UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LEONARDO MENDES SOUSA XAVIER

APLICAÇÕES DE APRENDIZADO DE MÁQUINA NA DETECÇÃO DE ANOMALIAS EM SISTEMAS EÓLICOS

Recife
2024

LEONARDO MENDES SOUSA XAVIER

APLICAÇÕES DE APRENDIZADO DE MÁQUINA NA DETECÇÃO DE ANOMALIAS EM SISTEMAS EÓLICOS

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Processamento de Energia.

Orientador: Prof. Dr. Pedro André Carvalho Rosas

Coorientador: Prof. Dr. Gustavo Medeiros de Souza Azevedo

Recife

2024

Catalogação na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

X3a Xavier, Leonardo Mendes Sousa.
Aplicações de aprendizado de máquina na detecção de anomalias em sistemas eólicos / Leonardo Mendes Sousa Xavier – 2024.
102 f.: il., figs., tabs., abrev. e siglas.

Orientador: Prof. Dr. Pedro André Carvalho Rosas.
Coorientador: Prof. Dr. Gustavo Medeiros de Souza Azevedo.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2024.
Inclui Referências.

1. Engenharia Elétrica. 2. Aprendizagem de máquina. 3. Detecção de falhas. 4. Análise preditiva. 5. Energia eólica. I. Rosas, Pedro André Carvalho (Orientador). II. Azevedo, Gustavo Medeiros de Souza (Coorientador). III. Título.

UFPE

621.3 CDD (22. ed.) BCTG/2024-62

LEONARDO MENDES SOUSA XAVIER

APLICAÇÕES DE APRENDIZADO DE MÁQUINA NA DETECÇÃO DE ANOMALIAS EM SISTEMAS EÓLICOS

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Processamento de Energia.

Aprovada em: 29 de fevereiro de 2024.

BANCA EXAMINADORA

Prof. Dr. Pedro André Carvalho Rosas
(Orientador e Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. José Filho da Costa Castro
(Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Jeydson Lopes da Silva
(Examinador Externo)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Inicialmente, gostaria de agradecer a Deus por ter saúde física, mental e espiritual, o que torna possível percorrer todos os desafios impostos até o momento. Agradeço imensamente aos meus pais, por todo o investimento escolar depositado em mim e por ter propiciado um ambiente em que eu pudesse estudar, sem a preocupação se terei onde dormir ou se terei o que comer todos os dias. Agradeço a todos os professores pelos ensinamentos teóricos e práticos das disciplinas, pelas lições de vida ensinadas e por me tornar uma pessoa extremamente resiliente às adversidades. À minha esposa, que sempre me apoiou nas decisões difíceis que apareceram ao longo da caminhada e que tem estado ao meu lado durante todo esse percurso acadêmico. Ao meu orientador Pedro Rosas e ao professor Gustavo Medeiros por todo suporte que me foi dado para que este trabalho pudesse ser produzido e por me acolher tão rapidamente como orientadores. Ainda, agradeço à sociedade como um todo, por pagar os impostos que me permitiram estar em uma Universidade Pública de referência. Por fim, agradeço à banca examinadora por aceitar e avaliar o presente trabalho e por realizar as devidas contribuições.

"Agora, aqui, sabe, é necessária toda a corrida que você tem para se manter no mesmo lugar. Se você quer ir a um lugar diferente, você deve correr pelo menos duas vezes mais rápido que aquilo!"

(Lewis Carroll)

RESUMO

Neste trabalho, é apresentada uma metodologia voltada para identificação de anomalias em aerogeradores, visando a antecipar variações de temperatura em componentes críticos da máquina. Os elementos submetidos à análise incluem o rolamento da caixa de engrenagem e o rolamento *drive-end* do gerador. O estudo fundamenta-se na modelagem e aplicação de algoritmos de aprendizagem de máquina com o intuito de prever a temperatura nesses componentes. Os algoritmos de regressão empregados abrangem a regressão múltipla linear, o aumento de gradiente extremo e uma rede neural recorrente denominada memória de curto longo prazo. Para modelagem das técnicas, são utilizados dados provenientes do sistema de supervisão de três aerogeradores pertencentes a um parque eólico brasileiro composto por doze máquinas. Foi optado pela escolha de máquinas com comportamento de temperatura distinto entre si, a fim de avaliar se há variações relevantes no desempenho dos modelos de aprendizado diante de comportamentos térmicos distintos. Os dados das máquinas são submetidos a uma etapa de pré-processamento para identificar valores atípicos da operação normal dos aerogeradores. Posteriormente, os dados são divididos em conjuntos específicos para aplicação do algoritmo. No caso do modelo de aumento de gradiente extremo, foi empregada uma técnica de otimização Bayesiana para encontrar os parâmetros ótimos que se adéquam ao conjunto de dados propostos. Os resultados dos algoritmos de regressão são analisados sob a ótica das métricas de desempenho, e, ainda, são realizadas comparações entre as temperaturas reais e previstas, dentro de limites de controle definidos, visando a identificação de anomalias na temperatura dos elementos estudados. Por fim, os modelos aplicados às três máquinas são comparados entre si para cada componente analisado. As principais vantagens deste modelo incluem sua capacidade em fornecer resultados excelentes para problemas complexos de previsão, baixo custo financeiro para implementação e alta adaptabilidade de implementação em outras máquinas.

Palavras-chave: aprendizagem de máquina; detecção de falhas; análise preditiva; energia eólica.

ABSTRACT

In this work, a methodology focused on identifying anomalies in wind turbines is presented, aiming to anticipate temperature variations in critical machine components. The elements subjected to analysis include the gearbox bearing and the drive-end bearing of the generator. The study is based on modeling and applying machine learning algorithms with the goal of predicting the temperature in these components. The employed regression algorithms encompass linear multiple regression, extreme gradient boosting, and a recurrent neural network called long short-term memory. For modeling these techniques, data from the supervision system of three wind turbines in a Brazilian wind farm consisting of twelve machines are utilized. Machines were chosen with distinct temperature behaviors to assess potential variations in the performance of learning models in the face of diverse thermal behaviors. The machine data undergo a preprocessing stage to identify outliers from the normal operation of wind turbines. Subsequently, the data is divided into specific sets for algorithm application. In the case of the extreme gradient boosting model, a Bayesian optimization technique was employed to find optimal parameters that suit the proposed dataset. The results of the regression algorithms are analyzed in terms of performance metrics, and comparisons between actual and predicted temperatures are conducted within defined control limits, aiming to identify anomalies in the temperature of the studied elements. Finally, the models applied to the three machines are compared for each analyzed component. The main advantages of this model include its ability to provide excellent results for complex prediction problems, low financial implementation costs, and high adaptability for implementation in other machines.

Keywords: machine learning; fault detection; predictive analysis; wind energy.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS

OPAEP	Organização dos Países Árabes Exportadores de Petróleo
ONS	Operador Nacional do Sistema Elétrico
UFPE	Universidade Federal de Pernambuco
IEA	*International Energy Agency*
GD	Geração Distribuída
GC	Geração Centralizada
LCOE	*Levelized Cost of Energy*
EPE	Empresa de Pesquisa Energética
PDE	Plano Decenal de Energia
CAPEX	*Capital Expenditure*
O&M	Operação e Manutenção
IoT	*Internet of Things*
SCADA	*Supervisory Control and Data Acquisition*
DFIG	*Doubly fed induction generator*
XGBoost	*Extreme Gradient Boost*
AM	Aprendizagem de Máquina
LSTM	Long Short-Term Memory
IA	Inteligência Artificial
RNA	Rede Neural Artificial
KNN	*k-Nearest Neighbors*
SVM	*Space Vector Machine*
IACs	*Department of Energy Industrial Assessments Centers*
MMQ	Método dos Mínimos Quadrados
ReLU	Rectified Linear Unit
AMA	Torre Anemométrica
PCA	*Principal Component Analysis*
RMSE	*Root Mean Squared Error*
MAE	*Mean Absolute Error*

MAPE	*Mean Absolute Percentage Error*
D.E.	*Drive-end*
MLR	*Multiple Linear Regression*

LISTA DE SÍMBOLOS

E_c	Energia cinética
m	Massa
v	Velocidade média
P	Potência
$\dot{m}$	Fluxo de massa
ρ	Densidade
A	Área
P_{disp}	Potência disponível
c	Parâmetro de escala
k	Fator de forma
v_{anual}	Velocidade média anual
r_{xy}	Coeficiente de Pearson
b	Coeficiente linear
a	Coeficiente angular
$r(b, a)$	Função objetiva de resíduo
$L^{(t)}$	Função objetiva do aumento de gradiente extremo
$\Omega(f_t)$	Termo de regularização
g_i	Gradiente
h_i	Hessiano
T	Nós folhas de uma árvore de decisão
λ	Hiperparâmetro da regularização de Ridge
w	Peso da árvore de decisão
γ	Hiperparâmetro de penalização
u_k	Neurônio de uma rede neural
w_{kn}	Pesos sinápticos
b_k	Viés da rede neural
H_t	Estado oculto atual
X_t	Entrada de dados SCADA

C_t	Estado atual da célula
F_t	Porta de esquecimento
I_t	Porta de entrada
$\tilde{C}_t$	Memória candidata
O_t	Porta de saída
p_d	Pressão parcial do ar seco
M_d	Massa molar do ar seco
p_v	Pressão de vapor da água
M_v	Massa molar do vapor da água
R	Constante do gás ideal
T_k	Temperatura em Kelvin
T_c	Temperatura em Celsius
X_{norm}	Normalização de mínimo e máximo

SUMÁRIO

1 INTRODUÇÃO

Neste capítulo, é apresentada uma visão geral sobre o panorama das energias renováveis no mundo, a produção de energia eólica e as ações visando o decrescimento do uso de combustíveis fósseis. Em seguida, são discutidos os conceitos da produção eólica, bem como os princípios de funcionamento de uma turbina eólica, as tecnologias das máquinas utilizadas nos aerogeradores e as principais falhas apresentadas nos aerogeradores. Posteriormente, são discutidos os principais desafios da operação e manutenção e os custos associados na manutenção de um parque eólico. Por fim, são descritos os objetivos do estudo proposto e a organização textual da dissertação.

1.1. Crises energéticas

A transição para fontes de energia renováveis tornou-se imperativa diante dos desafios associados ao uso intensivo de combustíveis fósseis. O investimento em fontes renováveis, como solar, eólica e a hidrelétrica traz uma série de benefícios, como o desenvolvimento tecnológico do setor de energia, a redução da poluição do ar, criação de novos empregos e, principalmente, diversificação da matriz energética, reduzindo a dependência de países em relação a importações de combustíveis fósseis, promovendo assim, a estabilidade no abastecimento de energia.

Em 1973, foi desencadeada a crise do petróleo devido ao embargo imposto pela Organização dos Países Árabes Exportadores de Petróleo (OPAEP) contra nações que apoiaram Israel durante a Guerra do Yom Kippur. A ação teve como objetivo retaliar o apoio ocidental a Israel durante o conflito no Oriente Médio e teve consequências imediatas. Com o aumento do preço do petróleo, diversos setores da sociedade foram impactados, como a indústria, agricultura e o transporte, causando uma desaceleração econômica global (BINI; GARAVINI; ROMERO, 2016).

Os países industrializados do Ocidente, bastante dependentes do petróleo, notaram sua fragilidade frente à crise energética. Nos Estados Unidos, cuja dependência do petróleo é elevada, a NASA[1], agência espacial americana buscou uma solução para o problema adotando o programa de energia eólica do governo americano, destinando um orçamento de duzentos milhões de dólares para o programa (PINTO, 2013). No mesmo período, na

[1] *National Aeronautics and Space Administration*

Dinamarca, uma comissão de especialistas declarou que com a utilização da energia eólica, seria possível prover 10% da necessidade energética do país, movimentando o setor eólico e contribuindo para o avanço tecnológico dos sistemas de captação do vento (PINTO, 2013).

Do ponto de vista energético, foi destacada a vulnerabilidade das economias dependentes do petróleo, o que incentivou esforços para diversificar as fontes de energia e promover a eficiência energética. Muitos países passaram a buscar alternativas, incluindo o desenvolvimento de fontes de energia renováveis, a fim de reduzir a dependência do petróleo e mitigar futuras crises similares.

Em 1990, verificou-se uma disparada nos valores do petróleo em decorrência da Guerra do Golfo, momento em que o Iraque procedeu à invasão do Kuwait. À época, ambas as nações representavam aproximadamente 9% da produção mundial de petróleo (HAMILTON, 2013). Uma vez mais, constatou-se o impacto abrangente sobre a economia global, principalmente aos países que possuíam elevada dependência de combustíveis fósseis (HUTCHISON, 1991).

Em 2022, a comunidade global foi testemunha da invasão russa no território ucraniano, causando instabilidades geopolítica em todo continente asiático e europeu. O acontecimento desencadeou uma intensa demanda por petróleo e gás, resultando em uma crise energética sem precedentes na Europa. Isso ocorreu devido a considerável influência que a Federação Russa exerce na oferta de gás natural para países membros da União Europeia (UE), visto que a Rússia é o segundo maior produtor global de gás (COUNCIL, 2023).

No Brasil, as principais crises energéticas ocorreram em 2001 e em 2021. A crise de 2001, conhecida como Crise do Apagão, ocorreu devido a elevada dependência do país com a fonte hidráulica, à escassez de chuvas, ao baixo índice de água nos reservatórios das hidrelétricas e a falta de políticas de planejamento energético. Para controlar a crise, o governo realizou diversos cortes controlados de energia e limitou eventos noturnos. De acordo com (TOLMASQUIM, 2000), a crise energética de 2001 teve como origem a falta de investimentos em geração e em transmissão, negligenciada muitos anos pelo poder público.

Em 2021, houve uma expressiva redução nos níveis dos reservatórios no Brasil, resultante da escassez de chuvas e da utilização dessas fontes hídricas para diversos fins.

Ainda, sabe-se, que as usinas precisam liberar uma vazão mínima de seus reservatórios, visando preservar as atividades essenciais que dependem da água, como agricultura, irrigação e transporte; o que afeta diretamente o nível do reservatório e pode comprometer a geração de energia elétrica, caso atinja níveis críticos.

E, de fato, o ano em questão foi marcado por uma severa escassez de chuvas, levando os principais reservatórios a atingirem níveis críticos. Conforme dados do Operador Nacional do Sistema Elétrico (ONS), datado de 1º de dezembro de 2021, as hidrelétricas do Sudeste e Centro-Oeste registraram volumes úteis extremamente baixos, com apenas 21,51% no reservatório de Furnas, 18,20% em Mascarenhas de Moraes, 15,26% em Marimbondo e 12,25% em Água Vermelha (ONS, 2021).

A questão da dependência energética ressalta a vulnerabilidade de muitas nações a flutuações nos preços e disponibilidade de combustíveis fósseis. A instabilidade geopolítica em regiões produtoras de petróleo pode ter impactos diretos nos custos e na disponibilidade desses recursos. As discussões sobre dependência energética visam explorar estratégias concretas para fortalecer a resiliência dos países diante dessas oscilações. Isso pode envolver não apenas a ampliação do uso de fontes renováveis, mas também a implementação de políticas eficazes de eficiência energética, armazenamento de energia em larga escala e a promoção de práticas de consumo sustentáveis.

No âmbito dos combustíveis fósseis, estudos indicam que o preço do petróleo teve relação direta com o desenvolvimento das tecnologias utilizadas na geração de energia elétrica renovável. (NUNES; CATALãO-LOPES, 2020), investigou o impacto dos preços do petróleo na inovação de fontes alternativas (renováveis) de energia, em particular o impacto da queda dos preços do petróleo utilizando um modelo de regressão binomial negativo. Os resultados mostram uma relação positiva entre os preços do petróleo e o número de pedidos de patentes no setor de energias alternativas.

De maneira análoga, (BAYER; DOLAN; URPELAINEN, 2013) realizou um estudo utilizando regressão binomial para identificar a relação de patentes de energia renovável com o preço do petróleo. O estudo mostrou que os elevados preços do petróleo têm fortes efeitos positivos na atividade de patenteamento.

Uma causa possível para esta resposta assimétrica é que os aumentos dos preços do petróleo forçaram países fortemente dependentes do petróleo, principalmente no âmbito da geração de energia elétrica, a investir em alternativas renováveis. De forma que os

investimentos gerem uma maior segurança energética para o país e ao mesmo tempo fortalece as políticas de alterações climáticas.

Dessa forma, fica evidente que as crises energéticas tiveram uma contribuição significativa no desenvolvimento de políticas e investimentos no setor renovável buscando alternativas diferentes das oferecidas pelo setor de óleo e gás.

1.2. **Desenvolvimento das energias renováveis**

A incorporação de projetos de energia renovável assume um papel estratégico fundamental na diversificação da matriz energética, atuando como elemento crucial na transição energética de fontes baseadas em combustíveis fósseis para a geração sustentável de energia. Globalmente, os empreendimentos voltados para energias renováveis destacam-se como líderes em desenvolvimento, projetando uma adição expressiva de 350 gigawatts (GW) em capacidade instalada para o ano de 2022. Desse montante, os projetos de energia solar fotovoltaica e eólica representam aproximadamente 90% (IEA, 2022).

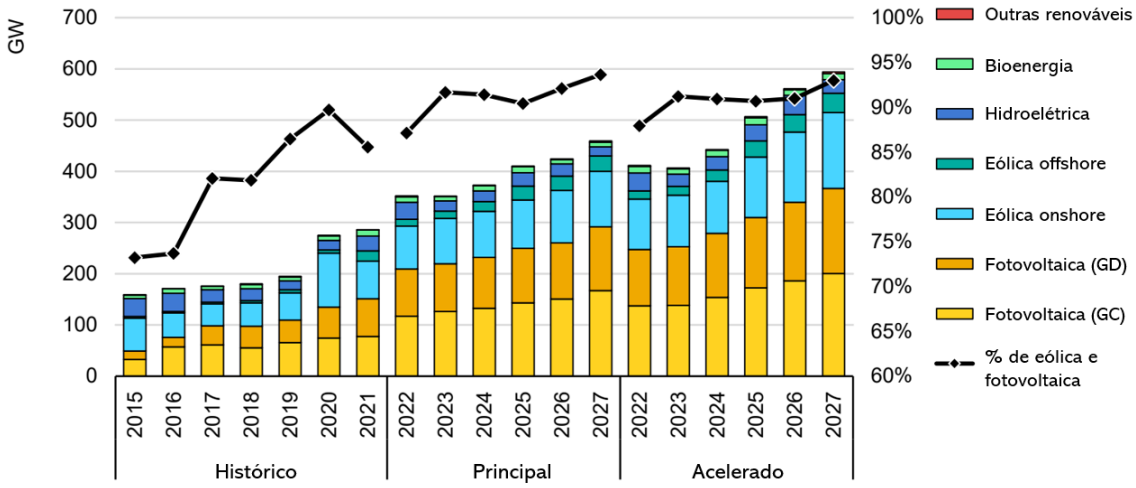
Os projetos de energia renovável têm um fator estratégico na diversificação da matriz energética do país e são cruciais para uma transição energética de combustíveis fósseis para a geração de energia limpa. No mundo, os projetos de energia renováveis são líderes em desenvolvimento com uma expectativa de adição de novos 350 GW em capacidade instalada em 2022, de forma que os projetos de energia solar fotovoltaica e eólica representam cerca de 90% desse montante (IEA, 2022).

A aceleração notável na expansão das energias renováveis ao longo dos últimos cinco anos pode ser atribuída a diversos fatores, porém destaca-se os elevados custos dos combustíveis fósseis e da eletricidade, decorrentes das recentes crises energéticas (IRENA, 2023). Essa conjuntura tornou as energias renováveis mais atraentes do ponto de vista financeiro. Além disso, como já mencionado, a invasão russa à Ucrânia, ocasionou impactos diretos nas exportações de gás e petróleo do governo russo para os países europeus dependentes.

De acordo com o (ONS, 2024), no Brasil, há cerca de 27,53 GW em capacidade instalada de usinas eólicas, com possibilidade de expansão ainda maior com os investimentos para produção de energia eólica *offshore*. No que concerna à energia solar fotovoltaica, a capacidade instalada atualmente é de 10,92 GW, considerando os projetos operacionais.

Além disso, existe uma expectativa de crescimento substancial nesse segmento até o ano de 2027, de acordo com as solicitações de acesso observadas no ONS.

Figura 1 – Histórico e previsão da adição de fontes renováveis, por tecnologia, para o período 2015 a 2027 no mundo.



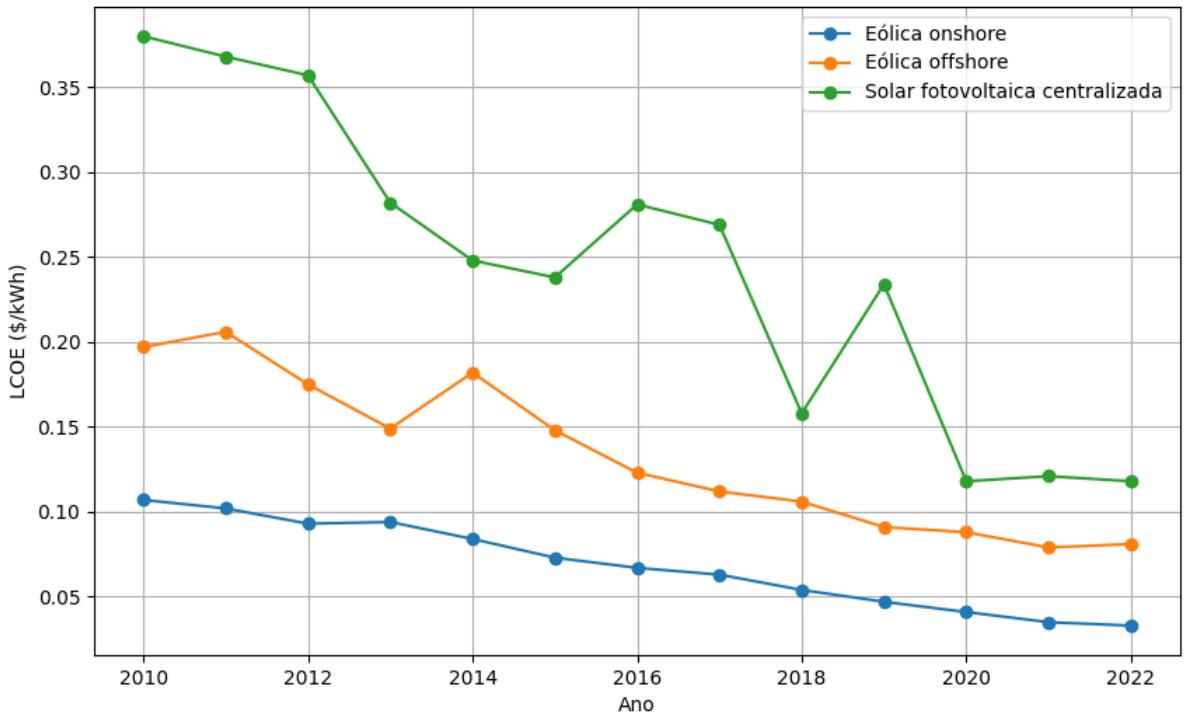
Fonte: Adaptado de (IEA, 2022).

A Agência Internacional de Energia (IEA) conduziu a simulação de dois cenários prospectivos, sendo um delineado como principal, adotando uma postura mais conservadora, e o segundo com uma projeção de crescimento acelerado. Conforme evidenciado pela Figura 1, observa-se uma expectativa de crescimento robusta em ambos os cenários para as fontes renováveis, incluindo a energia eólica em terra e fora da costa, bem como a energia solar, abrangendo tanto a Geração Distribuída (GD) quanto a Geração Centralizada (GC).

Um dos indicadores utilizados para determinar a viabilidade de um empreendimento renovável é o Custo Nivelado de Eletricidade (LCOE). É uma medida que avalia todos os custos associados à operação do parque ao longo de sua vida útil, como custo de implantação e operação e manutenção. O indicador é utilizado para planejar o nível de investimento e normalmente tem dimensão de unidade monetária por energia, como o $/kWh.

Verifica-se que, ao longo dos últimos dez anos, os projetos de fontes renováveis tiveram uma queda expressiva do LCOE, incluindo instalações eólicas fora da costa. Na Figura 2 é ilustrado o comportamento do LCOE mundial ao longo dos últimos dez anos para as fontes renováveis eólica e solar.

Figura 2 – Comportamento do LCOE ($/kWh) para as fontes renováveis eólica e solar (2010-2022).



Fonte: Adaptado de (IRENA, 2023).

Conforme apontado por (IRENA, 2023), no contexto financeiro, para o período compreendido entre 2021 e 2022, observou-se uma queda de 2% no LCOE para os novos empreendimentos de energia solar fotovoltaica centralizada e uma redução mais expressiva de 5% para os projetos de energia eólica em terra. Por outro lado, registrou-se um incremento de 5% no LCOE para as novas instalações de parques eólicos em localidades marítimas. Porém, no cenário de longo prazo, o LCOE desses projetos teve uma queda significativa conforme indicado na Tabela 3, que destaca a variação para as principais fontes.

Tabela 3 – Comparação do LCOE ($/kWh) para Fontes Renováveis (2010-2022).

Fonte	**LCOE 2010 ($/kWh)**	**LCOE 2022 ($/kWh)**	**Variação (%)**
Eólica onshore	0,107	0,033	-69,2
Eólica offshore	0,197	0,081	-58,9
Solar fotovoltaica centralizada	0,380	0,118	-68,9

Fonte: Adaptado de (IRENA, 2023).

Essas variações nos custos refletem uma dinâmica econômica significativa no setor de energia renovável, indicando a contínua busca por eficiência financeira e competitividade. Essas mudanças, especialmente as quedas nos custos do solar fotovoltaico centralizado e da energia eólica em terra, sugerem avanços tecnológicos e otimizações operacionais que contribuíram para a viabilidade econômica dessas formas de geração de energia.

Na realidade brasileira, a Empresa de Pesquisa Energética (EPE) empreendeu análises abrangentes para avaliar os custos associados à implementação, operação e manutenção de tecnologias renováveis. Essas investigações integram o Plano Decenal de Expansão de Energia 2030 (PDE 2030), proporcionando uma visão prospectiva dos custos dessas tecnologias ao longo do período considerado. Na Tabela 4 é apresentada a avaliação do investimento inicial (CAPEX) com exclusão dos encargos financeiros durante o período de construção, além da análise do custo operacional associado à manutenção dos ativos éolico (*onshore* e *offshore*) e solar.

Tabela 4 – Análises do PDE 2030 avaliando o investimento inicial (CAPEX) e custos de Operação e Manutenção (O&M).

Fonte	**Capex (R$/kW)**	**O&M (R$/kW/ano)**
Eólica onshore	4.500	90
Eólica offshore	12.250	490
Solar fotovoltaica	3.200	50

Fonte: Adaptado de ((EPE), 2021b).

Conforme (IRENA, 2018), os custos de O&M dos parques eólicos *onshore* podem representar até 30% do LCOE. Por outro lado, para parques eólicos *offshore*, devido aos elevados fatores de capacidade, os custos de O&M são amortizados pela maior produção de energia, resultando em uma variação de custo que pode situar-se entre 16% e 25% do LCOE. Essas estatísticas destacam a significativa influência dos custos de O&M na determinação da viabilidade econômica e eficiência operacional desses empreendimentos.

Com o intuito de mitigar os custos significativos associados à O&M dos parques eólicos, a gestão da manutenção direciona esforços para a redução dos custos globais do empreendimento ao longo de sua vida útil. Esta finalidade é alcançada mediante a implementação de um eficaz planejamento de paradas de máquinas, estrategicamente agendadas para períodos de baixa sazonalidade de ventos. Adicionalmente, a execução da

manutenção deve seguir rigorosamente as orientações do fabricante, evitando a substituição de componentes e lubrificantes não especificados, promovendo, assim, uma prolongada eficiência operacional.

No âmbito de uma revisão bibliográfica, (Pinar Pérez et al., 2013) identificou que a taxa média de falhas mais significativa ocorre nos componentes do sistema de controle, nas pás e no sistema elétrico de turbinas eólicas. Em uma análise complementar, (MCMILLAN, 2008) examinou as principais causas de inatividade de turbinas eólicas na Alemanha, constatando que a inatividade decorrente de falhas está predominantemente associada à caixa de engrenagem, ao gerador, ao rotor e ao mancal principal da turbina, representando 67% das ocorrências. Estas conclusões destacam a relevância crítica desses componentes na eficácia operacional de parques eólicos, ressaltando a importância de estratégias de manutenção específicas para otimizar a confiabilidade e a disponibilidade do sistema.

No cenário atual, o Brasil possui uma extensa rede de parques eólicos, muitos dos quais apresentam uma idade avançada e empregam tecnologias já obsoletas. Considerando que os contratos da fonte eólica nos Leilões têm, em geral, duração de 20 anos, prazo equivalente à vida útil de projeto dos equipamentos, percebe-se que os empreendimentos em operação desde os anos 90 já atingiram essa idade ((EPE), 2021a). Até 2030, mais de 50 parques alcançarão a faixa dos 20 anos de operação, representando mais de 600 aerogeradores e de 940 MW de potência ((EPE), 2021a).

A eficiência dessas instalações é diretamente afetada pela qualidade da manutenção empregada nas máquinas. Nesse contexto, a adoção de ferramentas preditivas pode trazer previsibilidade a gestão operacional, proporcionando uma abordagem proativa na identificação de falhas potenciais, permitindo intervenções antes que ocorram falhas catastróficas.

A busca pela eficiência na gestão da manutenção inclui, ainda, a provisão de redundância não apenas nos componentes associados aos aerogeradores, mas também na infraestrutura da rede de média tensão. Esta redundância engloba dispositivos como para-raios, cadeia de isoladores, linhas de média tensão e baterias, contribuindo para a confiabilidade e disponibilidade contínuas do parque eólico. Essas práticas não apenas visam a otimização dos custos, mas também a maximização do desempenho e da durabilidade do empreendimento, alinhando-se com as metas de sustentabilidade e eficiência energética.

Dessa maneira, é essencial a busca por ferramentas que maximizem a produção da energia eólica com segurança, confiabilidade e disponibilidade. Nesse contexto, a implementação de sistemas avançados de monitoramento, como soluções baseadas em Internet das Coisas (IoT) e análise de dados em tempo real, emerge como uma abordagem crucial. Essas tecnologias possibilitam a detecção precoce de falhas, permitindo a intervenção proativa e a redução do tempo de inatividade das máquinas.

