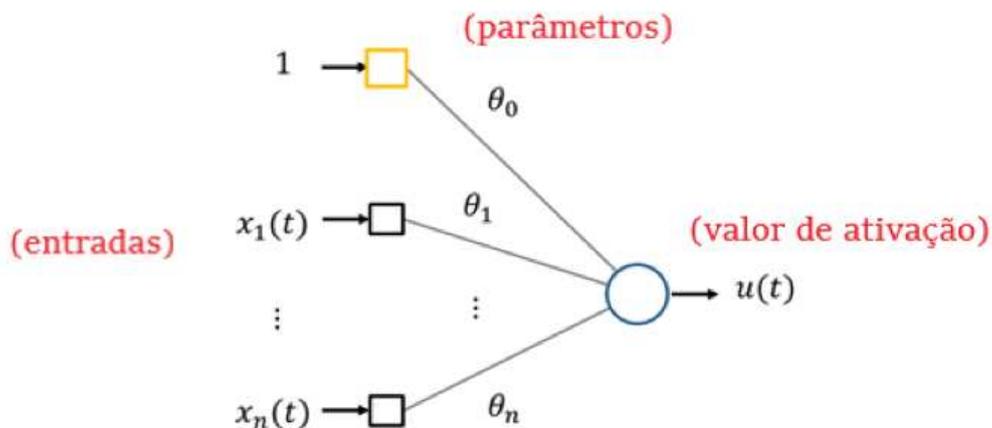
O presente capítulo aborda a fundamentação teórica que é necessária para a compreensão do estudo realizado. Na Seção 1, serão abordadas as Redes Neurais Artificiais (RNAs), com uma introdução ao conceito e à estrutura básica dessas redes. Na Seção 2 trata especificamente da Rede MLP (Perceptron Multicamadas), explicando sua arquitetura e funcionamento. Na Seção 3, será discutida a Rede LSTM (Memória de Longo Curto Prazo), um tipo de RNA projetada para lidar com séries temporais e dependências de longo prazo. A Seção 4 explora a Rede GRU (Unidade Recorrente com Portas), uma variação da LSTM com menor complexidade computacional. A Seção 5 aborda a Rede CNN (Rede Neural Convolutacional), amplamente utilizada em problemas de reconhecimento de padrões e processamento de imagens. Na Seção 6, será descrito o processo de treinamento de redes neurais, detalhando as etapas envolvidas na aprendizagem das redes. Finalmente, a Seção 7 apresenta o algoritmo de retropropagação de erro, fundamental para o ajuste dos pesos nas redes neurais durante o treinamento.

3.1 Redes Neurais Artificiais

Segundo Haykin (2008), as redes neurais artificiais (RNAs) são modelos de computação e de processamento de informação que é inspirado pelo modo como o cérebro humano realiza suas funções. Assim como um ser humano adquire conhecimento com base no aprendizado obtido no decorrer do tempo através de seus neurônios biológicos, uma RNA pode ser treinada para analisar dados, e usar o conhecimento adquirido a partir de um ambiente através de seus neurônios artificiais para chegar a um resultado.

Nos neurônios biológicos, o aprendizado ocorre através de adaptações nas sinapses, que são as conexões entre os neurônios. Quando um neurônio é estimulado repetidamente, a força das sinapses envolvidas nesse estímulo pode ser aumentada ou diminuída. Já nos neurônios artificiais, o procedimento utilizado para realizar o processo de aprendizagem ocorre por meio da modificação dos parâmetros da rede de uma forma ordenada com o intuito de se alcançar um objetivo de projeto desejado (Haykin, 2008). A Figura 4, ilustra a estrutura de um neurônio artificial.

Figura 4 - Neurônio artificial.



Fonte: Adaptado de Evsukoff (2020).

O modelo de neurônio apresentado foi proposto por McCulloch e Pitts, em 1943, e reproduz de forma simplificada o comportamento de um neurônio biológico. Nesse modelo, o processamento de cada unidade é calculado conforme a Equação 3.1. Nessa equação $u(t)$ é a saída, chamada de valor de ativação, $z(t)$ é o resultado do processamento linear, chamado de potencial de ativação e h é a função de ativação. As variáveis $x_1(t), x_2(t), \dots$, e $x_n(t)$ representam as entradas do neurônio. A primeira entrada, chamada de unidade *bias*, tem o valor de ativação constante e igual a 1 e o parâmetro θ_0 serve para ajustar o limiar da função de ativação. Os demais parâmetros $\theta_1, \theta_2, \dots$, e θ_n representam a intensidade das conexões (Evsukoff, 2020).

$$u(t) = h(z(t)) = h\left(\theta_0 + \sum_{i=1}^n x_i(t)\theta_i\right) \quad (3.1)$$

A função de ativação (h) é a componente principal do neurônio artificial, sendo útil para introduzir a característica não-linear nas redes neurais, que se formam com a interligação dos neurônios por camadas, permitindo que elas representem funções complexas que controlam o fluxo de informação entre os neurônios. Ela faz isso limitando a saída do neurônio a um intervalo de valor finito de acordo com o seu tipo (Goodfellow et. al., 2016). Dentre as diversas funções de ativação que existem apenas as três apresentadas na Tabela 1 serão utilizadas para a construção e treinamento dos modelos.

Tabela 1 - Funções de ativação.

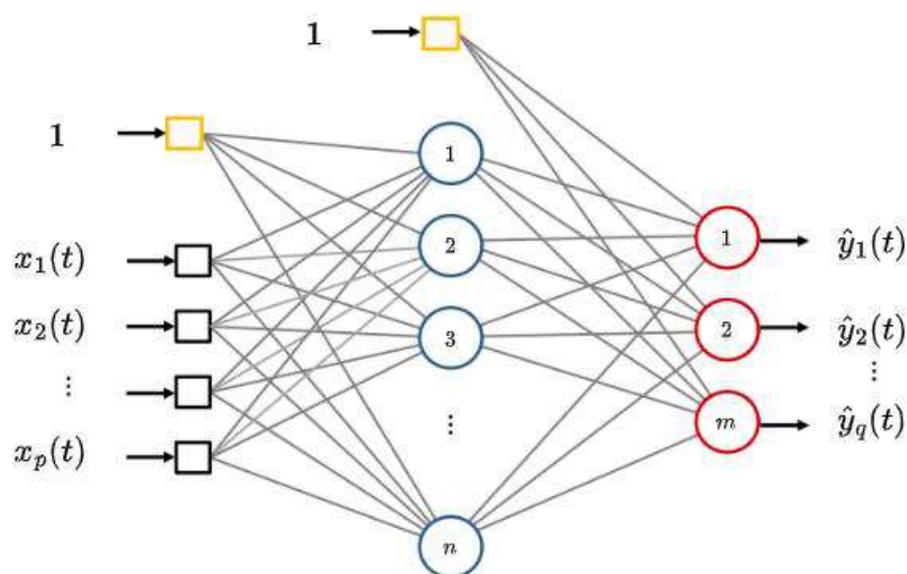
FUNÇÃO	$h(x)$	INTERVALO
Logística	$\frac{1}{1 + e^{-x}}$	(0,1)
Tangente Hiperbólica	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	(-1,1)
ReLU	$\max(0, x)$	$[0, \infty)$

Fonte: O próprio autor.

3.2 Rede de Perceptron Multicamadas (MLP)

Como visto na seção anterior, quando vários neurônios artificiais se interligam eles podem formar uma rede neural artificial (RNA) composta por camadas. A estrutura de RNA mais simples e que serve como ponto de partida para o adequado entendimento do funcionamento das demais redes neurais que serão exploradas nesse trabalho, é a rede neural do tipo MLP. Ela consiste numa rede formada por uma camada de entrada, uma ou mais camadas ocultas ou intermediárias e uma camada de saída. Ela é também conhecida por ser uma rede de múltipla camada alimentada adiante, pois o sinal de entrada se propaga para frente através da rede, camada por camada (Haykin, 2008). A Figura 5, ilustra uma rede do tipo MLP com uma camada intermediária.

Figura 5 - Rede MLP.



Fonte: Adaptado de Evsukoff (2020).

O exemplo de topologia apresentado da rede MLP é um modelo de uma camada intermediária com p entradas, n unidades intermediárias e m unidades na camada de saída. Como a camada de entrada não realiza processamento, esse modelo tem apenas duas camadas de processamento (Evsukoff, 2020). As Equações 3.2 e 3.3 representam de que forma ocorre a obtenção dos valores de saída para esse tipo de rede. Onde, $x(t) = (x_1(t), \dots, x_p(t))$ é o vetor de entradas, $u(t) = (u_1(t), \dots, u_n(t))$ é o vetor de ativação das unidades da camada intermediária e $\hat{y}(t) = (\hat{y}_1(t), \dots, \hat{y}_q(t))$ é o vetor de saída.

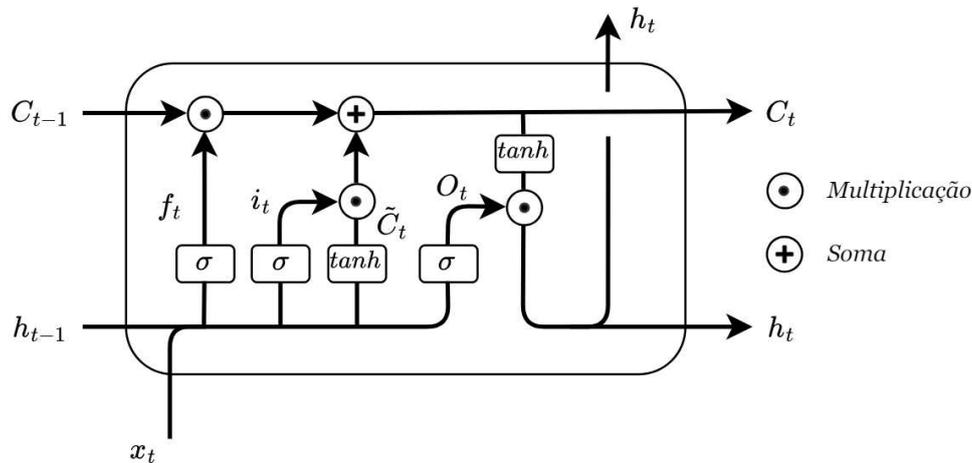
$$u(t) = h^{(1)}(z^{(1)}(t)) = h^{(1)}(b^{(1)} + x(t)W^{(1)}) \quad (3.2)$$

$$\hat{y}(t) = h^{(2)}(z^{(2)}(t)) = h^{(2)}(b^{(2)} + u(t)W^{(2)}) \quad (3.3)$$

3.3 Rede de Memória de Longo e Curto Prazo (LSTM)

O modelo de rede neural com memória de longo e curto prazo (LSTM) foi introduzido inicialmente por Hochreiter e Schmidhuber, e consiste em um tipo especial de rede neural recorrente, capaz de aprender dependências temporais de longo e curto prazo mantendo informações relevantes com o passar do tempo (Olah, 2015). A Figura 6, mostra a estrutura interna de uma camada LSTM, que é formada por uma unidade de memória e pelos “portões” de esquecimento, de entrada e de saída que se responsabilizam por controlar o fluxo de informação dentro da rede.

Figura 6 - Camada LSTM.



Fonte: Adaptado de Chandran et. al. (2021).

Nessa figura, C_{t-1} corresponde ao estado da unidade de memória no instante anterior. A unidade de memória é o componente principal da LSTM e é usada para controlar a transmissão da informação. O estado da camada no instante anterior é representado por h_{t-1} e a sequência de entrada no instante atual representado por x_t . O portão de entrada, o portão de esquecimento e portão de saída são representados respectivamente por i , f e o .

As Equações 3.4, 3.5 e 3.6 mostram como as saídas dos portões são calculadas. As matrizes W_i , W_f e W_o são as matrizes de pesos dos portões de entrada, esquecimento e saída, respectivamente. Os vetores de vieses b_i , b_f e b_o correspondem os vieses de seus respectivos portões. Como já visto σ representa a aplicação da função de ativação sigmoide.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3.4)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3.5)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (3.6)$$

Por fim, o novo estado da camada (h_t) é obtido a partir das Equações 3.7, 3.8 e 3.9. O símbolo \odot representa a multiplicação termo a termo dos vetores.

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3.7)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (3.8)$$

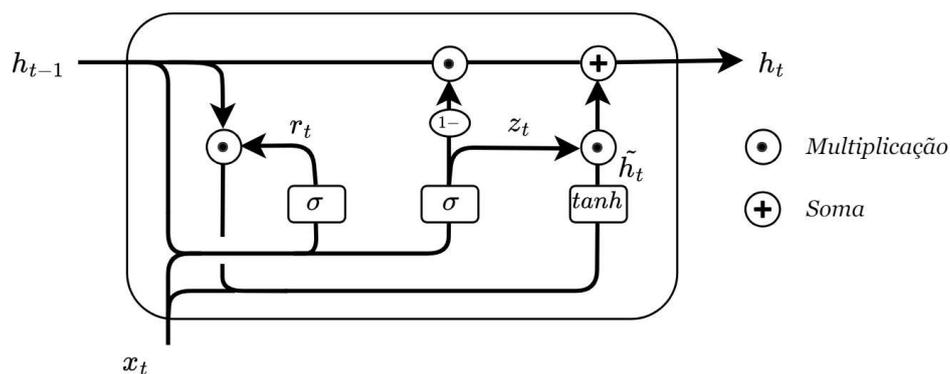
$$h_t = o_t \odot \tanh(C_t) \quad (3.9)$$

3.4 Rede de Unidade Recorrente com Portas (GRU)

A rede de unidade recorrente com portas (GRU) foi desenvolvida inicialmente por Cho et. al. (2014) e consiste em uma arquitetura mais simples e computacionalmente eficiente que a rede LSTM, pois possui menos parâmetros. Ela é também capaz de aprender dependências temporais de longo e curto prazo mantendo informações relevantes com o passar do tempo (Kostadinov, 2015). A

Figura 7, mostra a estrutura interna de uma camada GRU, que é formada por um portão de atualização e um portão de reinicialização. Esses dois portões são responsáveis por controlar o fluxo de informações, semelhante aos portões de entrada, esquecimento e saída na LSTM.

Figura 7 - Camada GRU.



Fonte: Adaptado de Chandran et. al. (2021).

Nessa figura, o portão de atualização e o portão de reinicialização são representados respectivamente por z e r . As Equações 3.14 e 3.15 mostram como as saídas dos portões são calculadas. As matrizes W_z e W_r são as matrizes de pesos,

associadas a sequência de entrada (x_t), dos portões de atualização e reinicialização, respectivamente. Já as matrizes U_z e U_r são as matrizes de pesos, associadas a sequência do estado anterior da camada (h_{t-1}), dos portões de atualização e reinicialização, respectivamente. Os vetores de vieses b_z e b_r correspondem aos vieses de seus respectivos portões. O estado atualizado da camada (h_t) é obtido a partir das Equações 3.12 e 3.13.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (3.10)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (3.11)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \quad (3.12)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (3.13)$$

A partir da verificação esquemática das camadas das redes recorrentes LSTM e GRU é possível notar que a rede GRU possui menor quantidade de parâmetros que a rede LSTM, devido à ausência de estado de memória. Essa menor quantidade de parâmetros permite que a rede GRU seja além de mais eficiente computacionalmente, mais interpretável que a rede LSTM.

