



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus Experimental de Rosana

MATHEUS PUSSAIGNOLLI DE PAULA

Aplicação de Técnicas de Aprendizado de Máquina e Decisões Orientadas a Dados no Setor Eólico Brasileiro

Rosana - SP
2022

Matheus Pussaignolli de Paula

**Aplicação de Técnicas de Aprendizado de Máquina e Decisões Orientadas a Dados
no Setor Eólico Brasileiro**

Trabalho de Conclusão de Curso apresentado à Coordenadoria de Curso de Engenharia de Energia do Campus Experimental de Rosana, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia de Energia.

Orientadores: Wallace Correa de Oliveira Casaca e Marilaine Colnago

Rosana - SP
2022

P324a

Paula, Matheus Pussaignolli de

Aplicação de técnicas de aprendizado de máquina e decisões orientadas a dados no setor eólico brasileiro / Matheus Pussaignolli de Paula. -- Rosana, 2022

58 f. : tabs., fotos

Trabalho de conclusão de curso (Bacharelado - Engenharia de Energia) - Universidade Estadual Paulista (Unesp), Faculdade de Engenharia e Ciências, Rosana

Orientador: Wallace Correa de Oliveira Casaca

Coorientadora: Marilaine Colnago

1. Aprendizado de Máquina. 2. Energia Eólica. 3. Energia. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Engenharia e Ciências, Rosana. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus Experimental de Rosana

MATHEUS PUSSAIGNOLLI DE PAULA

ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO
PARTE DO REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE
“GRADUADO EM ENGENHARIA DE ENERGIA”

APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE
GRADUAÇÃO EM ENGENHARIA DE ENERGIA

Prof. Dr. José Francisco Resende da Silva
Coordenador

BANCA EXAMINADORA:

Prof. Dr. Wallace Correa de Oliveira Casaca
Orientador/UNESP-Rosana

Prof. Dr. Kleber Rocha de Oliveira
UNESP-Rosana

Prof. Ms. Camila Faria Afonso Lages
Membro Externo

Janeiro 2022

*Dedico este trabalho,
de modo especial, à minha
família.*

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus, por me dar forças para finalizar o curso de Engenharia de Energia na UNESP. Agradeço à minha família por sempre me apoiar e dar todos os subsídios necessários para que eu pudesse finalizar o mesmo. Ao meu pai, que em todos os momentos de minha vida sempre apoiou minhas escolhas e incentivou que eu fizesse o meu melhor independente do lugar onde eu almejasse trilhar um caminho. A minha mãe, pela qual em todos os momentos ao decorrer da minha graduação sempre me apoiou quando preciso e trouxe diversos sermões, sendo o mais clássico: “Não fez mais do que a sua obrigação” (Claudia, 1998). A minha irmã, que sempre me apoiou e esteve ao meu lado em todo esse período além de minha cachorra Chivas.

Aos meus orientadores Prof. Dr. Wallace Correa de Oliveira Casaca e Prof. Dra. Marilaine Colnago pela orientação, apoio, amizade e todos os ensinamentos ao decorrer de toda minha graduação, os meus sinceros agradecimentos por toda confiança e oportunidades concedidas ao longo dos meus anos como graduando. Vocês são docentes exemplares que realmente possuem amor no que se diz à respeito ao ensino, pesquisa e extensão, além de serem verdadeiras inspirações como pessoas.

A minha querida amiga Dina de Lima, a qual vivenciei de perto diversos momentos de minha graduação e sempre esteve ao meu lado, me ajudando quando preciso: meus sinceros agradecimentos por esses 2 anos e 4 meses morando juntos, saiba que aprendi muito com você e espero que nossa amizade dure sempre.

A todos meus amigos que, ao longo desses 5 anos, compartilharam diversas experiências e ensinamentos em situações “desesperadoras”. Em especial, Arthur Pimentel, Lucas Gomes, Henrique Bergamo, Estevan Fauro, Dayara Pereira Basso, Thalita Botuem, Leonardo Fini, Felipe Augusto, Júlia Vilalva, João Pedro Jenson e João Leme. Além destes, agradeço a todos amigos que fiz durante o curso técnico de mecânica, aos quais em diversos momentos de minha graduação me ajudaram mesmo eu estando bem longe da capital, sendo eles: Matheus Rodrigues, Victor Hugo Freire, Henrique Albanez, Pablo Martins e Matheus Bazzana.

A todos os membros da empresa Júnior Fontes, uma corporação onde pude evoluir bastante como pessoa tanto em quesitos técnicos, bem como sociais de modo a colaborar diretamente com um Brasil mais empreendedor e sustentável.

A todas as amigas que fiz na cidade de Primavera, em especial, Christian Freire, Wesley Oliveira, João Victor, Pablo Arantes, Dener Oliveira e Vítinho, aos quais pude

passar diversos momentos memoráveis andando de skate que me serviram literalmente como uma terapia ao longo do período de graduação.

Agradeço por todas as plataformas digitais pelas quais compartilham conhecimento de maneira gratuita, tais como: canais do YouTube, Sci-Hub, Google Acadêmico, entre diversas outras que foram primordiais ao decorrer de meu período como graduando.

Agradeço a Logicalis Brasil e a Casa dos Ventos Energias Renováveis, por proporcionarem diversas experiências ao longo do estágio obrigatório onde pude utilizar diversos conceitos estudados durante a graduação para o meu desenvolvimento como profissional.

Agradeço a FAPESP (Proc. N° 2018/05341-4) pela bolsa de pesquisa concedida ao longo de três anos de minha graduação, além das duas bolsas aprovadas para avançar com minhas pesquisas em Portugal.

Por fim, agradeço a todos os funcionários e docentes da Faculdade de Engenharia e Ciências, UNESP Campus de Rosana por proporcionarem um ambiente único ao longo de minha graduação.

“Dream big, work hard, stay focused and surround yourself with good people.”

Desconhecido

RESUMO

Nos últimos anos tem-se observado uma forte ascensão das energias renováveis no cenário energético nacional. Dentre os principais sorvedouros desse tipo de energia, a energia eólica tem desempenhado um papel primordial, em especial, devido a sua intensa predominância na região Nordeste e Sul do Brasil. Entretanto, mesmo possuindo diversas vantagens, a energia eólica pode ser suscetível à altas variações de potência elétrica gerada, devido a características específicas climáticas e físicas como solo, temperatura, umidade do ar, entre outras. Assim, dentro desse contexto, este trabalho de conclusão de curso visa investigar a problemática de previsão de potência elétrica gerada no Parque Eólico de Praia Formosa a partir de dados disponibilizados pela Operador Nacional do Sistema Elétrico (ONS) e pelo Instituto Nacional de Meteorologia (INMET), com o intuito de colaborar na elaboração de planejamentos energéticos que sejam mais assertivos para o setor energético nacional. Nesse contexto, a fim de investigar tais conjuntos de dados para a realização da predição de potência elétrica, foram adotadas técnicas de Aprendizado de Máquina (AM), sendo elas: Máquinas Vetores de Suporte (*Support Vector Machines*), Aumento Extremo de Gradiente (*Extreme Gradient Boosting*), Aumento de Gradiente (*Gradient Boosting*), Redes Neurais Artificiais Multicamadas (*RNA Multilayer Perceptron*), e Árvores de Decisão via Floresta Aleatória (*Random Forest*).