Além disso, a proposta desse estudo visa utilização de sensores e equipamentos já existentes no parque eólico, reduzindo bastante o custo para implementação dos modelos de aprendizagem. As estratégias de manutenção preditiva, baseadas em algoritmos e aprendizado de máquina, podem então ser empregadas para antecipar potenciais problemas, otimizando os intervalos de manutenção e estendendo a vida útil dos componentes. A integração dessas ferramentas não apenas aprimora a eficiência operacional dos parques eólicos, mas também contribui para a sustentabilidade e competitividade contínuas do setor de energia renovável.

1.3. **Noções gerais de energia eólica**

Os sistemas eólicos são caracterizados pela sua capacidade de transformação da energia mecânica derivada do vento em energia elétrica. O vento pode ser definido como a deslocação de gases atmosféricos em grande escala causada por diferenças na pressão atmosférica e pelo aquecimento desigual do planeta. De acordo com (PINTO, 2013), somente aproximadamente 3% a 5% da radiação solar incidente é transformada em energia cinética capaz de impulsionar o movimento atmosférico por meio de variações de temperatura, estabelecendo assim o fundamento para a energia eólica.

Dessa corrente, apenas uma fração pode ser efetivamente capturada como energia eólica, utilizando aerogeradores instalados em altitudes elevadas. Nesse contexto, diversas forças atuam nas massas de ar na atmosfera como a força do gradiente de pressão, de Coriolis, centrífuga, atrito e a da gravidade. Dessa maneira, para facilitar as modelagens matemáticas que visam representar o comportamento dessas forças na natureza, o vento é modelado como correntes contínuas de parcelas de ar (PINTO, 2013).

Um dos interesses em desenvolver sistemas eólicos é a produção da energia elétrica de maneira eficiente, segura e confiável. No âmbito da questão energética, sabe-se que o ar em movimento contém energia cinética, que posteriormente é convertida em energia

elétrica com a utilização de aerogeradores. Considerando um aerogerador de três pás, pode-se supor que a massa de ar m irá atacar o bloco, com uma velocidade v de forma perpendicular, dessa maneira a energia cinética é dada por,

$$E_c = \frac{1}{2}mv^2, \tag{1.1}$$

Nota-se que a energia cinética aumenta com o quadrado da velocidade, de forma que se duplicarmos a velocidade, a energia cinética irá quadruplicar. Do ponto de vista da potência, é necessário avaliar a variação da energia cinética no tempo, ou seja, avaliar a taxa de variação.

$$P = \frac{\partial E_c}{\partial t} = \frac{1}{2}\dot{m}v^2, \tag{1.2}$$

Em que $\dot{m}$ pode ser definida como uma parcela da massa fluindo no tempo. Considerando que a vazão de um fluxo de massa pode ser calculada pela densidade do fluído (ar),

$$\dot{m} = \rho A v, \tag{1.3}$$

Então,

$$P = \frac{1}{2}(\rho A v)v^2, \tag{1.4}$$

Dessa maneira, a potência por unidade de área é calculada por,

$$P_{disp} = \frac{1}{2}\rho v^3. \tag{1.5}$$

Nesse momento, fica evidente que a potência disponível no vento é diretamente proporcional ao cubo da velocidade do ar, significando que a se a velocidade do vento dobrar, a potência disponível aumentará oito vezes.

Um aspecto importante no estudo da energia eólica consiste na compreensão da natureza estocástica do vento, ou seja, sua manifestação de maneira aleatória e sua susceptibilidade a diversas influências naturais, tais como temperatura e pressão local. Quando uma variável se manifesta de maneira contínua, mas apresenta um comportamento es-

tocástico, torna-se vantajoso realizar uma discretização para facilitar a análise do comportamento do vento (PINTO, 2013).

Para avaliar o comportamento do vento em um local, é comum a utilização da distribuição de Weibull, uma abordagem estatística amplamente empregada em estudos de recursos eólicos. A distribuição de Weibull é particularmente eficaz na modelagem da velocidade do vento, oferecendo uma representação flexível das características de sua distribuição de probabilidade. Além disso, a implementação prática desses modelos muitas vezes se beneficia de ferramentas computacionais, como evidenciado em abordagens contemporâneas que empregam linguagens de programação como Python para análise de dados e simulação de recursos eólicos.

A expressão analítica da distribuição de Weibull é dada por,

$$f(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left[\left(\frac{v}{c}\right)^{k}\right]}, \tag{1.6}$$

Em que,

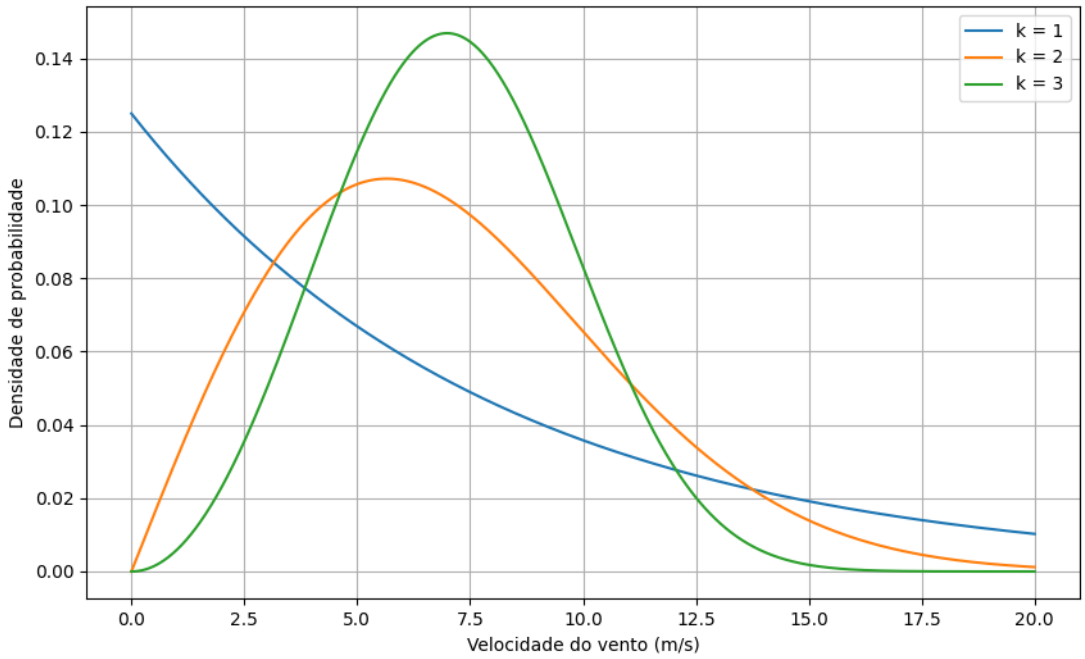
- v: velocidade média do vento (m/s);
- c: parâmetro de escala (m/s);
- k: fator de forma (adimensional);

Fixando o parâmetro de escala e variando o fator de forma, pode-se verificar o comportamento da curva de Weibull, conforme ilustrado na Figura 3. Para k=1, nota-se que a curva apresenta uma característica semelhante de uma exponencial em decaimento; para k=2, a curva apresenta um comportamento de distribuição distorcida com cauda tendendo para direita, com ventos moderados em que as velocidades mais frequentes estão em uma faixa de valores entre 4,5 e 7,5 m/s; já para k=3, a função apresenta um comportamento semelhante a uma função normal, com uma densidade de probabilidade mais intensa na faixa de velocidade de 5,0 a 11 m/s.

Com relação a velocidade média anual, o seu cálculo pode ser realizado utilizando a distribuição de Weibull da seguinte maneira,

$$v_{anual} = \int_0^\infty v f(v)\, dv, \tag{1.7}$$

Figura 3 – Função de Weibull analisando a variação do valor do fator de forma.



Fonte: próprio autor.

Na prática, têm-se distribuições discretas da velocidade média do vento em classes de 1 m/s. Dessa maneira, pode-se aproximar a velocidade média anual para,

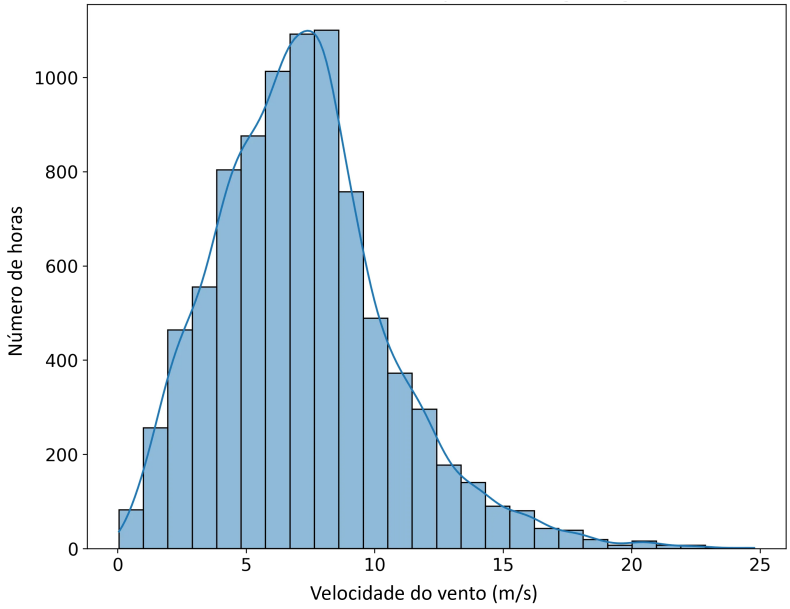
$$v_{anual} = \sum_{v=0}^{v} {}_{max} v f(v), \tag{1.8}$$

Na Figura 4 é ilustrado o comportamento da distribuição do vento em uma cidade da Holanda no ano de 2020.

É interessante, ainda, citar de maneira breve, o conceito do Limite de Betz. Esse conceito prova matematicamente a máxima eficiência teórica que uma turbina eólica pode atingir na conversão de energia do vento em energia mecânica. A prova matemática não será descrita por completo nesse estudo, porém pode-se provar que o valor máximo que uma turbina eólica pode retirar da potência P disponível do vento é de 59,3%. Os fundamentos para realizar a prova matemática são descritos a seguir.

Assume-se que a velocidade média do vento através do rotor é a média das velocidades μ_1, antes da turbina, e da velocidade μ_2, após a passagem pela turbina. A massa de ar através da seção plana do rotor é descrita por,

Figura 4 – Curva de distribuição do vento na cidade de Wageningen, Holanda, em 2020.



Fonte: Adaptado de (SHRESTHA, 2022).

$$m_r = \rho A \frac{\mu_1 + \mu_2}{2}, \tag{1.9}$$

Dessa forma, a potência do rotor é dada por,

$$P_r = \frac{1}{2}\rho A \left(\frac{\mu_1 + \mu_2}{2}\right)\left({\mu_1}^2 - {\mu_2}^2\right), \tag{1.10}$$

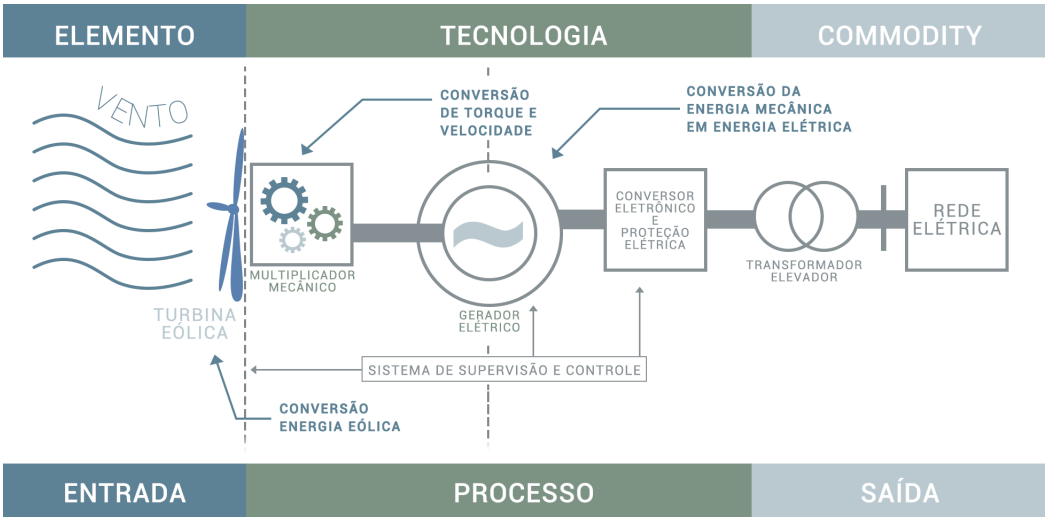
Após manipulações matemáticas diversas, têm-se que,

$$\frac{P_r}{P_{disp}} = \frac{16}{27} \approx 59,3\%. \tag{1.11}$$

1.3.1. **As turbinas eólicas**

Como descrito anteriormente, as turbinas eólicas são equipamentos desenvolvidos para converter a energia do vento em energia elétrica, utilizando um gerador elétrico acoplado. Este processo ocorre inicialmente nas pás da turbina, que capturam a energia cinética do vento e a converte em energia mecânica. A rotação resultante é então transferida ao gerador elétrico, onde é transformada em eletricidade e em seguida transmitida para rede elétrica. Na Figura 5 é ilustrado o princípio dessa conversão.

Figura 5 – Princípio da conversão de energia cinética em energia elétrica.



Fonte: (ENERGIA, 2024).

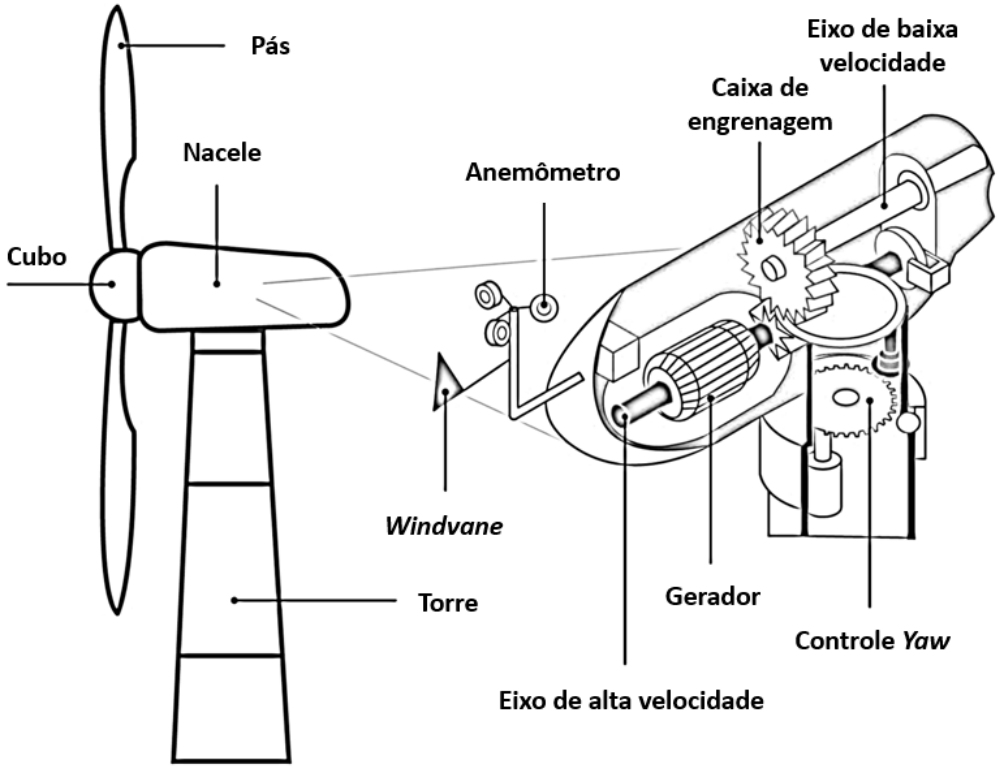
O design das turbinas eólicas é otimizado para maximizar a eficiência na conversão de energia, levando em consideração fatores como a velocidade do vento, a área de varredura das pás e a potência gerada. As turbinas podem ser do tipo eixo horizontal ou eixo vertical, sendo a primeira mais comum nos projetos eólicos de grande porte. A turbina de eixo horizontal tem diversas vantagens como o acesso a ventos com velocidades mais altas devido à altura da torre e uma elevada eficiência devido ao ângulo de ataque do vento nas pás (perpendicular).

O aerogerador propriamente dito, é a combinação de diversos componentes como a turbina eólica, a torre, nacele, unidades meteorológicas (anemômetro e *windvane*), gerador elétrico, sistemas de controle de *yaw* e *pitch*, cubo, pás e caixa de engrenagem. Na Figura 6, é ilustrado de maneira geral, os principais elementos que compõe um sistema de geração eólico.

As torres são responsáveis por prover a estrutura civil necessária para sustentar as demais partes do aerogerador, normalmente são construídas em concreto, aço ou um híbrido dos dois materiais. Esse tipo de torre recebe o nome de tubular cônica, devido ao seu formato geométrico. No passado, nas fases iniciais do emprego comercial de aerogeradores, as chamadas torres treliçadas foram empregadas em diversos projetos; entretanto, devido ao avanço da tecnologia estrutural, elas caíram em desuso nos projetos comerciais mais recentes.

A nacele é uma estrutura montada na parte superior da torre para abrigar o gerador e a caixa de engrenagem. Há aerogeradores que não dispõem de caixas de engrenagem

Figura 6 – Principais componentes de um aerogerador *on-shore*.



Fonte: Adaptado de (TYME, 2023).

devido a tecnologia do gerador e isso irá influenciar diretamente no tamanho da nacele. O sistema que controla a direção na qual a turbina está posicionada encontra-se dentro da nacele e é chamado de controle *yaw*. O material que constitui o freio da turbina é fabricado em aço e é utilizado em paradas de emergência ou de manutenção.

Na parte superior de nacele encontram-se as unidades meteorológicas, responsáveis por monitorar dados essenciais para o sistema de controle da turbina. Dentre os equipamentos, tem-se o anemômetro, responsável pela medição da velocidade do vento e o *windvane* que verifica a direção do vento.

Nos aerogeradores que necessitam elevar a velocidade de rotação do eixo do gerador, localizado na parte interna da nacele, há a chamada caixa de engrenagem, que é responsável por realizar a transição da velocidade baixa da turbina para a elevada velocidade requisitada pelo gerador. Em tais situações, o sistema de engrenagem requer adequada lubrificação e refrigeração para assegurar um desempenho eficiente. A lubrificação dessa unidade deve ser conduzida utilizando o óleo especificado pelo fabricante, seguindo rigorosamente as diretrizes delineadas no manual de manutenção. O uso de óleo divergente das recomendações ou a execução de procedimentos de manutenção inadequa-

dos podem resultar na redução da vida útil das engrenagens, aumento de temperaturas indesejadas e, em cenários mais adversos, danos substanciais à caixa de engrenagem.

A ocorrência de danos graves à caixa de engrenagem em sistemas eólicos pode ter impactos significativos para o empreendimento e resultar em prejuízos financeiros consideráveis. A caixa de engrenagem desempenha um papel central na transmissão eficiente de energia do rotor para o gerador, e danos a esse componente essencial podem levar a uma redução drástica na eficiência do sistema. O reparo ou substituição de uma caixa de engrenagem danificada geralmente envolve custos elevados, incluindo despesas com peças sobressalentes, mão de obra especializada e tempo de inatividade da turbina para a realização das operações de manutenção.

Além disso, a interrupção da geração de energia eólica durante o período de reparo pode resultar em perdas financeiras decorrentes da falta de produção de eletricidade, afetando diretamente os retornos financeiros do investimento na instalação. Portanto, é crucial a implementação de práticas de manutenção preditivas com a finalidade de minimizar os riscos do empreendimento.

Assim como a caixa de engrenagem, o gerador elétrico alocado na nacele também tem um elevado custo financeiro, afinal, é a máquina responsável pela produção da energia elétrica. As máquinas elétricas, fundamentais para a fabricação dos geradores empregados nos sistemas eólicos, podem ser classificadas como assíncronas ou síncronas. A máquina assíncrona, também conhecida como máquina de indução, manifesta suas características quando é acionada por uma força motriz externa, como a turbina, com velocidade superior à sua velocidade síncrona (n_{sinc}). Nesse momento, ocorre a inversão do sentido do conjugado, resultando em seu funcionamento como um gerador.

Na máquina de indução, o movimento relativo do rotor em relação ao campo magnético do estator que produz uma tensão induzida em uma barra do rotor (CHAPMAN, 2013). Dessa maneira, cada condutor do rotor percebe esse campo variável no tempo, permitindo a circulação de correntes, surgidas por indução quando os condutores estão em curto-circuito (PINTO, 2013). Há dois tipos de máquinas de indução, as com estrutura de rotor bobinado e as gaiolas de esquilo.

Nos sistemas eólicos, os geradores de indução com rotor bobinado e as máquinas síncronas são os tipos de geradores frequentemente utilizados no mercado. Dentre as tecnologias empregadas pelos fabricantes, destaca-se a do gerador de indução duplamente

alimentado (DFIG), caracterizado pela alimentação simultânea no estator e no enrolamento do rotor da máquina.

As máquinas elétricas e a caixa de engrenagem constituem o cerne deste estudo, que visa a avaliação preditiva da temperatura nos componentes cruciais desses equipamentos. A realização de análises preditivas em grandes montantes de dados não é algo trivial, dessa maneira, a aprendizagem de máquina emerge como uma ferramenta crucial na identificação de falhas nos aerogeradores. A complexidade dos aerogeradores e a quantidade volumosa de dados gerados durante sua operação tornam desafiador o monitoramento manual eficaz para detecção precoce de falhas.

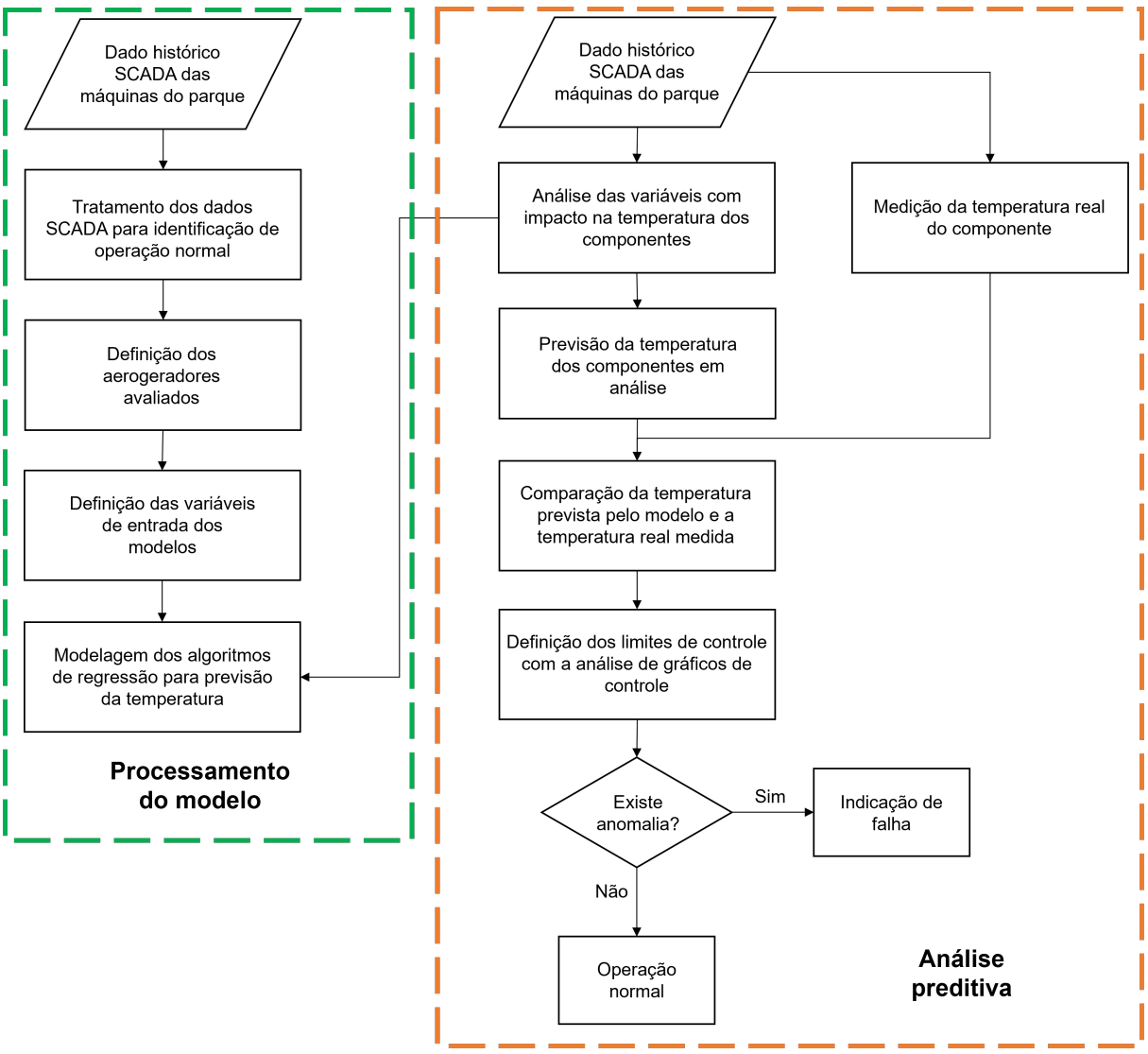
1.4. **Objetivos da dissertação de mestrado**

No contexto de uma gestão da manutenção mais eficiente, emerge a necessidade da utilização de técnicas preditivas de baixo custo de implementação visando a antecipação da temperatura em componentes críticos do aerogerador, como a caixa de engrenagem e o gerador. Os algoritmos de aprendizagem foram escolhidos como técnicas preditivas devido à sua comprovada eficácia em resolver problemas complexos, aliada à sua implementação em linguagem de programação de alto nível. A metodologia proposta visa empregar métodos de regressão múltipla linear, aumento de gradiente extremo (XGBoost) e rede neural recorrente como modelos preditivos, com auxílio dos gráficos de controle para avaliação de desvios de temperatura. Para isso são utilizados os dados do sistema de supervisão de três aerogeradores de um parque eólico localizado no nordeste do Brasil, os dados são tratados previamente visando a representação em regime de operação normal, sendo, em seguida, divididos em conjuntos de treino, teste e validação para a parametrização dos modelos de aprendizado.

Também é apresentada uma análise para escolha dos aerogeradores do estudo com base na produção global das máquinas, isso é feito com o cálculo corrigido da velocidade para densidade do fabricante e com a técnica de interpolação para encontrar o valor de potência associado a cada velocidade corrigida. A escolha de máquinas com comportamentos diferentes para a aplicação dos modelos construídos visa avaliar se há variações relevantes no desempenho dos modelos de aprendizagem dado os comportamentos de temperaturas distintos.

Realizada a aplicação dos modelos nos dados, obtém-se os resultados das métricas de desempenho e os modelos aplicados as três máquinas são comparadas entre si para cada componente de análise, no caso, o rolamento da caixa de engrenagem e o rolamento *drive-end* do gerador. A ideia é analisar o comportamento dos modelos perante sua aplicação nos componentes citados a fim de identificar qual modelo tem melhor desempenho sob a ótica do coeficiente de determinação e das métricas de erros. Este processo permite uma avaliação abrangente da eficácia dos modelos propostos na antecipação de anomalias de temperatura, contribuindo para aprimorar a gestão da manutenção em parques eólicos. De maneira geral, o diagrama de blocos representado pela Figura 7 ilustra as etapas para construção do algoritmo de detecção de faltas.

Figura 7 – Representação em diagrama de blocos do algoritmo desenvolvido para detecção de falhas em componentes do aerogerador.



Fonte: próprio autor.

Considerando as informações apresentadas até o momento, é possível resumir a seguir os objetivos e as proposições desta Dissertação de Mestrado:

1. Propor uma metodologia para identificação de anomalias em aerogeradores visando a antecipação da temperatura em componentes críticos da máquina, com baixo custo para desenvolvimento e aplicação;

2. Modelar e aplicar os algoritmos de aprendizagem nos dados tratados das máquinas, comparando os valores das temperaturas previstas com as temperaturas reais medidas a fim de identificar anomalias com o auxílio da teoria dos gráficos de controle;

3. Avaliar o resultado dos modelos de aprendizagem aplicados aos componentes em estudo, sob a ótica das métricas de desempenho, e, ainda, realizar a comparação entre os modelos aplicados nas três máquinas para cada componente analisado.

1.5. **Organização textual**

O trabalho desenvolvido está organizado nos seguintes capítulos:

- **Capítulo 2**: São apresentados os conceitos e aplicações gerais de Aprendizagem de Máquina (AM), juntamente com as suas categorias e características principais. Em seguida, são discutidos os principais métodos de regressão, destacando os fundamentos teóricos dos modelos aplicados; para finalizar, são explorados os fundamentos do modelo de *ensemble* e da rede neural recorrente.

- **Capítulo 3**: É apresentada uma nova proposta de metodologia para identificação de falhas em aerogeradores. São discutidos os aspectos que permeiam a manutenção preditiva, bem como a coleta, organização e limpeza dos dados SCADA visando uma execução correta dos modelos de previsão. Ainda, é realizado um breve estudo de produção de energia dos aerogeradores para definição das máquinas que serão avaliadas. Em seguida, são definidos os parâmetros de entrada dos modelos baseando-se em análises de correlação das variáveis e revisões da literatura.

- **Capítulo 4**: São aplicados os três modelos de aprendizagem para antecipação da temperatura no rolamento da caixa de engrenagem e no rolamento *drive-end* do gerador. É realizada uma avaliação dos resultados obtidos, avaliando o desempenho de

cada modelo sob a ótica do coeficiente de determinação e das métricas de avaliação. Em seguida, são definidos os limites de controle com base na teoria de gráficos de controle para identificação de anomalias nas máquinas.

- **Capítulo 5**: São apresentadas as conclusões gerais da dissertação, bem como sugestões de trabalhos futuros.