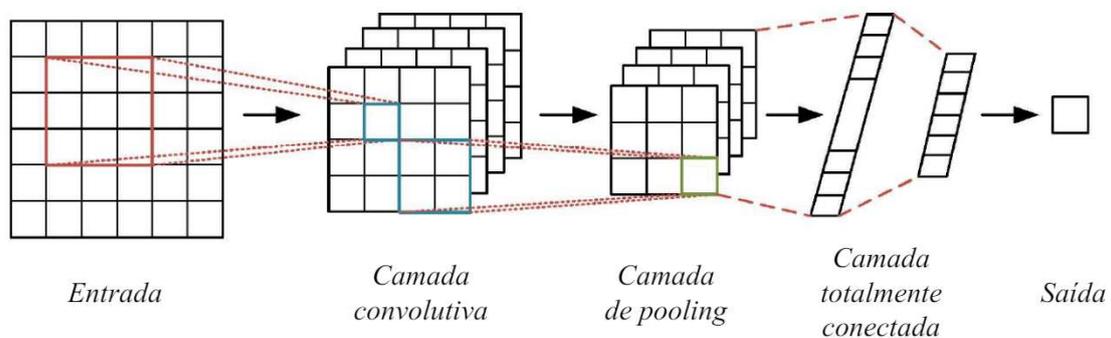
3.5 Rede Neural Convolutiva (CNN)

Segundo Evsukoff (2020), a rede neural convolutiva é um tipo de rede profunda alimentada adiante em camadas que aproveita as características de relações de vizinhança em que os pesos de conexões representam filtros convolucionais. Ela foi introduzida inicialmente por Lecun et. al. (1998) e possui parâmetros que são compartilhados em matrizes esparsas, que permite o processamento em alto desempenho. Como o próprio nome sugere, a rede CNN é baseada na operação matemática de convolução que é definida conforme a Equação 3.14 para o caso da convolução em uma dimensão.

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t - a) \quad (3.14)$$

Os modelos baseados na rede CNN foram desenvolvidos para problemas de visão computacional, mas também tem aplicações em outras áreas, como processamento de linguagem natural e previsão de séries temporais. Nessa rede existem três camadas entre a entrada e a saída que realizam o processamento dos dados conhecidas como, camada convolucional, camada de pooling e camada totalmente conectada. A Figura 8 ilustra como estas camadas estão interligadas.

Figura 8 - Rede CNN.



Fonte: Adaptado de Zhang et. al. (2021).

Na camada convolucional, após os dados serem dispostos em um formato de grade, pequenas matrizes chamadas de *kernels* (K) percorrem os dados de entrada (I) em uma vizinhança definida a partir de m e n , calculando convoluções com o intuito de extrair características locais. A camada de pooling é usada para amostrar os vetores de atributos gerados pela camada convolucional para reduzir os parâmetros da rede e evitar ajuste excessivo. A camada convolucional e a camada de pooling juntas constituem a camada de extração de características. A Equação 3.15 mostra como a convolução mencionada é calculada em duas dimensões em uma rede CNN. Nesta equação, o filtro K é movido sobre a entrada I , e em cada posição, uma soma ponderada é calculada multiplicando-se os elementos correspondentes da entrada e do filtro e somando os resultados (Goodfellow et. al., 2016).

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i - m, j - n) \cdot K(m, n) \quad (3.15)$$

De acordo com Zhang et. al. (2021) quanto maior a quantidade de qualquer uma das duas camadas (convolucional e pooling), mais características são extraídas. No entanto, se a rede for muito profunda a velocidade de treinamento diminuirá e até poderá causar ajuste excessivo. Ao mesmo tempo, se a profundidade da rede for muito rasa, os atributos do conjunto de dados não poderão ser bem extraídos e a relação não linear entre as variáveis não poderá ser obtida. Dessa forma, a escolha da quantidade dessas camadas é primordial para o desempenho da rede.

3.6 Treinamento

Segundo Silva et. al. (2010), o processo de treinamento de uma rede neural consiste na aplicação de passos ordenados que sejam necessários para ajuste dos pesos e vieses dos neurônios presentes na rede, tendo como objetivo final a generalização de soluções a serem produzidas pela sua saída. Existem 3 (três) tipos de treinamento possíveis de redes neurais de acordo com os dados disponíveis, sendo treinamento supervisionado, não supervisionado e com reforço. A escolha de cada um deles estará condicionada ao ambiente no qual será realizada a tarefa de previsão e se os dados de saída estão ou não disponíveis.

No treinamento supervisionado, além de se ter disponível os dados de entrada para a realização do treinamento, os dados de saída também são utilizados. Ou seja, nesse treinamento deve ser construída uma tabela com amostras contendo o par entrada-saída. No treinamento não supervisionado, por sua vez, os dados de saída não estão disponíveis, ficando a cargo da rede identificar os subconjuntos, conhecido como *clusters*, presentes que possuem relação entre si. Já no treinamento com reforço, um agente seleciona as ações com base em recompensas advindas dos incrementos ou decrementos no ajuste dos pesos e vieses da rede. Essa seleção é realizada de forma probabilística, isto é, as melhores ações serão escolhidas para a produção da maior recompensa acumulada (Silva et. al., 2010). O presente trabalho usa o treinamento supervisionado, visto que o problema de previsão de séries temporais é inerente desse tipo de treinamento, pois a rede é treinada com resultados de entradas e saídas já disponíveis.

3.7 Algoritmo de retropropagação de erro

Em cada neurônio os valores de entrada fornecidos são processados e geram um valor de saída e são comparados com a saída esperada por meio de uma função, denominada função perda, que informa a precisão do resultado. No presente trabalho a função de perda baseado no erro médio quadrático (MSE – Mean Squared Error) será a utilizada. O algoritmo de retropropagação do erro consiste em encontrar os melhores pesos e vieses de entrada para obter uma saída mais precisa ou minimizar a função perda, pois quanto menor é a perda de uma rede, mais precisa ela se torna.

O nível de ajuste dos pesos e vieses é determinado pelos gradientes da função de perda em relação a esses parâmetros conforme pode ser visto nas Equações 3.18 e 3.19. Nessas equações o novo parâmetro é calculado através do parâmetro antigo subtraído da derivada parcial da função perda (E_r), também chamada de função erro, em função do parâmetro antigo multiplicada pela taxa de aprendizagem (η) (Kostadinov, 2019). Esse processo de cálculo é conhecido também como Regra Delta.

$$w \leftarrow w - \eta \frac{\partial E_r}{\partial w} \quad (3.18)$$

$$b \leftarrow b - \eta \frac{\partial E_r}{\partial b} \quad (3.19)$$

O valor que será usado para somar ou subtrair do peso do neurônio é chamado de taxa de aprendizado sendo um número importante para o aprendizado da rede neural, onde um valor de taxa muito alto resultará em uma alteração de pesos alta, fazendo que o algoritmo ultrapasse facilmente o melhor valor de peso para cada neurônio, e por outro lado uma taxa muito baixa faz com que o aprendizado demore muito até encontrar um bom resultado (Kostadinov, 2019). A Tabela 2, mostra o pseudocódigo associado ao algoritmo de retropropagação de erro que é utilizado para atualizar os pesos das redes neurais do presente trabalho.

Tabela 2 - Pseudocódigo do algoritmo de retropropagação de erro.

Inicializa todos os pesos \mathbf{w} da rede e η de forma aleatória

Calcula a saída e a função perda E_r :

while $E_r < \epsilon$ (*valor desejado ou outro critério*) **do**

$\forall w \in \mathbf{w}: \Delta w = -\frac{\partial E_r}{\partial w}$ (Passo 1)

$w_{novo} \leftarrow w_{velho} + \eta \Delta w + \dots$ (Passo 2)

Calcule a saída e o E_r :

end

Fonte: Adaptado de Goglia (2021).

Uma vez concebida toda a fundamentação teórica necessária para realizar os treinamentos e teste dos modelos de redes neurais, é possível então seguir para a aquisição dos resultados. Nesta nova fase, o problema da previsão da velocidade do vento será abordado utilizando a metodologia selecionada, aplicada a um conjunto de dados reais. O objetivo é não apenas alcançar as metas estabelecidas neste trabalho, mas também fornecer ao leitor uma compreensão clara do processo de implementação de um modelo preditivo baseado em redes neurais. Além disso, espera-se que o estudo demonstre sua utilidade para futuras pesquisas e investigações relacionadas ao caso analisado.

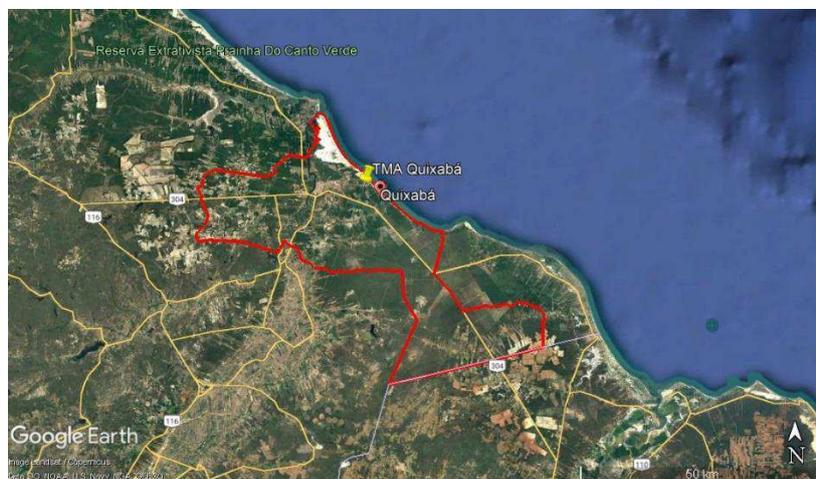
4. METODOLOGIA

O presente capítulo descreve a metodologia empregada para a realização do treinamento e validação dos modelos de previsão baseado nas redes neurais em estudo que foi inicialmente citada no capítulo 1 referente a introdução. Na Seção 1, será feita uma exploração do conjunto de dados que será utilizado de modo a identificar possíveis correções a serem feitas nos dados. Na seção 2, a série temporal de velocidade do vento já tratada será configurada adequadamente para a realização do treinamento dos modelos. Após isso, na Seção 3, será apresentada a estratégia de validação cruzada aliada a otimização de hiperparâmetros, que é útil para realizar a avaliação das redes neurais em estudo. Por fim, na Seção 4, as métricas de desempenho de modelos definidas e que serão utilizadas para traduzir os resultados são descritas.

4.1 Conjunto de dados utilizado

O conjunto de dados utilizado contém medições de velocidade do vento, em metros por segundo, coletadas através de um anemômetro instalado em uma TMA (Torre de Medição Anemométrica) localizada em Aracati-CE, próximo à praia de Quixabá. A referida TMA se trata de uma estação responsável por medir e registrar dados anemométricos e climatológicos do Parque Eólico Quixaba e possui as coordenadas geográficas 4,56334 S, 37,68846 W. A Figura 9 mostra no mapa onde está localizada a TMA.

Figura 9 - Localização da TMA.



Fonte: O próprio autor.

Os vários dados que a TMA do Parque Eólico Quixaba registra por meio de alguns sensores são integralizados a cada 10 minutos. Os dados de 2019 a 2023 foram concedidos pela empresa de geração de energia Energimp S.A e são fechados para disponibilização. A Tabela 3 descreve os sensores que estão presentes na torre e a altura na qual eles estão instalados. O sensor de medição de velocidade do vento escolhido foi o Anemômetro Superior que está localizado a 86,2 metros do nível do solo.

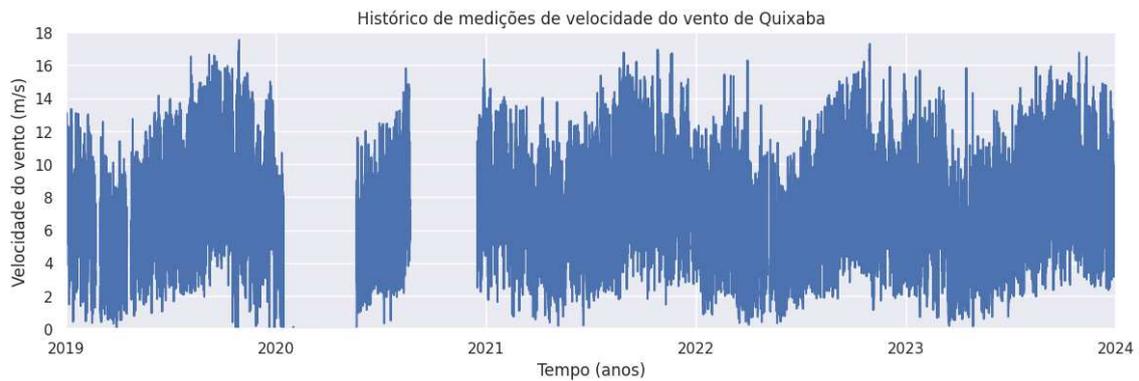
Tabela 3 - Sensores instalados na TMA.

SENSORES	DESCRIÇÃO	ALTURA (M)
Anemômetro Superior	Velocidade do vento	86,2
Anemômetro Intermediário	Velocidade do vento	85
Anemômetro Inferior	Velocidade do vento	61
Wind Vane Superior	Direção do Vento	84
Wind Vane Inferior	Direção do Vento	60
Termohigrômetro	Temperatura / Umidade	77
Barômetro	Pressão Atmosférica	77

Fonte: O próprio autor.

Dessa forma, foram coletados dados referentes a medições dos anos de 2019 a 2023 das médias de velocidade do vento a 86,2 metros de altura no intervalo de 10 minutos. Esses dados foram disponibilizados no formato CSV (*Comma-Separated Values*) e, posteriormente, lidos e tratados em Python no ambiente *Colaboratory Google*. Inicialmente, como pode ser visto na Figura 10, foi observado a ausência de uma grande quantidade de dados nos anos de 2019 e 2020, o que dificultou a consideração desses anos para a análise. Logo, foram considerados apenas as medições coletadas nos anos de 2021, 2022 e 2023 para o treinamento e teste dos modelos de previsão.

Figura 10 - Série histórica das medições de velocidade do vento (m/s).



Fonte: O próprio autor.

Após remover as ausências de medições da base de dados é importante ter uma visualização de como a velocidade do vento se comportou ao longo dos anos de análise. As Figura 11-13 apresentam as séries históricas das médias de velocidade do vento horária, diária e mensal, respectivamente, ao longo dos três anos de análise.

Figura 11 – Série histórica da média horária da velocidade do vento (m/s).



Fonte: O próprio autor.

Figura 12 – Série histórica da média diária da velocidade do vento (m/s).



Fonte: O próprio autor.

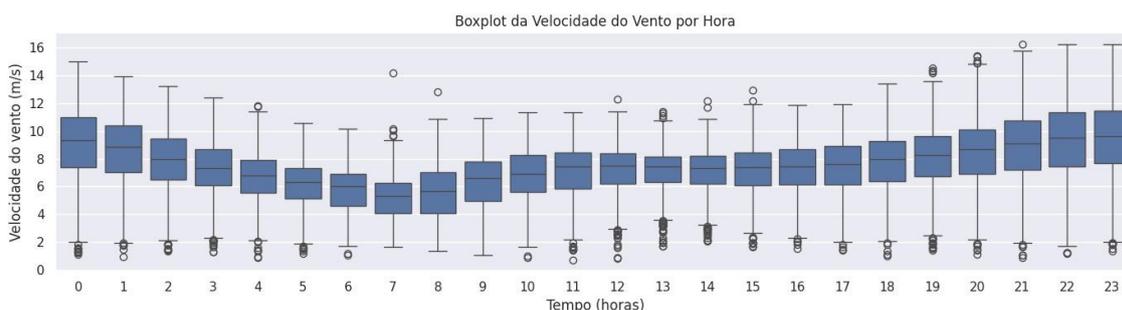
Figura 13 – Série histórica da média mensal da velocidade do vento (m/s).



Fonte: O próprio autor.

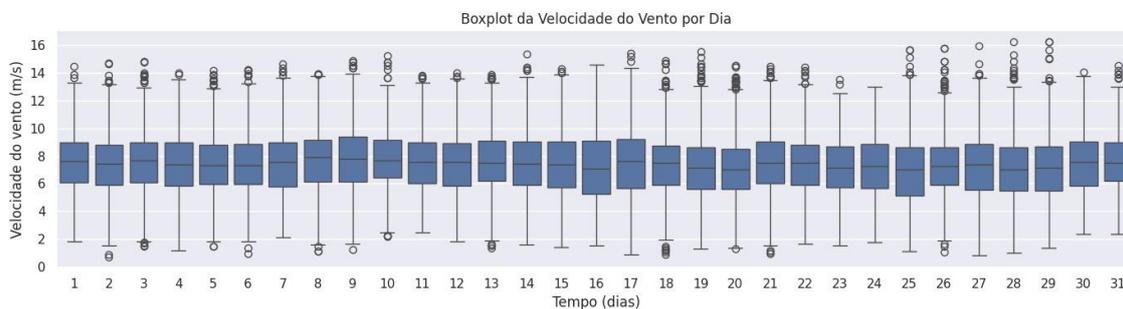
Além de uma visualização geral da série histórica da velocidade do vento também é importante saber como é a distribuição dessa variável nas diferentes médias temporais, de modo a constatar os períodos de baixa e alta variabilidade de velocidade do vento. As Figura 14-16 mostram as distribuições das séries históricas das médias de velocidade do vento horária, diária e mensal por meio do uso de gráficos do tipo *boxplot*. A linha horizontal no interior da caixa representa a mediana, que é o valor central dos dados e informa que 50% dos dados estão abaixo e 50% acima desse valor. A borda inferior da caixa define o primeiro quartil e representa o valor abaixo do qual estão 25% dos dados. A borda superior da caixa, por sua vez, define o terceiro quartil e representa o valor abaixo do qual estão 75% dos dados. A altura da caixa define o intervalo interquartil (IQR) e é útil para determinar os extremos inferior ($-1,5 \cdot IQR$) e superior ($+1,5 \cdot IQR$). Valores que ultrapassam esses limites são tidos como *outliers*, ou seja, valores discrepantes da distribuição.