PALAVRAS-CHAVE: Energia Eólica. Potência Elétrica. Aprendizado de Máquina.

ABSTRACT

In recent years, there has been a strong rise in renewable energies in the national energy scenario. Among the main types of energy generation, wind energy has played a key role, in particular, due to its intense predominance in the Northeast and South regions of Brazil. However, despite having several advantages, wind energy can also be susceptible to high variation of energy load due to specific climatic and physical features such as soil, temperature, air humidity, among and other factors. Therefore, this work aims at investigating the forecasting application of energy generation at Praia Formosa Wind Farm, based on available data taken from the National Electric System Operator (ONS) and National Institute of Meteorology (INMET), allowing for more assertive energy plans to support the national energy sector. In order to proceed with our study and analysis, the following Machine Learning (ML) techniques have been employed: Support Vector Machines, Extreme Gradient Boosting, Gradient Boosting, Artificial Neural Network Multilayer Perceptron and Random Forest.

KEYWORDS: Wind Energy. Energy Load. Machine Learning.

LISTA DE FIGURAS

Figura 1: Evolução da Capacidade Instalada da Potência Elétrica de Parques Eólicos Brasileiros em MW do período de 2005 até 2020.	15
Figura 2: Modelo de turbina horizontal upwind e downwind.	18
Figura 3: Regiões de Operação de uma Turbina Eólica.	19
Figura 4: Separação entre duas classes com o melhor hiperplano.	23
Figura 5: Representação de um neurônio artificial.....	26
Figura 6: Exemplo de árvore de regressão	27
Figura 7: Pipeline do trabalho.	29
Figura 8: Valores faltantes de cada variável em cada respectiva série. Na vertical: contagem dos dados totais e ruídos dos dados faltantes, e na horizontal variáveis da base de dados.	32
Figura 9: Mapa de Calor (grau de linearidade entre as variáveis).....	34
Figura 10: Grau de dependência entre a potência elétrica gerada e a velocidade do vento.	35
Figura 11: Distribuição da variável alvo.	36
Figura 12: Boxplot da variável alvo.	36
Figura 13: Boxplot variável alvo por mês.	37
Figura 14: Boxplot variável alvo por dia da semana.	37
Figura 15: Violin Plot da variável alvo por mês.	38
Figura 16: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) e sem ajuste dos mesmos (Previsto Default), além dos valores reais da série vs período (15 dias aleatórios) da série de teste estimada pelo modelo SVM.	41
Figura 17: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste. realizada pelo modelo aumento extremo de gradiente	43
Figura 18: Features mais importantes para a previsão do modelo Aumento Extremo de Gradiente	43
Figura 19: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo aumento de gradiente.	45
Figura 20: Ranking das features mais importantes para a previsão do modelo aumento de gradiente.....	46
Figura 21: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo redes neural MLP.....	47
Figura 22: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo árvores de decisão via floresta aleatória.	49
Figura 23: Ranking das features mais importantes para a previsão do modelo árvore de decisão via floresta aleatória.....	49

LISTA DE TABELAS

Tabela 1- Valores faltantes de cada variável na base de dados utilizada (quantidade e porcentagem).	31
Tabela 2 - Medidas de tendência central da base de dados.	33
Tabela 3 - VIF das variáveis preditoras.	40
Tabela 4 - Scores finais obtidos pelo modelo SVM (Default e Tuning).....	42
Tabela 5 - Scores finais obtidos pelo modelo aumento extremo de gradiente (Default e Tuning).	44
Tabela 6 - Scores finais obtidos pelo modelo aumento de gradiente (Default e Tuning).	46
Tabela 7 - Scores finais obtidos pelo modelo rede neural MLP (Default e Tuning).	48
Tabela 8 - Scores finais obtidos pelo modelo árvores de decisão via floresta aleatória (Default e Tuning).	50
Tabela 9 - Scores finais obtidos pelos modelos de aprendizado de máquina otimizados.	50

LISTA DE SÍMBOLOS

GW – Gigawatts

MW – Megawatts

m/s – Metros por segundo

ε – Erro permitido pela mensuração

Ω – Termo de regularização

AM – Aprendizado de Máquina

ONS – Operador Nacional do Sistema Elétrico

INMET – Instituto Nacional de Meteorologia

SIN – Sistema Interligado Nacional

°C – Graus celsius

mbar – Milibar

mm – Milímetros

MWmed – Potência elétrica média em megawatts

GB – Gigabyte

TB – Terabyte

VIF – Variance Inflation Factor (Fator de Inflação de Variância)