2 NOÇÕES INICIAIS DE APRENDIZAGEM DE MÁQUINA

Neste capítulo são apresentados conceitos e aplicações de Aprendizagem de Máquina, incluindo uma introdução à evolução da coleta de dados em massa, juntamente com as categorias de aprendizagem e suas características. Em continuidade, são descritos os principais métodos de regressão, abordando o conceito de correlação para avaliação das variáveis em um conjunto de dados. Adicionalmente, são discutidos os aspectos teóricos dos modelos de regressão utilizados no estudo, iniciando com uma breve apresentação do modelo de regressão múltipla linear, abrangendo tópicos como análise residual e o método dos mínimos quadrados. Posteriormente, são explorados os fundamentos do modelo XGBoost, destacando técnicas como *bagging*, *boosting* e *stacking*, e, por fim, são abordados os conceitos do método de memória curta de longo prazo[2] (LSTM).

2.1. Aspectos gerais

Durante a era digital, a coleta de dados tornou-se presente no dia a dia da população, ainda que de forma imperceptível. Nos grandes centros, usuários estão conectados por meio de celulares, computadores e tablets. Dados são coletados em proporções colossais de forma ininterrupta, as grandes empresas de tecnologias têm acesso aos gostos culinários, as escolhas, as doenças, as atividades de lazer, tudo que se compra ou que se pensa em comprar, organizado por faixa etária, sexo, localização, entre tantas outras características.

A utilização dos modelos de aprendizagem de máquina tem uma forte dependência da base de dados que está sendo empregada. Foi justamente devido à alta capacidade de coleta de dados em tempo real associada ao grande poder de computação e armazenamento de dados, que tornou a AM relevante nos últimos anos (NG, 2017).

A aprendizagem de máquina (em inglês, *machine learning*), é um ramo da inteligência artificial (IA) que utiliza algoritmos para imitar o comportamento do ser humano, possuindo regras programadas. E, através da grande quantidade de dados presente e disponível, é considerada uma tecnologia disruptiva devido a sua capacidade de solucionar problemas complexos.

[2] *Long Short Term Memory*

A primeira pessoa a utilizar o termo *machine learning* foi Arthur Samuel em 1959. De acordo com o cientista, a aprendizagem de máquina é a área de estudo que fornece aos computadores a capacidade de aprender sem serem explicitamente programados para tal. Nos primórdios do desenvolvimento de AM, havia cientistas com habilidades para tratar e trabalhar com os dados de entrada dos modelos e existiam máquinas com capacidade computacional para realizar testes, entretanto o conhecimento da área e as técnicas existentes ainda eram rudimentares (SAMUEL, 1959).

Atualmente, com o desenvolvimento das técnicas e linguagens de programação as técnicas de AM são utilizadas para identificar objetos em imagens, transcrever discursos em texto, descobrir publicações ou produtos de interesse do usuário e selecionar resultados relevantes em uma pesquisa (LECUN; BENGIO; HINTON, 2015).

As três principais categorias dos algoritmos de AM são (JORDAN; MITCHELL, 2015):

- Supervisionada: Nesta abordagem, o modelo é treinado em um conjunto de dados rotulados (com respostas corretas conhecidas) e, em seguida, é capaz de fazer previsões precisas em novos dados não vistos, com base no que aprendeu durante o treinamento.

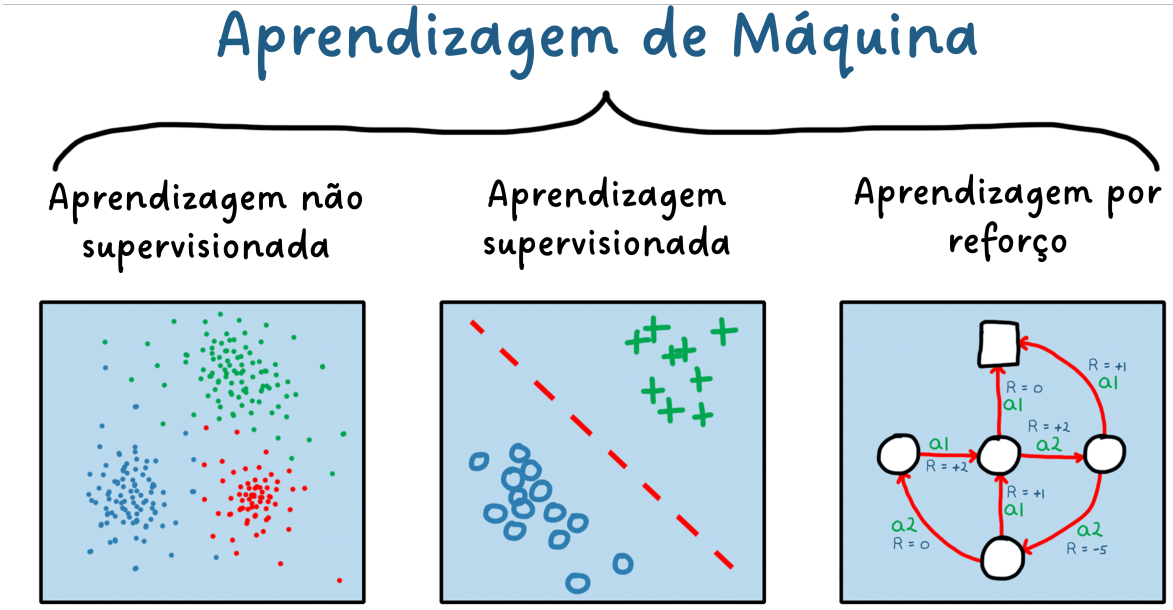
- Não supervisionada: Neste tipo, o modelo é treinado em um conjunto de dados não rotulados, sem respostas corretas conhecidas, e busca identificar padrões e estruturas nos dados por conta própria. É frequentemente usada em tarefas de análise exploratória de dados e pré-processamento.

- Reforçada: Neste caso, modelo é treinado a tomar decisões em um ambiente em constante mudança, com o objetivo de maximizar uma recompensa ao longo do tempo. O modelo é recompensado quando toma a decisão correta e punido quando toma a decisão errada (regularização), e ajusta sua estratégia de decisão com base nos resultados obtidos. É frequentemente usada em aplicações de jogos e robótica.

As divisões das categorias estão ilustradas na Figura 8, indicando uma aplicação de não supervisão com *clustering*, supervisão com uma fronteira bem definida e aplicação de aprendizagem por reforço.

A aprendizagem supervisionada é uma técnica que envolve treinar um modelo a aprtir de dados já rotulados. Nesse tipo de abordagem, o modelo é exposto a exemplos de

Figura 8 – Divisão das três categorias de aprendizagem de máquina: não supervisionada, supervisionada e reforçada.



Fonte: Adaptado de (REINFORCEMENT..., 2021).

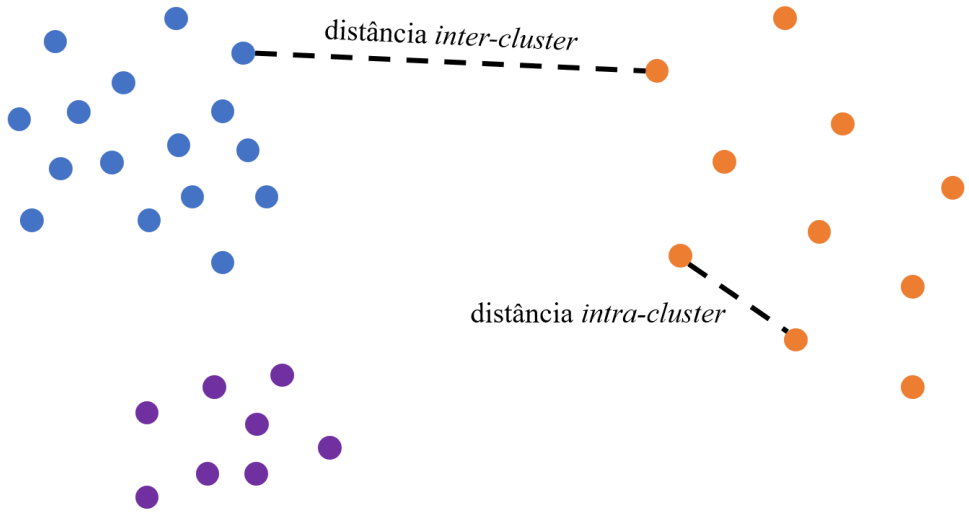
entrada e saída, para que possa aprender a mapear as entradas para as saídas desejadas. O objetivo é fazer com que o modelo generalize para novos exemplos não vistos anteriormente e seja capaz de produzir uma saída correta para essas entradas. Existem diversos modelos que são baseados em supervisão, como regressão linear, não linear e logística, Redes Neurais Artificiais (RNA), K-vizinho mais próximo (KNN), Máquina Vetor Suporte (SVM), árvores de decisão, entre outras aplicações.

A aprendizagem não supervisionada é bastante utilizada nos modelos de classificação, pois não há conhecimento a respeito das características ou rótulos dos dados que estão sendo trabalhados. A ideia da classificação é agrupar instâncias, de forma que os protótipos que ficarem no mesmo grupo (*cluster*) sejam semelhantes entre si e sejam distintos dos protótipos de outros *clusters*.

Para definir o pertencimento ou não de uma amostra de dados a um grupo e como os grupos podem ser divididos em um ou mais, as distâncias *inter-cluster* e a distância *intra-cluster* são utilizadas, conforme indicado na Figura 9.

Quanto menor a distância *intra-cluster*, melhor, pois os dados que pertencem a um mesmo grupo estão mais condensados; e quanto maior for a distância *inter-cluster*, melhor, pois os grupos estão separados o suficiente para não causar problemas de pertencimento.

Figura 9 – Distâncias utilizadas para definir pertencimento dos grupos em algoritmos de classificação.



Fonte: próprio autor.

Neste trabalho, o foco estará no emprego de métodos de regressão, que tem como base algoritmos supervisionados, aplicados na identificação de falhas em aerogeradores. Dessa forma, são conhecidos os rótulos dos dados de entrada e de saída desejados, sendo as previsões realizadas através da série histórica de dados disponíveis.

2.2. **Métodos de regressão**

A regressão é uma técnica estatística que visa identificar a relação entre variáveis dependente e independentes (SOTO, 2013). Com base nessas relações, busca-se realizar previsões ou inferências a respeito da variável dependente.

O uso da técnica de regressão pode ser usado tanto para entender a relação entre os rótulos (correlação), como para prever valores futuros com base nas variáveis de entrada. Dentre as principais técnicas de regressão, incluem-se a regressão linear simples, árvores de decisão, redes neurais, métodos de *boosting*, regressão múltipla linear, método de memória curta de longo prazo, entre outros.

Diversas são as aplicações de AM com foco em regressão, por exemplo, (SARSWATULA; PUGH; PRABHU, 2022) estudou a previsão do consumo de energia através de dados do Departamento de Energia dos Centros Industriais (IACs), utilizando técnicas como regressão múltipla linear, árvore de decisões randômicas e aumento de gradiente

extremo com a finalidade de realizar um levantamento geral do consumo de energia das indústrias do banco de dados e identificar novas oportunidades na eficiência energética.

Em (NG; LIM, 2022) foi proposta a aplicação de técnicas de AM para previsão da temperatura do óleo da *gearbox* de aerogeradores, nesse trabalho foram utilizadas três técnicas, k-vizinho mais próximo, XGBoost e redes neurais para realizar a previsão. As técnicas são comparadas entre si através de métricas e, ainda, é realizada a análise de componente principal e correlação de Pearson para avaliar quais variáveis tem maior importância para o modelo de previsão.

A vantagem de utilizar métodos de regressão está na acessibilidade de interpretação dos resultados, na alta capacidade de resolução de problemas complexos, na flexibilidade devido a quantidade de técnicas disponíveis e na utilização de dados históricos (rotulados) tornando capaz estimar saídas futuras.

Para definir quais serão as variáveis de entrada para um modelo de regressão é interessante a execução de análises para identificar as correlações que os rótulos possuem entre si, dessa forma as variáveis de entradas podem ser melhor alocadas para produzir um resultado de saída mais assertivo. Dentre as análises, incluem-se a covariância, correlação de Pearson, tabelas de contingência, estatística *Chi-square*, entre outras. Ainda, a escolha da técnica depende do tipo de variável que está sendo avaliada, podendo ser categórica ou quantitativa.

No caso dos aerogeradores, os dados obtidos através do sistema SCADA são em sua maioria quantitativos, como temperatura, pressão, velocidade, rotação do rotor, entre outros. Dessa forma, as análises de correlação podem ser direcionadas para associações quantitativas, através de gráficos de dispersões, análises de covariâncias e correlação simples.

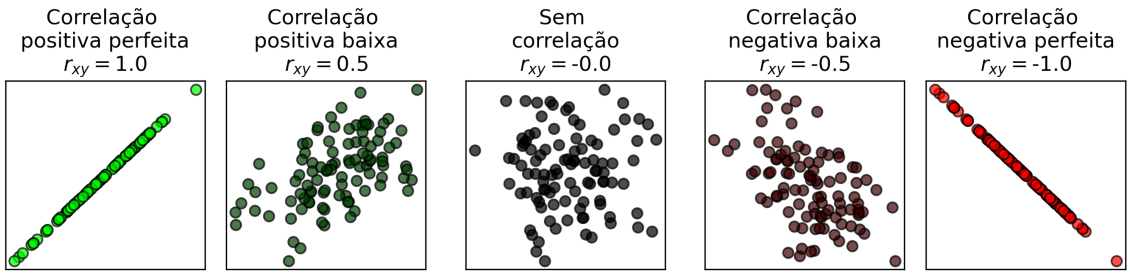
2.2.1. **Correlação de Pearson**

Uma das métricas de correlação mais utilizadas é o coeficiente de Pearson, que mede a associação linear entre duas variáveis (KIRCH, 2008). Este coeficiente pode ser definido por

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2}}, \tag{2.1}$$

onde n é o número total de amostras e $\bar{x}$ e $\bar{y}$ são as médias aritméticas de ambas as variáveis. Nessa métrica, o resultado do coeficiente pode variar de -1 (correlação inversamente proporcional) a 1 (correlação diretamente proporcional). Na Figura 10 são ilustrados diferentes tipos de correlação entre variáveis.

Figura 10 – Resultados da correlação de Pearson para duas variáveis.



Fonte: Adaptado de (HELLIWELL et al., 2020).

O uso da avaliação da correlação entre as variáveis para o emprego de métodos de aprendizagem voltados para regressão é essencial para identificar quais as melhores variáveis podem ser alocadas como de entrada buscando obter resultados com menores erros na saída do modelo. Ainda, o uso da correlação pode ser utilizado para preencher valores não existentes (ou não medidos) na amostra ou no conjunto de dados, permitindo que os dados possam ser melhor aproveitados mantendo um nível de assertividade com relação ao conjunto original.

2.3. **Regressão múltipla linear**

A regressão múltipla linear é um modelo que visa avaliar a relação de duas ou mais variáveis independentes e uma variável dependente, ajustando uma equação linear aos dados do conjunto. Cada variável de entrada x está relacionada com a variável de saída y, conforme indicado pela equação linear

$$y = b + a_1 x_1 + a_2 x_2 + ... + a_n x_n + \epsilon, \tag{2.2}$$

em que b é o coeficiente linear, a_1, a_2, ..., a_n são os coeficientes angulares e ϵ é o erro. O fundamento da regressão múltipla linear é encontrar o coeficiente linear e os coeficientes angulares, buscando minimizar o erro. O valor ótimo dos parâmetros a e b são encontrados quando é definido um valor aceitável para o erro.

Para avaliar o quão preciso é o modelo de regressão linear múltipla pode-se utilizar a técnica de Análise Residual. A diferença entre o valor real de y e a previsão do valor $\hat{y}$ é o valor residual e. O coeficiente de determinação do modelo preditivo R^2 é obtido por

$$R^2 = 1 - \frac{u}{v}, \tag{2.3}$$

em que

$$u = \sum_{i=1}^{N} (y - \hat{y})^2, \tag{2.4}$$

e

$$v = \sum_{i=1}^{N} (y - \overline{y})^2. \tag{2.5}$$

Nestas equações, $\overline{y}$ é a média do vetor y, N é quantidade de pontos de saída, o parâmetro u é o valor residual da soma dos quadrados e o parâmetro v é a soma total dos quadrados. De forma ideal, o interesse é que essa diferença seja mínima e que o valor previsto seja igual ao valor real, nesse caso ideal, $\mathbf{R^2}$ será igual a 1.

Duas técnicas são comumente utilizadas para implementação do modelo de regressão múltipla linear: o método dos mínimos quadrados e o gradiente descendente. Sendo o primeiro método mais preciso, porém custoso computacionalmente e o segundo mais flexível em termos de custo computacional, obtendo uma precisão aceitável para certas aplicações. Nesse estudo, será adotado o método dos mínimos quadrados utilizando a biblioteca do *scikit-learn* em Python para realizar as simulações.

2.3.1. **Método dos Mínimos Quadrados**

O método dos mínimos quadrados é uma técnica de otimização matemática que procura encontrar o melhor ajuste para um conjunto de dados buscando minimizar a soma dos quadrados das diferenças entre os valores estimados e os dados observados.

O Método dos Mínimos Quadrados (MMQ) tem sua origem no estudo dos valores máximos e mínimos de funções reais, mais precisamente, na determinação dos pontos mínimos de uma função que representa o desvio estimado na busca pelo ajuste (ALMEIDA, 2015). Na definição de (TORRES, 2018), dado um conjunto de n pontos $\{(x_i, y_i) \in \mathbb{R}^2\}_{i=1}^{n}$ e uma família de funções $\Gamma = \{f : \mathbb{R} \to \mathbb{R}; y = f(x)\}$, o problema de ajuste de curvas consiste em encontrar uma função da família Γ que melhor se ajusta aos pontos dados e

que não necessariamente os interpola. Dessa forma, busca-se encontrar uma função $f \in \Gamma$ tal que $f(x)$ resolva o seguinte problema de minimização,

$$r(b,a) = \sum_{i=1}^{n}(y_i - b - ax_i)^2 \to min, \tag{2.6}$$

A função objetiva é chamada de resíduo e consiste na soma dos quadrados das diferenças entre a ordenada y_i e o valor da função procurada $f(x_i)$. Dado que a função de resíduo é uma função quadrática, então seu mínimo irá ocorrer quando o gradiente for zero, dessa maneira, obtém-se as seguintes derivadas parciais.

$$\frac{\partial r}{\partial b} = -2\sum_{i=1}^{n}(y_i - b - ax_i) = 0, \tag{2.7}$$

$$\frac{\partial r}{\partial a} = -2\sum_{i=1}^{n}(y_i - b - ax_i)x_i = 0, \tag{2.8}$$

Tendo em vista que,

$$\sum_{i=1}^{n}(y_i - b - ax_i) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} b - \sum_{i=1}^{n} ax_i, \tag{2.9}$$

$$\sum_{i=1}^{n}(y_i - b - ax_i) = \sum_{i=1}^{n} y_i - nb - \left(\sum_{i=1}^{n} x_i\right) a, \tag{2.10}$$

E que,

$$\sum_{i=1}^{n}(y_i - b - ax_i)x_i = \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right) b - \left(\sum_{i=1}^{n} x_i^2\right) a, \tag{2.11}$$

Dessa forma, observa-se o seguinte sistema de equações a ser resolvido,

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i. \end{bmatrix} \tag{2.12}$$

Este sistema linear de duas equações e duas incógnitas admite uma única solução quando o determinante da matriz de coeficientes for não nulo (TORRES, 2018), ou seja,

$$n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \neq 0, \tag{2.13}$$

A solução do sistema pode ser obtida de forma simples ao utilizar a regra de Cramer para soluções de sistemas lineares, como segue,

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \tag{2.14}$$

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}. \tag{2.15}$$

Para fins de entendimento, observe uma variação mais simples da Equação 2.2. Note que para cada x, há um valor de y associado em que o objetivo é descobrir os valores de a e b de tal forma que a soma do erro quadrático é minimizada.

Por exemplo, considerando $i = 3$ é obtida a seguinte equação em formato matricial:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = a \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}, \tag{2.16}$$

que pode ser modificada para:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = m \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}, \tag{2.17}$$

ou, de forma mais compacta, tem-se

$$Y = X\beta + \epsilon. \tag{2.18}$$

Para isolar beta, que contém as informações de intersecção e coeficiente angular, é necessário realizar operações matriciais que envolvem transposição e inversão de matriz,

resultando em

$$\beta = (XX^T)^{-1}X^TY. \tag{2.19}$$

A razão por trás do Método dos Mínimos Quadrados ter um custo computacional mais elevado é devido a essa inversão matricial. E, ainda há a possibilidade da matriz XX^T não ter inversa, que pode acontecer quando houver uma relação linear exata entre as colunas de X.

2.4. **Fundamentos do aprendizado conjunto**

O aprendizado conjunto[3] é o processo de construção de um modelo complexo de aprendizado de máquina, combinando vários estimadores de base como blocos de construção. O principal objetivo desse método é obter um modelo de aprendizado de máquina resultante que seja mais eficiente e mais robusto do que seus componentes.

Nos métodos de aprendizado conjunto, muitas vezes os modelos básicos escolhidos tendem a ter um desempenho insatisfatório por conta própria e são normalmente referidos como aprendizes fracos. A preferência pela utilização de aprendizes fracos se fundamenta na eficiência computacional durante a implementação do modelo. A combinação de aprendizes fortes não necessariamente torna o modelo resultante mais eficiente.

Três métodos comuns de agrupamento serão discutidos a seguir: *bagging*, *boosting* e *stacking*. Esses métodos são a base da formação do algoritmo XGBoost.

2.4.1. ***Bagging***

O *Boostraping Aggregating*, conhecido como *bagging*, é um método de conjunto que combina os conceitos de *bootstrapping* e agregação. Os métodos de *bagging* usam aprendizes fracos como modelos básicos que são complexos e tendem a sofrer de alta variância (BREIMAN, 1996). Sua fraqueza como modelo se deve ao fato de serem construídos com apenas um subconjunto das variáveis de entrada disponíveis e em um subconjunto dos dados de treinamento devido ao *bootstrap*.

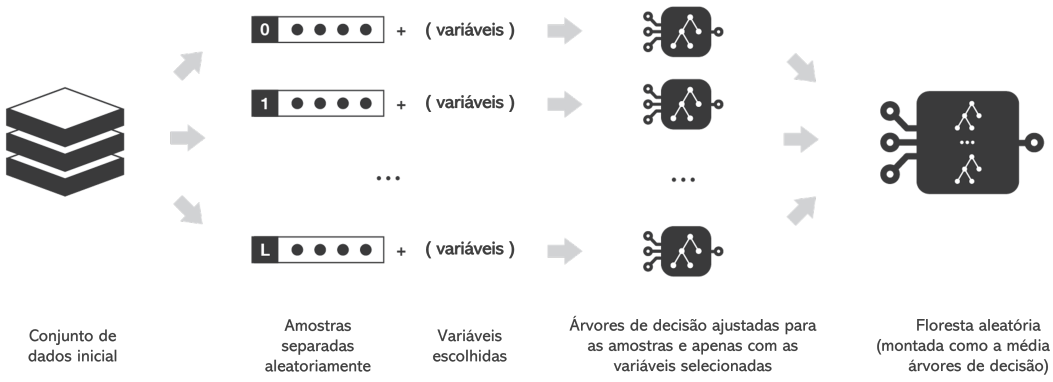
O *bootstrap* é uma técnica estatística de re-amostragem com reposição utilizada para estimar a distribuição de uma estatística amostral. Essa técnica permite obter uma

[3] *Ensemble Method*

estimativa da variabilidade e incerteza dos parâmetros estimados a partir de uma amostra de dados, sem assumir uma distribuição específica dos dados populacionais.

Por sua vez, o *bagging* utiliza o conceito do *bootstrap* para realizar as amostragens aleatórias em paralelo. Cada um dos modelos base é treinado independentemente dos outros e utilizando apenas um subconjunto das variáveis originais. A ideia conceitual do funcionamento do método de *bagging* pode ser observado na Figura 11.

Figura 11 – Diagrama explicativo do método de *bagging*.



Fonte: Adaptado de (ROCCA, 2019).

Como exemplo, pode-se pensar em um conjunto de treinamento com 1000 linhas para produzir o método de floresta aleatória. A floresta é composta por diversas árvores de decisão, em que toda vez que uma árvore de decisão é feita, ela é criada usando um subconjunto diferente dos pontos do conjunto de treinamento. As árvores podem ser criadas escolhendo 100 dessas linhas aleatoriamente para construí-las, dessa forma, cada árvore é diferente, mas todas as árvores ainda serão criadas a partir de uma parte dos dados de treinamento.

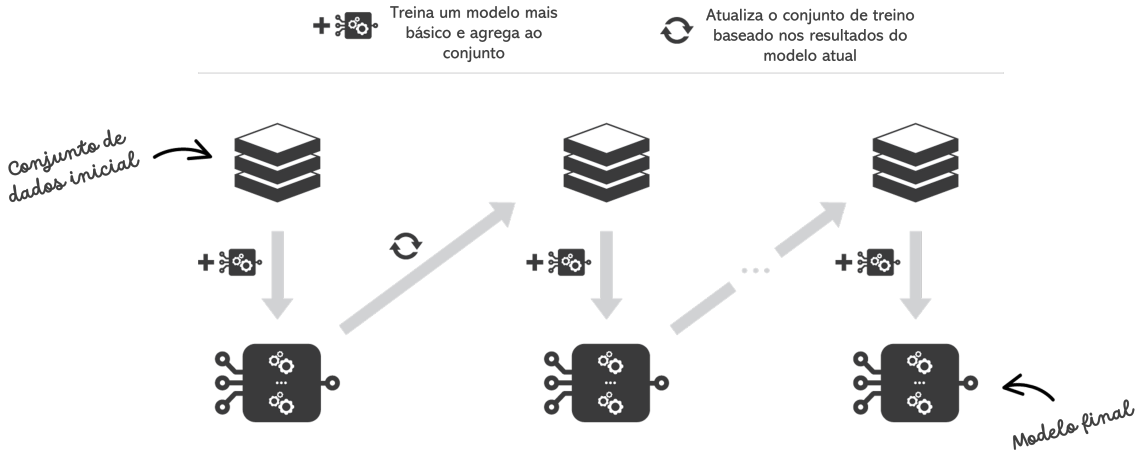
2.4.2. *Boosting*

Boosting é uma técnica de aprendizado conjunto em que os aprendizes fracos possuem baixa complexidade e tendem a sofrer alto viés, como árvores de decisão de apenas um nível. As árvores de decisão de um nível só podem tomar uma decisão com base em uma variável por vez, resultando em ajustes excessivos aos dados.

Essa técnica é sequencial, o que significa que cada um dos modelos básicos tem como base o modelo predecessor e cada modelo subsequente visa melhorar o desempenho

do modelo final, tentando corrigir os erros do estágio anterior. O funcionamento desse método pode ser observado na Figura 12.

Figura 12 – Diagrama explicativo do método de *boosting*.



Fonte: Adaptado de (ROCCA, 2019).

No diagrama, pode-se observar que, inicialmente um conjunto de dados será submetido ao primeiro modelo. Uma vez finalizado o treino, são determinados os erros residuais, que é a diferença entre os valores reais e previstos para cada uma das instâncias de dados de treinamento. Os erros serão maiores para as instâncias em que o modelo não fez boas previsões e serão menores nas situações em que o modelo se ajusta melhor aos dados.

$$h_1 = y_{atual} - y_{1(previsto)}, \tag{2.20}$$

O próximo modelo irá aprender com os erros do primeiro a fim de aprimorar o resultado de saída e minimizando o erro residual. Em um processo iterativo, pode-se definir o erro residual máximo aceitável para que o algoritmo seja finalizado e obtido o modelo final.

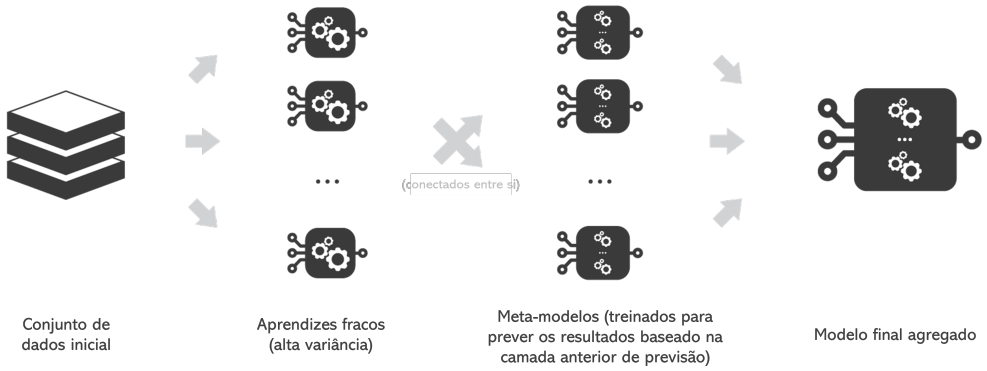
2.4.3. ***Stacking***

O método de *stacking*, ou empilhamento, é uma técnica de combinação flexível em que um modelo final é treinado para aprender como combinar melhor um conjunto de modelos de baixa complexidade para fazer melhores previsões. Os modelos combinados

não necessariamente precisam ser do mesmo tipo, e, diferente dos métodos apresentados, o *stacking* pode ser formado a partir de aprendizes fortes.

O método de *stacking* é útil quando deseja-se utilizar os benefícios de diferentes modelos de regressão ou classificação, combinando as suas melhores qualidades para formar uma ferramenta final mais robusta. Como por exemplo, utilizar um modelo de regressão logística e uma árvore de decisão para formar um modelo mais robusto, em que cada um terá um voto final por meio das suas previsões. A Figura 13, exibe um diagrama ilustrando o funcionamento do método.

Figura 13 – Diagrama explicativo do método de *stacking*.



Fonte: Adaptado de (ROCCA, 2019).

No diagrama da Figura 13, pode-se observar que, inicialmente um conjunto de dados é submetido aos modelos, que podem ser distintos entre si. Em seguida, têm-se os modelos treinados para prever as saídas desejadas. Por fim, os modelos são agregados formando um sistema final mais robusto, que terá diversas previsões que podem ser ponderadas de forma igualitária para formar uma previsão definitiva.

Não há uma regra sobre quais modelos podem ser agregados, a escolha irá depender da finalidade para qual a ferramenta está sendo desenvolvida. Pode-se utilizar, por exemplo, diversos modelos de árvores de decisão, porém com hiperparâmetros distintos, variando a profundidade ou o *alpha*.