Figura 14 - Distribuição da série histórica da média horária da velocidade do vento (m/s).



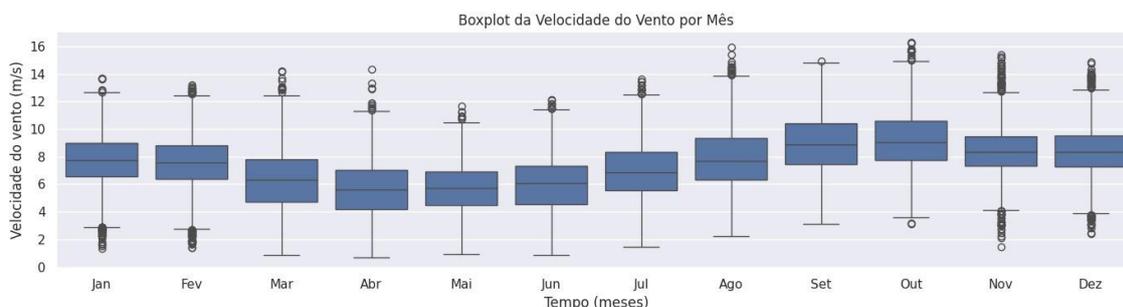
Fonte: O próprio autor.

Figura 15 - Distribuição da série histórica da média diária da velocidade do vento (m/s).



Fonte: O próprio autor.

Figura 16 - Distribuição da série histórica da média mensal da velocidade do vento (m/s).



Fonte: O próprio autor.

Com isso, é possível notar sob uma ótica mensal que o intervalo de agosto a dezembro apresenta maiores valores de velocidade do vento na localidade Quixabá, sendo uma característica intrínseca do nordeste brasileiro. Já sob uma ótica diária, é possível notar uma maior constância das médias de velocidade do vento. A partir de uma ótica horária, por sua vez, é notória a existência de maiores valores de velocidade do vento no período noturno.

Para o treinamento dos modelos foram utilizadas as medições de 10 minutos da velocidade do vento, totalizando 157.681 medições. Das quais, 105.121, que compreende os anos de 2021 e 2022, foram separadas para treinamento e otimização dos modelos e 52.560, que compreende o ano de 2023, foram separadas para a realização do teste desses modelos. A Tabela 4, apresenta o resumo estatístico do conjunto de dados.

Tabela 4 - Resumo estatístico das medições de velocidade do vento (m/s).

VARIÁVEL ESTATÍSTICA DA VELOCIDADE DO VENTO	VALOR
MÍNIMO	0,1582 m/s
MÁXIMO	17,3054 m/s
MÉDIA	7,3993 m/s
MEDIANA	7,4021 m/s
DESVIO PADRÃO	2,4462 m/s
VARIÂNCIA	5,9839 m ² /s ²
NÚMERO DE MEDIÇÕES	157.681

Fonte: O próprio autor.

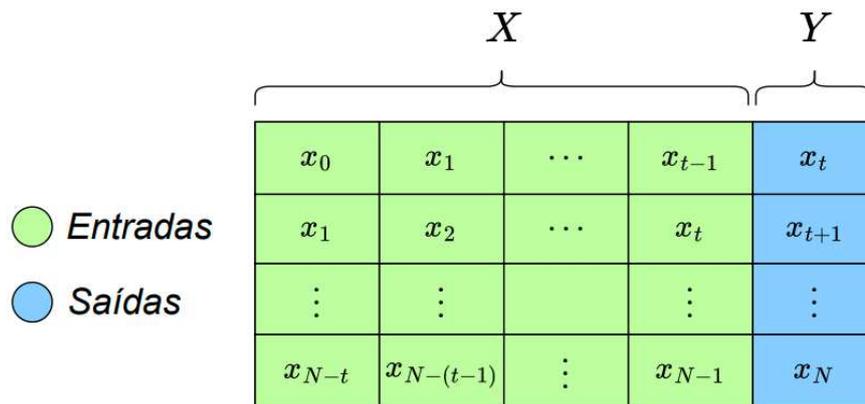
4.2 Preparação dos dados

A etapa de preparação dos dados envolve além da divisão desses dados em treino e teste, a normalização e a formato adequado deles para a realização do treinamento dos modelos baseado em redes neurais. Os dados foram normalizados utilizando a normalização linear, aplicando a Equação 4.1 e, deixando os dados entre 0 e 1, para serem aplicados as funções de ativação presentes nas redes neurais de forma adequada.

$$X_n = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (4.1)$$

Após isso, foi realizada a divisão dos dados em conjunto de treinamento e conjunto de teste. Os anos de 2021 e 2022 foram separados para treinamento e avaliação, e o ano de 2023 para teste. Além de preparar os dados para o adequado uso das funções de ativação, é necessário também preparar eles para o formato requerido das redes neurais. Como já exposto no Capítulo 3, o aprendizado que é requerido nesse tipo problema é o supervisionado. Sendo assim, a série histórica das medições de velocidade do vento foi transformada para esse tipo de treinamento conforme mostrado na Figura 17.

Figura 17 - Entradas e saídas para o treinamento das redes neurais.



Fonte: O próprio autor.

A matriz X tem dimensão $\mathbb{R}^{s-t \times t}$, enquanto o vetor Y possui dimensão \mathbb{R}^{s-t} , onde os índices N e t representam o tamanho da amostra e os passos de tempo atrás (atraso), respectivamente. Cada linha de X corresponde a um valor de Y , juntos eles constituem um par entrada-saída que é utilizado como entrada nas redes neurais do presente trabalho.

4.3 Otimização de hiperparâmetros

O *Keras* é um framework de código aberto que já vem integrado ao *TensorFlow* e é voltado para a construção e treinamento de modelos de aprendizado profundo. Desenvolvido em Python, o *Keras* oferece uma interface de alto nível que simplifica a criação de redes neurais complexas. Essas redes neurais são definidas a partir da determinação de hiperparâmetros, que são parâmetros definidos pelo usuário antes do treinamento do modelo, influenciando diretamente o desempenho e a eficácia da rede (Bergstra; Bengio, 2012). Exemplos de hiperparâmetros incluem a taxa de aprendizado, o número de camadas e neurônios, o tamanho do *batch* e o número de épocas de treinamento.

4.3.1 Busca aleatória de parâmetros

Vários métodos podem ser empregados para encontrar a melhor combinação de parâmetros que irão definir a rede neural. Contudo, os mais utilizados são os métodos de busca em grade (*Grid Search*) e busca aleatória (*Random Search*). Para a otimização dos hiperparâmetros das redes empregadas no presente trabalho, foi utilizada a estratégia

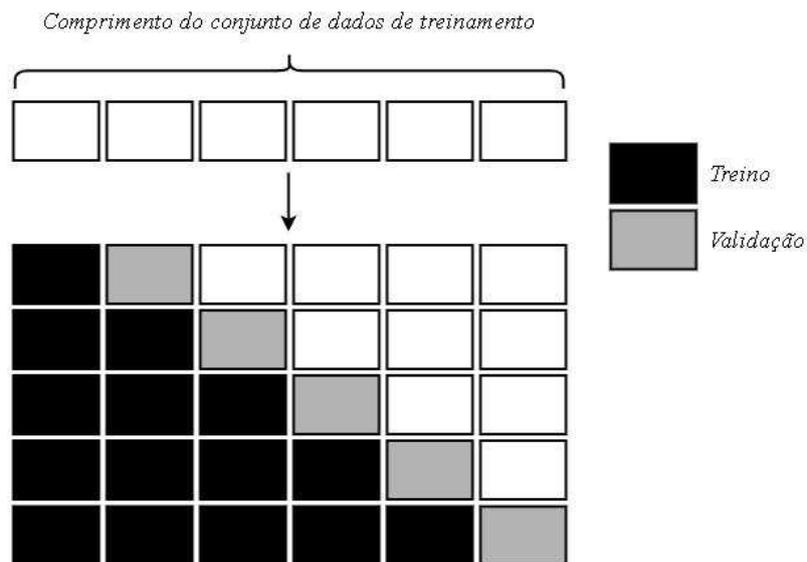
de busca aleatória que consiste em um método de ajuste de hiperparâmetros mais eficiente que o método de busca em grade (Bergstra; Bengio, 2012). Ao contrário deste outro que avalia de forma exaustiva todas as combinações possíveis de hiperparâmetros especificados, a busca aleatória realiza a busca sobre um espaço de hiperparâmetros definido de forma aleatória. Isso torna o processo de busca mais eficiente em termos de tempo e recursos (Bergstra; Bengio, 2012). A estratégia de busca aleatória de hiperparâmetros foi realizada utilizando a função *ParameterSampler* que está contida na biblioteca *Scikit-learn v.1.3.2*.

4.3.2 Validação cruzada com janela expansiva

Juntamente com a busca aleatória foi utilizada a validação cruzada para séries temporais com janela expansiva, que consiste em realizar o treinamento em diferentes configurações de conjunto de dados de treino e validação, chamadas de *folds* ou subconjuntos. Inicialmente o conjunto de dados é dividido em partes, mantendo a ordem temporal. No primeiro *fold* , uma pequena porção inicial dos dados é usada para treinamento e uma subsequente para validação. Em cada *fold* subsequente, a janela de treinamento é expandida, adicionando mais dados ao conjunto de treinamento e mantendo uma porção separada para validação. Esse processo imita a realidade, onde novos dados ficam disponíveis ao longo do tempo. (Poudel, 2023)

Para cada *fold* , o modelo é treinado com os dados disponíveis até aquele ponto e avaliado com os dados subsequentes. Isso resulta em múltiplas avaliações do modelo, cada uma refletindo seu desempenho em diferentes estágios de tempo. O desempenho do modelo em cada *fold* é registrado, permitindo a avaliação dele em vários cenários temporais. Após a execução de todos os *folds* , o desempenho médio dos modelos é comparado. A configuração de hiperparâmetros que obtiver o melhor desempenho médio é selecionada como a melhor. O processo de validação cruzada com janela expansiva utilizado nas redes neurais trabalhadas pode ser visualizado na Figura 18, onde foram utilizados 5 *folds* . Essa validação foi feita utilizando a função *TimeSeriesSplit* que está contida na biblioteca *Scikit-learn v.1.3.2*.

Figura 18 - Validação cruzada com janela expansiva.



Fonte: Adaptado de Victor et. al. (2020).

4.4 Métricas de avaliação

As métricas de avaliação de modelos de regressão como o erro médio absoluto (MAE), a raiz do erro quadrático médio (RMSE), o erro percentual absoluto médio (MAPE) e o coeficiente de determinação (R^2) foram utilizadas para analisar o desempenho das redes treinadas. As Equações a seguir apresentam a definição matemática de cada uma dessas métricas.

$$MAE = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \quad (4.1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2} \quad (4.2)$$

$$MAPE = \frac{1}{n} \sum_{i=0}^{n-1} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \quad (4.3)$$

$$R^2 = 1 - \frac{\sum_{i=0}^{n-1} y_i - \hat{y}_i}{\sum_{i=0}^{n-1} y_i - \bar{y}} \quad (4.4)$$

Nas Equações 4.2, 4.3 e 4.4, y_i representa o valor observado, \hat{y}_i o valor previsto pelo modelo treinado e \bar{y} a média dos valores observados. Ao passo que n representa a quantidade total de amostras no conjunto de dados de teste. Quanto menor os valores de MAE, RMSE e MAPE e maior o valor de R^2 (que varia entre 0 e 1), melhor é a performance o modelo.

Além dessas métricas, a análise de variância (ANOVA – Analysis of Variance) também foi utilizada para inferir o desempenho dos modelos do ponto de vista estatístico. Essa técnica é usada para verificar se há diferenças significativas entre as médias de três ou mais grupos independentes, no caso do trabalho os grupos são os quatro modelos. O teste ANOVA compara a variabilidade entre os grupos com a variabilidade dentro dos grupos (Montgomery, 2020). As Equações 4.5, 4.6 e 4.7 mostram de que forma é possível obter o valor de F que traduz a ANOVA.

$$F = \frac{MSB}{MSE} \quad (4.5)$$

$$MSB = \frac{SSB}{k - 1} \quad (4.6)$$

$$MSE = \frac{SSE}{N - k} \quad (4.7)$$

Onde,

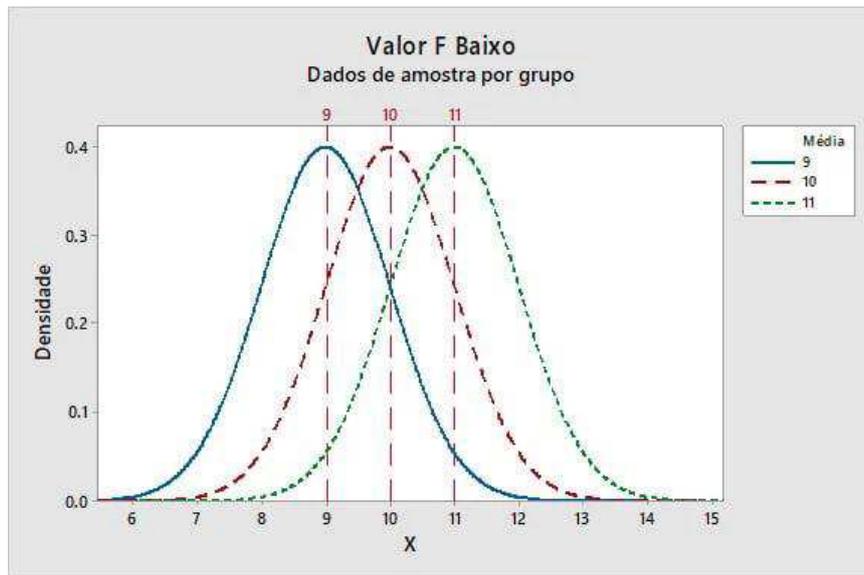
- MSB: Variância média entre os grupos;
- MSE: Variância média dentro dos grupos;
- SSB: Soma dos quadrados entre os grupos;
- SSE: Soma dos quadrados dentro dos grupos;
- N: Número de observações;
- k: Número de grupos ou modelos.

A Figura 19 ilustra um exemplo de como podemos interpretar a valor de F. Caso esse valor seja baixo (Figura 19-a) tem-se que as médias dos grupos estão próximas (baixa variabilidade) em relação a variabilidade dentro de cada grupo. Por outro lado,

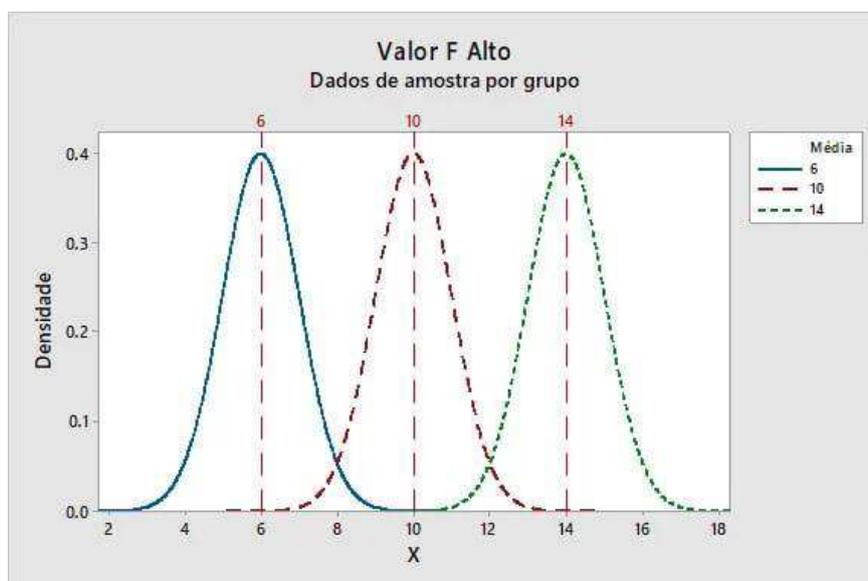
caso o valor de F seja alto (Figura 19-b) tem-se que a variabilidade das médias dos grupos é grande em relação à variabilidade intragrupo. Portanto, a hipótese nula de que as médias do grupo são iguais pode ser rejeitada (Montgomery, 2020).