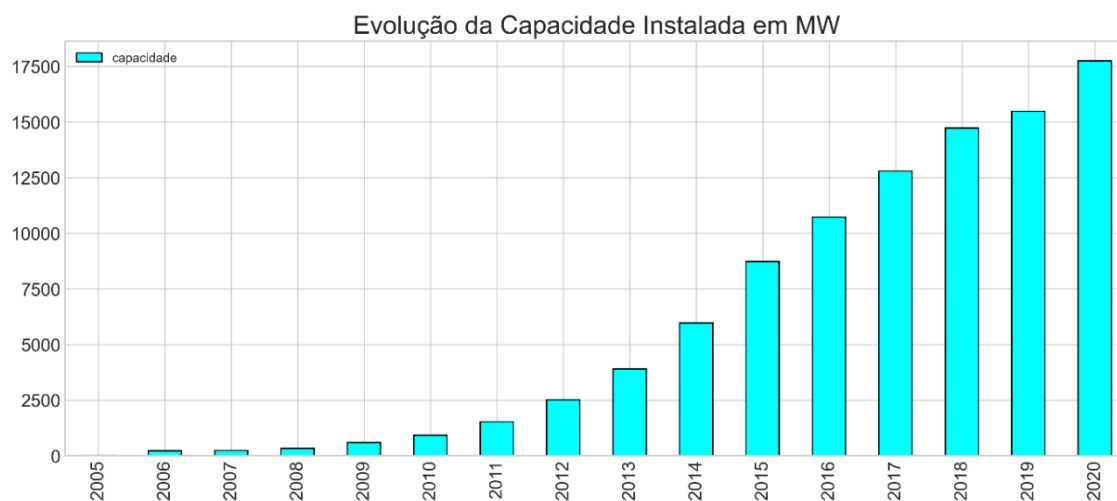
SUMÁRIO

1.	Introdução.....	15
2.	Objetivos	17
3.	Revisão Bibliográfica.....	18
3.1.	Principais Características das Turbinas Eólicas	18
3.2.	Aprendizado de Máquina no Cenário Eólico	19
4.	Materiais e Métodos.....	22
4.1.	Métodos de AM.....	22
4.1.1.	Máquinas Vetores de Suporte (<i>Support Vector Machines</i>)	22
4.1.2.	Aumento de Gradiente (<i>Gradient Boosting</i>).....	23
4.1.3.	Aumento Extremo de Gradiente (<i>Extreme Gradient Boosting</i>).....	24
4.1.4.	Redes Neurais Artificiais Multicamadas (<i>RNA Multi Layer Perceptron</i>).....	24
4.1.5.	Árvores de Decisão via Floresta Aleatória (<i>Random Forest</i>).....	26
4.2.	Base de Dados	27
4.3.	Recursos Computacionais	28
4.4.	Plataforma de Programação.....	28
4.5.	Material Bibliográfico	28
4.6.	Pipeline de Processamento dos Dados e Métricas de Avaliação	28
4.6.1.	Seleção de Hiperparâmetros dos Modelos.....	30
5.	Resultados e Discussões.....	31
5.1.	Pré-processamento e Análise Exploratória dos Dados	31
5.1.1.	Preparação dos Dados	31
5.1.2.	Análise Exploratória de Dados	32
5.1.3.	Criação de Novas Variáveis e Seleção dos Recursos.....	38
5.2.	Máquinas de Vetores de Suporte (<i>Support Vector Machines</i>)	40
5.3.	Aumento Extremo de Gradiente (<i>Extreme Gradient Boosting</i>)	42
5.4.	Aumento de Gradiente (<i>Gradient Boosting</i>)	44
5.5.	Redes Neurais Artificiais Multicamadas (<i>RNA Multi Layer Perceptron</i>)	46
5.6.	Árvores de Decisão via Floresta Aleatória (<i>Random Forest</i>).....	48
5.7.	Comparação de Desempenho dos Modelos.....	50
6.	Conclusão.....	52
7.	Referências.....	54

1. Introdução

Atualmente, a geração eólica é uma das formas de geração de energia elétrica de maior ascensão no mundo. No ano de 2020, no mundo, foi registrado um crescimento da potência instalada dessa forma de geração de 53% em relação ao ano anterior, o que mostra uma forte tendência dessa fonte híbrida, mesmo no período pandêmico da Covid-19 (GLOBAL WIND REPORT, 2021). No Brasil, foi alcançada, na metade de 2021, uma potência instalada de 19GW, o que representa cerca de 10% da matriz elétrica nacional, sendo a segunda maior fonte renovável presente no país, perdendo apenas para a fonte de origem hidrelétrica (CANAL ENERGIA, 2021). A Figura 1 mostra a evolução da capacidade instalada (MW) dessa forma de energia no Brasil ao longo dos últimos anos.

Figura 1: Evolução da Capacidade Instalada da Potência Elétrica de Parques Eólicos Brasileiros em MW do período de 2005 até 2020.



Fonte: Adaptado de (CANAL ENERGIA, 2021).

Ao contrário de outras fontes de energia como o petróleo e, até mesmo, a geração hídrica, o vento origina-se continuamente do ar em circulação e, em razão disso, pode ser considerada uma fonte “inesgotável” de energia. Além disso, os efeitos da instalação de uma usina de energia eólica não afetam, de forma drástica, a fauna e a flora das regiões locais, o que faz com que haja a geração de energia de forma limpa e, também, viável economicamente (AMBIENTE ENERGIA, 2019). Os custos de instalação e manutenção desse tipo de geração energética vem diminuindo ao longo dos anos, principalmente, devido ao surgimento de novas tecnologias no processo de produção e ao custo associado

de seus maquinários de operação, fatores que contribuíram para a diminuição do custo total de produção desse tipo de energia (AMBIENTE ENERGIA, 2019).

Em contrapartida, a energia elétrica gerada por cada parque ou conjunto eólico, pode vir a sofrer altas variações, em razão do vento não apresentar um comportamento regular ao longo do tempo, o que faz com que as provedoras de energia possuam intensa dificuldade em dimensionar a exatidão de suas produções em meio à demanda energética de cada local consumidor. Outro fator determinístico nesse sentido diz respeito ao modelo das turbinas utilizadas pelos parques eólicos, que podem possuir aerogeradores com velocidade variáveis, dificultando, desta forma, a geração de energia elétrica de forma constante (TARNOWSKI et al., 2006.)

Assim, dado o contexto apresentado, o presente trabalho tem como principal objetivo o de realizar a predição da potência elétrica em um complexo eólico, em um curto prazo de tempo, considerando a intensa volatilidade da energia eólica na localidade, colaborando assim para melhorar a tomada de decisões no setor.

2. Objetivos

O presente Trabalho de Conclusão de Curso (TCC) tem como finalidade explorar metodologias baseadas em Aprendizado de Máquina (AM) dado o contexto de previsão da potência elétrica no Parque Eólico de Praia Formosa (Ceará). As técnicas de AM foram estudadas e implementadas computacionalmente, além de aplicadas para se estimar a quantidade de potência elétrica conforme o objetivo proposto.

Na intenção de adquirir maior embasamento teórico da problemática em questão, primeiramente foi necessário estudar o tema previsão de potência elétrica aplicado à energia eólica, para em seguida explorar os dados do parque eólico proposto. Ademais, foi realizada uma análise exploratória dos dados obtidos para, posteriormente, aplicar as metodologias de AM propostas, utilizando diferentes técnicas estatísticas para um estudo mais aprofundado perante os dados disponibilizados pelo ONS e INMET.

Assim, podemos destacar que o trabalho tem como objetivo principal estimar a potência elétrica média produzida por um parque eólico em uma dada janela futura de tempo, a partir de técnicas de Inteligência Artificial, com o intuito de auxiliar a tomada de decisões no que se diz respeito ao despacho de energia realizado por parte do complexo eólico para o Sistema Interligado Nacional (SIN).

3. Revisão Bibliográfica

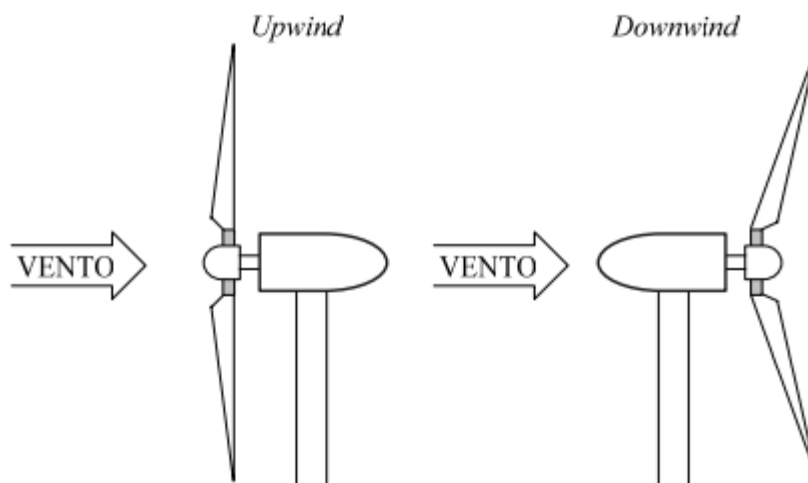
3.1.Principais Características das Turbinas Eólicas

As turbinas eólicas modernas possuem diferentes características, fator que torna necessário ter-se um conhecimento prévio sobre as mesmas para, assim, realizar análises mais detalhadas e entender qual sua influência no contexto de predição.