2.5. **Aumento extremo de gradiente**

O método de Aumento extremo de gradiente (XGBoost) é uma técnica de aprendizagem de máquina desenvolvida pelos pesquisadores Tianqui Chen e Carlos Guestrin. Neste trabalho, os cientistas de dados apresentaram um novo algoritmo com reconheci-

mento de dispersão para dados esparsos e um esboço de quantil ponderado para aprendizado aproximado de árvores (CHEN; GUESTRIN, 2016). O método surgiu como um aperfeiçoamento do algoritmo de aumento de gradiente[4], que é uma técnica utilizada em aplicações de regressão e classificação.

O método XGBoost tem como base os algoritmos de aprendizado conjunto, que são algoritmos voltados para construção de modelos de aprendizagem complexos combinando e integrando diversos estimadores de base. O principal objetivo para empregar métodos de aprendizado em conjunto é obter um modelo resultante que seja mais eficiente e mais robusto que seus componentes, conforme discutido anteriormente.

O modelo matemático do XGBoost é voltado para minimização da função objetiva $\mathcal{L}^{(t)}$. A função objetiva depende dos valores alvos nos dados de treinamento (y_i) e é função de diferentes árvores de decisão em cada iteração sucessiva (CHATTERJEE; DETHLEFS, 2020). O resumo da base matemática do algoritmo é descrito a seguir.

Inicialmente, a equação da função objetiva do método é dada por:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t), \tag{2.21}$$

Em que o primeiro termo do somatório se trata da função de perda, sendo y_i o valor real visto no conjunto de dados, seguido do $\Omega(f_t)$ que é o termo de regularização.

Dado que uma função pode ser aproximada por uma série infinita de polinômios de Taylor, pode-se definir a seguinte função suave.

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}f''(a)h^2 + ... + \frac{1}{n!}f^n(a)h^n, \tag{2.22}$$

Pode-se obter uma versão simplificada da função objetiva realizando um truncamento de segunda ordem no polinômio de Taylor.

$$f(a+h) \approx f(a) + f'(a)h + \frac{1}{2}f''(a)h^2, \tag{2.23}$$

E, definindo,

[4] *Gradient Boosting*

$$a = \hat{y}_i^{(t-1)}, \tag{2.24}$$

$$h = f_t(x_i), \tag{2.25}$$

Desse modo,

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)}) + \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}} f_t(x_i) + \frac{1}{2} \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)^2}} f_t(x_i)^2 \right] + \Omega(f_t), \tag{2.26}$$

As derivadas na função truncada em segunda ordem são estatísticas do gradiente da função perda denominadas de gradiente (g_i) e hessiano (h_i). Substituindo as derivadas pelos termos citados,

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)} + g_i f_t(x_i) + \frac{1}{2} h_i (f_t(x_i))^2 \right] + \Omega(f_t), \tag{2.27}$$

Como o objetivo é encontrar o f_t que minimiza essa equação podemos desprezar o termo l, visto que ele é constante na equação. Por conseguinte,

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i (f_t(x_i))^2 \right] + \Omega(f_t). \tag{2.28}$$

A Equação 2.28 pode ser aplicada para qualquer modelo genérico f_t, entretanto a ideia é utilizar o modelo construído com base nas árvores de decisão. Dessa forma, a função objetiva pode ser reescrita considerando o comportamento das árvores. Sabe-se que para cada amostra x_i, teremos uma associação com uma folha j.

$$f_t(x_i) = w_j, \tag{2.29}$$

Dessa maneira, a Equação 2.28 é reformulada como segue,

$$\sum_{i=1}^{n} g_i f_t(x_i) = \sum_{j=1}^{T} w_j \sum_{i \in I_j}^{T} g_i, \tag{2.30}$$

$$\sum_{i=1}^{n} h_i f_t(x_i)^2 = \sum_{j=1}^{T} w_j^2 \sum_{i \in I_j}^{T} h_i, \tag{2.31}$$

Substituindo ambos os termos na função objetiva, dispomos da seguinte expressão matemática,

$$\mathcal{L}^{(t)} \approx \sum_{j=1}^{T} w_j \sum_{i \in I_j}^{T} g_i + \frac{1}{2} \sum_{j=1}^{T} w_j^2 \sum_{i \in I_j}^{T} h_i + \Omega(f_t), \tag{2.32}$$

A equação que descreve o termo de regularização, $\omega(f_t)$, é dada por,

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda ||w||^2, \tag{2.33}$$

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{2.34}$$

Em que,

- γ: hiperparâmetro que controla o impacto da penalização da complexidade das árvores;
- T: são os "T" nós folhas que a árvore possui;
- λ: hiperparâmetro da regularização de Ridge (L2);
- w: é o peso atribuído a cada folha da árvore;

Dessa forma, substituindo 2.34 em 2.32,

$$\mathcal{L}^{(t)} \approx \sum_{j=1}^{T} w_j \sum_{i \in I_j}^{T} g_i + \frac{1}{2} \sum_{j=1}^{T} w_j^2 \sum_{i \in I_j}^{T} h_i + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{2.35}$$

Realizando manipulações matemáticas triviais, é obtida a seguinte equação,

$$\mathcal{L}^{(t)} \approx \sum_{j=1}^{T} \left[\sum_{i \in I_j} (g_i) w_j + \frac{1}{2} \sum_{i \in I_j} (h_i + \lambda) w_j^2 \right] + \gamma T, \tag{2.36}$$

Logo, é possível realizar a otimização da função considerando a observação apenas para uma folha da árvore. A ideia é encontrar o conjunto de pesos w que irá minimizar L.

$$\mathcal{L}^{(t)} \approx \sum_{i \in I_j} (g_i) w_j + \frac{1}{2} \sum_{i \in I_j} (h_i + \lambda) w_j^2 + \gamma T. \tag{2.37}$$

Ao observar a Equação 2.37, nota-se que a função do erro para apenas uma folha possui comportamento quadrático, logo, o mínimo da função é obtido encontrando o ponto de inflexão da curva.

$$\frac{d\mathcal{L}^{(t)}}{dw_j} = \sum_{i \in I_j}(g_i) + 2 \cdot \frac{1}{2}\sum_{i \in I_j}(h_i + \lambda)w_j = 0, \tag{2.38}$$

Logo, o ponto mínimo é obtido isolando w_j,

$$w_j = -\frac{\sum_{i \in I_j}(g_i)}{\sum_{i \in I_j}(h_i + \lambda)}, \tag{2.39}$$

Substituindo a Equação 2.39 na função objetiva 2.36 é obtida a função otimizada para uma folha arbitrária.

$$\mathcal{L}^{(t)}(q) = -\frac{1}{2}\sum_{j=1}^{T}\frac{\left(\sum_{i \in I_j} g_i\right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \tag{2.40}$$

De acordo com (CHATTERJEE; DETHLEFS, 2020), essa equação é denominada função de pontuação q, que avalia a redução de perdas para cada aprendiz, iterando todas as variáveis predominantes nos dados de treinamento e avaliando a redução de perdas em cada nó sucessivo.

A função de pontuação será utilizada para avaliar a divisão de cada nova árvore. Sabe-se que a cada divisão de um nó de uma árvore, duas folhas são geradas, a direita e a esquerda; e, que o ganho da divisão é dado pela soma dos erros das duas divisões subtraído do erro anterior.

$$\mathcal{L}_{div} = L_{antes} - L_{esq} - L_{dir}, \tag{2.41}$$

Avaliando a equação do ganho, tem-se,

$$\mathcal{L}_{div} = \frac{1}{2}\left[\frac{\left(\sum_{i \in I_{esq}} g_i\right)^2}{\sum_{i \in I_{esq}} h_i + \lambda} + \frac{\left(\sum_{i \in I_{dir}} g_i\right)^2}{\sum_{i \in I_{dir}} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i\right)^2}{\sum_{i \in I} h_i + \lambda}\right] - \gamma. \tag{2.42}$$

Por fim, o modelo XGBoost calcula uma probabilidade (t) para cada previsão sucessiva no intervalo de 0 a 1 usando uma função de perda para classificação binária (CHATTERJEE; DETHLEFS, 2020), conforme Equação 2.43.

$$yln(p) + (1 - y)ln(1 - p), \tag{2.43}$$

Em que,

$$p = \frac{1}{(1 + e^{-x})}. \tag{2.44}$$

Dessa forma, fica descrita a fundamentação matemática do modelo de aumento de gradiente extremo. O estudo das equações do modelo é importante, pois alguns dos parâmetros manipulados nas equações anteriores estão presentes nos hiperparâmetros do XGBoost, como o *gamma* (controla a regularização) e o *lambda* (regularização de Ridge). A escolha dos hiperparâmetros são de suma importância na construção do modelo de aprendizagem, pois controlam o funcionamento do algoritmo, custo computacional, penalização, entre outros.

Ainda, é possível realizar uma busca pelos melhores hiperparâmetros com base nos dados de entrada com a utilização de algoritmos de busca, como a otimização Bayesiana. Para esse estudo, foi aplicado a essa otimização para encontrar os parâmetros ótimos que se adéquam ao conjunto de dados proposto.

2.6. **Introdução às redes neurais artificiais**

Uma Rede Neural Artificial é um modelo computacional inspirado no funcionamento do sistema nervoso biológico, composto por camadas de unidades interconectadas chamadas neurônios artificiais (HAYKIN, 2001). As RNAs são amplamente utilizadas em tarefas de aprendizado de máquina, incluindo classificação, regressão, reconhecimento de padrões e processamento de dados (LECUN; BENGIO; HINTON, 2015). Em 1943, (MCCULLOCH; PITTS, 1943) publicou um trabalho que consistia em uma rede neural simples modelada com circuitos elétricos. Em sua tese, foi evidenciado que era possível construir uma rede neural apenas com matemática e algoritmos. A arquitetura de uma rede neural é constituída de três elementos básicos:

- **Pesos sinápticos**: os neurônios de cada camada são conectados por pesos sinápticos, que são os parâmetros de aprendizagem do modelo, eles determinam a força entre as conexões de cada neurônio e podem possuir valores positivos ou negativos;

- **Junção aditiva**: responsável por somar os sinais de entrada, ponderados pelos respectivos pesos do neurônio;
- **Função de ativação**: utilizada para restringir a amplitude de saída de um neurônio, normalmente no intervalo unitário fechado [0, 1] ou [-1, 1];

A representação da arquitetura de um neurônio artificial é ilustrada na Figura 14, em que, x_n são os sinais de entrada do neurônio, w_{kn} são os pesos sinápticos do neurônio k e b_k é o viés, que tem como efeito aumentar ou diminuir a entrada líquida da função de ativação;

Em termos matemáticos, um neurônio k pode ser expresso da seguinte forma:

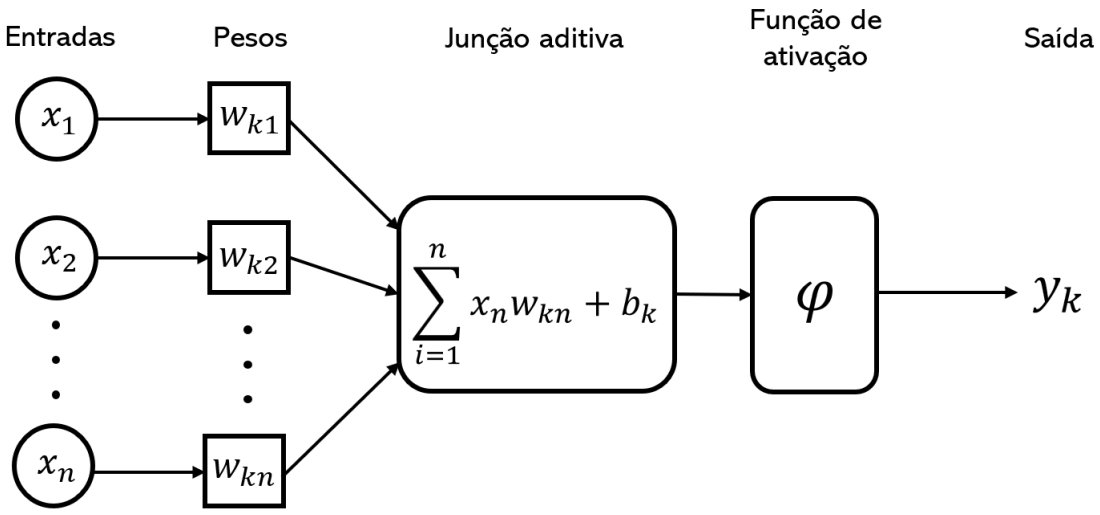
$$u_k = \sum_{i=1}^{n} w_{ki} x_i. \tag{2.45}$$

E a sua saída é descrita por

$$y_k = \varphi(u_k + b_k). \tag{2.46}$$

Ainda, a função de ativação foi escrita da forma φ devido a existência de diversas funções restritivas que podem ser utilizadas nessa etapa. No caso dos parâmetros monitorados dos aerogeradores, devido ao comportamento não-linear dos dados, é indicado

Figura 14 – Representação da estrutura de um neurônio artificial.



Fonte: próprio autor.

que a função de ativação para restringir a saída também seja não-linear, como a Sigmoide ou a Ativação Linear Retificada (ReLU) [5].

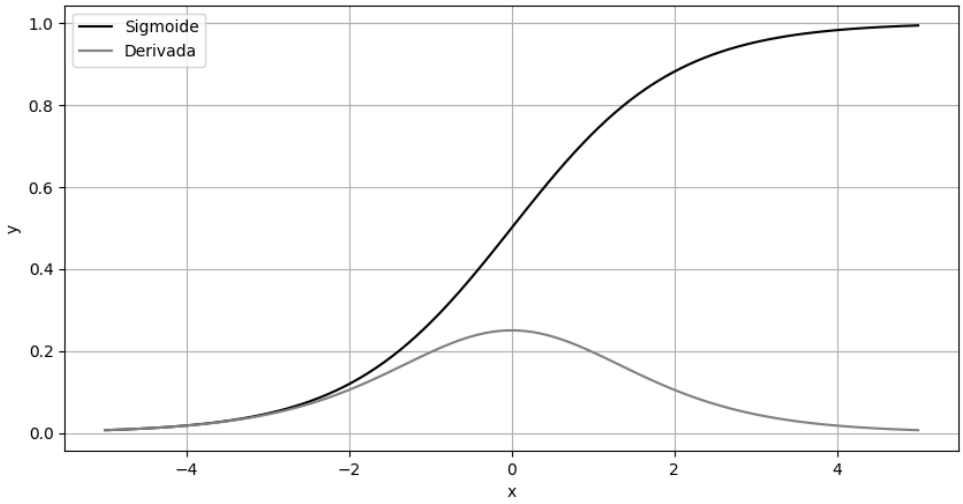
A função de ativação sigmoide dada pela Equação 2.47, é uma das principais funções de ativação utilizadas no estudo das redes neurais, devido a sua facilidade em modelar o comportamento binário do neurônio. Entretanto, a função possui desvantagens, como por exemplo, ao observar a derivada da função, indicada pela Equação 2.48, nota-se que há uma saturação para valores acima de 5 e abaixo de -5. Logo, com as derivadas tendendo a zero, a propagação do gradiente desvanece nessas regiões, causando dificuldades no treinamento (GOODFELLOW; BENGIO; COURVILLE, 2016).

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{2.47}$$

$$f'(x) = f(x)(1 - f(x)). \tag{2.48}$$

Ainda, a derivada da função sigmoide é sempre menor que a unidade, causando um desvanecimento do produto da regra da cadeia na propagação dos gradientes (GOODFELLOW; BENGIO; COURVILLE, 2016). Na Figura 15, é ilustrado o comportamento da função sigmoide e sua derivada.

Figura 15 – Comportamento da função sigmoide no intervalo x=[-5, 5]



Fonte: próprio autor.

A função de ativação ReLU dada pela Equação 2.49 tem sido utilizada nos modelos de redes neurais pela sua simplicidade e eficiência. A função restritiva retorna zero para

[5] *Rectified Linear Unit*

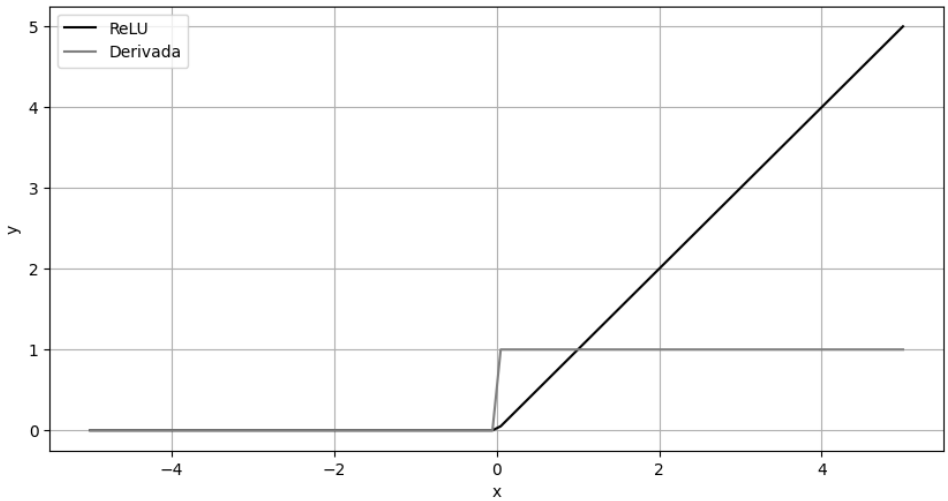
quaisquer valores de entrada negativos e o próprio valor para valores positivos. Uma das desvantagens dessa função é a possibilidade de desvanecimento da rede durante o treinamento, que ocorre quando há entradas negativas ocasionando na produção de zeros. Ainda, para essa região, a derivada também é zero, o que compromete as atualizações dos pesos sinápticos com o gradiente descendente.

$$f(x) = \begin{cases} x, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases} \tag{2.49}$$

$$f'(x) = \begin{cases} 1, & \text{se } x > 0 \\ 0, & \text{se } x < 0 \end{cases} \tag{2.50}$$

O comportamento da função restritiva está ilustrado na Figura 16. Nota-se, que a função ReLU tem valor igual a zero na metade do seu domínio e para quaisquer valores maiores que zero possui derivadas crescentes. Ainda, a derivada para $x = 0$ é indefinida, entretanto pode-se utilizar artifícios para implementá-la como 0 ou 1, dependendo da aplicação.

Figura 16 – Comportamento da função ReLU no intervalo x=[-5, 5]



Fonte: próprio autor.

Nesse estudo, devido a eficiência e simplicidade, a função de ativação ReLU foi adotada no modelo de rede neural recorrente.

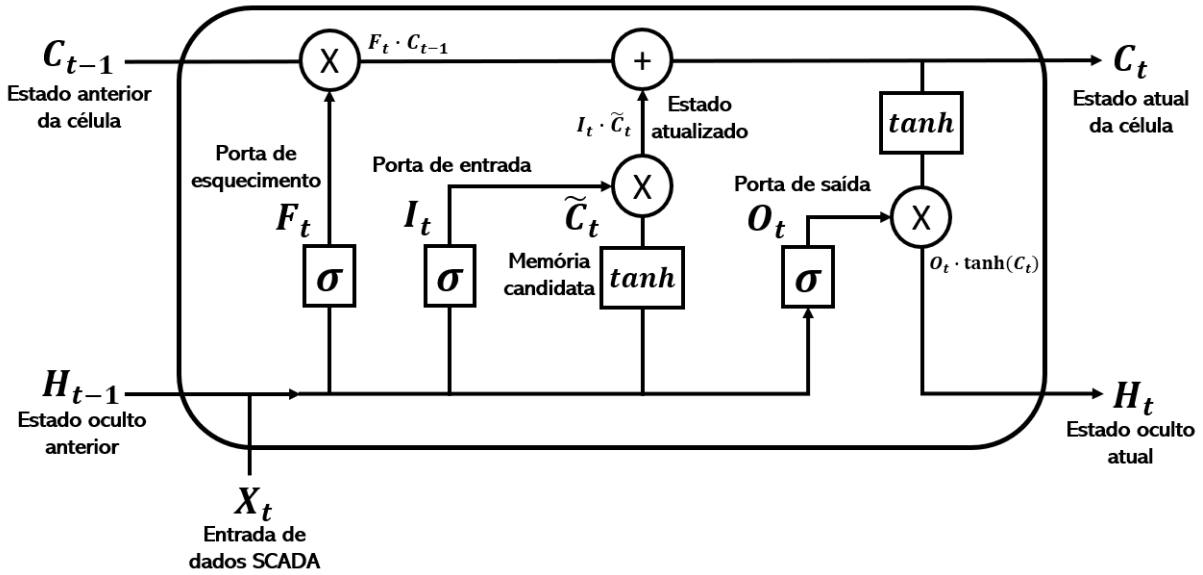
2.6.1. **Memória longa de curto prazo**

O método de aprendizagem de memória longa de curto prazo foi desenvolvido pelos cientistas Sepp Hochreiter e Jürgen Schmidhuber em 1997. No estudo, foi proposta uma nova arquitetura para lidar com problemas frequentes de redes neurais recorrentes, como o desaparecimento do gradiente, que dificulta o treinamento para longas sequências de dados (HOCHREITER; SCHMIDHUBER, 1997).

A principal inovação do modelo LSTM é a capacidade de lembrar informações de longo prazo e lidar de forma eficaz com dependências temporais. Por esta razão, a escolha deste modelo de aprendizado se justifica, uma vez que os dados adquiridos do sistema supervisório dos aerogeradores têm natureza temporal.

A arquitetura do LSTM é indicada na Figura 17. De forma geral, a tarefa desempenhada pelo algoritmo é realizar previsões da saída y a partir de um vetor de entrada X, de maneira análoga aos outros métodos de regressão. Em que o vetor X é um conjunto variáveis provenientes dos dados SCADA, como velocidade do rotor, temperatura ambiente e potência ativa.

Figura 17 – Arquitetura do modelo de aprendizado de memória longa de curto prazo com dados SCADA de entrada.



Fonte: Adaptado de (CHATTERJEE; DETHLEFS, 2020).

O modelo de aprendizagem LSTM tem como base o mapeamento das entradas X para uma saída y a partir da estimação do estado oculto h, indicado na Figura 17, em que o estado oculto é função dos parâmetros de entrada do sistema supervisório. Ao

estado oculto, é aplicado uma função restritiva que pode ser do tipo sigmoide, tangente hiperbólica ou outra que atenda a aplicação do projeto. Ainda, como as redes neurais recorrentes tem como principal aplicação séries temporais para realizar as previsões, o estado oculto é na verdade estimado recursivamente, observando passos de tempos anteriores.

$$H_t = f(X_t, H_{t-1}). \tag{2.51}$$

No diagrama, F_t é a porta de esquecimento e I_t a porta de entrada, ambas são utilizadas para controlar o estado atual da célula. O parâmetro σ é a função sigmoide que é aplicada para restringir quais valores serão utilizados no modelo, normalmente em um intervalo de 0 a 1. Quanto mais próximo da unidade, significa que o modelo irá manter as informações, do contrário, a informação poderá ser descartada.

A porta de esquecimento controla o estado anterior da célula e é utilizada para determinar quais informações devem ser transmitidas para o estado atual da célula. Note que, W_f e b_f é o peso sináptico da ligação e o viés atribuído a porta de esquecimento, respectivamente.

$$F_t = \sigma(W_f \cdot [H_{t-1}, X_t] + b_f). \tag{2.52}$$

De maneira similar, a porta de entrada determina quanto de informação deverá ser adicionada ao estado atual da célula.

$$I_t = \sigma(W_i \cdot [H_{t-1}, X_t] + b_i). \tag{2.53}$$

Ainda, a função tangente hiperbólica é utilizada para modular os valores, $[h_{t-1}, x_t]$, entre -1 e 1 e criar a memória candidata $\tilde{C}$. O estado atualizado é obtido ao multiplicar a porta de entrada e a memória candidata.

$$\tilde{C}_t = tanh(W_c \cdot [H_{t-1}, X_t] + b_c). \tag{2.54}$$

Em seguida, o estado atual da célula pode ser obtido com a soma da porta de esquecimento multiplicada pelo estado anterior da célula e o estado atualizado.

$$C_t = F_t \cdot C_{t-1} + I_t \cdot \tilde{C}_t. \tag{2.55}$$

Por fim, a porta de saída é obtida de maneira análoga a porta de esquecimento e de entrada. Nota-se que a porta de saída controla também o estado atual da célula e o estado oculto, indicando a recorrência da rede neural.

$$O_t = \sigma(W_o \cdot [H_{t-1}, X_t] + b_o). \tag{2.56}$$

E, mais uma vez, a função tangente hiperbólica é utilizada para modular os valores do estado oculto, conforme segue.

$$H_t = O_t \cdot tanh(C_t). \tag{2.57}$$

Dessa forma, fica dissecada a toda a estrutura arquitetural do modelo de aprendizagem de memória longa de curto prazo. É interessante ressaltar, que a escolha do modelo LSTM foi fundamentada na natureza complexa e não linear dos dados de entrada e devido a sua característica temporal, o que resulta em uma aplicação adequada do algoritmo de previsão.

3 METODOLOGIA PROPOSTA PARA IDENTIFICAÇÃO DE FALHAS EM AEROGERADORES

Neste capítulo, são abordados os diversos aspectos que permeiam a manutenção preditiva, a coleta de dados e os modelos de AM aplicados para identificação de falhas em aerogeradores. Inicialmente, é descrito o método para coleta, organização e limpeza dos dados SCADA da máquina para que a aplicação dos modelos de previsão seja bem executada. Em seguida, são definidas as máquinas que serão avaliadas no estudo, com base em uma análise de produção de energia dos aerogeradores. Após, são definidos os parâmetros de entrada dos modelos, baseado em análises de correlação, componentes principais e artigos da literatura.

3.1. Metodologia para identificação de falhas em aerogeradores

O estudo visa apresentar uma metodologia para identificação de falhas em componentes dos aerogeradores, através de dados SCADA de monitoração das máquinas. Serão descritos os passos necessários para construir um modelo preditivo com a finalidade de suportar o setor de manutenção de parques em operação.

Em parques em operação, é comum termos posse dos relatórios de falhas, alertas de manutenção e alerta de alarmes dos aerogeradores (UDO; MUHAMMAD, 2021). Entretanto, os dados obtidos para realizar esse estudo não contém quaisquer relatórios ou indicadores de falhas do parque eólico, dessa forma, o estudo visa a identificação de possíveis sinistros nas máquinas com apenas o uso dos dados SCADA. A escolha das técnicas utilizadas no estudo foi baseada em revisões literárias (CHATTERJEE; DETHLEFS, 2020), (UDO; MUHAMMAD, 2021), (NG; LIM, 2022) e estudos de linguagem de programação. No estudo, serão aplicadas três técnicas de aprendizagem de máquina: regressão múltipla linear, XGBoost e LSTM.

A metodologia consiste na coleta dos dados, organização e limpeza dos dados SCADA, normalização dos dados, aplicação dos modelos preditivos e avaliação dos possíveis pontos de falhas das máquinas com o uso dos conceitos de controle estatístico do processo.

Os controles estatísticos são utilizados para monitorar a variação da temperatura no tempo. A técnica consiste em calcular uma linha central que representa a média da

variável e limites superiores e inferiores, chamados de limites de controle. Os limites de controle são calculados com base no desvio padrão da série histórica avaliada. Nesse estudo, os limites foram definidos como sendo três vezes o desvio padrão da temperatura analisada ($\pm$ 3 σ).

Caso os pontos do gráfico de controle estejam dentro desses limites, o processo é considerado dentro do padrão histórico. Se um ponto cair fora, isso pode indicar uma situação especial de variação que requer investigação para identificar se é o início de uma falha.

3.1.1. **Coleta e tratamento dos dados**

Os dados coletados são de um complexo eólico, localizado no Rio Grande do Norte (RN). As máquinas do parque eólico dispõem de sistema supervisório que é responsável pela coleta de diversos dados como velocidade do vento, potência ativa, temperatura ambiente, temperatura no óleo da caixa de engrenagem, rotação do gerador, tensão na rede, entre outras variáveis. A turbina analisada no estudo é da fabricante Gamesa, modelo G114-2.1 MW, com potência nominal de 2,1 MW, diâmetro do rotor de 114 m, altura da nacele 95,25 m (referência no solo) e pás do aerogerador do tipo Gamesa 56.

Os dados (XAVIER, 2024) são coletados, com o uso de sensores, em intervalos de 10 em 10 minutos e representam uma média do período. Visto que a coleta ocorre com a utilização de um sensor, podem existir problemas com a aquisição de dados. Os erros nos sensores podem surgir devido à falta de calibração ou degradação do equipamento ao longo do tempo (ZIEGLER et al., 2018). Ainda, as usinas eólicas estão sujeitas às reduções de geração devido a manutenções ou à solicitações do Operador Nacional do Sistema Elétrico, resultando em medições que não representam o funcionamento normal da turbina.

Dessa maneira, se faz necessário realizar uma análise exploratória dos dados com a finalidade de identificar possíveis valores anômalos. Há grande importância na qualidade dos dados de entrada para os modelos de previsão, pois em caso de utilização de dados anômalos é possível que o modelo não retorne bons resultados na saída.

Para esse estudo, busca-se conjunto de dados que represente a operação normal da turbina eólica. Dessa maneira, são realizados diversos filtros no banco de dados original para determinar o conjunto de dados que trará essa representatividade. Os critérios

que foram utilizados para realizar a eliminação dos dados SCADA do conjunto base são descritos a seguir.