Figura 19 – (a) Distribuição das amostras por grupo com valor F baixo. (b) Distribuição das amostras por grupo com valor de F alto.

(a)



(b)



Fonte: Entendendo Análise de Variância (ANOVA) e o teste F. Disponível em: <https://blog.minitab.com/pt/entendendo-analise-de-variancia-anova-e-o-teste-f>

Obtendo o valor F, é possível obter o valor-p a partir da tabela de distribuição F. Com esse valor também é possível inferir se as médias dos grupos são estatisticamente iguais ou não. Caso o valor-p seja menor que o nível de significância de 0,05 a hipótese nula de que as médias dos grupos são iguais pode ser rejeitada. Caso contrário, é possível assumir a hipótese nula e inferir que as médias dos grupos são estatisticamente iguais (Montgomery, 2020). No presente trabalho a função *f_oneway*, que está na biblioteca *Scipy* do Python, foi utilizada para obter de forma rápida os valores de F e valor-p para todas as métricas de desempenho e averiguar se os modelos possuem desempenho semelhante.

5. RESULTADOS

Após preparar o conjunto de dados adequadamente, os dados separados para treino e validação, que compreende valores horários de velocidade do vento dos anos de 2021 e 2022, são utilizados para encontrar a melhor combinação de hiperparâmetros de cada arquitetura. Posteriormente, os dados separados para teste, compreendendo os valores de velocidade do vento do ano de 2023, são utilizados para realizar a previsão, computar as métricas de avaliação e inferir aquele modelo com o melhor desempenho.

5.1 Otimização dos hiperparâmetros

A otimização dos hiperparâmetros de cada rede neural foi feita utilizando a validação cruzada com janela expansiva aliado a estratégia de busca aleatória. Isso foi feito utilizando as funções *TimeSeriesSplit*, utilizada para realizar a validação cruzada, e a função *ParameterSampler* para escolher de forma aleatória a combinação de hiperparâmetros em cada iteração sobre o subconjunto de dados de treino definido pela função *TimeSeriesSplit*. Mais detalhes de como essas funções foram complementadas e implementadas podem ser consultados no link que direciona para o código salvo em um repositório do *GitHub*. As Tabelas 5 a 8 mostram os hiperparâmetros os melhores valores encontrados pelo algoritmo de acordo com o desempenho nos dados de validação.

Tabela 5 – Hiperparâmetros da rede MLP.

HIPERPARÂMETROS	MELHOR VALOR
Atraso	18
Units	50
Learning rate	0.01
Batch size	32

Fonte: O próprio autor.

Tabela 6 - Hiperparâmetros da rede LSTM.

HIPERPARÂMETROS	MELHOR VALOR
Atraso	18
Units	50

Learning rate	0.001
Batch size	16

Fonte: O próprio autor.

Tabela 7 - Hiperparâmetros da rede GRU.

HIPERPARÂMETROS	MELHOR VALOR
Atraso	18
Units	50
Learning rate	0.001
Batch size	16

Fonte: O próprio autor.

Tabela 8 - Hiperparâmetros da rede CNN.

HIPERPARÂMETROS	MELHOR VALOR
Atraso	18
Filters	64
Kernel size	3
Learning rate	0.01
Batch size	64

Fonte: O próprio autor.

O *Atraso* determina o formato adequado dos dados para utilizar na entrada das redes, em especial, define quantos passos de tempo para trás serão utilizados. *Units* define a quantidade de neurônios ou unidades que serão utilizadas na camada da rede neural. *Learning rate*, a taxa de aprendizado utilizada que determina a velocidade com que o treinamento ocorrerá para uma mesma quantidade definida de episódios. Já o *Batch size* representa o número de amostras que serão processadas antes do modelo atualizar os pesos.

5.2 Previsão de velocidade do vento

Após a realização da otimização dos hiperparâmetros, as redes neurais foram treinadas com um número de épocas igual a 150, que determina a quantidade de vezes em que os dados de treino são completamente apresentados a rede neural. Para o modelo MLP o tempo de treinamento foi de aproximadamente 14 minutos. Já para os modelos

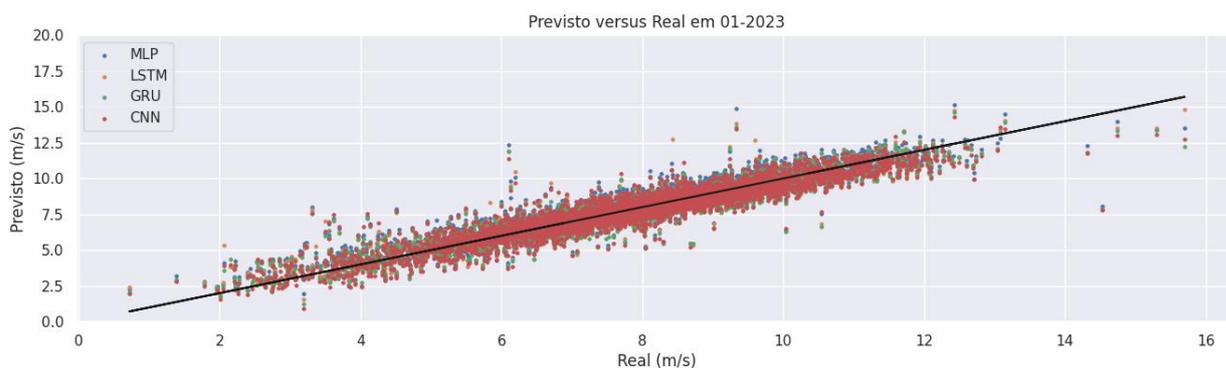
LSTM e GRU, que são as redes recorrentes em estudo, o tempo de treinamento foi de aproximadamente 166 minutos e 195 minutos, respectivamente. Por fim, para o modelo CNN o tempo de treinamento foi de 20 minutos, aproximadamente.

Os experimentos foram realizados em uma máquina equipada com CPU Intel(R) Xeon(R) com velocidade de clock de 2.00 GHz. O sistema contava com uma GPU Nvidia Tesla T4 com 12 GB de memória e suporte à arquitetura CUDA v. 2.2. A máquina possuía 12 GB de RAM disponível. O ambiente de execução utilizado foi o ambiente em nuvem *Google Colaboratory*.

Com a finalização do treinamento dos modelos e, portanto, com a obtenção dos pesos e vieses ótimos inerentes de cada arquitetura, a previsão de velocidade do vento no ano de 2023 do local onde se encontra a TMA foi realizada. Essa previsão foi feita mês a mês de modo a verificar o desempenho dos modelos ao longo de todo o ano de 2023, observando os meses com pior e melhor desempenho.

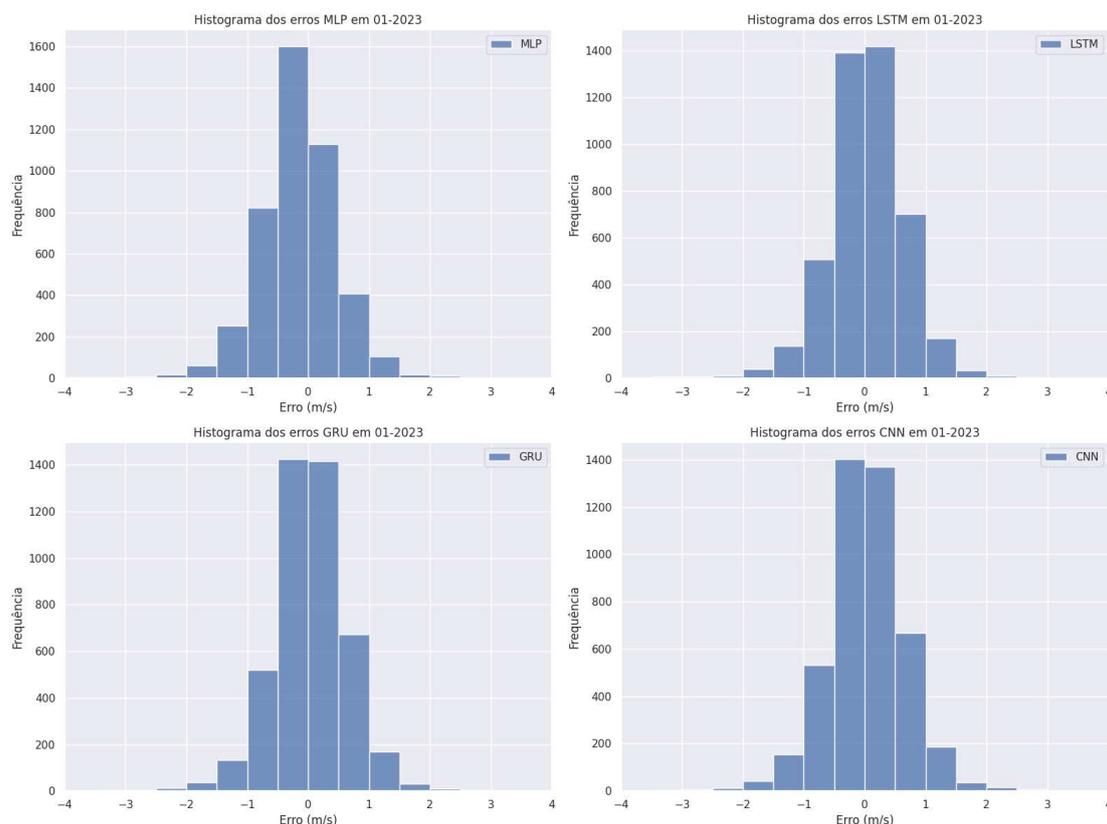
Nas Figuras 20 e 21 são mostrados os resultados da previsão de velocidade do vento do mês de janeiro de 2023. A Figura 20, mostra ao longo do mês os registros das previsões que foram feitas pelos modelos em confronto com a velocidade do vento real verificada. Já a Figura 21, apresenta na forma de histogramas a distribuição dos erros para cada modelo no mês de janeiro de 2023. Os resultados das previsões obtidos para os demais meses podem ser consultados no APÊNDICE A – RESULTADOS DAS PREVISÕES

Figura 20 - Velocidade do vento Prevista versus Real em janeiro de 2023.



Fonte: O próprio autor.

Figura 21 - Distribuição dos erros de previsão em janeiro de 2023.



Fonte: O próprio autor.

5.3 Observações

Os resultados apresentados para a previsão de velocidade do vento na localidade onde se encontra a TMA do Parque Eólico Quixaba, com os modelos baseados em redes neurais no ano de 2023, obtiveram boa precisão. As Tabela 9-12 mostram as métricas de desempenho MAE, RMSE, MAPE e R^2 , respectivamente, coletadas nos meses do ano de 2023 para cada um dos modelos que foram treinados.

Tabela 9 - Erro Médio Absoluto (MAE) em 2023.

ANO-MÊS	MLP	LSTM	GRU	CNN
2023-01	0,4991	0,4748	0,4695	0,4840
2023-02	0,5080	0,4782	0,4790	0,4993
2023-03	0,5109	0,4693	0,4675	0,4896
2023-04	0,4754	0,4244	0,4260	0,4471

2023-05	0,4610	0,4247	0,4258	0,4425
2023-06	0,4841	0,4536	0,4537	0,4739
2023-07	0,5024	0,4754	0,4733	0,4880
2023-08	0,5130	0,4838	0,4855	0,5007
2023-09	0,5475	0,5202	0,5213	0,5370
2023-10	0,5440	0,5258	0,5258	0,5362
2023-11	0,5064	0,4841	0,4831	0,4937
2023-12	0,4982	0,4705	0,4707	0,4857
MÉDIA	0,5042	0,4737	0,4734	0,4898

Fonte: O próprio autor.

Tabela 10 - Raiz do Erro Quadrático Médio (RMSE) em 2023.

ANO-MÊS	MLP	LSTM	GRU	CNN
2023-01	0,6772	0,6571	0,6426	0,6631
2023-02	0,6886	0,6499	0,6517	0,6785
2023-03	0,7067	0,6704	0,6617	0,6889
2023-04	0,6548	0,6104	0,6085	0,6393
2023-05	0,6392	0,5974	0,6064	0,6262
2023-06	0,6447	0,6127	0,6131	0,6370
2023-07	0,6631	0,6402	0,6363	0,6544
2023-08	0,6721	0,6401	0,6420	0,6592
2023-09	0,7064	0,6779	0,6802	0,6987
2023-10	0,7136	0,6838	0,6849	0,6994
2023-11	0,6682	0,6386	0,6387	0,6558
2023-12	0,6487	0,6144	0,6159	0,6319
MÉDIA	0,6736	0,6411	0,6402	0,6610

Fonte: O próprio autor.

Tabela 11 - Erro Percentual Médio Absoluto (MAPE) em 2023.

ANO-MÊS	MLP	LSTM	GRU	CNN
2023-01	7,2897	6,8273	6,7206	6,9409
2023-02	8,6210	7,9290	7,9577	8,2348
2023-03	12,6829	10,6686	10,6999	10,8920
2023-04	12,9733	10,8089	10,9097	11,0288
2023-05	9,1856	8,2191	8,2403	8,4453
2023-06	8,8159	7,9824	8,0315	8,2893
2023-07	7,6058	7,0045	6,9622	7,1772

2023-08	6,7922	6,2607	6,2807	6,5019
2023-09	6,5155	6,0574	6,0674	6,2993
2023-10	6,1696	5,8357	5,8332	5,9842
2023-11	6,3417	5,9435	5,9391	6,0746
2023-12	6,5234	6,0378	6,0434	6,2633
MÉDIA	8,2930	7,4646	7,4738	7,6776

Fonte: O próprio autor.

Tabela 12 - Coeficiente de Determinação (R^2) em 2023.

ANO-MÊS	MLP	LSTM	GRU	CNN
2023-01	0,8818	0,8888	0,8936	0,8867
2023-02	0,9117	0,9214	0,9209	0,9143
2023-03	0,9228	0,9306	0,9323	0,9267
2023-04	0,8774	0,8935	0,8941	0,8832
2023-05	0,8564	0,8746	0,8708	0,8622
2023-06	0,8972	0,9072	0,9070	0,8997
2023-07	0,9045	0,9110	0,9121	0,9070
2023-08	0,9125	0,9206	0,9201	0,9158
2023-09	0,8961	0,9043	0,9037	0,8984
2023-10	0,8796	0,8894	0,8891	0,8843
2023-11	0,8844	0,8944	0,8944	0,8886
2023-12	0,8960	0,9067	0,9062	0,9013
MÉDIA	0,8934	0,9035	0,9037	0,8973

Fonte: O próprio autor.

É possível notar que no mês de março os modelos obtiveram melhor desempenho que nos demais meses. Contudo, do ponto de vista de um horizonte maior de tempo, é possível notar um pior desempenho geral dos modelos no primeiro semestre do ano. Enquanto no segundo semestre é possível notar um desempenho melhor e mais constante. Isso está associado não apenas a construção e treinamento dos modelos em si, mas também a característica mais estacionária do vento nesse período do ano na região cearense.

A Tabela 13 apresenta os resultados do teste de análise de variância (ANOVA) realizado. É possível observar que, para as métricas Erro Médio Absoluto (MAE), Erro Percentual Absoluto Médio (MAPE) e Coeficiente de Determinação (R^2) a hipótese nula de que não existe diferença dessas métricas entre os modelos é aceita, visto que foi obtido um valor-p menor que 0,05 em todos os casos. A exceção fica para a métrica Raiz do Erro

Quadrático Médio (RMSE), que apresentou um valor-p de 0,0172 e dessa forma, pode-se inferir que para essa métrica existem diferenças entre os modelos do ponto vista estatístico.