Em relação à posição do eixo de rotação, as turbinas eólicas podem ser categorizadas em dois subgrupos: turbinas de eixo vertical e turbinas de eixo horizontal (POPE et al., 2010). As mais comuns encontradas em parques eólicos brasileiros são as horizontais, já que elas são conectadas diretamente à rede elétrica para que ocorra a distribuição de potência elétrica.

As turbinas de eixo horizontais são classificadas de acordo com a posição do rotor em relação à torre, podendo ser: *upwind* ou *downwind* (WANG et al., 2018). Nas turbinas *upwind* (ver Figura 2), o vento incide na área de varredura do rotor, isto é, pela frente da turbina, sendo que a principal vantagem desse feito é a possibilidade de se evitar distúrbios causados pela torre no vento, motivo pelo qual esse modelo é o mais utilizado no mercado atual. Já em relação as turbinas *downwind*, o vento incide na área de varredura por trás da turbina eólica, o que pode causar mais flexibilidade no rotor quando ocorrem fortes rajadas de vento (WANG et al., 2018).

Figura 2: Modelo de turbina horizontal *upwind* e *downwind*.



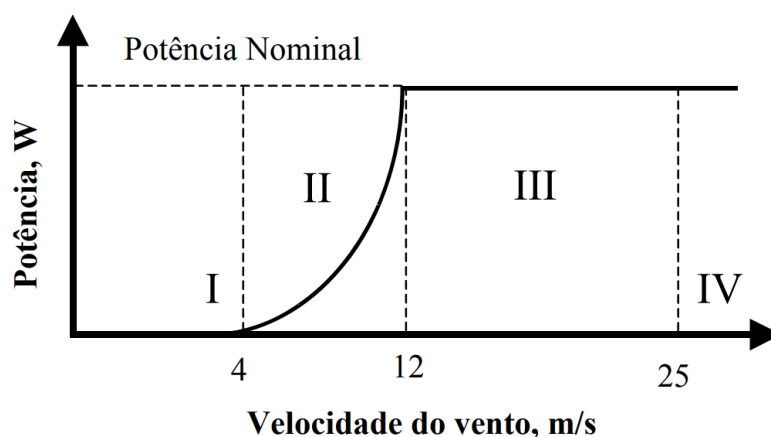
Fonte: MARQUES et al., 2004.

Em relação ao número de pás eólicas, vale ressaltar que esse fator é determinado dependendo da aplicação eólica para qual será realizada, entretanto vale dizer que quase

sempre são utilizadas três pás, o que garante diversos benefícios, tais como: melhor distribuição dos pesos na área de varredura do rotor, maior estabilidade para a torre, redução das cargas mecânicas para os distintos componentes da turbina, além de colaborar diretamente com um melhor funcionamento do sistema no geral (PAO et al., 2009).

É importante relatar sobre os modos de operação de uma turbina eólica podem ser divididos em: velocidade fixa e velocidade variável, dependendo exclusivamente das limitações de potência do sistema. Normalmente, a potência máxima gerada se limita ao intervalo de 12 até 25 m/s de velocidade do vento, e a mínima para a geração no intervalo de 3 a 5 m/s. Assim, como podemos observar na Figura 3, pode-se afirmar que no intervalo de 4 m/s até 12 m/s, as turbinas podem operar tanto em velocidade constante como também variável, dependendo do sistema utilizado (MARQUES et al., 2004).

Figura 3: Regiões de Operação de uma Turbina Eólica.



Fonte: MARQUES et al., 2004.

Para esse trabalho, foram exemplificados e explicados os resultados obtidos na Análise Exploratória de Dados (AED) levando em conta tais aspectos apresentados e considerando que as turbinas utilizadas pelos parques eólicos são do tipo horizontais, com três pás, *upwinds*, e levando em consideração as regiões de operações mais comuns das turbinas.

3.2. Aprendizado de Máquina no Cenário Eólico

Dentro do cenário de predição de potência em parques eólicos, metodologias relacionadas ao tema Aprendizado de Máquina (AM) vêm sendo cada vez mais usadas como forma de auxílio na tomada de decisões e no suporte à elaboração de planejamentos energéticos que sejam mais otimizados e seguros. A intuição por trás do AM é realizar a

construção de sistemas que, de forma automatizada (ou ainda, semiautomatizada), acabam “adquirindo conhecimento”, isto é, venham a “aprender” a partir dos dados disponíveis e da interação entre eles. Assim, o sistema computacional originado (Modelo de AM) toma as melhores decisões a partir dos dados de entrada, ou ainda, a partir de problemas e experiências já avaliadas anteriormente (MONARD et al., 2003).

Embora as técnicas existentes na literatura apresentem resultados bastante satisfatórios quando aplicadas aos mais variados tipos de problemas, vale mencionar que não existe um algoritmo “ótimo” (ideal) para solucionar todos os problemas de predição existentes. Portanto, é necessário entender as limitações iniciais relacionadas a cada procedimento de aprendizado empregado nos problemas, uma vez que eles podem ser aprimorados de forma a aumentar gradativamente seus respectivos controles e gerais de desempenho.

Na literatura, são apresentadas diferentes propostas de modelos de AM ajustados para atuar sob diversos segmentos, incluindo a área de predição de potência elétrica em conjuntos eólicos. Um dos trabalhos nessa direção é apresentado em (ZHENG et al., 2019), onde os autores realizaram uma etapa de Engenharia de Seleção de Recursos utilizando a técnica de agrupamento de dados *K-means* para, posteriormente, testar o modelo de predição *XGBoosting* em um horizonte de curto prazo para a previsão de potência, utilizando diversas variáveis, tais como: temperatura, umidade do ar, precipitação e diversas outras. Já no trabalho de (DEMOLLI et al., 2019), os autores utilizam diversas técnicas de aprendizado de máquina, sendo elas: *XGBoosting*, *Random Forest*, SVM e uma regressão linear adaptada (*Lasso Regression*) para realizar a previsão de potência diária utilizando dados de velocidade do vento.

Outro trabalho interessante é o descrito por (OPTIS et al., 2019), em que os autores utilizaram diversas técnicas de AM para realizar a previsão de potência elétrica, dentre elas: Rede Neural Multi Layer Perceptron, Gradient Boosting, e Aumento Extremo de Randomização de Árvores (*Extremely Randomized Trees*), além das Máquinas de Vetores de Suporte (SVM) utilizando a biblioteca *scikit-learn*, da linguagem Python. No trabalho de (LI et al., 2020), os autores utilizaram as Máquinas Vetores de Suporte (SVM) e um algoritmo chamado *DragonFly* para realizar a previsão de potência elétrica em um período de curto prazo.