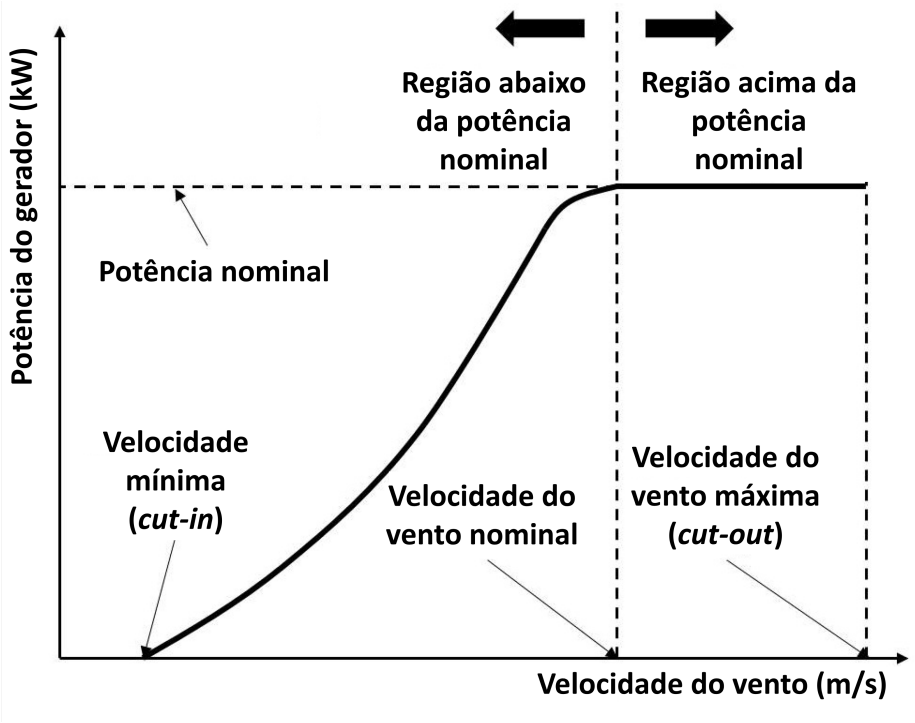
- Conjunto de dados que possua potência ativa zero ou menor, porém com velocidades abaixo da *cut-in* ou acima da *cut-out*.
- Conjunto de dados que possua potência ativa fora da faixa do mínimo e máximo permitido pela máquina, dentro da faixa operativa de ventos. Para esse critério, a faixa operativa de vento foi de: vento mínimo de 5 m/s e vento máximo de 40 m/s.
- Conjunto de dados em que a potência *versus* vento esteja fora de blocos pré-determinados, chamados de *bins*. Para esse critério, o comprimento do bloco foi definido em 40 kW e o valor mínimo para 20 kW, com o centro definido na mediana.
- Conjunto de dados em que o vento não sofreu modificação em um período de medição t-1 e t+1 (dados congelados).

A curva de potência de um aerogerador ilustra a relação entre a potência da turbina e a velocidade do vento. As curvas de potência são geradas para ajudar na estimação do potencial de energia eólica para um determinado local (UDO; MUHAMMAD, 2021). Na Figura 18, está ilustrado um modelo de curva de potência indicando a velocidade mínima do vento para entrar em operação (*cut-in*), velocidade máxima do vento de operação (*cut-out*), velocidade e potência nominal e regiões de operação acima e abaixo da nominal.

Entretanto, quando coletados, os dados SCADA não tem uma distribuição conforme a curva do fabricante, na realidade verifica-se diversos pontos operativos de vento e potência. Durante o funcionamento da máquina, os dados são coletados por sensores que enviam os dados ao sistema supervisório; esses dados por muitas vezes possuem valores que não representam a operação normal da turbina. Na Figura 19, estão indicados os pontos de velocidade *versus* potência de uma das máquinas do parque éolico coletados entre 2018 e 2020.

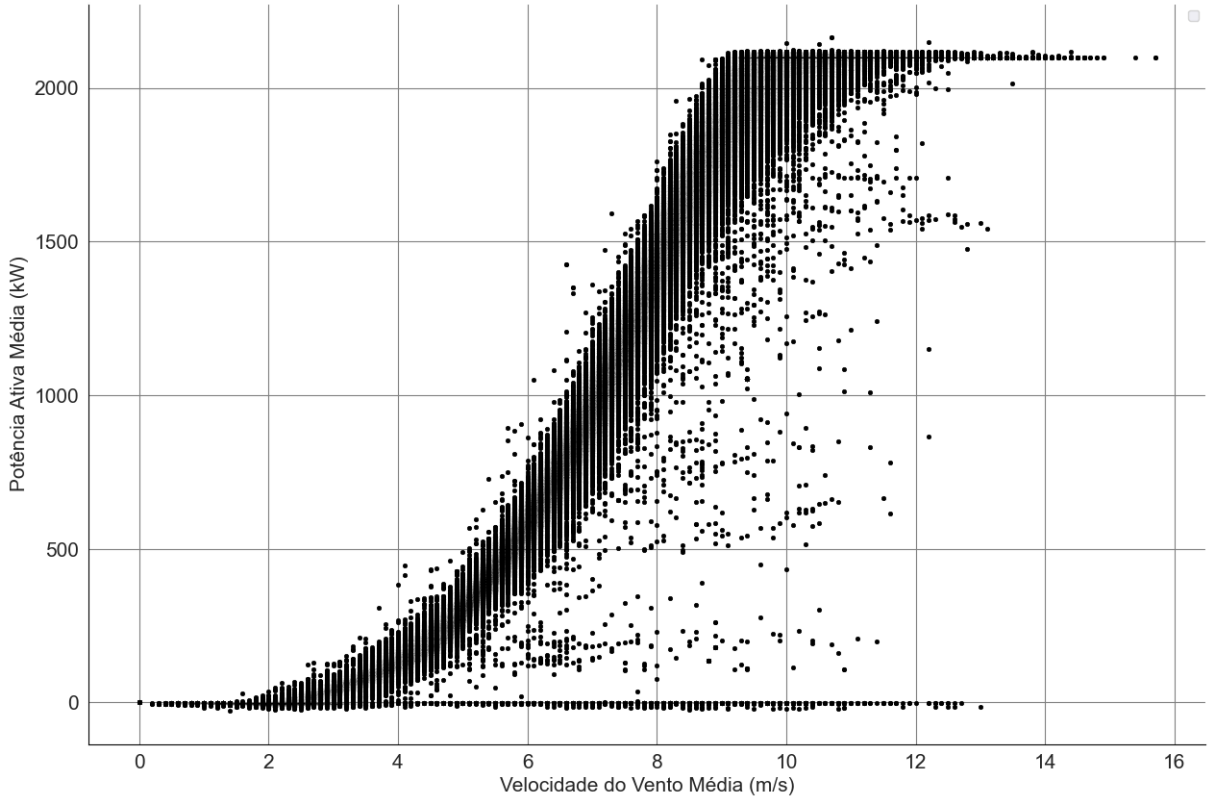
Note que existem diversos pontos que não representam a operação normal da máquina, como os pontos em que há velocidades superiores a 5 m/s e potências próximas a zero. A ideia de realizar o tratamento de dados é eliminar todos os pontos que não representem operação normal, baseado nos critérios previamente citados. Dessa maneira, a curva de potência da máquina terá maior similaridade à curva do fabricante de forma

Figura 18 – Curva de potência típica de uma turbina eólica.



Fonte: Adaptado de (CHEON et al., 2019).

Figura 19 – Curva de potência de um aerogerador do parque éolico.

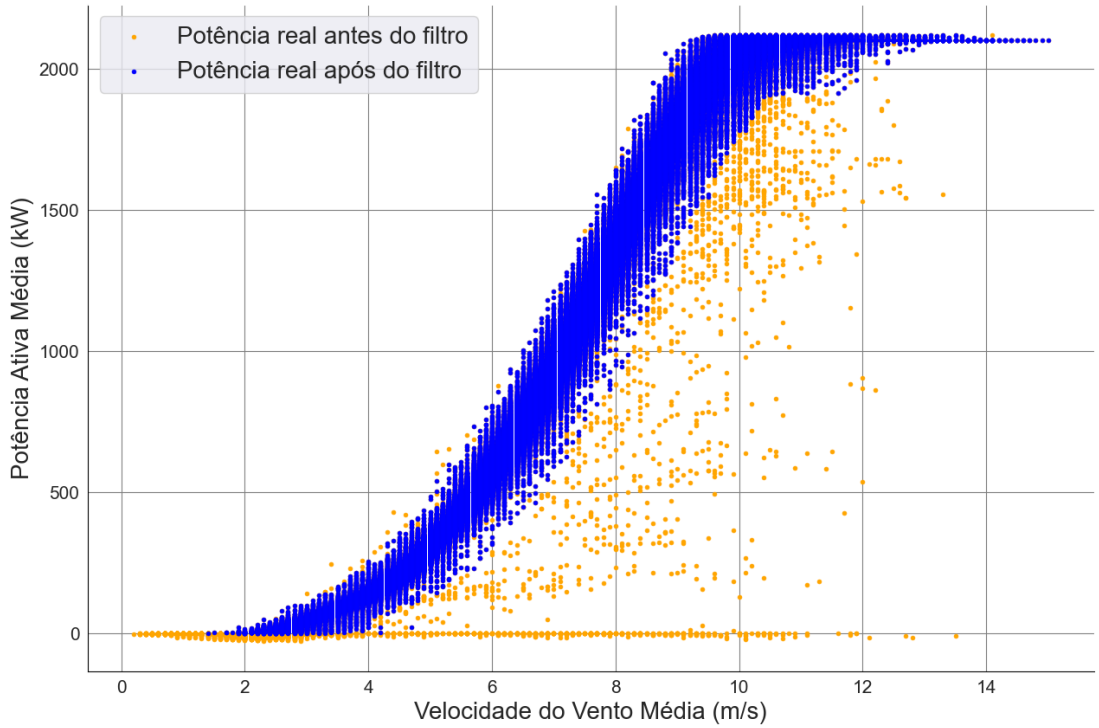


Fonte: próprio autor.

que os dados que entrem nos modelos de aprendizagem sejam dados de referência para operação normal do aerogerador.

O modelo de AM será montado com base nos dados após a filtragem com a premissa de que todos os pontos restantes são pontos operativos sem falhas, o que não necessariamente é verdade, essa incerteza é conhecida e será incorporada ao modelo, pois não há informações a respeito das manutenções nas máquinas para realizar uma melhor classificação nos dados. Na Figura 20 são ilustrados os dados antes e depois da aplicação das premissas de limpeza aos dados SCADA em uma das máquinas avaliadas no estudo.

Figura 20 – Curva de potência do aerogerador filtrada de acordo com as premissas de limpeza dos dados SCADA.



Fonte: próprio autor.

3.1.2. **Definição dos aerogeradores monitorados**

O complexo eólico dispõe de diversos grupos de aerogeradores, no estudo foram obtidos dados para 12 aerogeradores participantes de um único grupo. Para avaliar quais aerogeradores seriam escolhidos para criação dos modelos de aprendizagem, foi realizada uma análise da potência ativa média de cada uma das máquinas com relação à potência esperada, informada pelo fabricante.

Para realizar a análise, foram utilizados os dados da curva de potência do fabricante e os dados da torre anemométrica (AMA) do grupo de aerogeradores. A referência para execução da avaliação da produção das máquinas é a curva de potência da fabricante Gamesa, fornecida para uma densidade do ar de 1,225 kg/m^3. Entretanto, os dados obtidos pela torre AMA *in loco* estão associados a diferentes valores de densidade do ar, sendo assim necessário realizar a correção da velocidade do vento para a densidade do fabricante.

A torre AMA não dispõe do valor da densidade do ar, entretanto são coletados os parâmetros temperatura, pressão e umidade pelo sistema supervisório. A densidade do ar úmido pode ser calculada através das equações revisadas por (PICARD et al., 2008).

$$\rho_{\text{ar úmido}} = \frac{p_d M_d + p_v M_v}{R T_k}, \tag{3.1}$$

Em que,

- p_d: pressão parcial do ar seco (Pa);
- M_d: massa molar do ar seco, 0,028964 kg/mol;
- p_v: pressão de vapor da água (Pa);
- M_v: massa molar do vapor da água, 0,018016 kg/mol;
- R: constante do gás ideal, 8,314 J/(K·mol);
- T_k: temperatura em Kelvin.

A pressão de vapor da água pode ser estimada a partir da úmidade relativa e da equação de Magnus-Tetens (MAGNUS, 1844; ATKINS; ATKINS; PAULA, 2014), em que a temperatura T_c deve estar na unidade de graus Celsius e UR é a umidade relativa.

$$p_v = UR \cdot p_{sat}, \tag{3.2}$$

$$p_{sat} = 6,1078 \cdot 10^{\frac{7,5 T_c}{T_c + 237,3}}. \tag{3.3}$$

Por fim, a pressão parcial do ar seco p_d é a diferença entre a pressão medida no SCADA e a pressão do vapor de água.

$$p_d = p - p_v. \tag{3.4}$$

Com os valores de densidade do ar para cada ponto de medição do SCADA, pode-se realizar a correção da velocidade para densidade do fabricante e utilizar a interpolação para encontrar o valor de potência associado a cada velocidade corrigida. Dessa maneira, é possível avaliar a produção global das máquinas durante o período de 2018 a 2020.

No âmbito deste estudo, foram selecionadas três turbinas eólicas com base em sua performance avaliada na geração de energia, denominadas pelas nomenclaturas CV109, CV111 e CV104. A escolha deliberada de máquinas com características distintas para a construção dos modelos de aprendizado teve como objetivo avaliar se há variações significativas nos comportamentos das temperaturas dos rolamentos do gerador e da caixa de engrenagens, mesmo sendo turbinas da mesma marca, Gamesa.

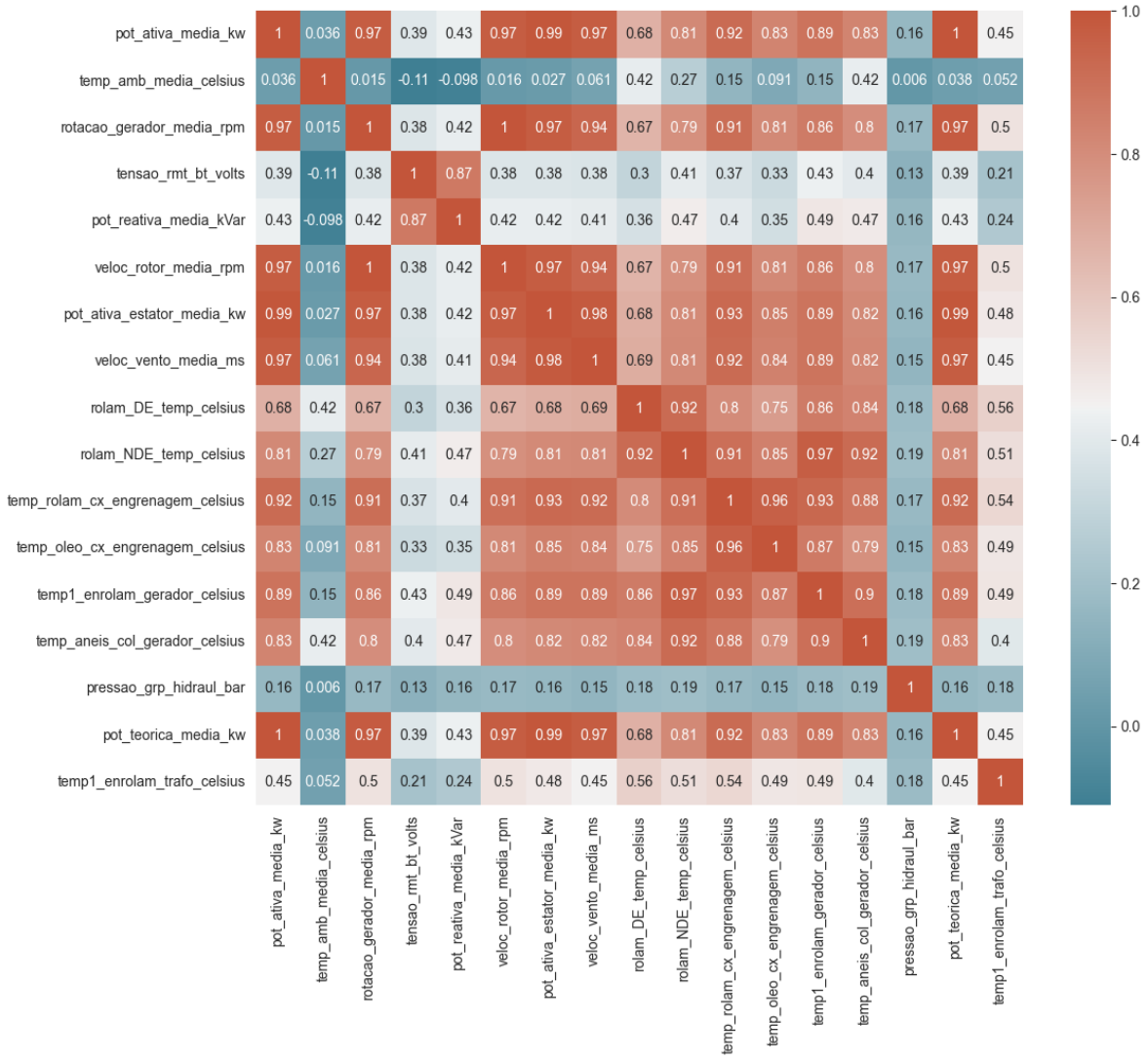
3.1.3. **Seleção das variáveis de entrada**

As variáveis de entrada de um modelo de regressão são fundamentais para um resultado satisfatório na previsão. Para avaliação de falhas em aerogeradores, a literatura aborda diversos parâmetros que tem impacto na temperatura do óleo da caixa de engrenagem e nos enrolamentos do gerador.

(BILAL; ADJALLAH; SAVA, 2019) desenvolveu um modelo para detectar falhas em aerogerador e utilizou como parâmetros de entrada a temperatura da caixa de engrenagem, temperatura da turbina, ângulo de inclinação das pás, velocidade do vento e rotação da turbina. O estudo de (ZAHER et al., 2009), utilizou uma rede neural de multicamadas para identificar anomalias de temperatura na caixa de engrenagem de aerogeradores. Nesse estudo o autor utiliza a temperatura do rolamento do gerador, potência ativa, temperatura da nacelle e o *status* do ventilador de refrigeração do aerogerador como parâmetros de entrada do modelo.

No presente estudo, para definição dos parâmetros de entrada para os modelos foi considerado o histórico observado nas revisões da literatura e, ainda, foi feita uma análise de correlação de Pearson entre as variáveis disponíveis no sistema SCADA, com foco nas correlações ligadas aos parâmetros de temperatura do aerogerador, conforme ilustrado pela Figura 21.

Figura 21 – Análise de correlação de Pearson entre as variáveis presentes no conjunto de dados SCADA da máquina CV109.



Fonte: próprio autor.

Ainda, foi verificada a possível existência de multicolinearidades, o que deve ser evitado devido a não contribuição ao modelo de regressão e aumento do custo computacional nas simulações. Essa avaliação foi realizada considerando tanto a temperatura do rolamento da caixa de engrenagem, como a temperatura nos anéis coletores do gerador no lado *drive-end* (D.E.) da máquina. Isso porque para diferentes saídas desejadas, o comportamento correlacional das variáveis de entrada podem mudar, sendo necessário realizar essa análise de correlação de forma separada.

Com a avaliação da matriz de correlação das variáveis e considerando os estudos da literatura, foram definidos os parâmetros de entrada para o modelo que tem a temperatura do rolamento da caixa de engrenagem e a temperatura do rolamento *drive-end* do gerador

como saída. A Tabela 5, indica os componentes avaliados na máquina, os parâmetros de entrada e de saída adotados.

Tabela 5 – Parâmetros de entrada e saída utilizados para modelar os componentes

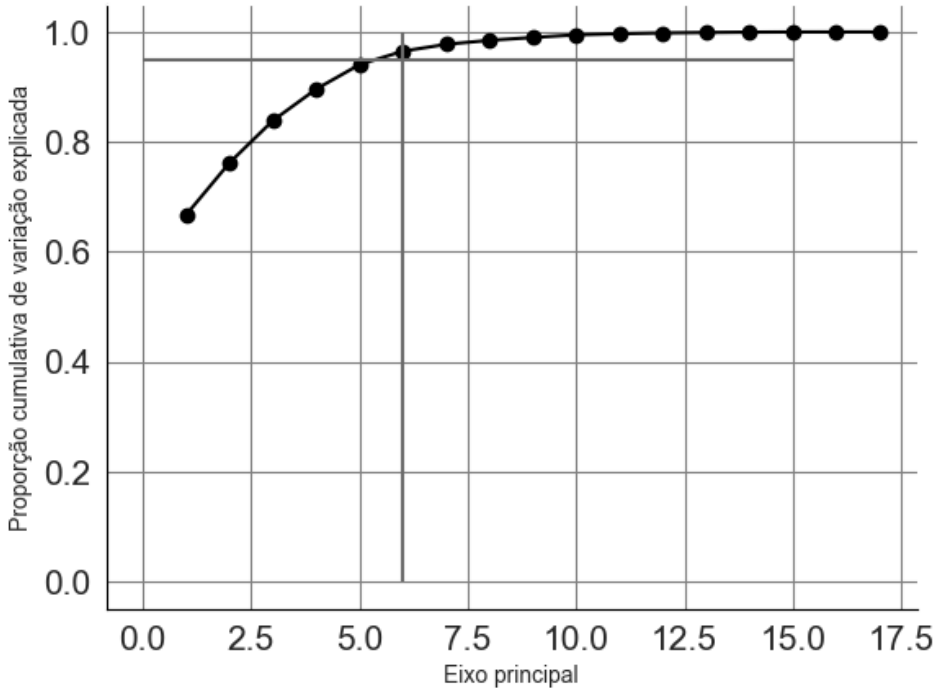
Componente	Entrada	Saída
Caixa de engrenagem	Velocidade do rotor Potência ativa Temperatura ambiente Temperatura do óleo	Temperatura do rolamento
Gerador	Rotação do gerador Potência ativa Temperatura do enrolamento Temperatura dos anéis coletores Temperatura ambiente	Temperatura do rolamento drive-end

Ainda, foi realizada uma Análise de Componente Principal de forma superficial para avaliar qual o número de variáveis que acumulam a maior porcentagem de informação relevante dos dados. Na análise PCA, os elementos com maior variância têm maior importância. A análise inicial foi realizada utilizando a biblioteca *numpy*, e foi observado que cerca de 6 componentes, respondem por cerca de 95% da proporção cumulativa de variação. Na Figura 22, é ilustrada a saída da simulação realizada para verificar qual o número de variáveis que representam maior parte da informação contida no *dataset* original.

Na Figura 23, é ilustrado o comportamento da variância explicada com relação aos componentes principais. Nota-se que a primeira componente representa parte significativa das informações e que as demais são menos informativas quando comparadas com a primeira componente.

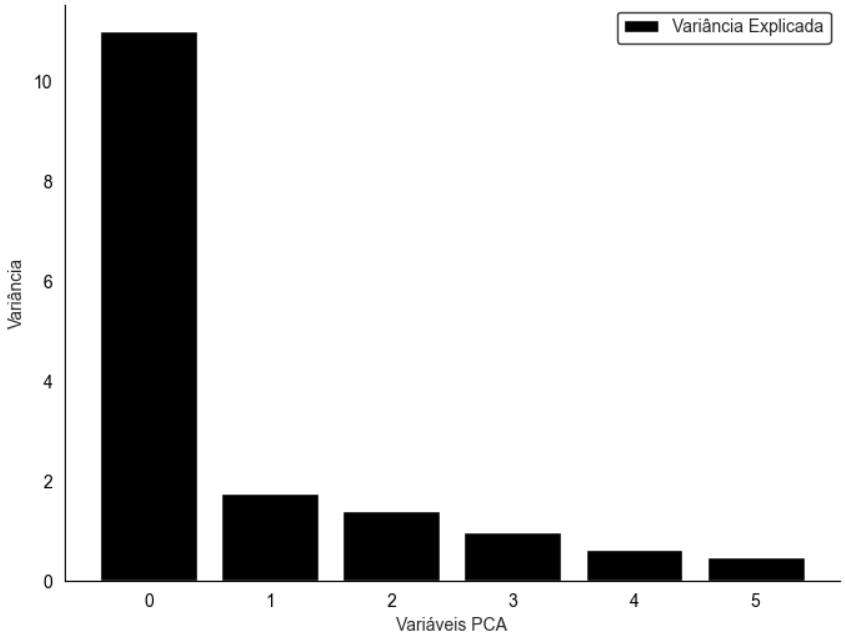
Ainda, a análise PCA pode ser bastante útil quando há um número excessivo de variáveis, gerando um alto custo computacional para realizar a simulação de modelos preditivos, nesses casos pode-se reduzir o número de variáveis a fim de otimizar o modelo. Por fim, a análise de componente principal nesse estudo foi apenas de caráter introdutório,

Figura 22 – Visão inicial do número de componentes principais com a informação cumulativa em percentual.



Fonte: próprio autor.

Figura 23 – Variância explicada em termos das componentes principais.



Fonte: próprio autor.

com a finalidade de visualizar o comportamento das principais componentes e qual seria o número de variáveis que teríamos em caso de uma redução de variáveis.

Dessa maneira, fica claro, que para determinar o funcionamento normal da máquina, é necessário conhecer os parâmetros que serão utilizados na entrada do modelo e qual a saída desejada. A combinação de uma revisão literária, com a análise de correlação e de componentes principais fomentou a base necessária para uma escolha mais assertiva dos dados que serão utilizados como entrada nesse estudo.

3.1.4. **Aplicação dos dados nos modelos de regressão**

De posse das variáveis de entrada, segue-se para construção do modelo para identificar falhas nos aerogeradores. Para estruturar a entrada dos modelos, os dados foram divididos em dados de treino (64%), validação (16%) e teste (20%). A escolha dessa divisão foi com base na estrutura de 80% para treino e 20% para teste. Entretanto, foi destinado 20% dos 80% dos dados de treino para validação, o que gera os 16% especificado., E, com relação a saída desejada, os componentes avaliados são a temperatura do rolamento da caixa de engrenagem e a temperatura do rolamento *drive-end* do gerador.

Considerando que os dados de entradas possuem escalas diferentes, para utilizá-los como entrada dos modelos de regressão é necessário normalizá-los para uma faixa de valores bem definidos. Nesse caso, a normalização de mínimo e máximo é realizada, conforme destaca a Equação 3.5. Essa normalização foi escolhida em detrimento de outras devido a sua compatibilidade com o tipo e escala dos valores das variáveis que compõem o conjunto de dados.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}. \tag{3.5}$$

Na equação, X são os dados de entrada de cada variável do conjunto de dados, X_{max} é o máximo valor da variável e X_{min}, o valor mínimo.

Os modelos irão utilizar as variáveis de entrada normalizadas do conjunto de treinamento para prever a variável de saída desejada, buscando evidenciar que há relações entre elas. Em seguida, são utilizados os dados de validação para verificar que o modelo está operando de maneira satisfatória, para então ser aplicado os dados de testes para verificar a acurácia do modelo, ou seja, o número total de previsões classificadas corretamente.

Na sequência, as saídas previstas pelo modelo são comparadas com as saídas reais da máquina. Ainda, para verificar o desempenho dos modelos são calculados os seguintes

parâmetros: a raiz quadrada da média dos quadrados dos erros (RMSE), média dos valores absolutos dos erros (MAE) e média das porcentagens absolutas dos erros (MAPE). As equações das métricas de avaliação são descritas abaixo.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{3.6}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{3.7}$$

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{3.8}$$

O coeficiente de determinação (R^2) também é calculado, porém sua equação e discussão será vista no tópico de regressão múltipla linear. No python, as equações são acessadas por meio da biblioteca *scikit-learn*.

Ao final, serão utilizados gráficos de controle para visualizar os desvios dos valores previstos comparados com os valores reais. Os desvios que estiverem fora do limite superior ou inferior, são uma possível indicação de que há um comportamento anômalo por parte da máquina, o que pode gerar um alerta de temperatura para a cabine de controle do parque a fim de atuar na máquina de maneira preditiva. A ideia é comparar o comportamento das três máquinas e avaliar quais modelos tem melhor desempenho com relação a temperatura do rolamento da caixa de engrenagem e a temperatura do rolamento *drive-end* do gerador.

4 APLICAÇÃO DOS MODELOS DE APRENDIZAGEM

Neste capítulo procede-se à avaliação dos resultados obtidos mediante a aplicação de três modelos de aprendizado de máquina aos aerogeradores, com o propósito de antecipar a temperatura tanto do rolamento da caixa de engrenagem quanto do rolamento *drive-end* do gerador da máquina. Inicialmente, será empreendida uma análise do desempenho dos modelos em relação aos elementos investigados no âmbito deste estudo. Subsequentemente, serão abordadas as métricas de avaliação, acompanhadas de uma análise dos gráficos de controle, a fim de identificar possíveis indicações de falhas nas máquinas.

4.1. Processamento dos modelos

Os modelos desenvolvidos nesse estudo visam obter uma previsão das temperaturas em componentes críticos de um aerogerador. Para o treinamento dos modelos XGBoost e LSTM, é necessário definir os hiperparâmetros que estarão no núcleo do algoritmo. Determinar quais hiperparâmetros devem ser adotados em um modelo de AM não é uma tarefa fácil e encontrar os parâmetros ótimos de maneira manual é extremamente complexo, devido as diversas combinações possíveis entre os hiperparâmetros.

Dessa maneira, existem algumas técnicas que são utilizadas para encontrar a melhor combinação de hiperparâmetros que reduza os erros do modelo, como o algoritmo de busca randômica, busca em grade, otimização Bayesiana, entre outros. Para esse estudo, foi aplicado a otimização Bayesiana para encontrar os parâmetros ótimos que se adequam ao conjunto de dados proposto, bem como testes com hiperparâmetros utilizados na literatura. O espaço de busca dos hiperparâmetros é indicado na Tabela 6.

Vale ressaltar que, apesar de serem máquinas do mesmo fabricante e mesmo modelo os resultados da otimização Bayesiana pode diferir devido à natureza dos dados, pois cada máquina não apresenta comportamento exatamente igual. Devido a isso, foi necessário realizar uma ligeira adequação dos hiperparâmetros calculados, com a finalidade de utilizar um mesmo conjunto de hiperparâmetros para evitar comparações imprecisas entre os modelos propostos. Os parâmetros adotados para o modelo de previsão das temperaturas rolamentos da caixa de engrenagem e do gerador são indicados na Tabela 7.

De maneira semelhante, para o modelo LSTM, é necessário definir o número de camadas ocultas da rede neural, a função de ativação da saída, a função de perda, o

Tabela 6 – Espaço de busca dos hiperparâmetros para a modelagem do XGBoost.

Hiperparâmetros	Espaço de busca
Profundidade máxima da árvore	Inteiros de 1 a 19
Taxa de aprendizado	Flutuantes distribuídos uniformemente na escala logarítmica entre 0,0001 e 1
Taxa de subamostragem	Flutuantes uniformemente distribuídos entre 0,3 e 1
Taxa de subamostragem de colunas por árvore	Flutuantes uniformemente distribuídos entre 0,3 e 1
Taxa de subamostragem de colunas por nível	Flutuantes uniformemente distribuídos entre 0,3 e 1
Peso mínimo da folha	Inteiros de 1 a 9
Perda mínima por divisão	Flutuantes uniformemente distribuídos entre 0 e 5
Regularização de Ridge	Flutuantes distribuídos uniformemente na escala logarítmica entre 0,00001 e 100
Número de estimadores	30, 40, 50, 60, 70, 80

algoritmo otimizador, o número de lotes e o número de épocas. Para esse modelo foi adotada a configuração com os parâmetros indicados na Tabela 8.