Tabela 13 - Teste ANOVA.

MÉTRICA	F	VALOR-P	RESULTADO
MAE	3,1471	0,0565	Não há diferença entre os grupos/modelos
RMSE	4,5358	0,0172	Há diferença entre os grupos/modelos
MAPE	0,4914	0,6901	Não há diferença entre os grupos/modelos
R²	1,0251	0,3907	Não há diferença entre os grupos/modelos

Fonte: O próprio autor.

Mesmo com desempenhos próximos como verificado a partir do teste ANOVA, o modelo baseado na rede neural GRU foi o que mais se destacou entre todos, com um Erro Médio Absoluto (MAE) em média de 0,47 m/s, uma Raiz do Erro Quadrático Médio (RMSE) em média de 0,64 m/s, um Erro Percentual Médio Absoluto em média de 7,47% e um Coeficiente de Determinação R^2 em média de 0,90. A superioridade verificada dos resultados desse modelo para o modelo baseado na rede MLP ocorreu devido ao fato de que as redes MLP são limitadas em sua habilidade de capturar dependências temporais, enquanto que redes GRU são projetadas especificamente para lidar com sequências temporais.

A mesma justificativa pode ser dada para a superioridade frente ao modelo baseado na rede CNN, pois esta é mais adequada para trabalhar com imagens, podendo extrair características espaciais (Zhang et. al., 2021), mas não tão útil onde há uma dependência intrínseca entre os dados temporais, como é o caso da velocidade do vento. Já a leve superioridade frente ao modelo baseado na rede LSTM se deu pelo fato de que a rede GRU, mesmo tendo a capacidade de capturar dependências de longo prazo, similar ao LSTM, ela realiza isso de maneira mais eficiente, sendo benéfico no caso em estudo onde os dados de velocidade do vento possuem padrões temporais complexos (Kostadinov, 2017).

Mesmo com essas ressalvas os demais modelos também apresentam desempenho satisfatório. Ao analisar o tempo de treinamento foi possível verificar que,

em contrapartida a superioridade frente aos outros em termos do erro de previsão, o modelo baseado na rede GRU demorou mais que 9 vezes o tempo de treinamento dos modelos MLP e CNN, por exemplo. Sendo assim, do ponto de vista computacional o custo de implementação do modelo GRU foi bem alto frente aos demais e, com isso, a aplicação dos modelos MLP ou CNN em produção para horizontes de previsão de curto período seria mais adequado em vista do desempenho próximo ao GRU.

Cabe ressaltar que na etapa de pré-processamento alguns dados faltantes, que refletem a inoperância do anemômetro instalado na torre, tiveram que ser substituídos. Essa substituição afetou o desempenho dos modelos, visto que a alteração foi feita por intermédio de interpolação linear com os demais dados disponíveis. Além disso, a quantidade de dados de entrada utilizada de apenas 3 horas (18 amostras de 10 minutos), que está associada ao número de passos de tempo para trás, pode ter sido um valor pequeno, porém, bem otimizado para o espaço de busca definido.

6. CONCLUSÃO

Os objetivos pretendidos foram alcançados com os resultados apresentados na seção anterior, sendo possível tanto realizar a coleta e preparação dos dados como realizar o estudo comparativo almejado entre as redes neurais que mais são utilizadas em previsões de velocidade do vento. Os resultados verificados sugerem que o modelo preditivo baseado na rede GRU é mais eficaz na captura de padrões complexos e não lineares nos dados de velocidade do vento. Ele apresentou em média ao longo dos meses do ano de 2023 um MAPE de 7,47%, o que significa que em média a previsão realizada pelo modelo irá divergir do real em 7,47%, para mais ou para menos, e um R^2 de 0,90, o que revela que o modelo tem uma qualidade de previsão maior comparado aos demais.

A linguagem Python utilizada tanto para a etapa de pré-processamento como para a construção, treinamento e teste dos modelos, foi simples de utilizar. Isso se deve ao fato de que essa linguagem oferece uma vasta gama de bibliotecas e ferramentas que facilitam o pré-processamento de dados, a visualização de resultados e a implementação de modelos complexos. A vantagem dela se dá por possuir uma comunidade ativa e uma extensa documentação que garantem que soluções para problemas comuns estejam prontamente disponíveis. O framework Keras possui grande simplicidade e capacidade de construir protótipos de forma rápida e eficiente. Dessa forma, foi possível criar os modelos com poucas linhas de código, além de integrá-lo facilmente com outras bibliotecas.

6.1 Sugestões para trabalhos futuros

Existe algumas sugestões que podem ser destacadas para aprimorar o estudo realizado sobre a previsão de velocidade do vento utilizando dados reais de um parque eólico, dentre as quais podemos citar:

- Explorar o uso de outras técnicas de pré-processamento de dados, além da normalização dos dados e remoção de outliers, para melhorar a qualidade dos dados de entrada;
- Combinar diferentes modelos de redes neurais em uma abordagem de *ensemble* pode ser investigada para potencialmente aumentar a precisão das previsões;

- Realizar a integração de dados meteorológicos adicionais, como temperatura, pressão atmosférica e umidade, pode também contribuir para melhorar a performance dos modelos;
- Aplicar técnicas modernas de aprendizado profundo, como *Transformers*, poderia ser explorada para capturar dependências temporais de longo alcance nos dados de velocidade do vento;
- Aprofundar o teste estatístico ANOVA para casos em que o teste aponta que há diferenças entre o desempenho dos modelos. Isso pode ser feito utilizando algum teste adicional que seja capaz de realizar comparações múltiplas entre os modelos como o teste de Tukey.

REFERÊNCIAS

- AZANK, Felipe. **Como avaliar seu modelo de regressão: As principais métricas para avaliar seus modelos de regressão**, 2020. Disponível em: <https://medium.com/turing-talks/como-avaliar-seu-modelo-de-regress%C3%A3o-c2c8d73dab96>. Acesso em: 24/05/2024.
- BERGSTRA James; BENGIO Yoshua. **Random Search for Hyper-Parameter Optimization**. Journal of Machine Learning Research. v. 13, p. 281-305, 2012.
- BROWNLEE, Jason. **How to Develop LSTM Models for Time Series Forecasting**. 2020. Disponível em: <https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/>. Acesso em: 12/03/2024.
- CARNEIRO, Tatiane C.; ROCHA, Paulo A.C.; CARVALHO, Paulo C.M.; FERNÁNDEZ-RAMÍREZ, Luis M. **Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain**. Applied Energy, v. 314, p. 118936, 2022.
- CARNEIRO, Tatiane C.; SANTOS, Heron A.; BRAGA Arthur P. S.; CARVALHO, Paulo C. M. **Redes Neurais Artificiais para Previsão de Velocidade do Vento: Estudo de caso para Maracanaú – CE**. XX Congresso Brasileiro de Automática - CBA, 2014.
- CHANDRAN V.; PATIL Chandrashekhar K.; MANOHARAN Anto Merline, GHOSH Aritra, SUMITHRA M.G., KARTHICK Alagar, RAHIM Robbi. **Wind power forecasting based on time series model using deep machine learning algorithms**. Materials Today: Proceedings, v.47, p. 115-126, 2021.
- DA SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes neurais artificiais para engenharia e ciências aplicadas – curso prático**. São Paulo: Artliber, 2010.
- DE SOUZA Manoel V. F.; RODRIGUEZ Luis C.; VARELLA Carlos T. **Previsão de Geração de Energia em uma Usina Eólica a partir de séries temporais**, v.11, 2020.
- Entendendo Análise de Variância (ANOVA) e o teste F**. Minitab Blog Editor, 2019. Disponível em: <https://blog.minitab.com/pt/entendendo-analise-de-variancia-anova-e-o-teste-f>. Acesso em 03/09/2024.
- EVSUKOFF, Alexandre G. **Inteligência Computacional**. 1. ed. Rio de Janeiro: E-papers, 2020.
- FILHO, Mario. **How To Do Time Series Cross-Validation In Python**, 2023. Disponível em: <https://forecastgy.com/posts/time-series-cross-validation-python/>. Acesso em: 16/04/2024.

FOLEY Aoife M.; LEAHY Paul G.; MARVUGLIA Antonino; MCKEOGH Eamon J. **Current methods and advances in forecasting of wind power generation.** Renewable Energy, v. 37, p. 1-8, 2012.

GOGLIA, Diletta. **Backpropagation for Dummies**, 2021. Disponível em: <https://medium.com/analytics-vidhya/backpropagation-for-dummies-e069410fa585>. Acesso em: 14/03/2024.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning.** MIT press, 2016.

HAYKIN, Simon. **Redes neurais: princípios e prática.** 2. ed. Bookman Editora, 2001.

KOSTADINOV, Simeon. **Understanding GRU Networks**, 2017. Disponível em: <https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be>. Acesso em: 22/03/2024.

KOSTADINOV, Simeon. **Understanding Backpropagation Algorithm: Learn the nuts and bolts of a neural network's most important ingrediente**, 2019. Disponível em: <https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd>. Acesso em: 14/03/2024.

MANWELL, J. F.; MCGOWAN, J. G.; ROGERS, A. L. **Wind Energy Explained: Theory, Design, and Application.** Wiley, 2009.

MONTGOMERY, D. C. **Design and Analysis of Experiments.** 10. ed. Wiley, 2020.

MULLER, Andreas C.; GUIDO Sarah. **Introduction to Machine Learning with Python.** 1. ed. 2016.

OLAH, Christopher. **COLAH'S Blog: Understanding LSTM Networks**, 2015. Disponível em: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acesso em: 12/03/2024.

POUDEL, Sushmita. **Cross Validation in Time Series Forecasting**, 2023. Disponível em: <https://medium.com/@poudelsushmita878/cross-validation-in-time-series-forecasting-db2bc7601875> Acesso em: 16/04/2024.

SINGH Upma; Rizwan M. **SCADA system dataset exploration and machine learning based forecast for wind turbines.** Results in Engineering, v. 16, 2022.

Sobre o SIN: O sistema em números. Disponível em: <https://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>. Acesso em 12/03/2024.

SOLAS Margarida; CEPEDA Nunes; VIEGAS Joaquim L. **Convolutional Neural Network for Short-term Wind Power Forecasting**. 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Bucharest, Romania, pp. 1-5, 2019.

TSAI Wen-Chang; HONG Chih-Ming; TU Chia-Sheng; LIN Whei-Min; CHEN Chiung-Hsing. **A Review of Modern Wind Power Generation**. Sustainability 2023, v. 15, 10757, 2023

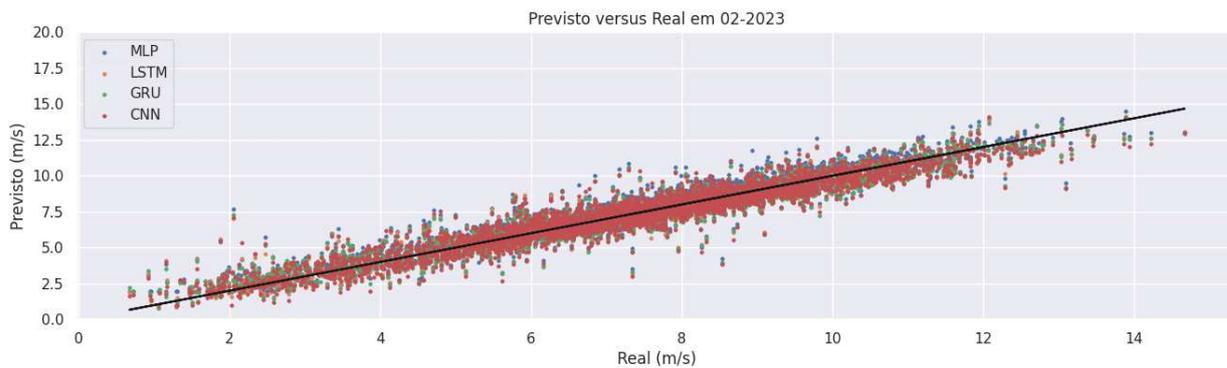
UNIVERSIDADE FEDERAL DO CEARÁ. Biblioteca Universitária. **Guia de normalização de trabalhos acadêmicos da Universidade Federal do Ceará**. Fortaleza, 2013.

VITOR, Manoel Victor Florencio de Souza; CUEVAS RODRIGUEZ, Luis; TRICHET VARELLA, Carlos. **Previsão de geração de energia em usina eólica utilizando séries temporais**. XI Computer on the Beach, 2020, Santa Catarina. v.11, 2020.

ZHANG Huiting; ZHAO Lixun; DU Zhipeng. **Wind power prediction based on CNN-LSTM**. 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2), Taiyuan, China, pp. 3097-3102, 2021.

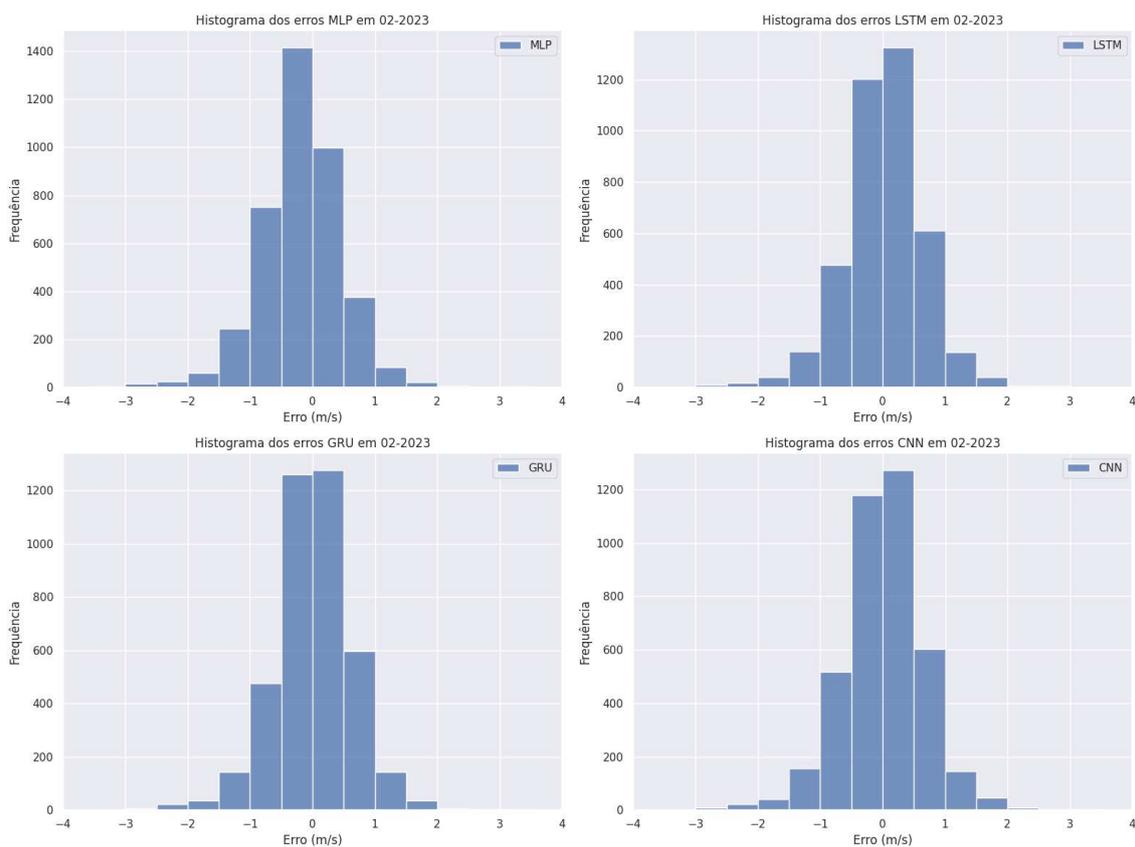
APÊNDICE A – RESULTADOS DAS PREVISÕES

Figura 22 - Velocidade do vento Prevista versus Real em fevereiro de 2023.