Ainda dentro do contexto de predição de potência elétrica em parques eólicos, no trabalho de (KARINIOTAKIS et al., 1996), os autores utilizaram uma arquitetura de Rede Neural Recorrente para realizar a predição de potência elétrica em um parque eólico a

partir de técnicas avançadas de Estatística. Outro trabalho que apresenta técnicas avançadas nessa mesma linha para a predição de potência elétrica pode ser visto em (ZAMMER et al., 2015), cujo estudo teve como objetivo realizar previsões de curto prazo utilizando uma modelagem híbrida de Redes Neurais Artificiais.

Diante do exposto, podemos afirmar que são muitas as técnicas e metodologias específicas de AM que foram propostas e que podem ter sucesso no estudo aplicado na predição de potência elétrica em parques eólicos. Assim, é válido afirmar que no cenário energético em geral, o uso de técnicas de AM é de suma importância, tendo em vista que tais técnicas acabam trazendo maior confiabilidade, segurança e precisão em diversas aplicações do setor elétrico, tais como: determinação do preço da energia, despacho de energia no SIN, previsão das possíveis perdas de transmissão, etc.

4. Materiais e Métodos

Nessa seção do trabalho, são apresentados os materiais e métodos utilizados e todo o *pipeline* computacional adotado na condução de nossa pesquisa.

4.1. Métodos de AM

4.1.1. Máquinas Vetores de Suporte (*Support Vector Machines*)

As técnicas baseadas nas Máquinas de Vetores de Suporte (ou *Support Vector Machines* – SVMs) constituem um novo paradigma na área de Aprendizado de Máquina, sendo conhecidas, principalmente, por possuírem grande respaldo em um campo chamado Aprendizado Estatístico. Essa teoria (de Aprendizado Estatístico) tem como premissa resolver problemas quando dispomos de uma quantidade de dados reduzida ou ainda nula, circunstância encontrada em várias aplicações atualmente (VELASCO et al., 2018; WERON et al., 2014).

Tais técnicas induzem à criação de superfícies para segregar instâncias de classes distintas. Assim, é gerado um hiperplano “ótimo” em um espaço de alta dimensão de dados, que maximiza a chamada “margem” nesse espaço, que é a distância entre os vetores suporte de classes distintas. Esses vetores recebem essa denominação em virtude da sua proximidade com a superfície de corte, contribuindo de maneira decisiva para a definição de tal superfície. Além disso, podemos afirmar que, para criar uma superfície de decisão em problemas que não podem ser separados linearmente, o SVM dispõe do uso de funções do tipo *kernel*, para que seja possível realizar transformações lineares no vetor de atributos a fim de criar dimensões maiores nas quais as classes podem ser separadas por meio de um hiperplano \vec{w} (IZMAILOV et al., 2013; MARMARAS et al., 2017).

Para problemas relacionados a regressão, como o apresentado neste trabalho, podemos utilizar o SVM a partir do uso de uma função de perda, a qual é minimizada com um *regularizador*. Assim, frequentemente, o método SVM é denominado de SVR (*Support Vector Regression*) para a tarefa de previsão em séries temporais. Dessa forma, adaptando para o contexto do presente TCC, temos como objetivo: encontrar uma função não linear f que minimize, ao máximo, o erro da previsão, quando estamos ao longo do processo de treinamento do algoritmo, segundo a equação:

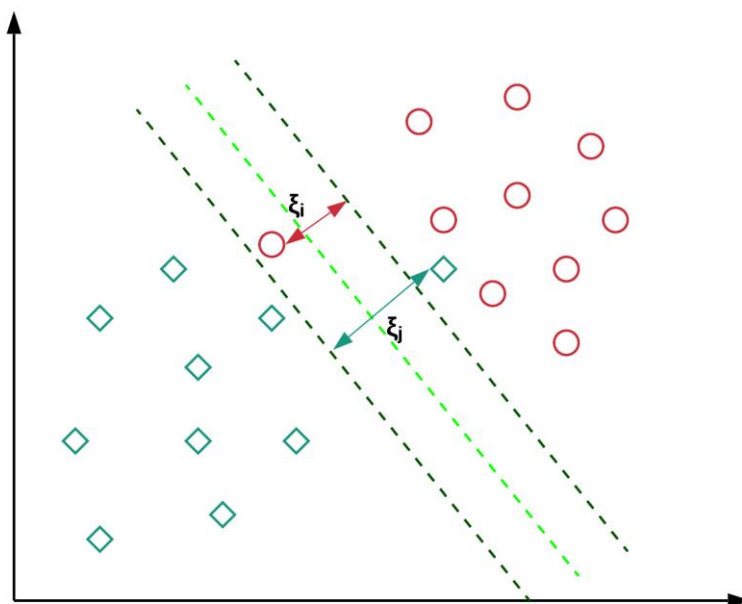
$$f(X_i) = \max_i(0, |f(X_i) - y_i|) - \varepsilon \quad (1)$$

onde ε representa o erro permitido pela mensuração (QIU et al., 2017).

A Figura 4 demonstra um exemplo de separação do hiperplano, considerando duas

classes distintas, além do erro de mensuração que pode vir a ocorrer na subdivisão.

Figura 4: Separação entre duas classes com o melhor hiperplano.



Fonte: Misra et al., 2019.

4.1.2 Aumento de Gradiente (*Gradient Boosting*)

Aumento de Gradiente (*Gradient Boosting*), GTB, é um método baseado em árvores de regressão. Tal metodologia emprega a técnica de *boosting*, a qual consiste em treinar diversos submodelos, em distintas amostras aleatórias dos dados de treino, combinando-as com o intuito de se obter um melhor desempenho e com menor de sobreajuste (*overfitting*¹) (FRIEDMAN et al., 2002; NATEKIN et al., 2013). Esse algoritmo é amplamente utilizado em problemas de classificação e regressão, e vem apresentando resultados bastante satisfatórios em problemas relacionados à previsão de geração de energias renováveis, como pode ser visto em (ANDRADE et al., 2017), onde foram realizadas previsões numéricas aplicadas no contexto de energia eólica.

O Aumento de Gradiente cria basicamente uma árvore de regressão simples e, após isso, utiliza uma variação do gradiente descendente para ir otimizando as árvores nas interações a partir de uma função de custo. Assim, a decisão tomada pelo GTB é dada a partir da soma das estimativas da árvore (LI et al., 2016).

¹ *Overfitting*: é um termo usado em estatística para descrever quando um modelo estatístico se ajusta de forma absoluta ao conjunto de dados anteriormente observado, mas se mostra ineficaz para prever novos resultados.

4.1.3 Aumento Extremo de Gradiente (*Extreme Gradient Boosting*)

O Aumento Extremo de Gradiente (*XGBoosting*), XGB, é um algoritmo de aprendizado de máquina baseado em árvores decisórias, que tem como objetivo modificar o aumento de gradiente de forma que ele funcione com qualquer função de perda diferenciável (MANDOT et al., 2019). Assim como o modelo *Gradient Boosting* (GB), no *XGBoosting* são aplicados os chamados “*weak learners*” (modelos fracos) por intermédio da estruturação do gradiente descendente, porém, ele potencializa propriedades básicas do GB, através da otimização dos sistemas e aprimoramentos dos algoritmos utilizados no processo de aprendizado (GOMES et al., 2020). Dessa forma, as árvores acabam não possuindo os mesmos pesos em suas configurações, sendo possível diminuir o *overfitting* por parte do modelo.