A configuração das camadas ocultas e os números de neurônios adotados estão descritos na Tabela 9. Ainda, foi utilizado o artifício de retrochamada com parada prematura[6]. É uma técnica comum usada durante o treinamento de modelos de aprendizado de máquina para interromper o treinamento prematuramente se determinado critério não estiver melhorando. Essa técnica é particularmente útil para evitar sobreajustes (*overfitting*) e economizar tempo de computação. Para o modelo em questão, foi adotado um critério de parada para o treinamento no caso em que não há melhoria do modelo LSTM por 10 épocas consecutivas.

Após a definição dos hiperparâmetros de cada modelo, foram realizadas as simulações e obtidos os resultados que serão discutidos nos tópicos seguintes. A apresentação dos resultados será dividida para cada elemento em análise e ao fim serão comparados a performance dos modelos aplicados em cada uma das três máquinas (CV109, CV104 e CV111).

[6] *Em inglês, essa técnica é chamada de 'callback' utilizando o método EarlyStopping.*

Tabela 7 – Hiperparâmetros adotados para no modelo XGBoost.

Hiperparâmetros	Valores adotados
Profundidade máxima da árvore	10
Taxa de aprendizado	0,10
Taxa de subamostragem	0,65
Taxa de subamostragem de colunas por árvore	0,70
Taxa de subamostragem de colunas por nível	0,70
Peso mínimo da folha	1,0
Perda mínima por divisão	1,0
Regularização de Ridge	0,1
Número de estimadores	60

Tabela 8 – Hiperparâmetros adotados para o modelo LSTM.

Hiperparâmetros	Valores adotados
Número de camadas ocultas	4
Função de ativação de saída	ReLU
Função de perda	Erro médio quadrático (MSE)
Otimizador	Adam com taxa de aprendizado de 0,001
Épocas	100

Tabela 9 – Configuração das camadas ocultas do modelo LSTM.

Camadas ocultas	Número de neurônios
Camada LSTM 1	50
Camada LSTM 2	50
Camada LSTM 3	25
Camada Densa 4	1

4.2. **Gerador**

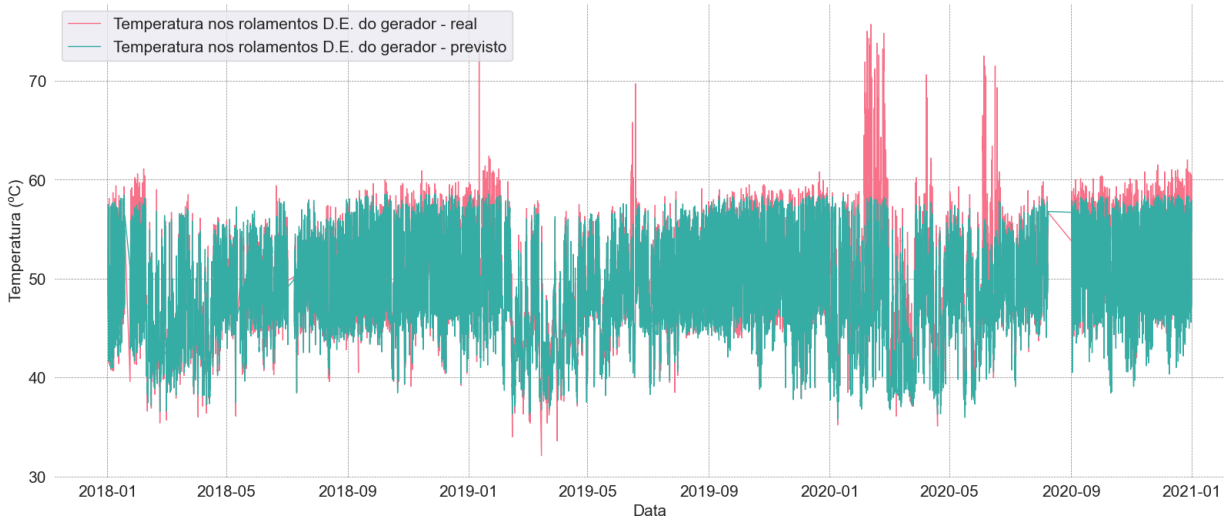
Os modelos de regressão múltipla linear, aumento extremo de gradiente e memória longa de curto prazo foram aplicados para prever a temperatura do rolamento do gerador das três máquinas em estudo. Para a máquina CV109, o modelo LSTM obteve um melhor desempenho com um R-quadrado de 0,81727, RMSE de 2,06194°C, MAE de 1,17476°C e MAPE de 0,02236%.

É interessante ressaltar, que na construção do modelo de rede neural recorrente foi utilizado o erro médio quadrático como função de erro, em que a métrica penaliza erros com valores mais elevados designando um maior peso durante as iterações do modelo.

Ainda, sabe-se que a proximidade do coeficiente de determinação (R-quadrado) à unidade indica uma melhor adequação dos dados ao modelo de aprendizado.

Na Figura 24 é ilustrado o comportamento do modelo aplicado para prever os valores da temperatura do rolamento do gerador para a máquina CV109 em escala de dados SCADA, 10 em 10 minutos. Na Figura 25 está indicado o comportamento dos desvios da temperatura com a avaliação do gráfico de controle. O limite superior e inferior calculado para a escala SCADA é de ±4,99°C.

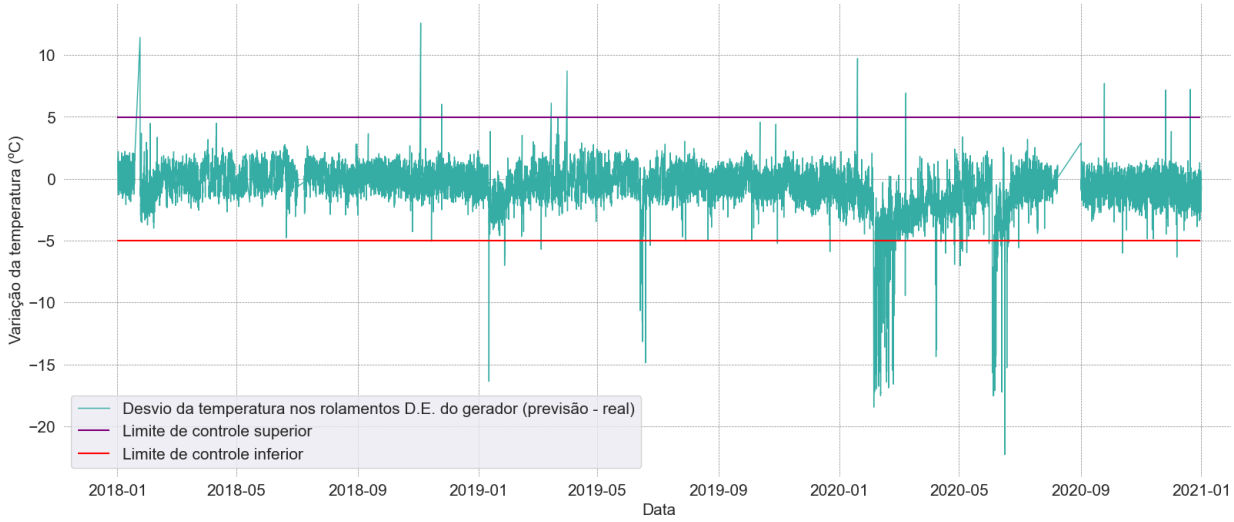
Figura 24 – Resultado do modelo preditivo LSTM para o gerador da máquina CV109 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Constata-se que o modelo do gerador, aplicado a esta máquina específica, apresentou dificuldades em antecipar os valores de temperatura para todo o intervalo de tempo.

Figura 25 – Gráfico de controle para desvios da temperatura do rolamento para o gerador da máquina CV109 em escala de 10 em 10 minutos.

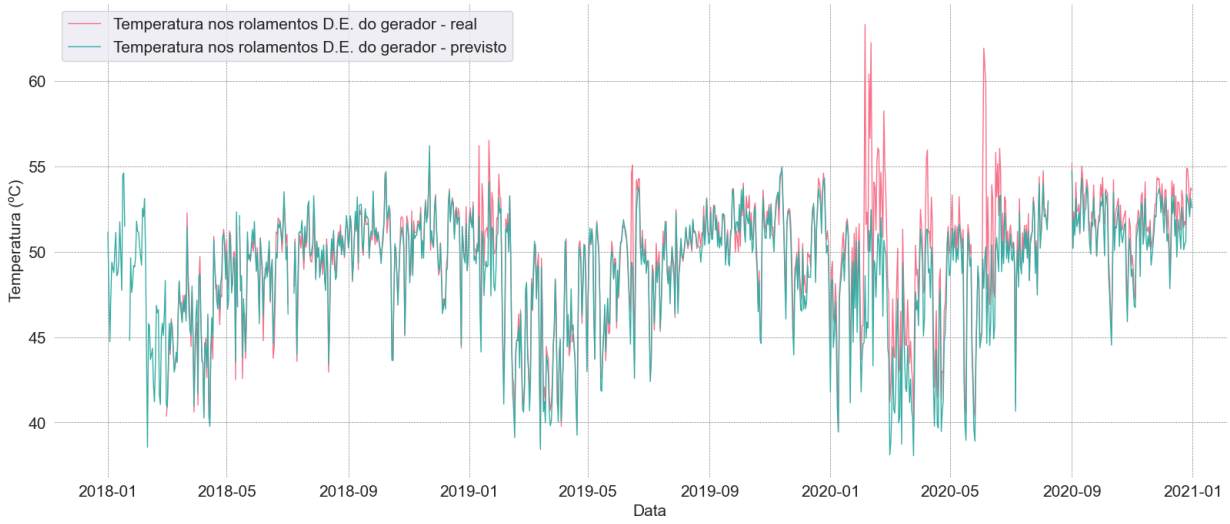


Fonte: próprio autor.

Na Figura 24, nota-se que isso ocorre devido ao comportamento anômalo da temperatura do rolamento do gerador durante um intervalo específico de tempo.

Na Figura 26, é ilustrado o gráfico das previsões, e na Figura 27, os desvios da temperatura do rolamento *drive end* do gerador, ambos em escala de média diária. O limite superior e inferior calculado para a escala de médias diárias é de ±4,233°C.

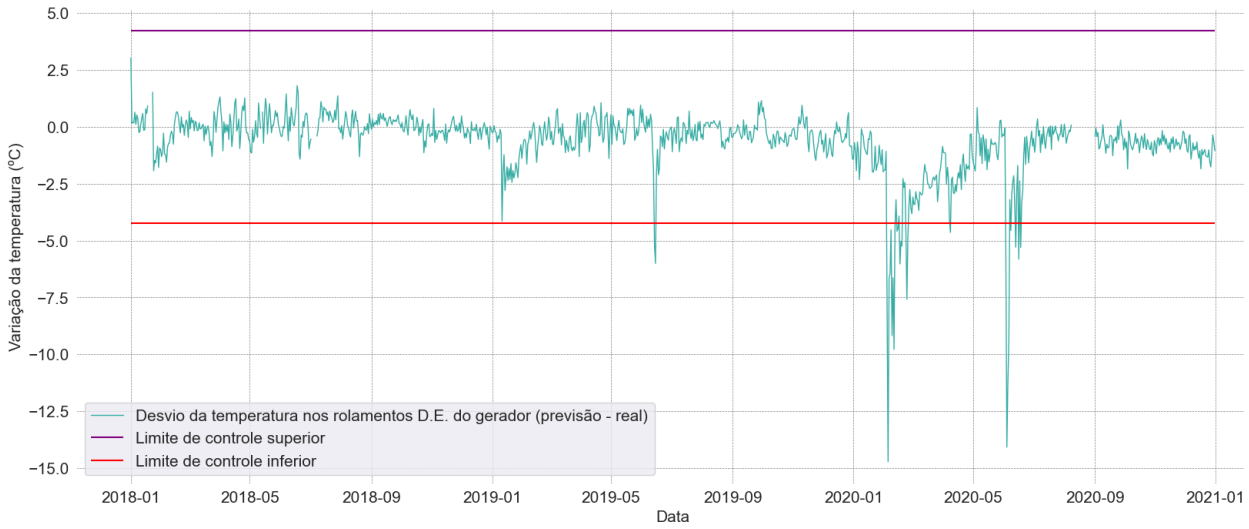
Figura 26 – Resultado do modelo preditivo LSTM para o gerador da máquina CV109 em escala diária.



Fonte: próprio autor.

Percebe-se que em três momentos há a violação do limite inferior do gráfico de controle. O ponto em que o desvio cruza com o limite inferior pode ser utilizado como uma

Figura 27 – Gráfico de controle para desvios da temperatura do rolamento do gerador da máquina CV109 em escala diária.



Fonte: próprio autor.

ferramenta de sinalização com respeito ao início de temperaturas anômalas no rolamento D.E. da máquina.

O primeiro ponto abaixo do limite inferior ocorre em 14 de junho de 2019, entretanto não há tendência de maior violação e a máquina retoma rapidamente aos níveis normais de operação. Porém, vale ressaltar, que essa primeira violação de controle pode sugerir, preliminarmente, uma possível falha no componente em análise. Para validar essa hipótese, seria crucial avaliar outros parâmetros operacionais, além de examinar os registros de manutenção correspondentes a essa máquina.

O segundo episódio de desvio da temperatura do rolamento D.E. do gerador corta o limite inferior ocorre em 4 de fevereiro de 2020, com uma temperatura média diária de 52,29°C. No dia seguinte, 5 de fevereiro de 2020, a máquina operou com uma temperatura média de 63,32°C, destacando-se da média observada durante o período analisado.

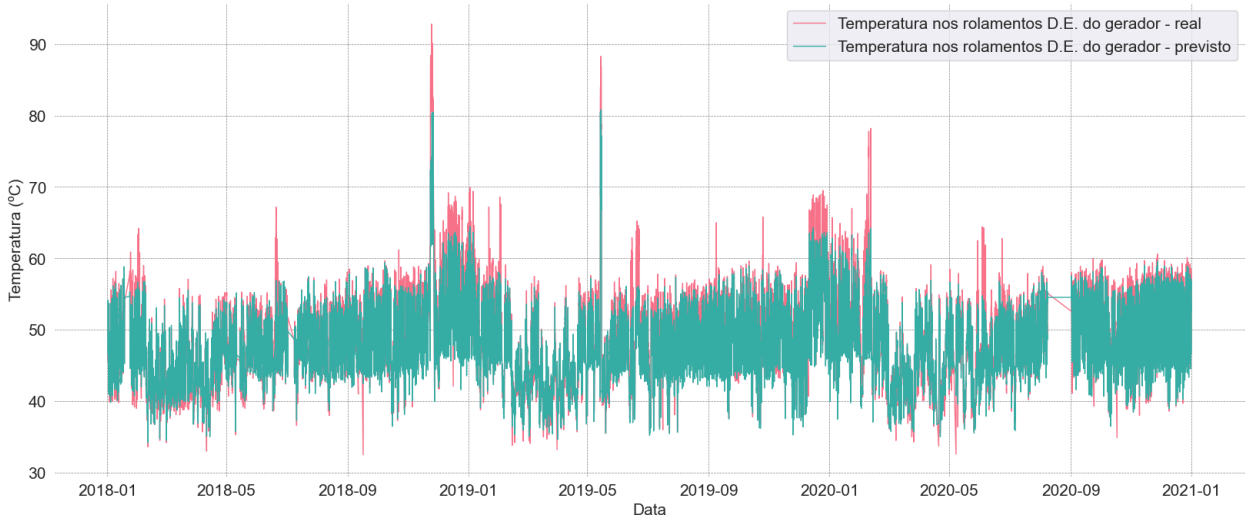
De maneira análoga, o terceiro ponto abaixo do limite inferior ocorre em 3 de junho de 2020, com uma temperatura média diária de 55,5°C. Nos três dias seguintes, a máquina opera com uma média de temperatura de 61°C e ao longo desses dias, atinge picos de temperatura superiores a 70°C, semelhante ao que ocorreu no ponto analisado anteriormente.

Neste contexto, a identificação prévia de uma anomalia na temperatura desse componente proporcionaria a mitigação dos picos de temperatura nos dias subsequentes à detecção. Assim sendo, caso o departamento de manutenção tivesse recebido um alerta

preditivo em 4 de fevereiro de 2020 e em 3 de junho de 2020, seria possível mitigar os picos de temperatura superiores a 70°C que se manifestam nos dias posteriores.

Para a máquina CV111, o modelo LSTM obteve um melhor desempenho com um R-quadrado de 0,92452, RMSE de 1,33960°C, MAE de 1,08216°C e MAPE de 0,02171%. Na Figura 28 é ilustrado o comportamento do modelo aplicado para prever os valores da temperatura do rolamento do gerador para a máquina em escala de dados SCADA, 10 em 10 minutos. Na Figura 29 está indicado o comportamento dos desvios da temperatura com a avaliação do gráfico de controle. O limite superior e inferior calculado para a escala SCADA é de ±4,99°C.

Figura 28 – Resultado do modelo preditivo LSTM para o gerador da máquina CV111 em escala de 10 em 10 minutos.



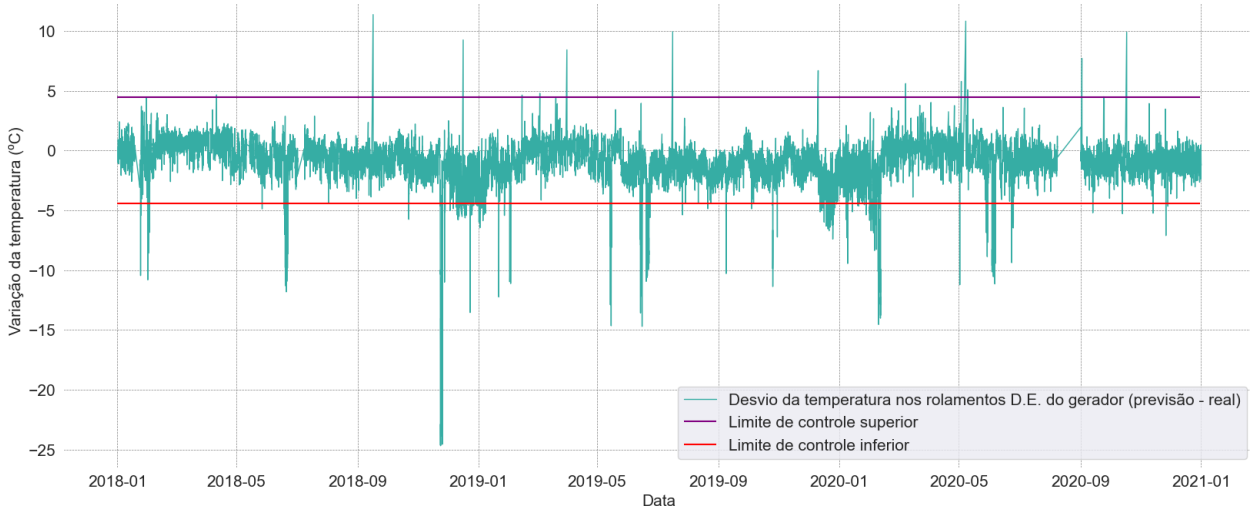
Fonte: próprio autor.

Nota-se a presença de picos significativos de temperatura, em que um deles chega a atingir valores superiores a 90°C. É possível observar também que o modelo apresenta leve dificuldade em prever temperaturas mais altas, quando comparado com a previsão de pontos de temperatura mais baixas.

Na Figura 30, é ilustrado o gráfico das previsões, e na Figura 31, os desvios da temperatura do rolamento *drive end* do gerador, ambos em escala de média diária. O limite superior e inferior calculado para a escala de médias diárias é de ±4,012°C.

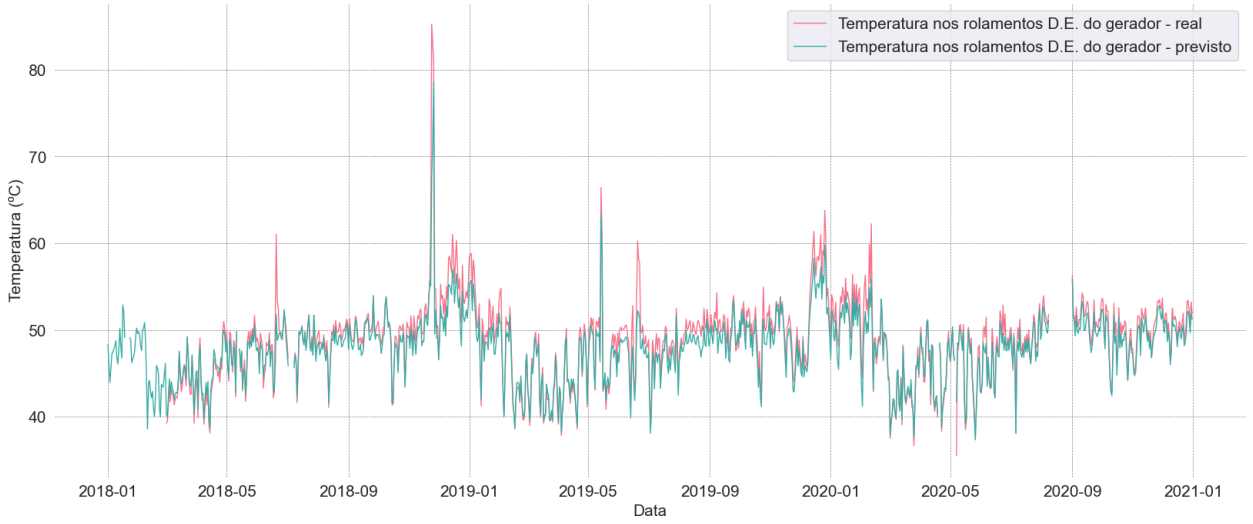
Percebe-se que para o gráfico de médias diárias em quatro momentos há a violação do limite inferior do gráfico de controle, sendo três dessas violações significativas devido aos níveis de temperatura apresentados. O primeiro ponto abaixo do limite inferior ocorre

Figura 29 – Gráfico de controle para desvios da temperatura do rolamento para o gerador da máquina CV111 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Figura 30 – Resultado do modelo preditivo LSTM para o gerador da máquina CV111 em escala diária.

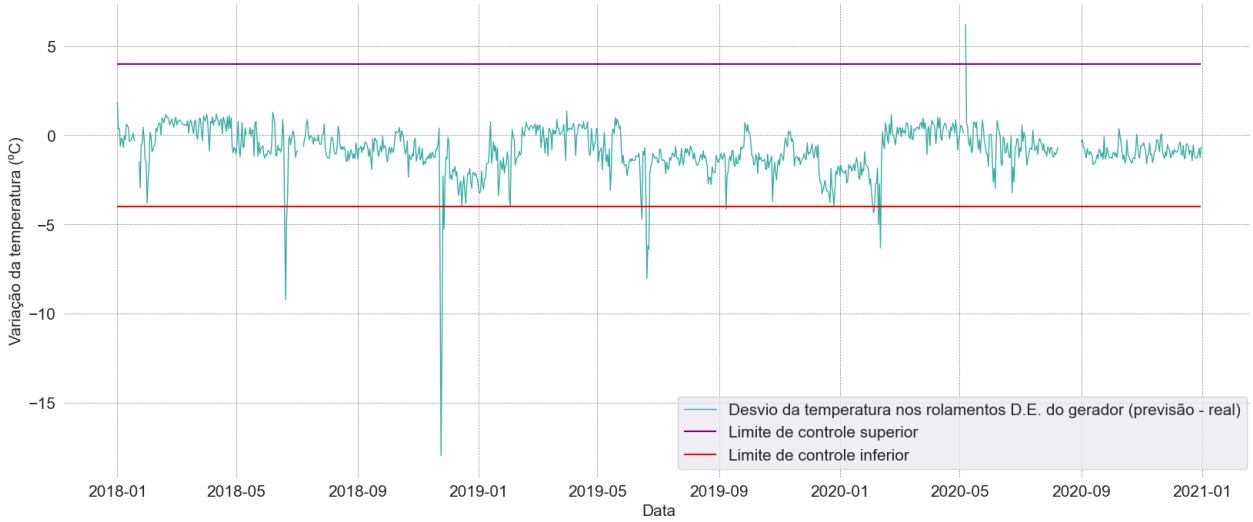


Fonte: próprio autor.

em 19 de junho de 2018, entretanto não há tendência de maior violação e a máquina retoma rapidamente aos níveis normais de operação.

O segundo episódio de desvio da temperatura do rolamento D.E. do gerador corta o limite inferior em 23 de novembro de 2018, com uma temperatura média diária de 61,86°C. Nos dois dias seguintes, a máquina operou com uma temperatura do rolamento D.E. do gerador com uma média de 85,24°C e 82,99°C, respectivamente. Esse episódio é o mais grave, pois em escala de 10 minutos a máquina chega a atingir temperaturas acima de 90°C, podendo resultar em danos graves a máquina.

Figura 31 – Gráfico de controle para desvios da temperatura do rolamento do gerador da máquina CV111 em escala diária.



Fonte: próprio autor.

O terceiro ponto abaixo do limite inferior ocorre no dia 15 de junho de 2019, com uma temperatura média diária de 52,8°C. Nos dias seguintes a máquina retorna a temperaturas de operação dentro dos limites do gráfico de controle. Entretanto, no dia 20 de junho de 2019, a máquina apresenta um novo pico de temperatura detectada pelo modelo. Nesse dia, a máquina opera com uma temperatura média de 60,2°, porém ao longo do dia tem picos de temperatura da ordem de 65°C.

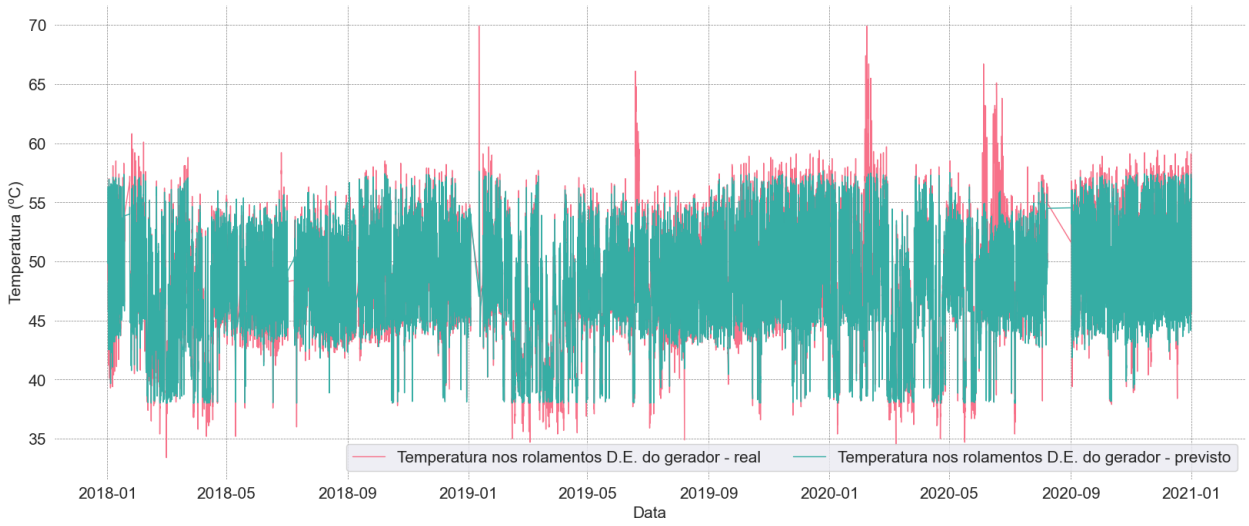
O quarto episódio ocorre no dia 4 de fevereiro de 2020, entretanto não há tendência de maior violação e a temperatura logo retorna para os níveis normais, com os desvios dentro do gráfico de controle.

Dessa maneira, a identificação prévia de uma anomalia na temperatura desse componente proporcionaria a mitigação dos picos de temperatura nos dias subsequentes à detecção. Assim sendo, caso o departamento de manutenção tivesse recebido um alerta preditivo nos primeiros dias em que houve a violação do limite indicado no gráfico de controle, seria possível mitigar os picos de temperatura apresentados.

Para a máquina CV104, o modelo LSTM obteve um melhor desempenho com um R-quadrado de 0,88935, RMSE de 1,43152°C, MAE de 0,78683°C e MAPE de 0,01547%. Na Figura 32 é ilustrado o comportamento do modelo aplicado para prever os valores da temperatura do rolamento do gerador para a máquina em escala de dados SCADA, 10 em 10 minutos. Na Figura 33 está indicado o comportamento dos desvios da temperatura

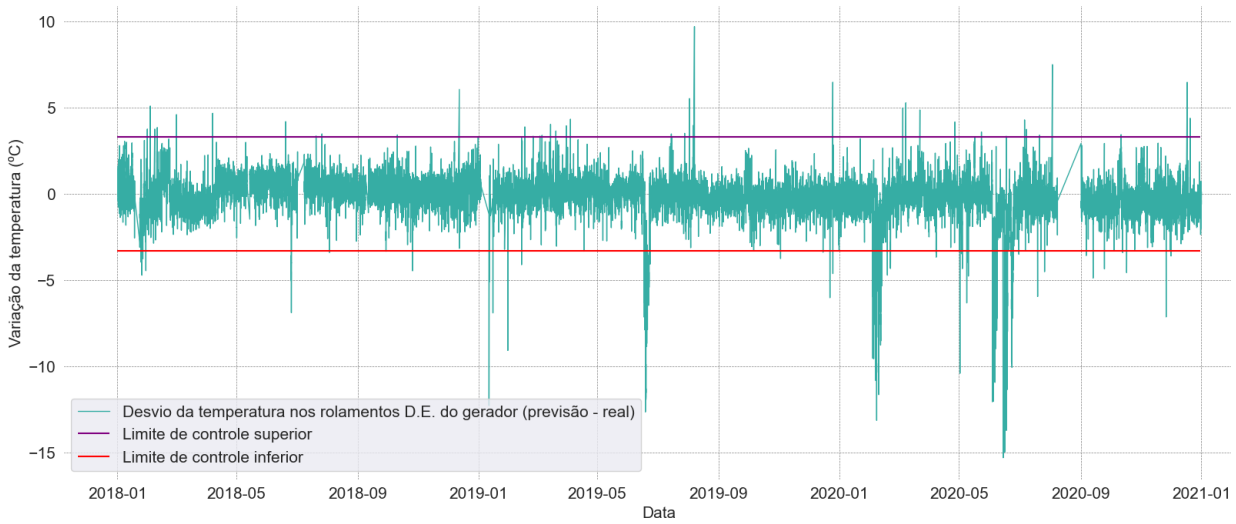
com a avaliação do gráfico de controle. O limite superior e inferior calculado para a escala SCADA é de ±3,29°C.