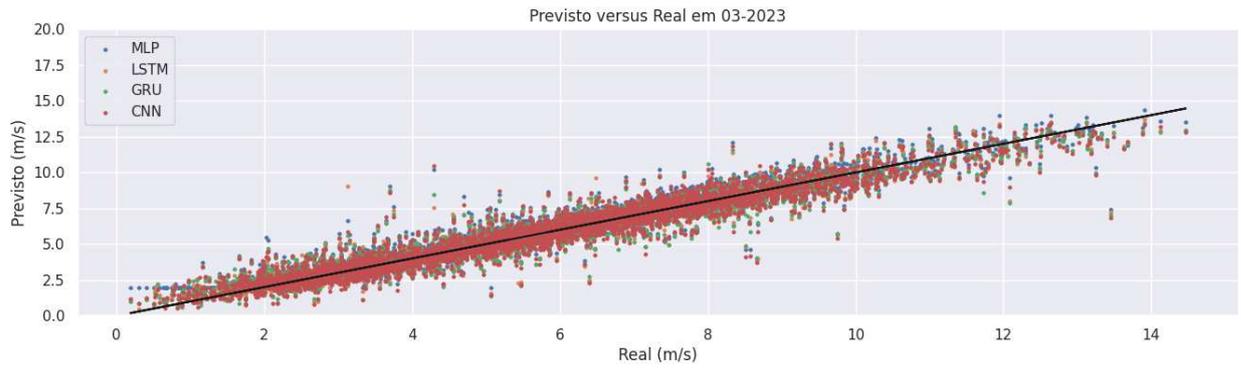
Fonte: O próprio autor.

Figura 23 Distribuição dos erros de previsão em fevereiro de 2023.



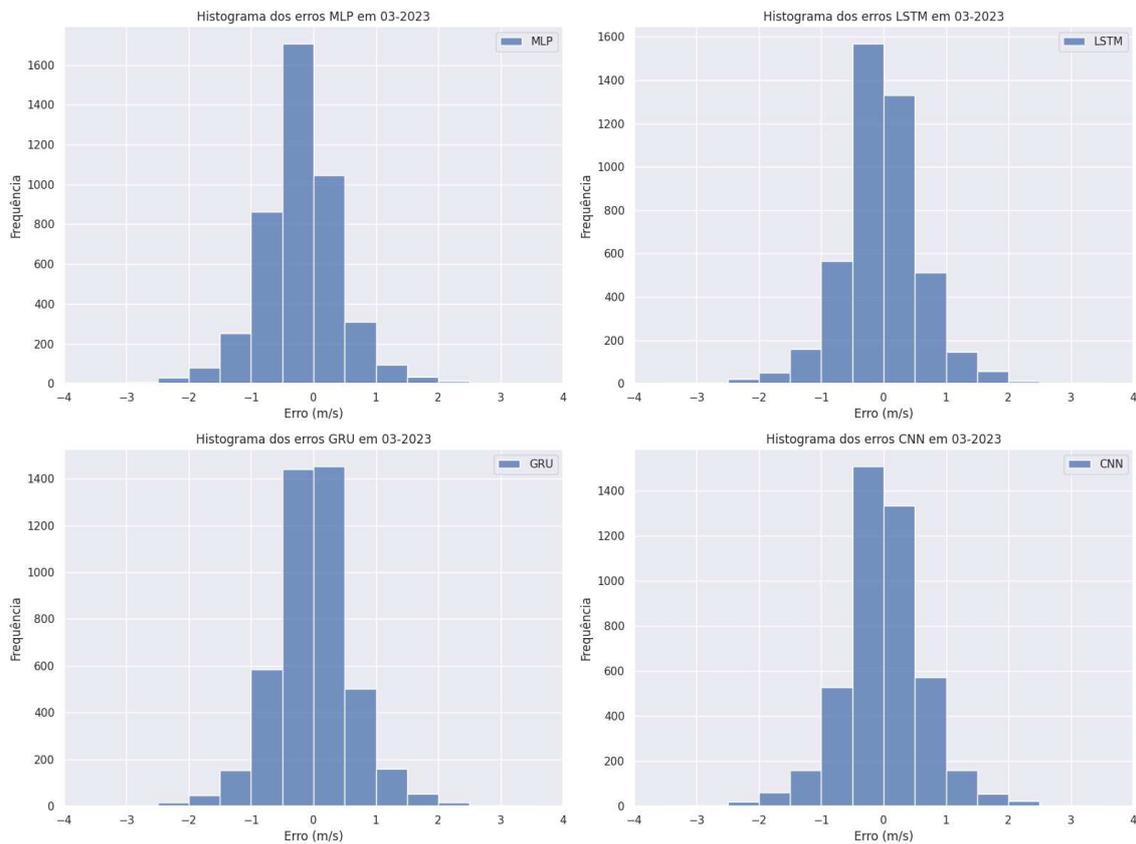
Fonte: O próprio autor.

Figura 24 - Velocidade do vento Prevista versus Real em março de 2023.



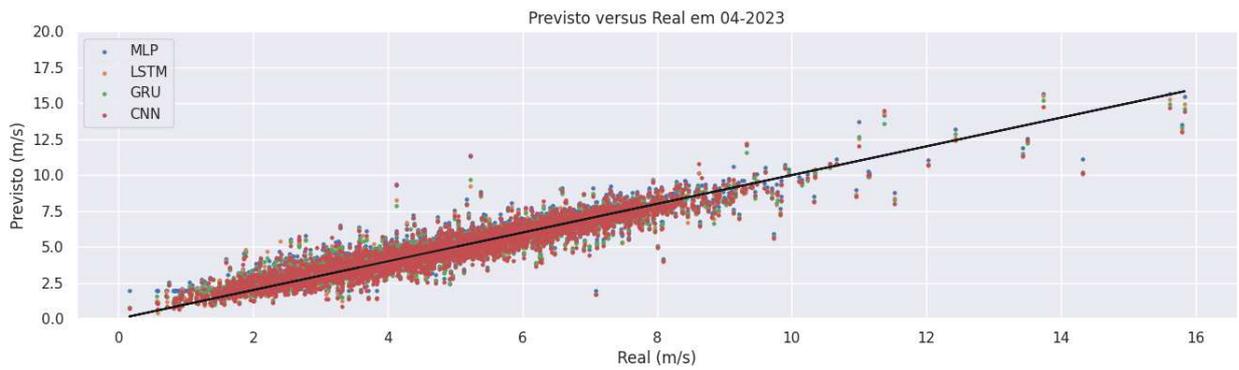
Fonte: O próprio autor.

Figura 25 - Distribuição dos erros de previsão em março de 2023.



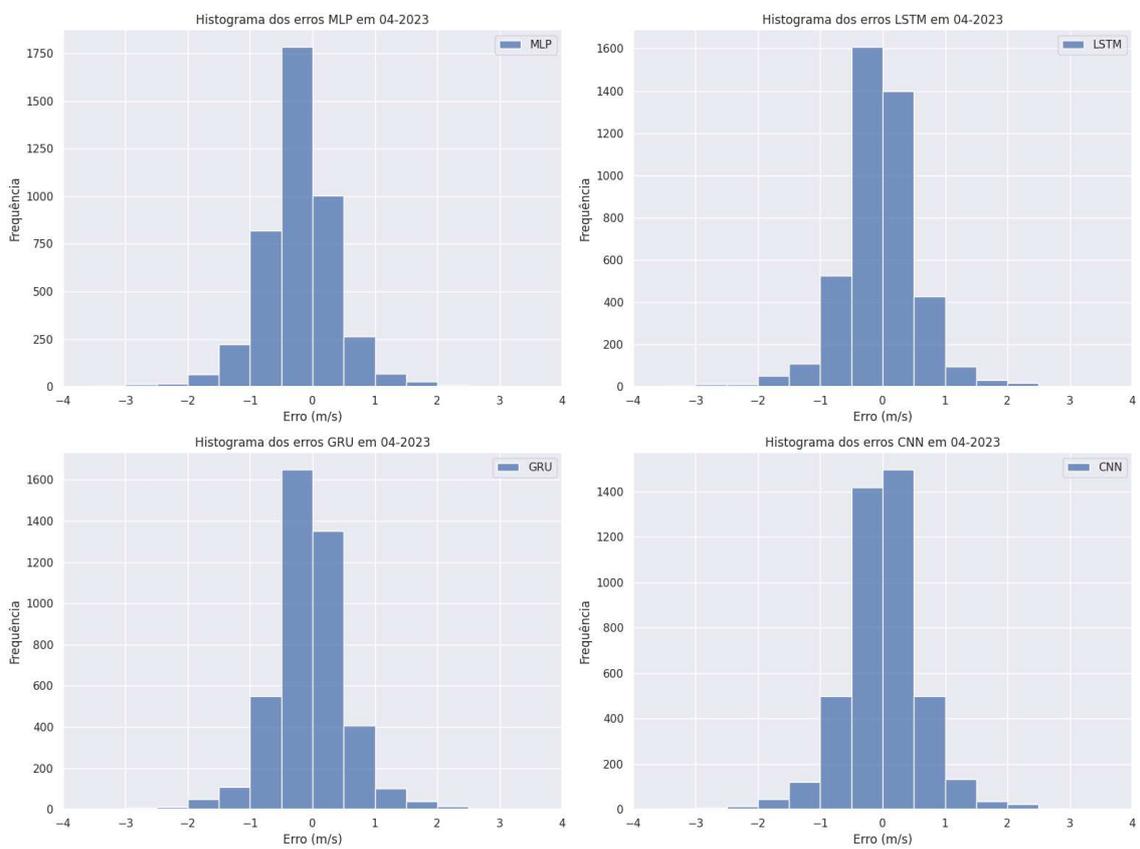
Fonte: O próprio autor.

Figura 26 - Velocidade do vento Prevista versus Real em abril de 2023.



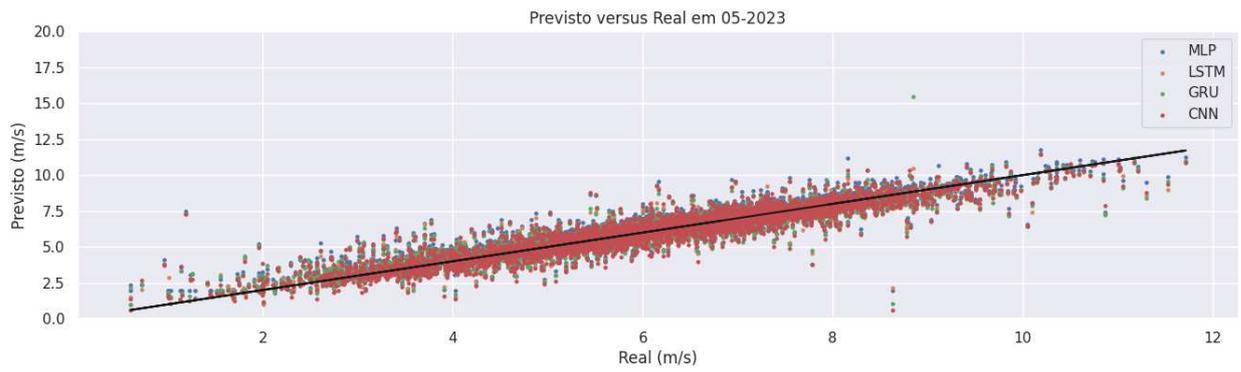
Fonte: O próprio autor.

Figura 27 - Distribuição dos erros de previsão em abril de 2023.



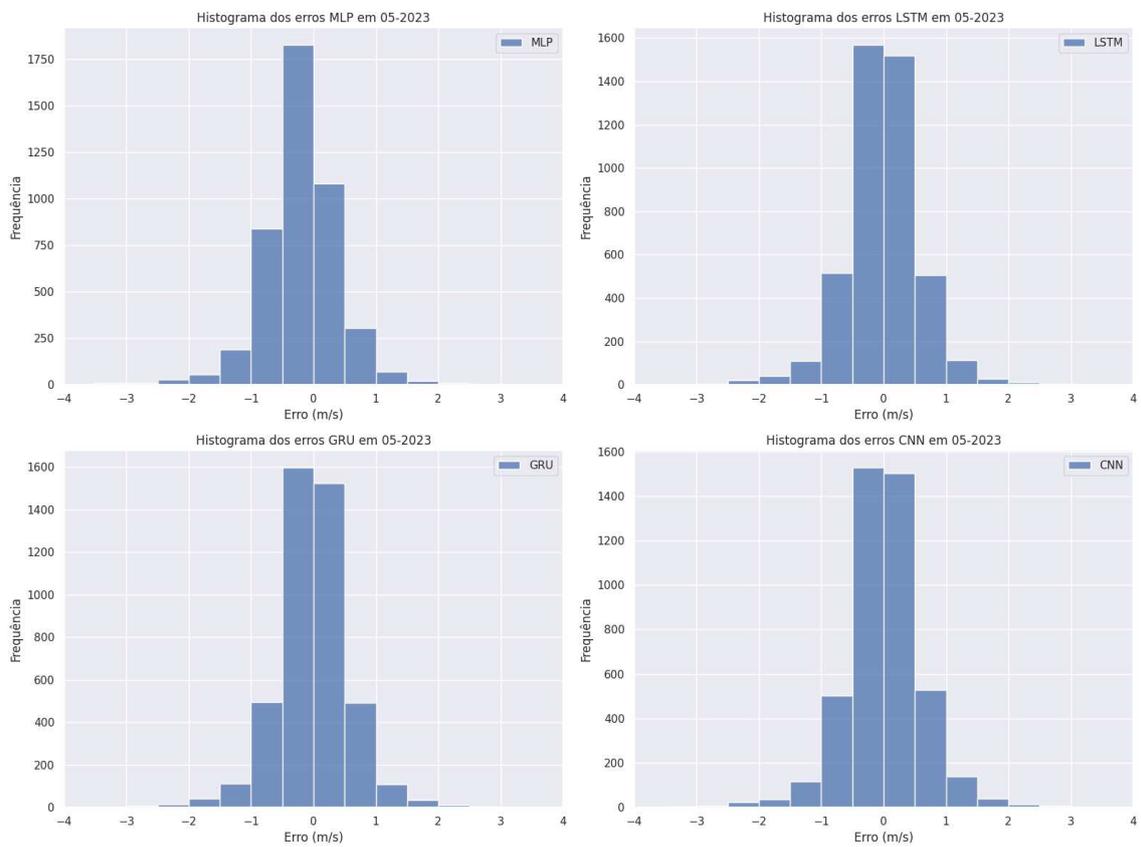
Fonte: O próprio autor.

Figura 28 - Velocidade do vento Prevista versus Real em maio de 2023.



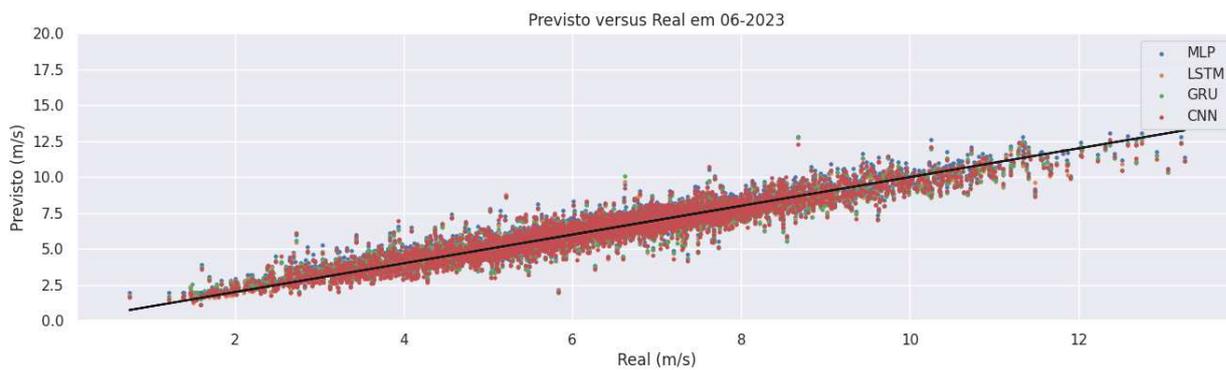
Fonte: O próprio autor.

Figura 29 - Distribuição dos erros de previsão em maio de 2023.