Assim, para que o modelo possa alcançar resultados mais acurados, temos como objetivo diminuir ao máximo a função de perda (*loss function*) do modelo XGB, dada pela equação a seguir (AGUIAR et al., 2020):

$$L_t = \sum_{i=1}^n l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t), \quad (2)$$

em que y_i é a variável target, \hat{y}_i^{t-1} é a predição no instante $t - 1$ somada à predição de um novo modelo $f_t(x_i)$ e $\Omega(f_t)$ é o termo referente a regularização utilizada para controlar a complexidade do modelo.

Assim, podemos utilizar os termos referentes a regularização como o *reg_lambda* e *reg_alpha*, para atribuir mais ou menos peso para as folhas, além de realizar a remoção, ou “*splits*”, de árvores com pouca importância ao longo da etapa de treinamento do algoritmo, com o objetivo de obter melhores scores (AGUIAR et al., 2020). Dessa forma, podemos afirmar que o princípio básico de funcionamento do algoritmo consiste nos seguintes fatores principais:

- Uma função de custo L_t a ser otimizada (*loss function*).
- Modelos preditivos baseados em árvores de decisão.
- Modelos adicionais de “*weak learners*” para minimizar a função de custo.
- Termos de Regularização.

4.1.4 Redes Neurais Artificiais Multicamadas (*RNA Multi Layer Perceptron*)

As Redes Neurais Artificiais (RNA’s) são metodologias computacionais idealizadas com o intuito de representar um paradigma matemático inspirado em uma estrutura neural

inteligente. Realizando uma analogia com a neurologia biológica, podemos comparar o cérebro humano com um sistema computacional, e os neurônios artificiais como sendo uma simplificação dos neurônios reais (HAYKIN et al., 2007).

A partir de um intensivo processo de treinamento, o modelo treinado passa a adquirir “conhecimento” por meio da experiência contínua. Com isso, é válido salientar a extrema importância desse tipo de técnica, principalmente quando não é possível determinar um modelo mais simplificado do problema. De uma forma mais geral, esse tipo de artifício de aprendizado é empregado quando queremos realizar uma previsão a partir de um problema de difícil modelagem ou, ainda, para analisar padrões complexos de serem identificados na prática (CASTRO et al., 2001).

Com uma separabilidade de dados bastante customizável e acurada, as RNA's passaram a armazenar tal “conhecimento adquirido” nos chamados pesos da rede, os quais são encarregados de realizar as conexões entre os neurônios. Desse modo, o processo de aprendizagem irá relacionar um conjunto apropriado de pesos de forma que a rede se comporte de maneira personalizada. Entretanto, existem algumas redes que podem possuir aspectos mais incompreensíveis devido à acentuada complexidade das relações entre seus pesos (CASTRO et al., 2001; RAUBER et al., 2005).

Para que se possa projetar uma RNA de maneira apropriada, devemos seguir os seguintes passos (HAYKIN et al., 2007):

- Definição dos neurônios artificiais que serão utilizados, isto é, o gerador dos dados coletados.
- Escolha de um padrão para a rede que será criada.
- Utilização de um algoritmo de aprendizagem a partir dos parâmetros estabelecidos.

Uma RNA, a partir da execução das etapas básicas acima, pode alcançar a funcionalidade pretendida por meio de estratégias que irão atuar em constante sinergia com os padrões identificados na rede neural. Vale ressaltar que cada rede em particular irá representar uma arquitetura específica, onde cada configuração irá representar funcionalidades específicas inerentes aos dados explorados do problema (CASTRO et al., 2001).

A partir do chamado “*threshold*” (limiar), o neurônio passa a lançar uma série de “impulsos nervosos” que, quando acumulados pelo sistema, acaba por representar uma saída que pode ser um impulso ou não. Para uma melhor visualização sobre o que

acontece em uma RNA, podemos definir a rede a partir de combinações lineares que representam o somatório de todas as entradas multiplicadas pelos seus respectivos pesos (FILHO et al., 2018).

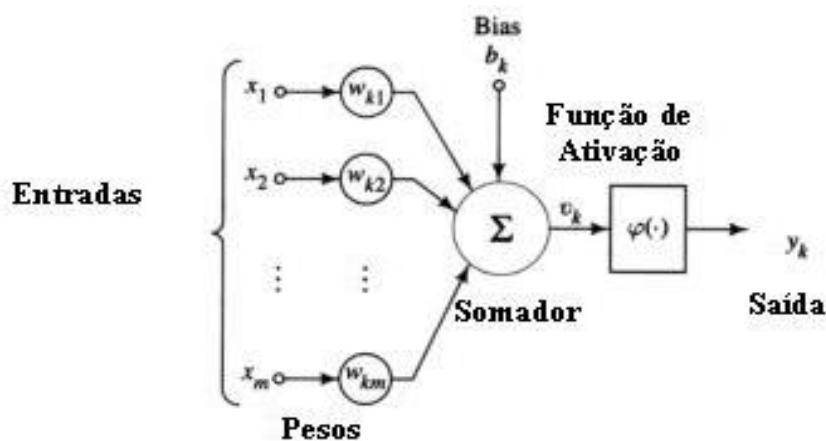
$$net_i(t) = \sum_{j=1}^n x_j w_{ij}(t) \geq \theta \quad (3)$$

Na equação acima, n é o número de entradas do neurônio net_i , w_{ij} são os pesos associados as entradas x_j 's, e θ é um valor de limiarização (de *thresholding*).

É importante salientar sobre o papel das funções de ativação $\varphi(\cdot)$, as quais fazem com que o estado atual de algum neurônio analisado seja adaptado a partir do ‘net’ da entrada. Outro fator que pode ocorrer é o uso de funções constantes que acabam não modificando comportamentos anteriores da rede, sendo estes denominados de “neurônios estáticos”. Por essa razão, os modelos de saídas mais comuns utilizados são as funções lineares, logísticas e hiperbólicas (VELLASCO et al., 2007).

A Figura 5 demonstra, de maneira simplificada, uma rede neural artificial, em que são ilustradas as camadas de entrada (x_1, x_2, x_n), os seus respectivos pesos ($w_{k1} w_{k2} w_{km}$), o “somador” de tais entradas, a função de ativação (φ) e a saída (y_k).

Figura 5: Representação de um neurônio artificial



Fonte: ZUBEN et al., 2010.