Figura 32 – Resultado do modelo preditivo LSTM para o gerador da máquina CV104 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Figura 33 – Gráfico de controle para desvios da temperatura do rolamento para o gerador da máquina CV104 em escala de 10 em 10 minutos.

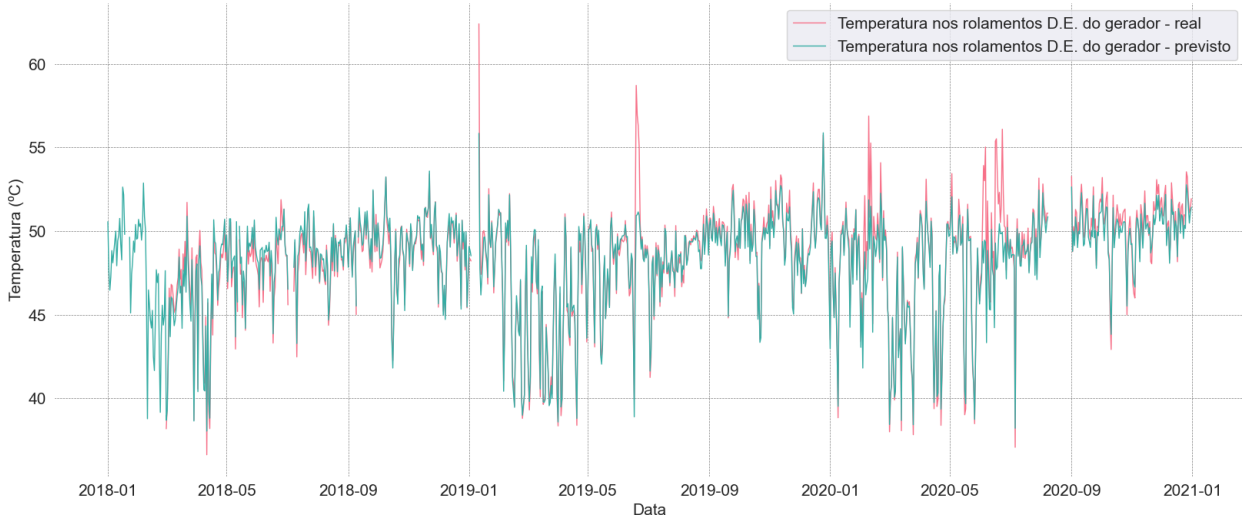


Fonte: próprio autor.

Para essa máquina, nota-se que o modelo do gerador obteve um resultado satisfatório na previsão da temperatura do rolamento *drive-end* e, observando o gráfico de controle, houve a identificação de picos de temperatura não convencionais durante a operação.

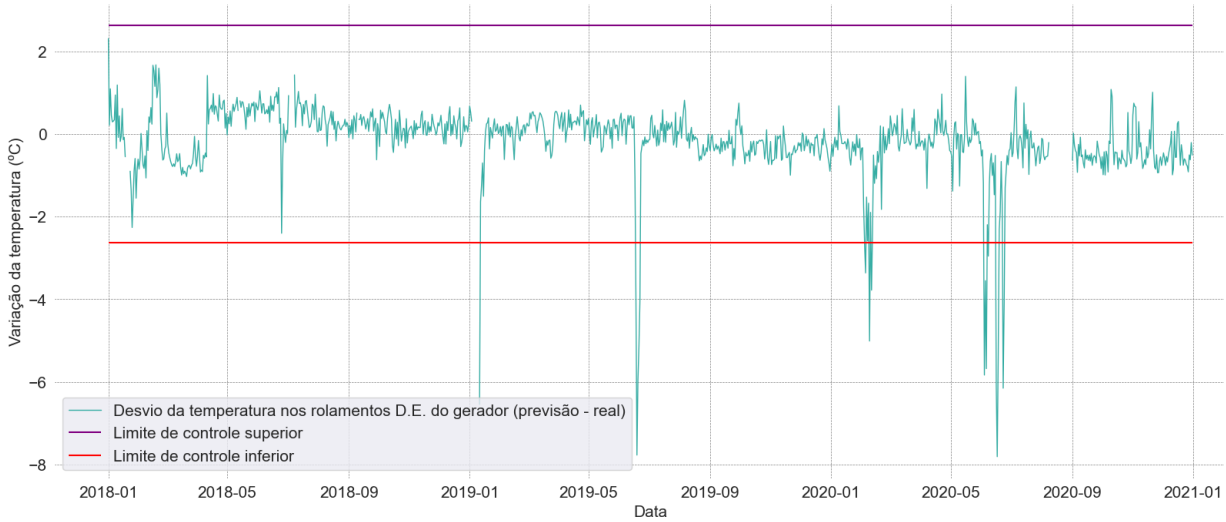
Na Figura 34, é ilustrado o gráfico das previsões, e na Figura 35, os desvios da temperatura do rolamento *drive end* do gerador, ambos em escala de média diária. O limite superior e inferior calculado para a escala de médias diárias é de ±2,63°C.

Figura 34 – Resultado do modelo preditivo LSTM para o gerador da máquina CV104 em escala diária.



Fonte: próprio autor.

Figura 35 – Gráfico de controle para desvios da temperatura do rolamento do gerador da máquina CV104 em escala diária.



Fonte: próprio autor.

Percebe-se que para o gráfico de médias diárias em quatro momentos há a violação do limite inferior do gráfico de controle, sendo três dessas violações significativas devido a permanência da temperatura em patamar não convencional. O primeiro ponto abaixo

do limite inferior ocorre em 11 de janeiro de 2019, entretanto não há tendência de maior violação e a máquina retoma rapidamente aos níveis normais de operação.

O segundo episódio de desvio da temperatura do rolamento D.E. do gerador corta o limite inferior em 18 de junho de 2019, com uma temperatura média diária de 50,64°C. No dia seguinte, 19 de junho de 2019, a máquina operou com uma temperatura do rolamento D.E. do gerador com uma média de 58,70°C e apresentou picos de temperatura superior a 60°C. No dia 20 de junho de 2019, a temperatura média de operação é de 56,92°C, indicando uma tendência de queda, conforme exibido no gráfico de controle.

Salienta-se que este pico de temperatura poderia ter sido antecipado um dia antes devido ao primeiro alerta em 18 de junho de 2019. E, ao contrário do que foi observado na máquina CV111, este equipamento não apresentou elevações críticas de temperatura.

O terceiro ponto situado abaixo do limite inferior manifesta-se no dia 4 de fevereiro de 2020, com uma temperatura média diária de 49,69°C. No dia 9 de fevereiro de 2020, o rolamento *drive-end* da máquina atinge uma temperatura média de 56,87°C. Analisando o gráfico em escala SCADA, nota-se a presença de diversos pontos de temperaturas acima de 65°C no início de fevereiro de 2020, que podem ser prejudiciais para uma operação a longo prazo.

Dessa forma, ressalta-se a importância de uma avaliação abrangente de ambos os gráficos, uma vez que a visualização ponto-a-ponto permite a identificação dos valores mais elevados de temperatura. E, de maneira análoga, os picos de temperatura acima de 65°C poderiam ser evitados com um alerta no dia 4 de fevereiro de 2020, quando a máquina ainda estava operando com uma temperatura da ordem de 50°C.

Ainda, note que a avaliação apenas do gráfico de médias diárias poderia indicar um comportamento médio de temperatura dentro do padrão operacional da máquina, desprezando picos de temperatura que ocorrem na operação com a escala de 10 em 10 minutos. Esta ocorrência pode ser crítica, pois pode não identificar um problema a curto-prazo, que poderá se manifestar posteriormente. Ainda, mesmo que a violação do gráfico de controle não seja significativa, não é indicado desconsiderar transgressões dos limites definidos, mesmo quando as temperaturas não atingem patamares significativamente elevadas.

O quarto episódio manifesta-se no dia 16 de junho de 2020 com uma temperatura média de 55,36°C. Nesse caso, o rolamento *drive-end* da máquina inicia temperaturas

elevadas, em seguida tem uma queda momentânea na temperatura, porém retornar a pontos de temperatura elevada. De acordo com o gráfico em escala SCADA, é possível verificar picos recorrentes de temperatura, superiores a 60°C, ao longo dos dias seguintes à 16 de junho de 2020. A temperatura média do rolamento atinge a marca de 56,08° no dia 23 de junho de 2020.

Esse comportamento da temperatura não é convencional e é previsto pelo modelo logo no dia 16 de junho de 2020, indicando que houve uma violação do limite inferior do gráfico de controle. Conforme já comentado, essa indicação pode servir como alerta para o setor de manutenção a fim de evitar picos de temperatura que possam prejudicar o desempenho da máquina e causar prejuízos catastróficos.

Os resultados apresentados na Tabela 10 fornecem informações sobre o desempenho dos modelos aplicados à previsão da temperatura do rolamento da caixa de engrenagem em diferentes turbinas. Ao analisarmos os resultados para as máquinas CV109, CV111 e CV104, observamos que o modelo LSTM se destacou, apresentando os menores valores de RMSE, MAE e MAPE, além do maior valor de R-quadrado em comparação com os modelos de Regressão Múltipla Linear (MLR) e XGBoost. Esse resultado indica que, para essas turbinas específicas, o algoritmo LSTM teve um melhor desempenho ao capturar os padrões de comportamento nos dados e realizar previsões precisas da temperatura do rolamento *drive-end* do gerador.

Tabela 10 – Métricas de desempenho do modelo aplicado para previsão de temperatura do rolamento *drive-end* do gerador.

Turbina	Modelo	R-quadrado	RMSE	MAE	MAPE
CV109	MLR	0,77216	2,29991	1,55748	0,02995
	XGBoost	0,80973	2,10174	1,27333	0,02412
	LSTM	0,81727	2,06194	1,17476	0,02236
CV111	MLR	0,86730	1,78123	1,42008	0,02884
	XGBoost	0,92365	1,35105	1,02946	0,02091
	LSTM	0,92452	1,33960	1,08216	0,02171
CV104	MLR	0,79558	1,94350	1,46839	0,02929
	XGBoost	0,85905	1,61384	1,05149	0,02075
	LSTM	0,88935	1,43152	0,78683	0,01547

Note que o modelo de rede neural recorrente teve uma maior dificuldade em se ajustar aos dados da máquina CV109 e por este apresentou o menor valor de R-quadrado e os maiores valores de RMSE, MAE e MAPE quando comparado com o mesmo modelo aplicado às outras máquinas. Na avaliação dos gráficos de previsão, torna-se evidente que essa discrepância foi influenciada pela presença de picos de temperatura anômalos que surgiram no ano de 2020, os quais o modelo não conseguiu reproduzir com precisão em relação aos valores de temperatura reais da máquina.

O modelo LSTM aplicado a máquina CV111 teve o melhor ajuste aos dados com o maior R-quadrado e menores valores de RMSE, MAE e MAPE quando comparado com as outras duas máquinas. Um aspecto interessante na previsão da temperatura do rolamento *drive-end* do gerador é que o modelo LSTM demonstrou ser a escolha mais eficiente para todas as máquinas em análise. Esse resultado sugere que a antecipação do comportamento térmico desse componente tem melhores resultados com o emprego de redes neurais.

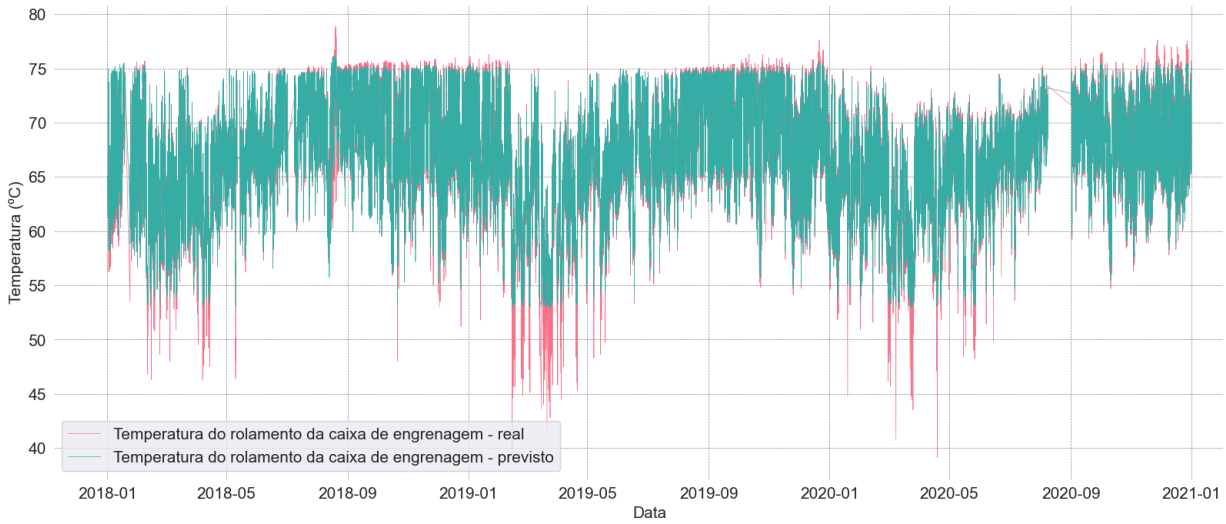
4.3. **Caixa de engrenagem**

De forma análoga, os modelos de aprendizagem já mencionados foram aplicados para prever a temperatura do rolamento da caixa de engrenagem das três máquinas em estudo. Para a máquina CV109, o modelo LSTM obteve um melhor desempenho com um R-quadrado de 0,98999, RMSE de 0,41603°C, MAE de 0,31790°C e MAPE de 0,00470%. Dessa maneira, nota-se que para essa máquina o modelo LSTM obteve um resultado robusto do ponto de vista do R-quadrado, indicando que cerca de 99% da variabilidade é explicada pelo modelo construído.

Na Figura 36 é ilustrado o comportamento do modelo aplicado para prever os valores da temperatura do rolamento da caixa de engrenagem para a máquina CV109 em escala de dados SCADA, 10 em 10 minutos. Na Figura 37 está indicado o comportamento dos desvios da temperatura com a avaliação do gráfico de controle. O limite superior e inferior calculado para a escala SCADA é de ±1,67 °C.

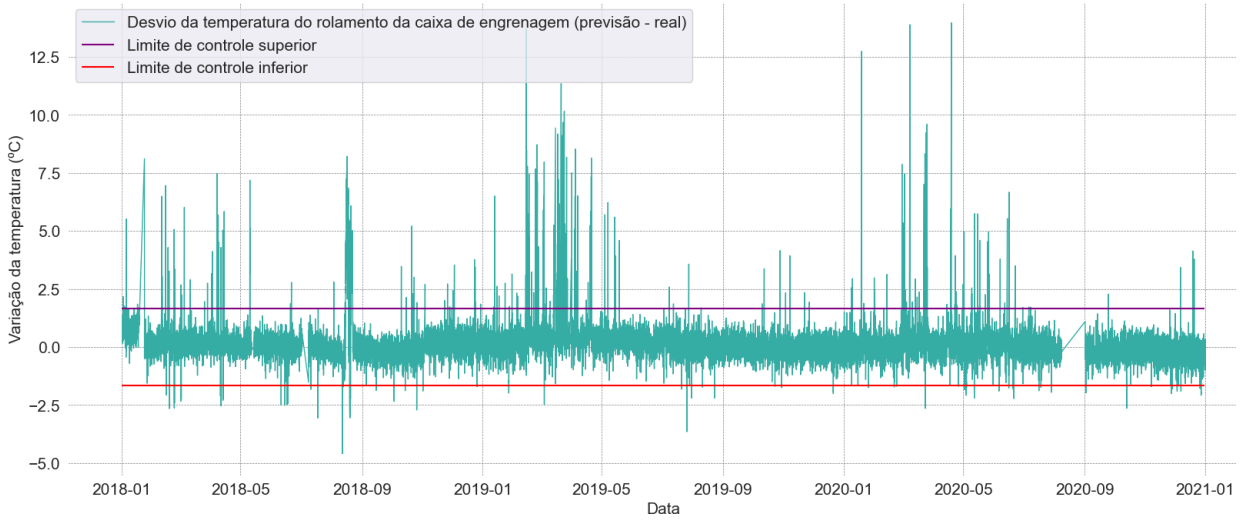
Observando o gráfico de controle da Figura 37, é possível notar que a maioria dos pontos fora dos limites de controle encontram-se no eixo positivo da variação da temperatura. Isso indica que há uma variação significativa entre o valor previsto e o valor real da temperatura na máquina. No caso da máquina CV109, nota-se que como as violações

Figura 36 – Resultado do modelo preditivo LSTM para caixa de engrenagem da máquina CV109 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Figura 37 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV109 em escala de 10 em 10 minutos.



Fonte: próprio autor.

são do limite superior de controle, significa que o modelo previu uma temperatura maior do que real medida pelo supervisório.

Isso pode direcionar a análise para um comportamento normal[7], visto que a máquina estava operando em temperaturas baixas. Entretanto, é importante avaliar o porquê a máquina sofreu variações bruscas na temperatura, ainda que para valores de operação

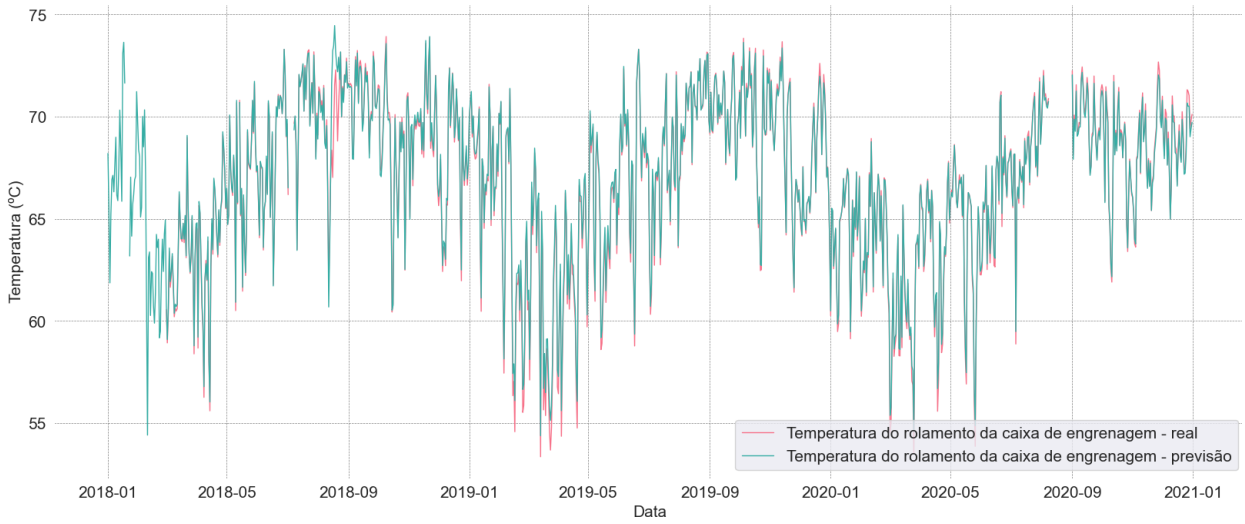
[7] *A referência ao comportamento normal é com relação a série histórica de temperaturas apresentadas pela máquina, visto que não há logs de manutenção. Nesse caso, a anormalidade é tratada como quaisquer desvios com relação a série histórica da modelagem em operação normal da máquina.*

não críticos (temperaturas mais baixas). As variações mais significativas ocorrem nos períodos de fevereiro a junho, em que o potencial eólico é reduzido.

Outro ponto importante a ser observado é o comportamento sazonal da temperatura, que pode sofrer influência do comportamento climático da região. Note que os pontos de temperatura mais altos ocorrem justamente nos períodos em que o potencial da energia eólica é maior (junho a dezembro).

De maneira adicional, foi realizada um aglutinamento dos dados para uma média diária das previsões e dos valores reais com a finalidade de ter uma visualização mais limpa do comportamento da temperatura, e, ainda, avaliar seu comportamento médio diário. Na Figura 38 é ilustrada a previsão em escala diária e na Figura 39 é ilustrado o gráfico de controle para médias diárias. O limite superior e inferior calculado para a escala de médias diárias é de ±0,99 °C.

Figura 38 – Resultado do modelo preditivo LSTM para caixa de engrenagem da máquina CV109 em escala diária.

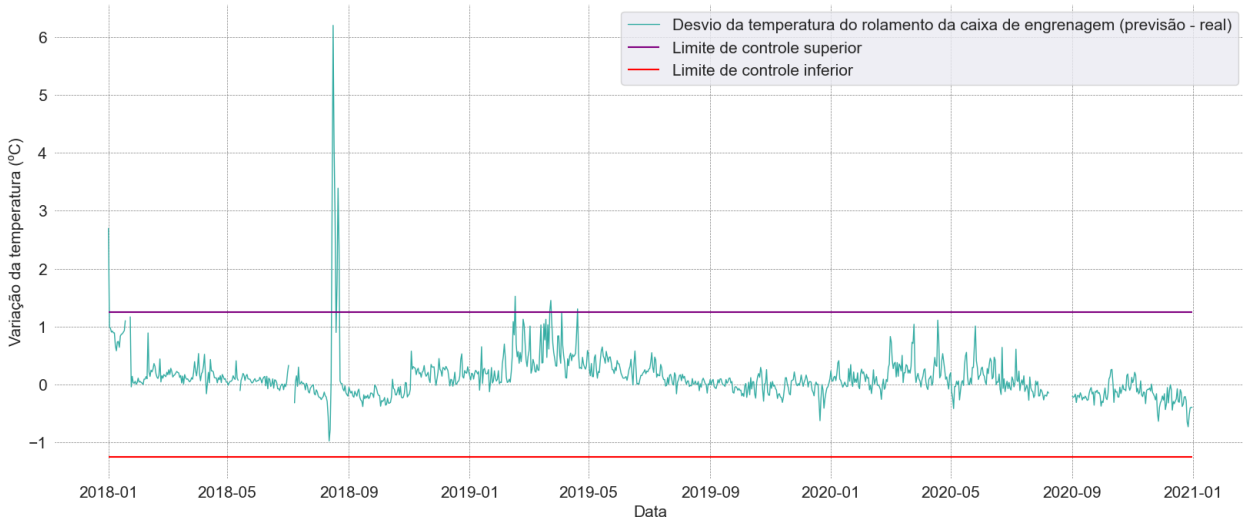


Fonte: próprio autor.

Analisando o gráfico é possível notar que ao realizar médias diárias com os valores de previsão e reais temos uma diminuição dos picos de temperatura da máquina. E, que, do ponto de vista do gráfico de controle, o comportamento médio da máquina não apresentou desvios significativos na temperatura do rolamento da caixa de engrenagem. Ainda, vale ressaltar que os maiores valores de temperatura observados não superaram o limite de 80°C.

Para a máquina CV111, o modelo XGBoost obteve um melhor desempenho com um R-quadrado de 0,96987, RMSE de 0,72626°C, MAE de 0,57090°C e MAPE de 0,00835%.

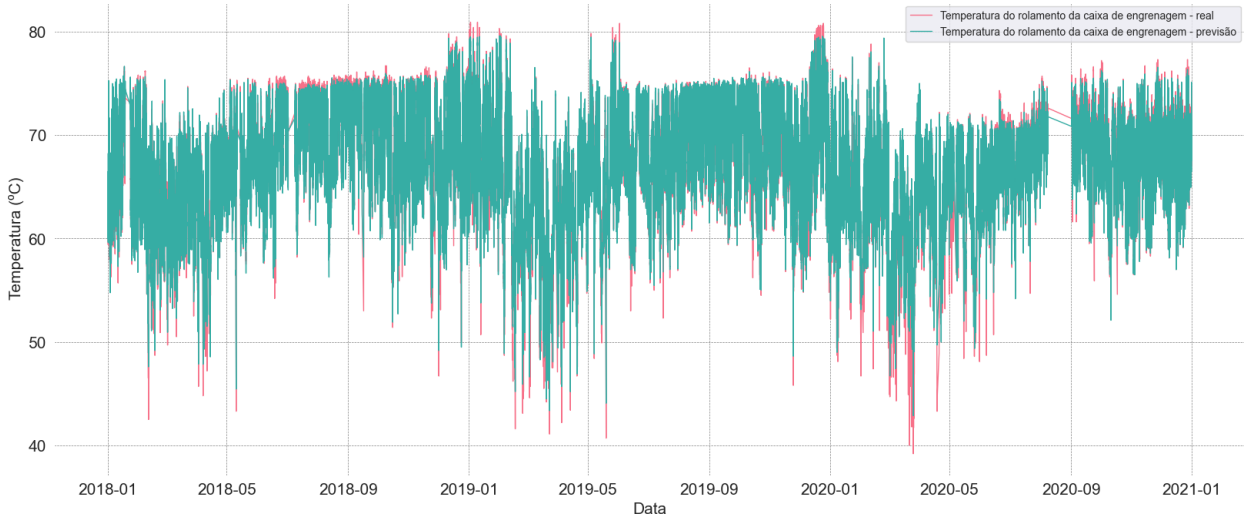
Figura 39 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV109 em escala diária.



Fonte: próprio autor.

Na Figura 40 é ilustrado o comportamento do modelo aplicado para prever os valores de temperatura do rolamento da caixa de engrenagem para a máquina CV111 em escala de dados SCADA, 10 em 10 minutos. Na Figura 41 está indicado o comportamento dos desvios da temperatura do rolamento da caixa de engrenagem com a avaliação do gráfico de controle. O limite superior e inferior calculado para a escala SCADA é de ±1,931 °C.

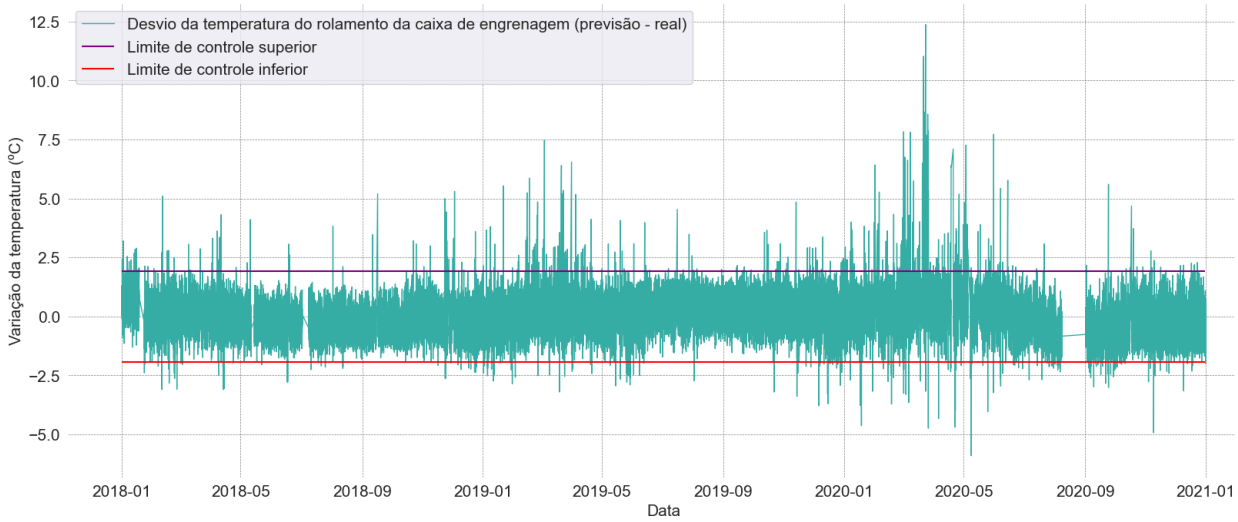
Figura 40 – Resultado do modelo preditivo XGBoost para caixa de engrenagem da máquina CV111 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Nota-se que os dados estão em sua maioria dentro dos limites do gráfico de controle, indicando um comportamento normal. No período de fevereiro a abril de 2020, houve uma

Figura 41 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV111 em escala de 10 em 10 minutos.



Fonte: próprio autor.

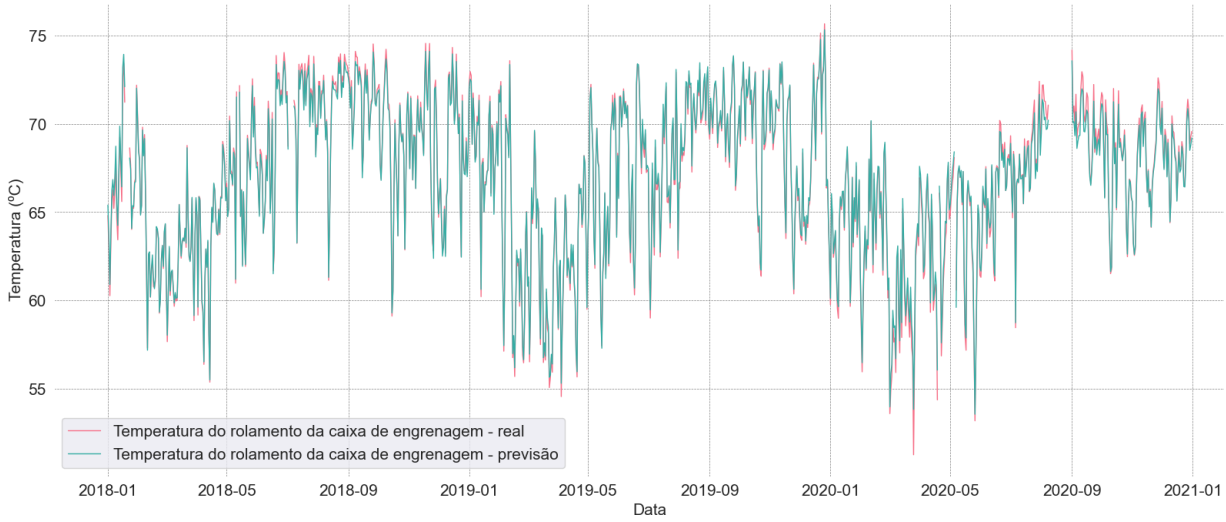
queda acentuada na temperatura do rolamento da caixa de engrenagem gerando pontos acima do limite superior do gráfico de controle, ou seja, indicando que há uma variação significativa entre o valor previsto e o valor real da temperatura na máquina. Nesse caso, a temperatura real apresentou valores menores que a temperatura prevista pelo modelo.