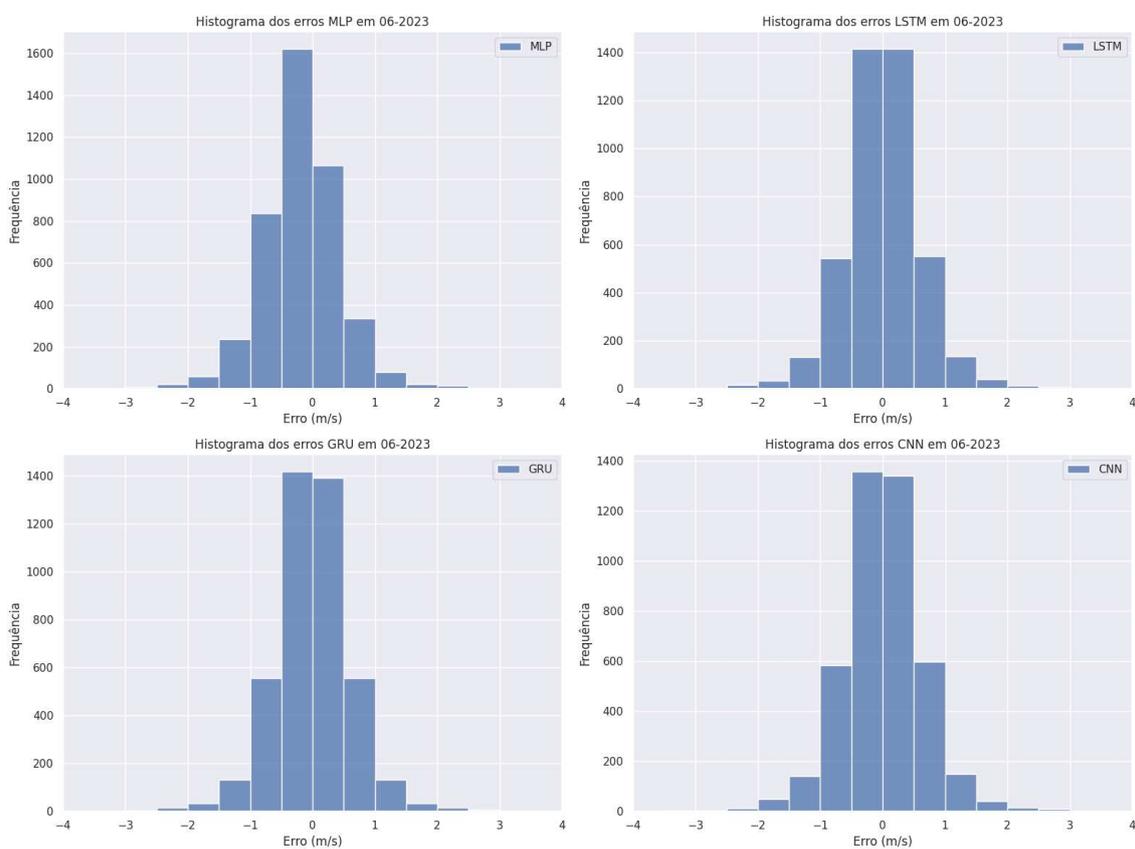
Fonte: O próprio autor.

Figura 30 - Velocidade do vento Prevista versus Real em junho de 2023.



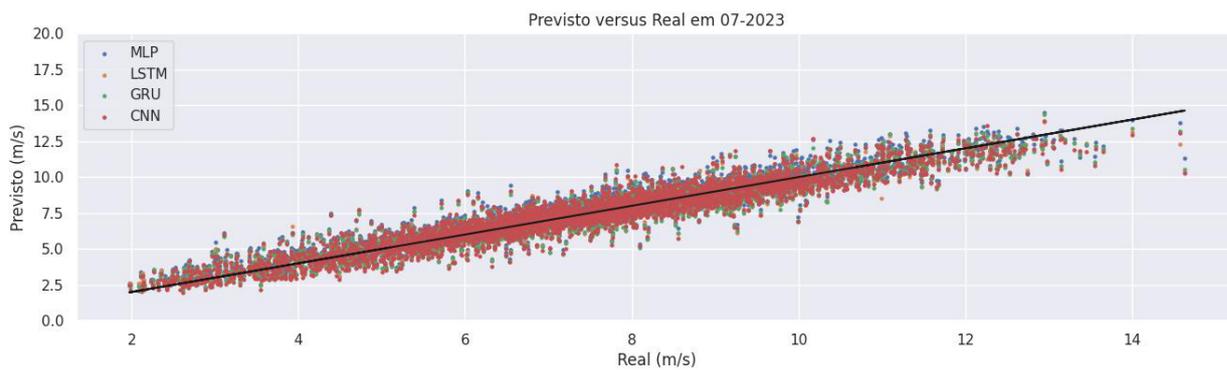
Fonte: O próprio autor.

Figura 29 - Distribuição dos erros de previsão em junho de 2023.



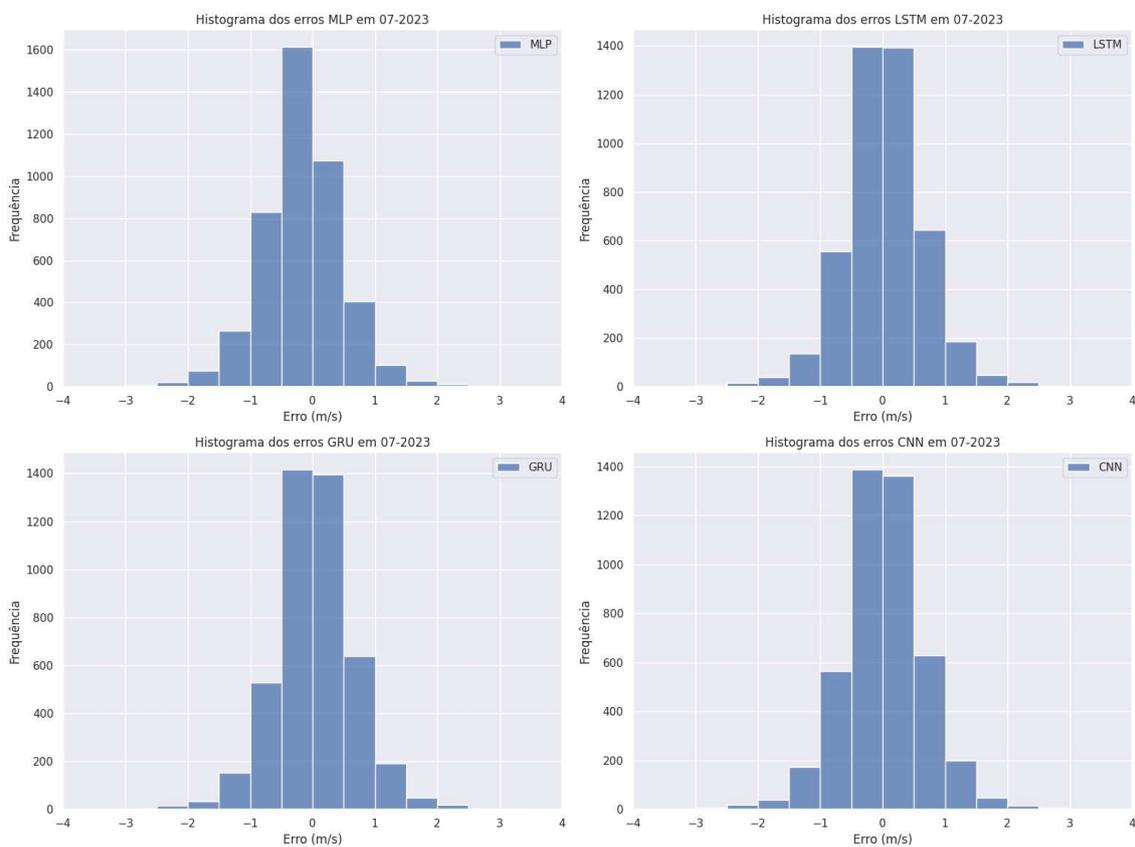
Fonte: O próprio autor.

Figura 30 - Velocidade do vento Prevista versus Real em julho de 2023.



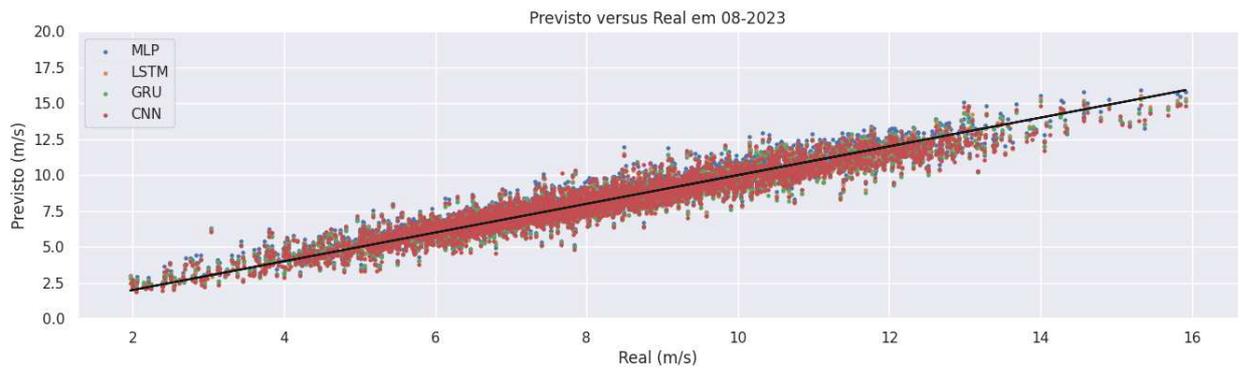
Fonte: O próprio autor.

Figura 31 - Distribuição dos erros de previsão em julho de 2023.



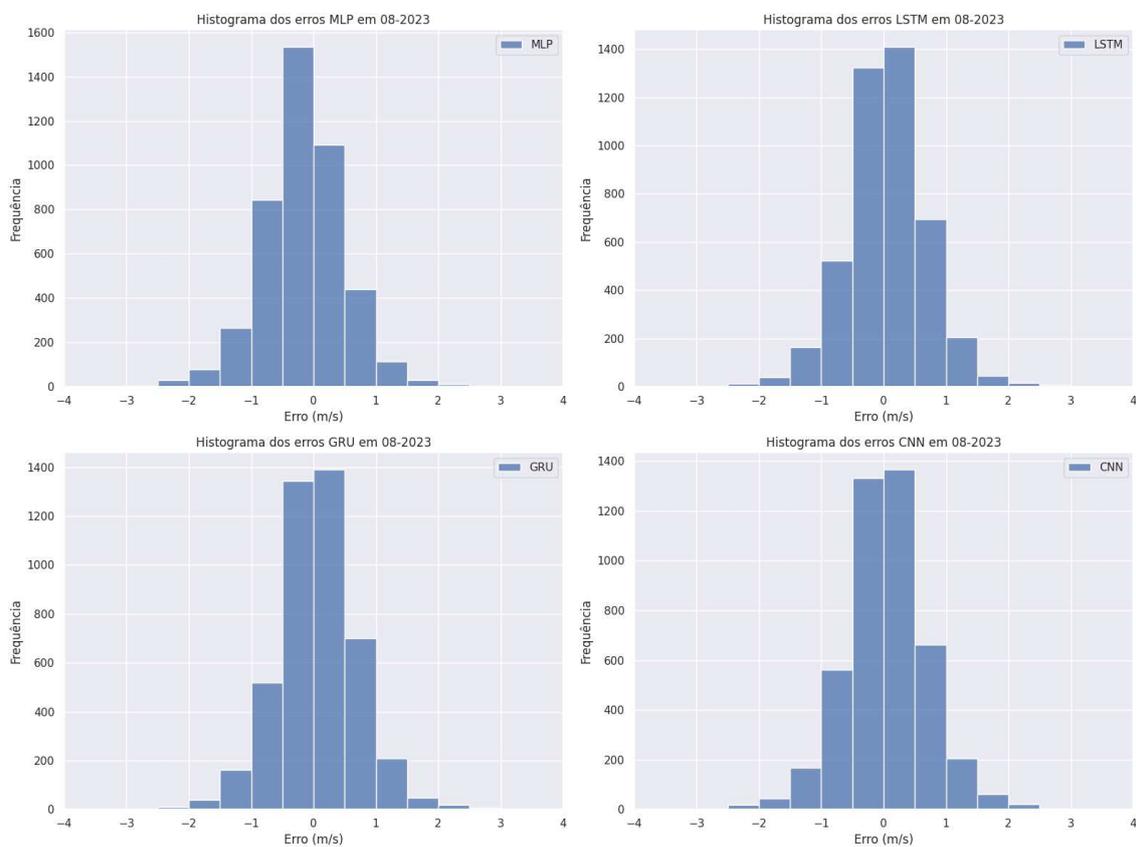
Fonte: O próprio autor.

Figura 32 - Velocidade do vento Prevista versus Real em agosto de 2023.



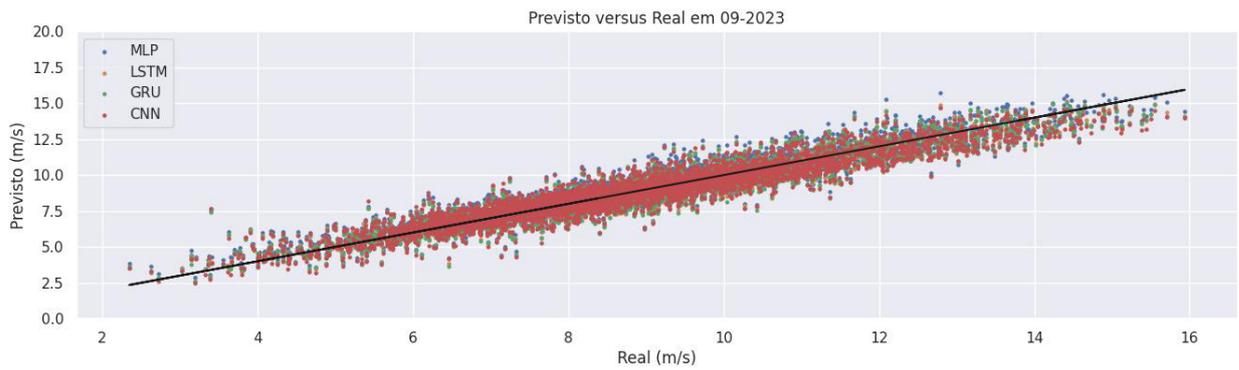
Fonte: O próprio autor.

Figura 33 - Distribuição dos erros de previsão em agosto de 2023.



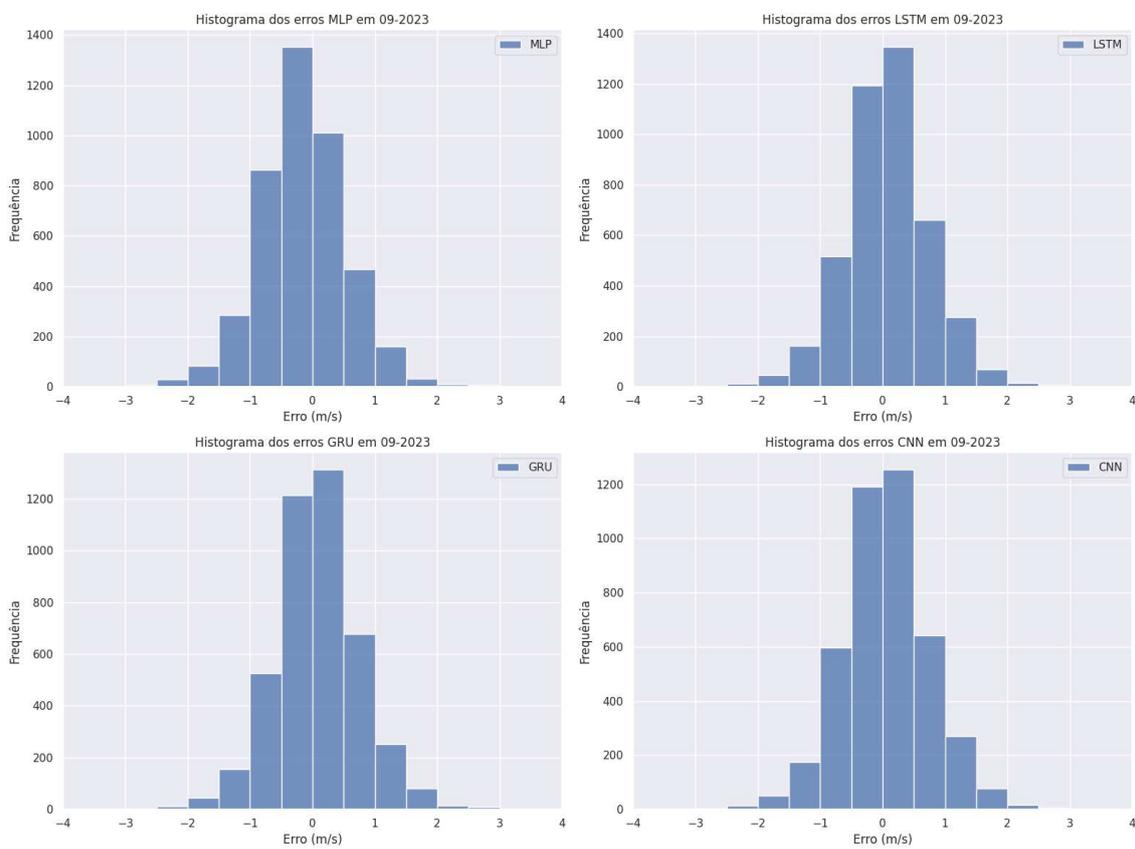
Fonte: O próprio autor.