4.1.5 Árvores de Decisão via Floresta Aleatória (*Random Forest*)

A floresta aleatória (*Random Forest*), RF, é um método pelo qual realiza-se a previsão a partir de um conjunto de estimadores que induzem à estratégias específicas de aprendizagem, onde as mesmas são feitas com base em árvores de regressão ou

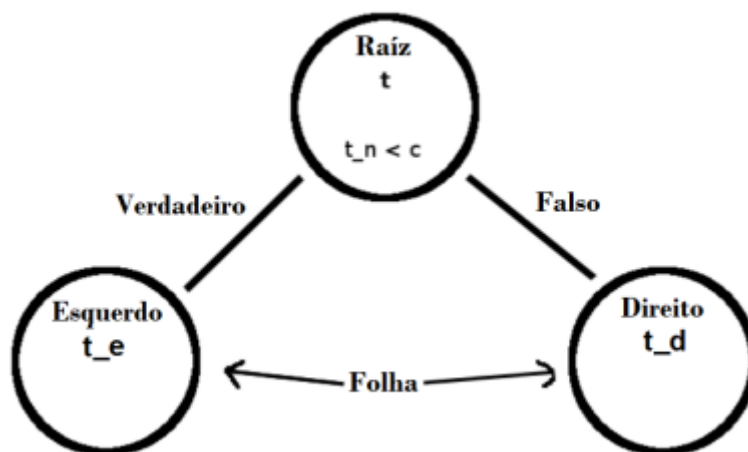
classificação (BREIMAN et al., 2001). Essa metodologia tem como princípio utilizar um subconjunto aleatório no decorrer da seleção dos nós da árvore. Dessa forma, podemos afirmar que a randomização pode fazer com que seja possível levar em consideração as variáveis que mais influenciam no modelo independente da correlação e, assim, melhorar o desempenho do algoritmo (SVETNIK et al., 2003).

De uma forma geral, o método do *Random Forest* consiste em três etapas:

1. Gerar um número N de conjuntos da amostra *bootstrap* (base de treinamento).
2. Para cada amostra, gerar uma árvore de regressão (sem ajuste) com a seguinte modificação: em cada nó, deve-se gerar uma amostra aleatória P das variáveis de entrada do dataset de treino, e escolher a melhor divisão das mesmas, onde $P < F$, sendo F o número de variáveis (*features*) do dataset.
3. Prever a nova saída, a partir do cálculo da média das saídas de M árvores de regressão quando novas variáveis são inseridas no modelo.

A Figura 6 apresenta um exemplo de uma árvore de regressão, em que é possível observar o nó raiz, o qual contém os dados de treinamento t , além da regra estabelecida na subdivisão $t_n < c$. Os nós resultantes da subdivisão (folhas) são tais que t_e e t_d que constituem as amostras de treinamento para os próximos níveis da árvore.

Figura 6: Exemplo de árvore de regressão



Fonte: Adaptado de (LIER et al., 2015).

4.2 Base de Dados

A base de dados utilizada no presente TCC foi construída a partir de dados disponíveis pelo ONS (Operador Nacional do Sistema Elétrico) e pelo INMET (Instituto Nacional de Meteorologia), e relaciona os dados referentes ao Parque Eólico de Praia

Formosa (Ceará). Tais conjuntos de dados correspondem às coletas médias diárias do período de 01/01/2016 – 31/12/2020. As variáveis disponíveis, no formato de séries temporais, são:

- Temperatura: dada em °C.
- Umidade relativa do ar: dada em %.
- Velocidade do vento: dada em m/s.
- Pressão: dada em mbar.
- Precipitação: dada em mm.
- Potência elétrica: dada em MWmed.
- Período: referente ao intervalo de 01/01/2016 – 31/12/2020.

4.3 Recursos Computacionais

Foi utilizado o notebook pessoal do discente, cuja configuração é um *Dell Inspiron 15 (5000) – Core i5 (7ª Geração)*, rodando em sistema operacional Windows 10 com memória de 8 GB e disco rígido de 1 TB.

4.4 Plataforma de Programação

A plataforma de programação utilizada foi o *Jupyter Notebook (Python 3.8)*, o qual possui diversas *bibliotecas* disponíveis para a limpeza, análise dos dados descritos acima, bem como ferramentas básicas para a implementação dos protótipos de AM estudados.

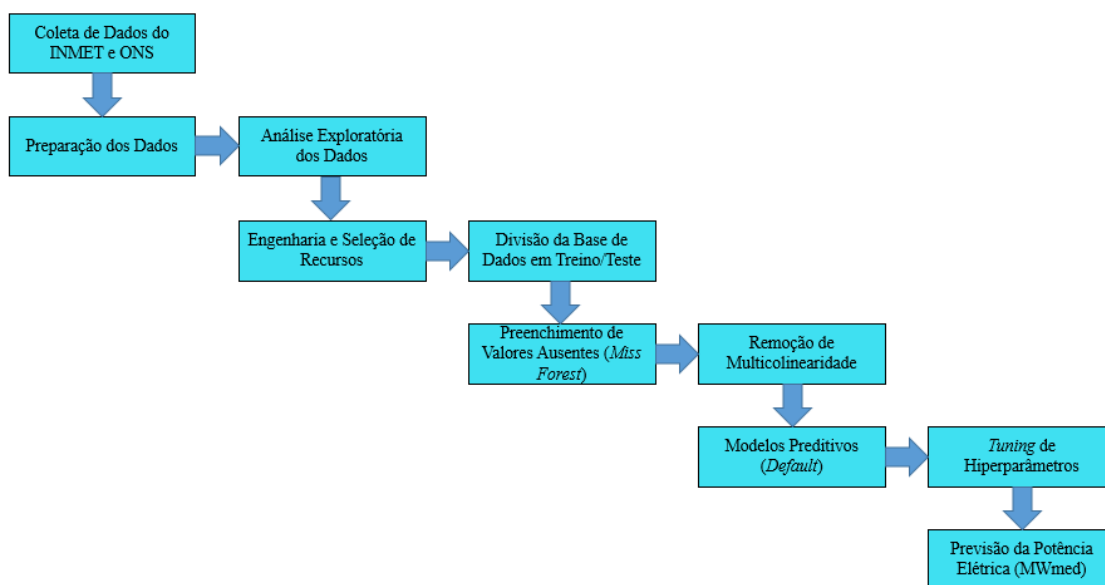
4.5 Material Bibliográfico

Foram realizados diversos acessos à periódicos e publicações da área, mediante ao vínculo com o portal CAPES, além de material disponível em repositórios de instituições de ensino brasileiras e do exterior disponibilizados pela UNESP.

4.6 Pipeline de Processamento dos Dados e Métricas de Avaliação

De forma a contextualizar o trabalho realizado, foi realizado um *pipeline* para conduzir as análises e testes realizados, apresentado na Figura 7.

Figura 7: Pipeline do trabalho.



Fonte: Próprio Autor.

Já em relação às métricas empregadas para avaliar os desempenhos dos modelos construídos, foram utilizadas: MSE (*Mean Squared Error*) e o MAPE (*Mean Absolute Percentage Error*) (PAULA et al., 2020). Tais métricas foram escolhidas por já serem amplamente empregadas em problemas relacionados a regressão, tais como podemos ver em (LEME et al., 2020), (MYTTENAERE et al., 2016) e (RAMOS et al., 2022).