Na Figura 42, é ilustrado o gráfico das previsões, e na Figura 43, os desvios da temperatura do rolamento da caixa de engrenagem, ambos em escala de média diária. Fica claro que, na média, o rolamento da caixa de engrenagem apresenta níveis de temperatura dentro dos limites do gráfico de controle, indicando uma operação normal. O limite superior e inferior calculado para a escala de médias diárias é de ±1,029 °C.

Ainda, foram observados pontos de temperatura acima de 80°C em janeiro, junho e dezembro de 2019, porém a permanência nesse patamar de temperatura não foi mantida por muito tempo. E, do ponto de vista do gráfico de controle, não foram observados desvios de temperatura significativos a ponto de indicar um possível problema no rolamento da caixa de engrenagem dessa máquina.

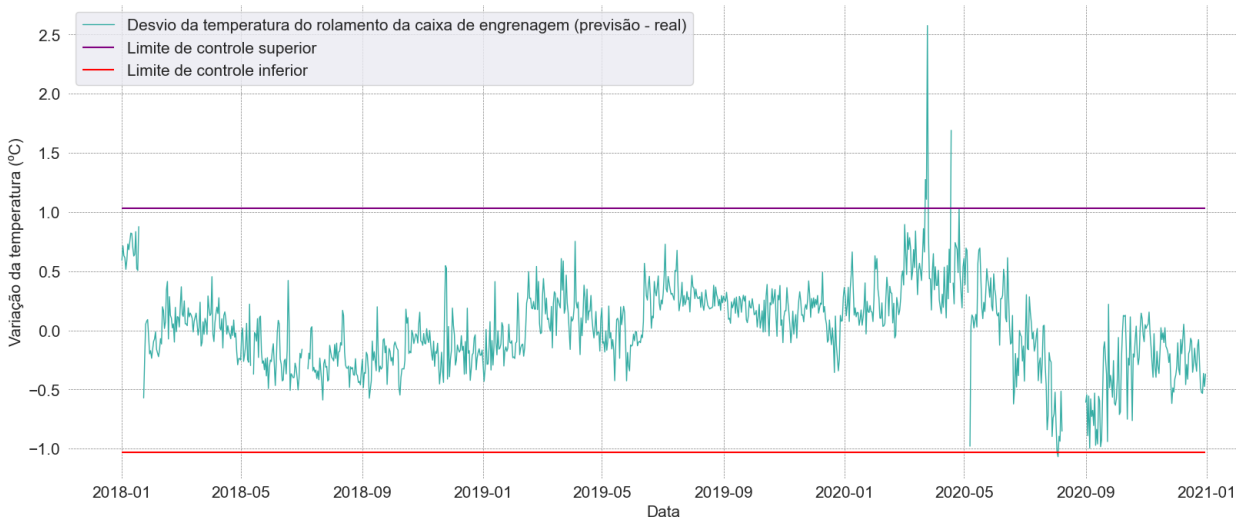
Para a máquina CV104, o modelo XGBoost obteve um melhor desempenho com um R-quadrado de 0,97133, RMSE de 0,60708°C, MAE de 0,47823°C e MAPE de 0,00701%. Na Figura 44 é ilustrado o comportamento do modelo aplicado para prever os valores de temperatura do rolamento da caixa de engrenagem para a máquina CV104 em escala de dados SCADA, 10 em 10 minutos. Na Figura 44 está indicado o comportamento dos

Figura 42 – Resultado do modelo preditivo XGBoost para caixa de engrenagem da máquina CV111 em escala diária.



Fonte: próprio autor.

Figura 43 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV111 em escala diária.

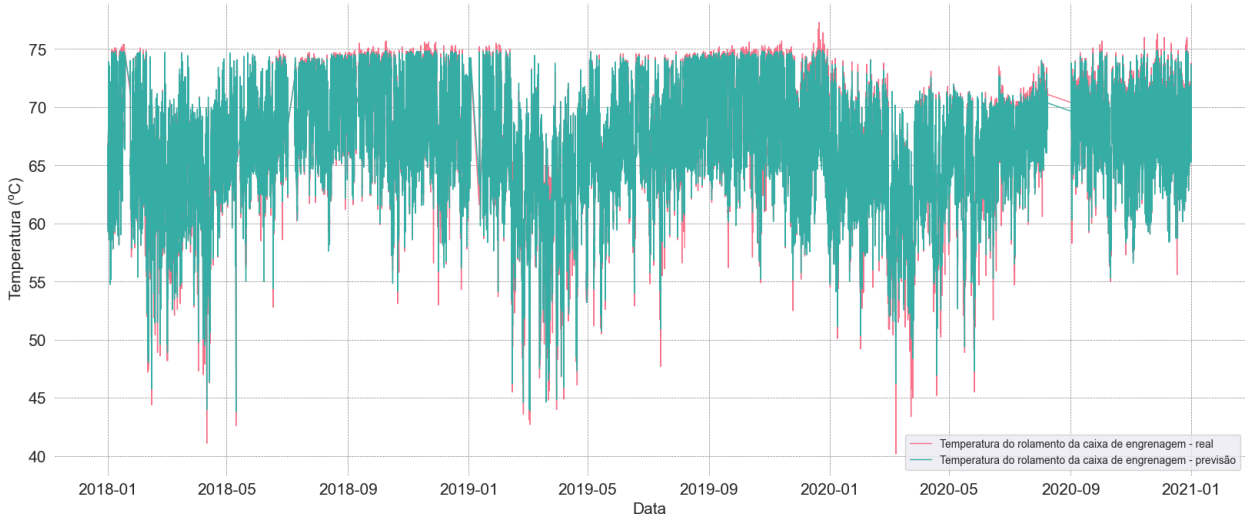


Fonte: próprio autor.

desvios da temperatura do rolamento da caixa de engrenagem com a avaliação do gráfico de controle. O limite superior e inferior calculado é de ±1,601 °C.

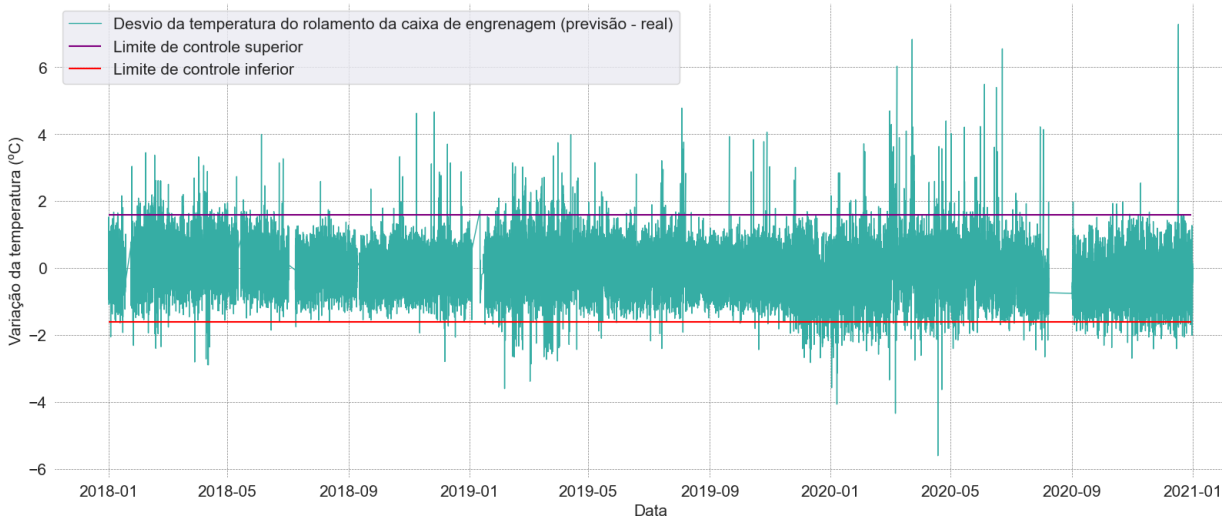
Dessa maneira, procedemos a uma abordagem análoga à discutida em relação a máquina anterior para analisar as temperaturas do rolamento na caixa de engrenagem. Uma inspeção do gráfico de controle evidencia que a maioria dos dados que ultrapassam os limites situa-se na região superior. Tal observação sugere que a temperatura real da máquina é inferior àquela prevista, não constituindo, portanto, uma possível indicação para falha.

Figura 44 – Resultado do modelo preditivo XGBoost para caixa de engrenagem da máquina CV104 em escala de 10 em 10 minutos.



Fonte: próprio autor.

Figura 45 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV104 em escala de 10 em 10 minutos.



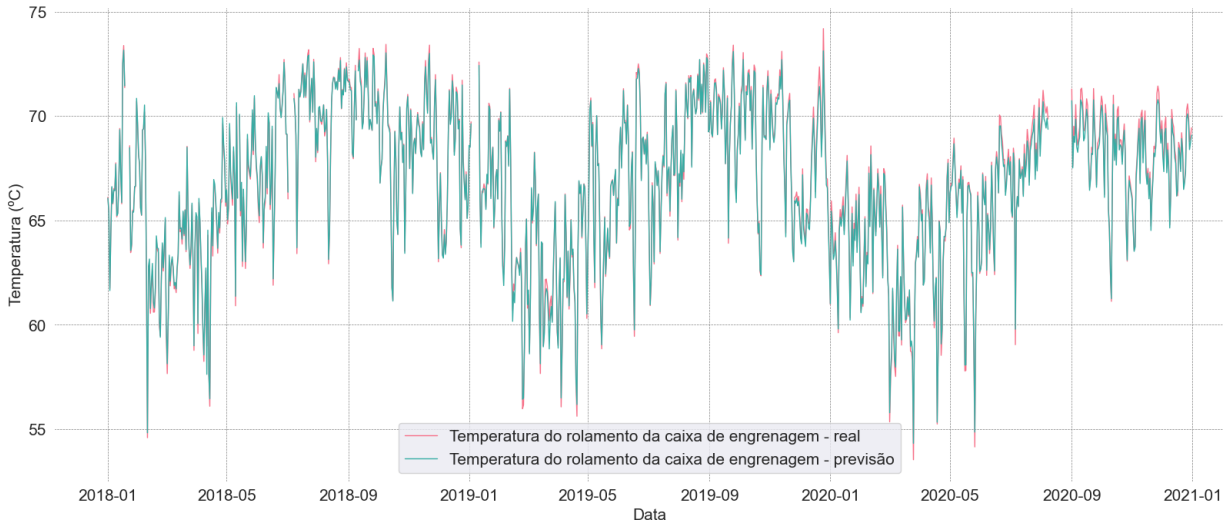
Fonte: próprio autor.

Na Figura 46, é ilustrado o gráfico das previsões, e na Figura 47, os desvios da temperatura do rolamento da caixa de engrenagem, ambos em escala de média diária. A análise revela que, em média, o rolamento da caixa de engrenagem mantém-se em níveis de temperatura contidos nos limites preestabelecidos no gráfico de controle, indicando um funcionamento dentro dos padrões normais. Os cálculos efetuados para a determinação dos limites superior e inferior, considerando a escala de médias diárias, resultaram em $\pm 0{,}705$ °C.

Ainda, analisando a Figura 47, destaca-se a relevância de observar que apenas em três ocasiões ocorreram violações do limite inferior do gráfico de controle. Nas duas primeiras instâncias, registradas em abril de 2018 e abril de 2019, não foram identificadas temperaturas que pudessem ser consideradas proibitivas no rolamento da caixa de engrenagem da máquina.

No que tange à terceira violação, ocorrida em dezembro de 2019, destaca-se um pico aproximado de 80°C (em escala SCADA). Importante notar que este evento não se perpetuou ao longo do tempo, sendo seguido por um retorno às temperaturas normais de operação.

Figura 46 – Resultado do modelo preditivo XGBoost para caixa de engrenagem da máquina CV104 em escala diária.

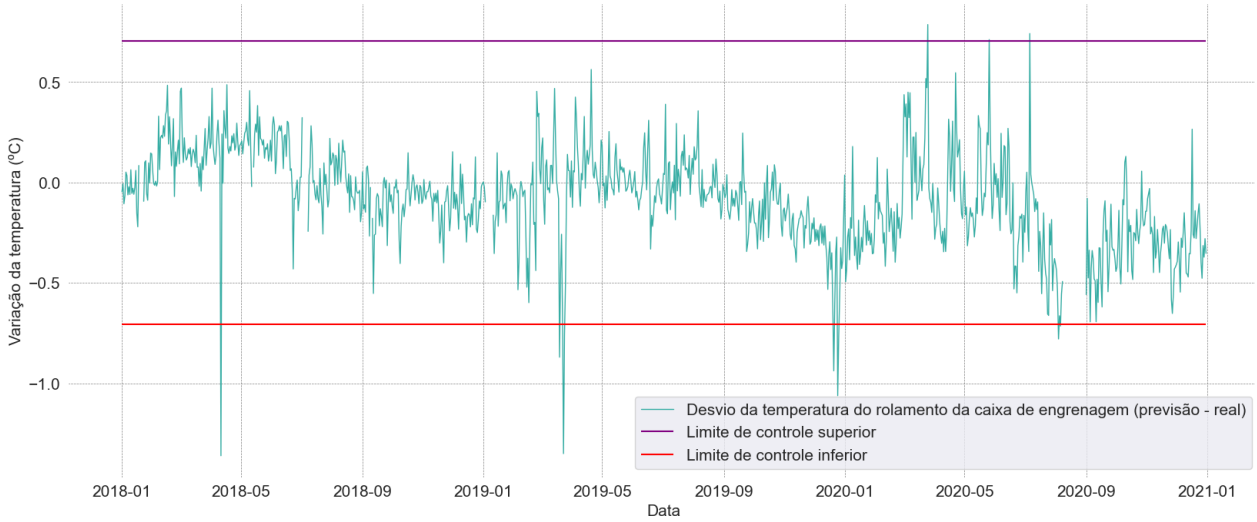


Fonte: próprio autor.

Desta forma, torna-se evidente que os modelos desenvolvidos para a previsão da temperatura do rolamento na caixa de engrenagens apresentaram resultados robustos com relação ao coeficiente de determinação (R-quadrado), indicando que mais de 95% da variabilidade está sendo explicada pelos modelos construídos. Estes modelos demonstraram uma capacidade robusta de antecipar com precisão os valores da temperatura do rolamento para as três máquinas objeto de análise.

Adicionalmente, constatou-se que, de maneira geral, durante a avaliação da temperatura na caixa de engrenagens, as três máquinas manifestaram um comportamento em conformidade com os padrões de operação para temperatura do componente. Em raros momentos, foram identificadas elevações de temperatura significativas, sugerindo a possibilidade de ocorrência de anomalias no componente da máquina.

Figura 47 – Gráfico de controle para desvios da temperatura do rolamento para caixa de engrenagem da máquina CV111 em escala diária.



Fonte: próprio autor.

Os resultados apresentados na Tabela 11 fornecem detalhes importantes sobre o desempenho dos modelos aplicados à previsão da temperatura do rolamento da caixa de engrenagem em diferentes turbinas. Ao analisarmos os resultados para a máquina CV109, observamos que o modelo LSTM se destacou, apresentando os menores valores de RMSE, MAE e MAPE, além do maior valor de R-quadrado em comparação com os modelos MLR e XGBoost. Isso indica que o modelo LSTM demonstrou um melhor resultado ao ajustar os dados associados à CV109 ao modelo preditivo.

Tabela 11 – Métricas de desempenho do modelo aplicado para previsão de temperatura do rolamento da caixa de engrenagem.

Turbina	Modelo	R-quadrado	RMSE	MAE	MAPE
CV109	MLR	0,96721	0,75349	0,61327	0,00892
	XGBoost	0,97255	0,68937	0,53039	0,00776
	LSTM	0,99000	0,41603	0,31790	0,00470
CV111	MLR	0,95968	0,84012	0,68598	0,00995
	XGBoost	0,96987	0,72627	0,57090	0,00835
	LSTM	0,95653	0,86948	0,76732	0,01108
CV104	MLR	0,95673	0,74592	0,61089	0,00888
	XGBoost	0,97134	0,60708	0,47823	0,00701
	LSTM	0,95963	0,72117	0,67486	0,01004

Em contraste, para as máquinas CV111 e CV104, o método XGBoost demonstrou superioridade em termos de desempenho preditivo. Ele obteve os menores valores de RMSE, MAE e MAPE e o maior R-quadrado em comparação com os modelos MLR e LSTM. Esse resultado indica que, para essas turbinas específicas, o algoritmo XGBoost teve um melhor desempenho ao capturar os padrões subjacentes nos dados e realizar previsões precisas da temperatura do rolamento.

5 CONCLUSÕES

Este trabalho buscou desenvolver uma metodologia para identificação de anomalias na temperatura de componentes vitais de um aerogerador. Diante do elevado montante de geração eólica no Brasil e no mundo, o desenvolvimento de ferramentas inovadoras com o auxílio de algoritmos de aprendizagem de máquina se faz interessante devido ao baixo custo e elevado potencial preditivo.

Inicialmente, buscou-se entender as raízes da motivação do crescente investimento em energias renováveis no mundo, bem como uma avaliação geral do crescimento das energias renováveis nos últimos anos. Em seguida, foi discutido o aspecto financeiro do custo nivelado da energia elétrica, despesa de capital e custos de operação e manutenção para as usinas eólicas. Na revisão literária, verificou-se que existem estudos evidenciando que a inatividade dos aerogeradores decorre devido a falhas na caixa de engrenagem, no gerador, rotor e no mancal principal da turbina. Por fim, foi realizada uma introdução geral a respeito da conversão de energia eólica em elétrica, principais características e seus componentes.

Em seguida, foi segmentada toda a base teórica dos fundamentos de aprendizagem de máquina visando o detalhamento matemático dos modelos de regressão utilizados no estudo. O conhecimento a fundo das ferramentas é de suma importância para a correta parametrização e aplicação dos modelos preditivos aos dados das máquinas.

Com esse trabalho, é evidenciado que o uso de ferramentas de aprendizagem tem um potencial enorme na avaliação de falhas e previsão de variáveis, em específico a temperatura. Foi realizada uma análise exploratória dos dados a fim de limpar e organizar os dados brutos SCADA das máquinas em estudo com o propósito de modelar uma operação normal da máquina.

Os modelos, desenvolvidos em linguagem de alto nível, foram ajustados aos dados tratados das máquinas e com isso foi possível realizar a previsão da temperatura dos componentes estudos. Em seguida, com as análises dos gráficos de controle foi possível identificar desvios anômalos de temperatura com um certo tempo de antecedência à ocorrência crítica de temperatura.

Os modelos aplicados foram comparados entre si, evidenciando discrepâncias quando aplicados a componentes distintos. No estudo, ficou claro que, o algoritmo de rede neu-

ral recorrente teve um melhor desempenho quando aplicado para previsão do rolamento *drive-end* do gerador, enquanto o modelo XGBoost teve um melhor desempenho para a previsão da temperatura do rolamento da caixa de engrenagem.

Essas discrepâncias nos resultados entre os aerogeradores sugerem que a escolha do modelo mais adequado pode depender das características específicas de cada máquina. A aplicação de diferentes algoritmos de aprendizagem de máquina pode levar a resultados distintos, destacando a importância de considerar as particularidades do sistema em questão ao escolher uma abordagem de modelagem.

Em geral, os resultados indicam que a escolha do modelo mais apropriado pode variar entre as máquinas, ressaltando a importância de uma abordagem personalizada para a previsão de temperatura. Essa análise comparativa dos modelos fornece uma base sólida para a otimização futura e refinamento dos métodos de previsão em cada aerogerador específico.

No contexto geral, os seguintes objetivos foram alcançados durante o desenvolvimento desta Dissertação de Mestrado:

1. Foi desenvolvida uma metodologia para identificação de anomalias em aerogeradores com a finalidade de antecipar a temperatura em componentes vitais e evitar ocorrências catastróficas;

2. Foram apresentados os conceitos, a modelagem e a aplicação dos algoritmos de aprendizagem de máquina para previsão da variável temperatura;

3. Foram utilizados gráficos de controle para identificar desvios de temperatura, comparando os valores das temperaturas previstas com as temperaturas reais dos componentes em estudo;

4. Os modelos foram avaliados entre si considerando as métricas de desempenho, e, também, realizada a comparação dos modelos aplicados nas três máquinas para cada componente analisado.

Conclui-se, então, que esta pesquisa busca oferecer uma contribuição para o entendimento e aprimoramento das técnicas de manutenção preditiva em parques eólicos, colaborando para a previsão da variabilidade térmica em componentes críticos dos sistemas eólicos.

5.1. Trabalhos futuros

Em seguida, são enumeradas algumas sugestões de trabalhos futuros relacionadas à esta dissertação:

1. Utilizar combinações de algoritmos de aprendizagem para aprimorar o modelo base, combinando suas melhores qualidades para formar um modelo mais robusto;

2. Utilizar uma entrada de dados SCADA em tempo real de um parque eólico, por meio de uma Interface de Programação de Aplicação, alimentando um banco de dados que posteriormente é utilizado para alimentar os dados de entrada do modelo de previsão;

3. Construir uma interface gráfica para visualização dos sinais de alerta emitidos pelo resultado da análise dos gráficos de controle em tempo real;

4. Adaptar a metodologia criada para identificação de anomalias em sistemas eólicos para sistemas fotovoltaicos, visando antecipar problemas na operação do parque.

REFERÊNCIAS

ALMEIDA, R. N. de.

O MÉTODO DOS MÍNIMOS QUADRADOS: ESTUDO E APLICAÇÕES PARA O ENSINO MÉDIO — Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, Brazil, Maio 2015.

ATKINS, P.; ATKINS, P. W.; PAULA, J. de. *Atkins' physical chemistry*. [S.l.]: Oxford university press, 2014.

BAYER, P.; DOLAN, L.; URPELAINEN, J. Global patterns of renewable energy innovation, 1990–2009. *Energy for Sustainable Development*, v. 17, n. 3, p. 288–295, 2013. ISSN 0973-0826. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0973082613000094>.

BILAL, B.; ADJALLAH, K. H.; SAVA, A. Data-driven fault detection and identification in wind turbines through performance assessment. In: *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. [S.l.: s.n.], 2019. v. 1, p. 123–129.

BINI, E.; GARAVINI, G.; ROMERO, F. *Oil shock: The 1973 crisis and its economic legacy*. [S.l.]: Bloomsbury Publishing, 2016.

BREIMAN, L. Bagging predictors. *Machine Learning*, v. 24, n. 2, p. 123–140, 1996.

CHAPMAN, S. J. *Fundamentos de máquinas elétricas*. [S.l.]: AMGH editora, 2013.

CHATTERJEE, J.; DETHLEFS, N. Deep learning with knowledge transfer for explainable anomaly prediction in wind turbines. *Wind Energy*, v. 23, n. 8, p. 1693–1710, 2020. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/we.2510>.

CHEN, T.; GUESTRIN, C. Xgboost: A scalable and accurate implementation of gradient boosting machines. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016. Disponível em: <https://dl.acm.org/doi/10.1145/2939672.2939785>.

CHEON, J. et al. Development of hardware-in-the-loop-simulation testbed for pitch control system performance test. *Energies*, v. 12, n. 10, 2019. ISSN 1996-1073. Disponível em: <https://www.mdpi.com/1996-1073/12/10/2031>.

COUNCIL, G. W. E. *Global Wind Report 2023*. Rue de Commerce 31, Brussels, Belgium, 2023.

ENERGIA, V. *Conversão de energia cinética em elétrica*. 2024. Acesso em: 28/01/2024. Disponível em: <https://voltaenergia.com.br/>.

(EPE), E. de P. E. *Empreendimentos Eólicos ao Fim da Vida Útil: Situação Atual e Alternativas Futuras*. Brasília, Distrito Federal, 2021.

(EPE), E. de P. E. *Estudos do Plano Decenal de Expansão de Energia 2030 - Parâmetros de Custos – Geração e Transmissão*. [S.l.], 2021.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

HAMILTON, J. D. Historical oil shocks. In: *Routledge handbook of major events in economic history*. [S.l.]: Routledge, 2013. p. 239–265.

HAYKIN, S. *Redes Neurais: Princípios e Prática*. Bookman Editora, 2001. ISBN 9788577800865. Disponível em: <https://books.google.com.br/books?id=bhMwDwAAQBAJ>.

HELLIWELL, J. F. et al. Social environments for world happiness. *World happiness report*, JSTOR, v. 2020, n. 1, p. 13–45, 2020.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, 11 1997. ISSN 0899-7667. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>.

HUTCHISON, M. M. Aggregate demand, uncertainty and oil prices: The 1990 oil shock in comparative perspective. 1991.

IEA, I. E. A. *Renewables 2022*. Paris, 2022. License: CC BY 4.0. Disponível em: <https://www.iea.org/reports/renewables-2022>.

IRENA, I. R. E. A. *Renewable power generation costs in 2017*. Abu Dhabi, 2018.

IRENA, I. R. E. A. *Renewable Power Generation Costs in 2022*. Abu Dhabi, 2023.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, American Association for the Advancement of Science, v. 349, n. 6245, p. 255–260, 2015.

KIRCH, W. Pearson's correlation coefficient. In: KIRCH, W. (Ed.). *Encyclopedia of Public Health*. Dordrecht: Springer Netherlands, 2008. p. 1090–1091. Disponível em: <https://doi.org/10.1007/978-1-4020-5614-7_2569>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

MAGNUS, G. Versuche über die spannkräfte des wasserdampfs. *Annalen der Physik*, Wiley Online Library, v. 137, n. 2, p. 225–247, 1844.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, Springer, v. 5, p. 115–133, 1943.

MCMILLAN, D. Condition monitoring benefit for onshore wind turbines: sensitivity to operational parameters. *IET Renewable Power Generation*, Institution of Engineering and Technology, v. 2, p. 60–72(12), March 2008. ISSN 1752-1416. Disponível em: <https://digital-library.theiet.org/content/journals/10.1049/iet-rpg_20070064>.

NG, A. Machine learning yearning. *URL: http://www. mlyearning. org/(96)*, v. 139, 2017.

NG, E. Y.-K.; LIM, J. T. Machine learning on fault diagnosis in wind turbines. *Fluids*, MDPI, v. 7, n. 12, p. 371, 2022.

NUNES, I. C.; CATALãO-LOPES, M. The impact of oil shocks on innovation for alternative sources of energy: Is there an asymmetric response when oil prices go up or down? *Journal of Commodity Markets*, v. 19, p. 100108, 2020. ISSN 2405-8513. Disponível em: <https://www.sciencedirect.com/science/article/pii/S240585131930073X>.

ONS. *Capacidade Instalada*. 2024. Acessado em 18/01/2024. Disponível em: <https://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/capacidade_instalada.aspx>.

ONS, O. N. do S. E. *Resultados da Operação - Dados Hidrológicos Volumes*. 2021. Acesso em: 09 de janeiro de 2024. Disponível em: <https://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/dados_hidrologicos_volumes.aspx>.

PICARD, A. et al. Revised formula for the density of moist air (cipm-2007). *Metrologia*, IOP Publishing, v. 45, n. 2, p. 149, 2008.

Pinar Pérez, J. M. et al. Wind turbine reliability analysis. *Renewable and Sustainable Energy Reviews*, v. 23, p. 463–472, 2013. ISSN 1364-0321. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1364032113001779>.

PINTO, M. de O. *Fundamentos de energia eólica*. Grupo Gen - LTC, 2013. ISBN 9788521621607. Disponível em: <https://books.google.com.br/books?id=xrK9NAEACAAJ>.

REINFORCEMENT Learning - MATLAB Simulink. 2021. <https://au.mathworks.com/discovery/reinforcement-learning.html>. Acesso em 23 de abril de 2023.

ROCCA, J. *Ensemble methods: bagging, boosting and stacking*. 2019. Publicado em Medium. Disponível em: <https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205>.

SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, IBM, v. 3, n. 3, p. 210–229, 1959.

SARSWATULA, S. A.; PUGH, T.; PRABHU, V. Modeling energy consumption using machine learning. *Frontiers in Manufacturing Technology*, v. 2, p. 12, 2022. Disponível em: <https://www.frontiersin.org/articles/10.3389/fmtec.2022.855208>.

SHRESTHA, H. B. *Wind Energy Physics and Resource Assessment with Python*. 2022. Accessed on: 28/01/2024. Disponível em: <https://towardsdatascience.com/wind-energy-physics-and-resource-assessment-with-python-789a0273e697>.

SOTO, T. Regression analysis. In: *Encyclopedia of Autism Spectrum Disorders*. New York, NY: Springer New York, 2013. p. 2538–2548. Disponível em: <https://doi.org/10.1007/978-1-4419-1698-3_320>.

TOLMASQUIM, M. *As origens da crise energética brasileira*. [S.l.]: SciELO Brasil, 2000.

TORRES, G. L. *Métodos Computacionais*. [S.l.]: Laboratório de Otimização Aplicada a Sistemas de Potência, Universidade Federal de Pernambuco, 2018.

TYME, J. *Understanding the Key Parts of a Wind Turbine*. 2023. Acessado em 15/02/2024. Disponível em: <https://titanww.com/understanding-the-key-parts-of-a-wind-turbine>.

UDO, W.; MUHAMMAD, Y. Data-driven predictive maintenance of wind turbine based on scada data. *IEEE Access*, IEEE, v. 9, p. 162370–162388, 2021.

XAVIER, L. *Dados SCADA G114-2.1 MW*. 2024. Repositório do GitHub. Disponível em: <https://github.com/knopfzangwer/dados_scadaG114-2.1-MW>.

ZAHER, A. et al. Online wind turbine fault detection through automated scada data analysis. *Wind Energy*, v. 12, n. 6, p. 574–593, 2009. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/we.319>.

ZIEGLER, L. et al. Lifetime extension of onshore wind turbines: A review covering germany, spain, denmark, and the uk. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 82, p. 1261–1271, 2018.