Figura 34 - Velocidade do vento Prevista versus Real em setembro de 2023.



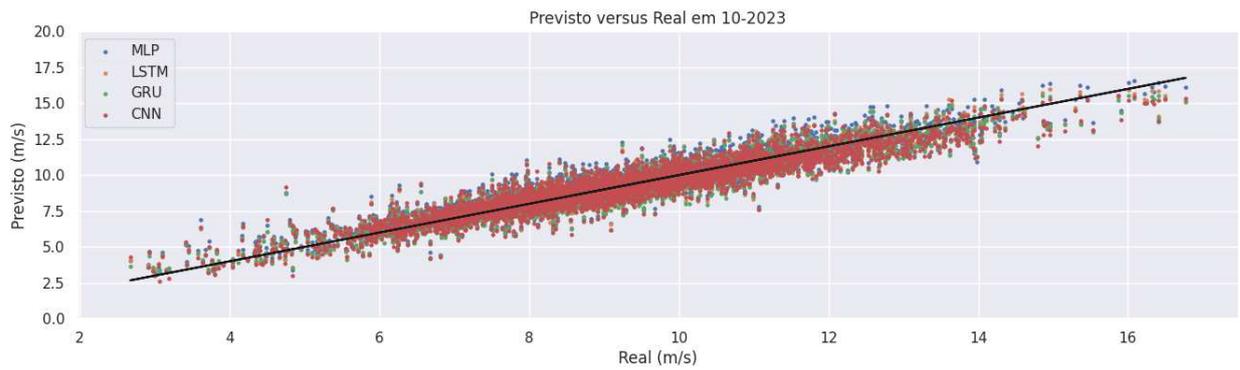
Fonte: O próprio autor.

Figura 35 - Distribuição dos erros de previsão em setembro de 2023.



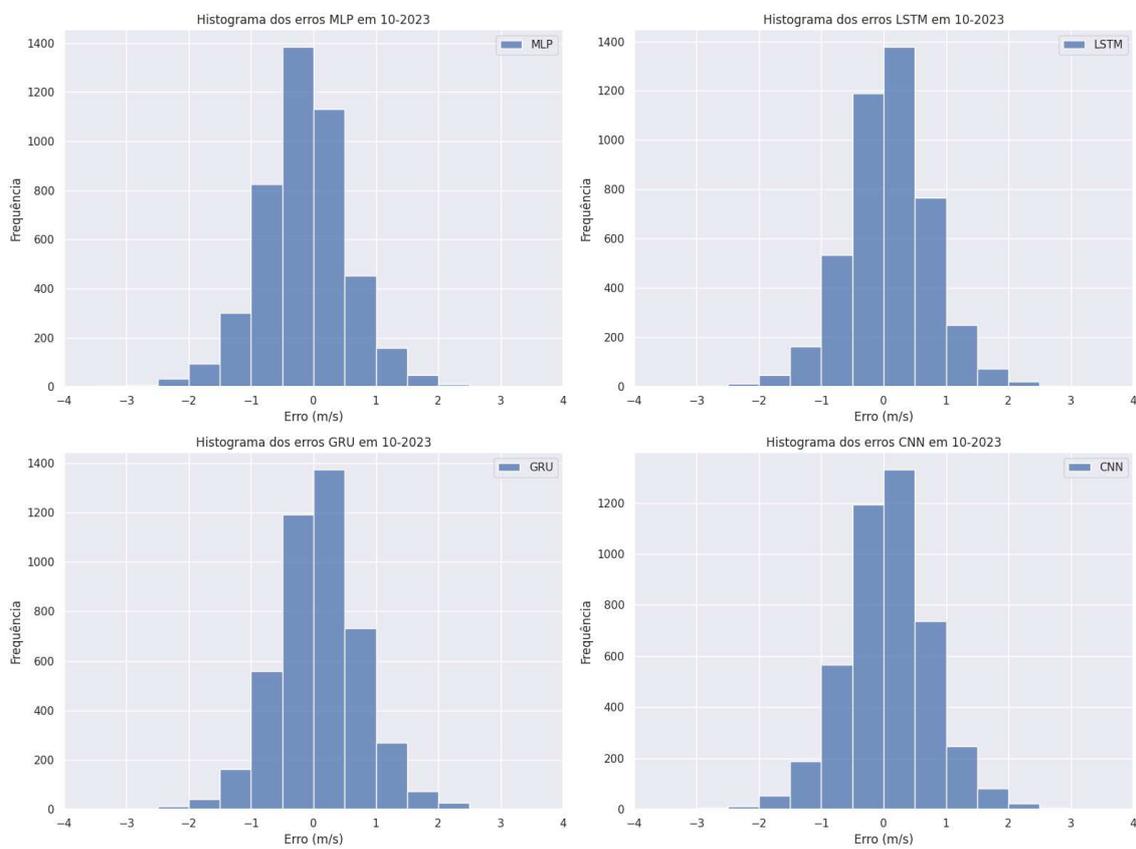
Fonte: O próprio autor.

Figura 36 - Velocidade do vento Prevista versus Real em outubro de 2023.



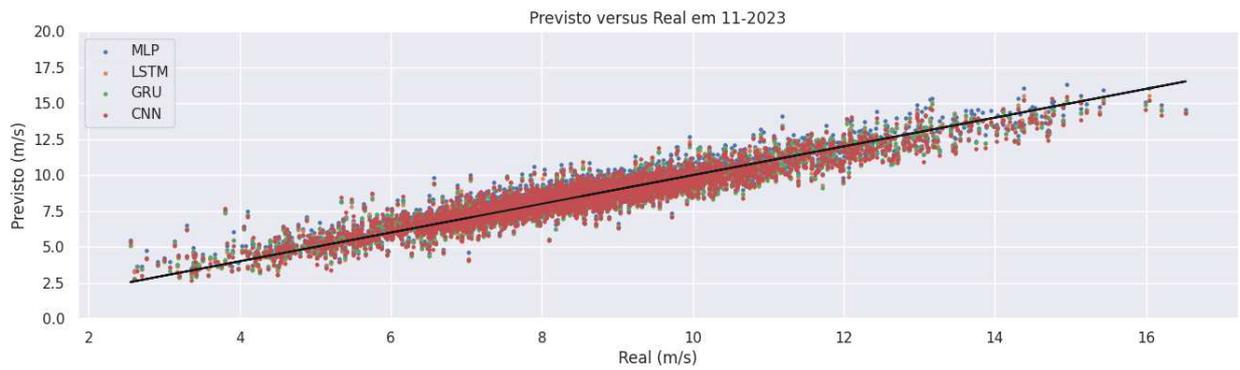
Fonte: O próprio autor.

Figura 37 - Distribuição dos erros de previsão em outubro de 2023.



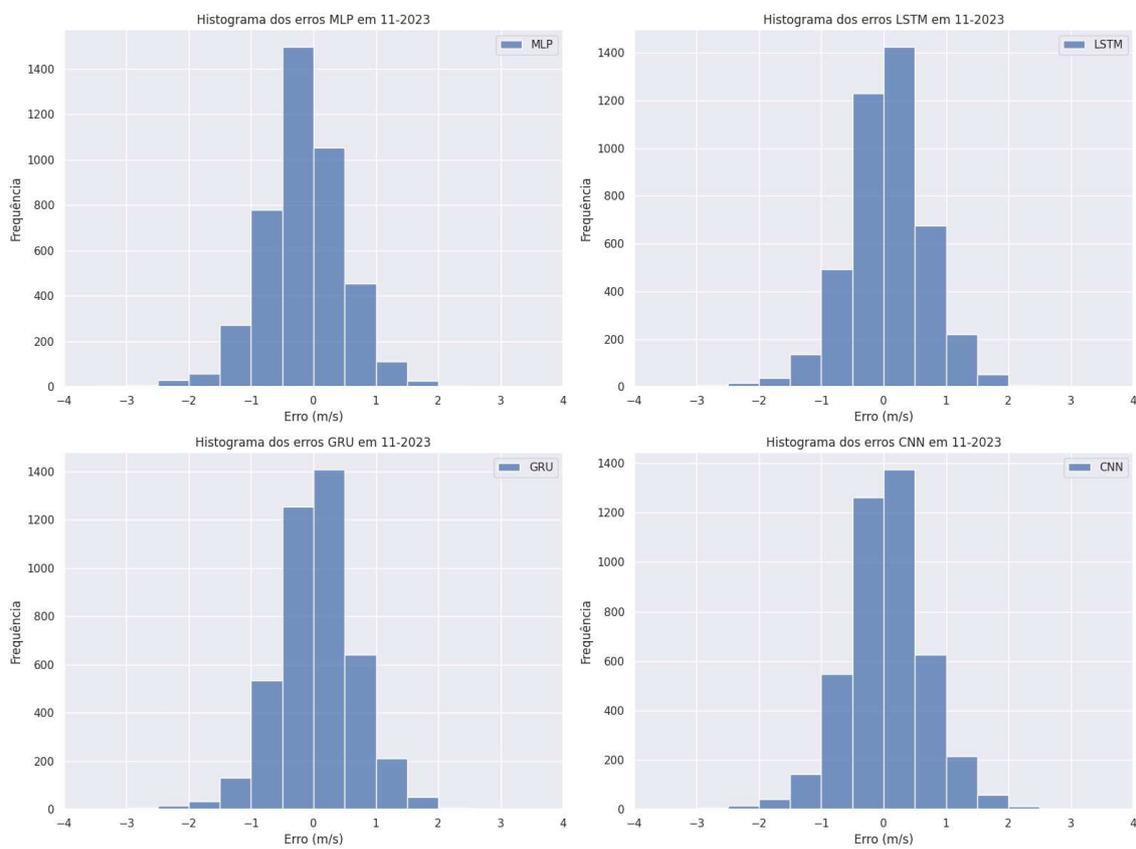
Fonte: O próprio autor.

Figura 38 - Velocidade do vento Prevista versus Real em novembro de 2023.



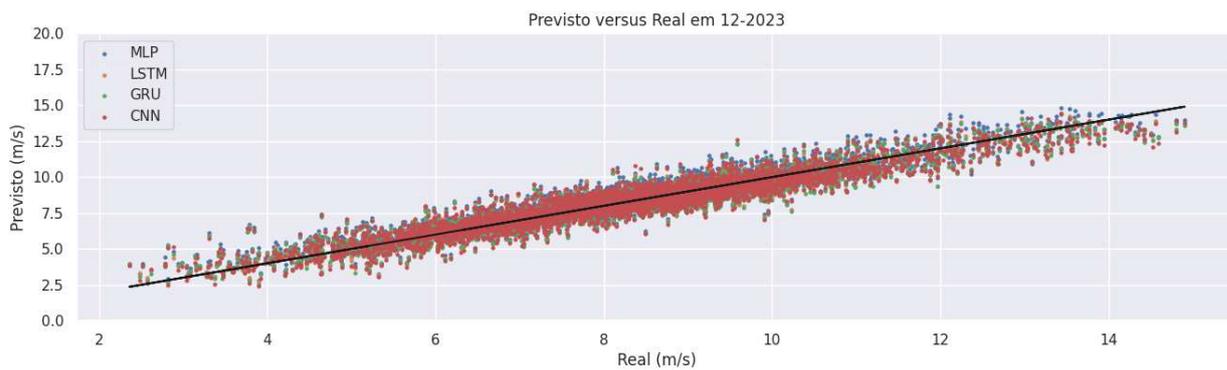
Fonte: O próprio autor.

Figura 39 - Distribuição dos erros de previsão em novembro de 2023.



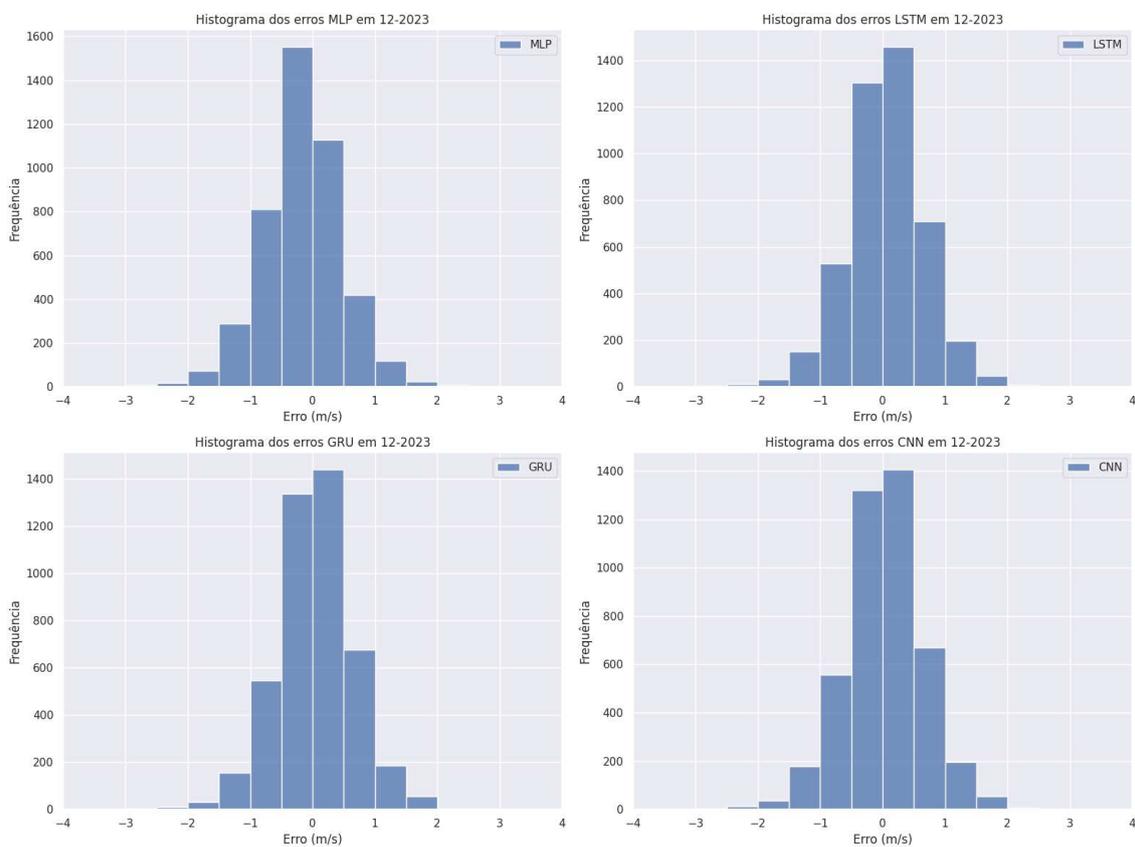
Fonte: O próprio autor.

Figura 40 - Velocidade do vento Prevista versus Real em dezembro de 2023.



Fonte: O próprio autor.

Figura 41 - Distribuição dos erros de previsão em dezembro de 2023.



Fonte: O próprio autor.