O MSE é utilizado para realizar a comparação entre dois estimadores, principalmente para verificar se algum dos mesmos está viciado. Sendo assim, quanto menor a variância, mais eficaz é o modelo preditivo. Sua fórmula pode ser descrita como:

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2, \quad (4)$$

em que n é o número de amostras, Y_i valores reais e \hat{Y}_i valores preditos.

Outra métrica bastante interessante no contexto AM é o MAPE, pela qual é possível computar o erro do modelo preditivo na forma de porcentagem, sendo calculado a partir da média dos erros percentuais, como pode ser visto na equação:

$$\frac{1}{n} \sum_{i=0}^n \frac{Y_i - \hat{Y}_i}{Y_i}. \quad (5)$$

Assim como na Equação (4), os parâmetros da expressão acima são: n (amostras), Y_i valores reais e \hat{Y}_i valores preditos.

4.6.1 Seleção de Hiperparâmetros dos Modelos

Nos algoritmos que foram testados em no trabalho, existem diversos hiperparâmetros que podem ser ajustados a fim de se obter uma melhor aderência aos dados. Na maioria dos casos, esses hiperparâmetros são determinados a partir de uma série de ajustes manuais por parte do analista, a fim de se obter o melhor score possível. Entretanto, existe uma técnica que vem sendo bastante utilizada a fim de maximizar a eficácias dos modelos, o chamado “*Random Search*”.

Tal técnica basicamente funciona explorando o aspecto de aleatoriedade, de um número limitado de possibilidades, para a construção de *pipelines* de árvores de regressão. Dessa forma, estima-se um número k de interações a partir de alguns parâmetros adicionados, e o *Random Search* retorna o melhor *pipeline* para a amostragem testada (BERGSTRA et al., 2012).

5 Resultados e Discussões

Nesse capítulo, são apresentados os resultados obtidos a partir do uso das metodologias aplicadas, isto é: Máquinas Vetores de Suporte (*Support Vector Machines*), Aumento Extremo de Gradiente (*Extreme Gradient Boosting*), Aumento de Gradiente (*Gradient Boosting*), Redes Neurais do tipo MLP (*Multi Layer Perceptron*), e Árvores de Decisão via Floresta Aleatória (*Random Forest*). Além disso, são apresentadas comparações entre os *scores* obtidos pelos modelos a partir das métricas descritas na Seção 4.6.

Para todos os modelos preditivos aplicados, foram adotados 80% dos dados para o processo de treinamento e 20% para teste, em formato aleatório de amostragem, sendo essa uma das técnicas mais clássicas da literatura chamada *Hold-Out* (KIM, et al., 2009).

5.1 Pré-processamento e Análise Exploratória dos Dados

5.1.1 Preparação dos Dados

Para que se possa aplicar os modelos preditivos, a primeira etapa a ser realizada é a “limpeza” dos dados, onde são utilizadas técnicas para o tratamento de valores inconsistentes ou nulos das bases. A partir da Tabela 1, é possível observar os valores que deixaram de ser tabulados (nulos) durante a coleta dos dados na base original.

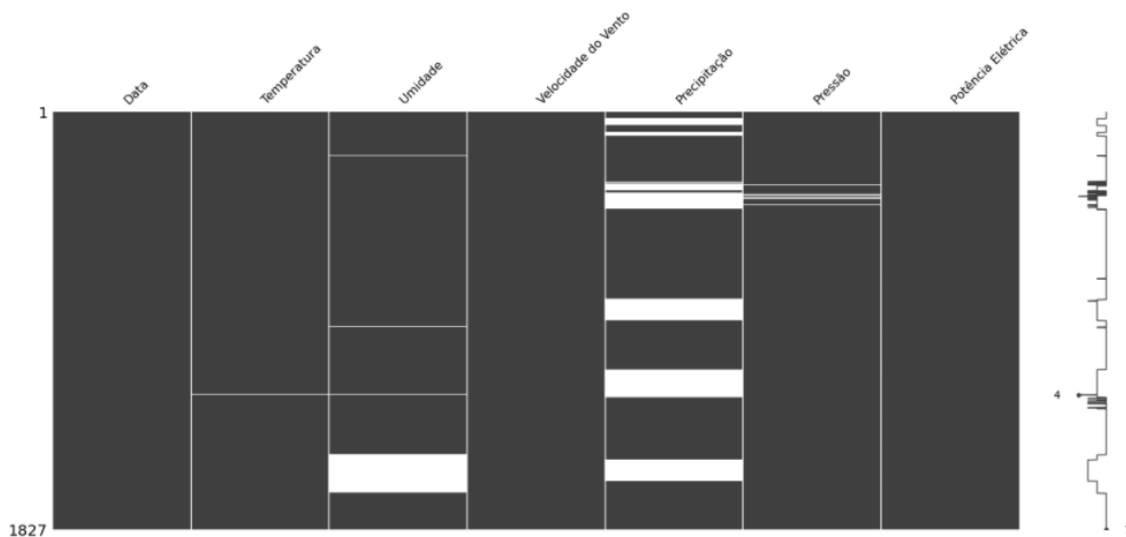
Tabela 1- Valores faltantes de cada variável na base de dados utilizada (quantidade e porcentagem).

<i>Feature</i>	Quantidade	Porcentagem (%)
Data	0	0
Umidade Relativa do Ar	180	9,85
Velocidade do Vento	0	0
Precipitação	461	25,23
Pressão	20	1,09
Potência Elétrica	0	0

Fonte: Próprio Autor.

Na Figura 8, é possível observar onde estão situados os valores faltantes ao longo de cada série da respectiva variável (valores em branco). Fica evidente que a variável de precipitação apresentou um conjunto significativo de valores faltantes ao longo do período analisado. Além disso, é possível observar que umidade relativa acabou apresentando mais valores ausentes justamente no período final da série.

Figura 8: Valores faltantes de cada variável em cada respectiva série. Na vertical: contagem dos dados totais e ruídos dos dados faltantes, e na horizontal variáveis da base de dados.



Fonte: Próprio Autor.

Em um primeiro momento do trabalho, foram mantidas todas as variáveis da base de dados, tendo em vista que a quantidade de valores faltantes não foi tão elevada (menos que 10%), exceto a precipitação. Assim, para realizar o preenchimento de preenchimento dos dados faltantes, optou-se por usar a biblioteca *Miss Forest* que, em síntese, utiliza o *Random Forest* para estimar os dados ausentes a partir dos dados presentes na série de cada variável. Com isso, vale dizer que essa etapa será realizada mais adiante, antes da fase de testes dos modelos, após a divisão da base de dados para que possamos evitar o fenômeno de *data leakage*² (vazamento dos dados).

5.1.2 Análise Exploratória de Dados

Para que se possa obter melhores *scores* por parte dos modelos preditivos, é necessário que haja uma etapa prévia de análise estatística e exploratória dos dados disponíveis no banco de dados, de forma a observar as melhores práticas e estratégias de manuseio de dados. Nessa seção, são apresentadas diversas visualizações com o objetivo de identificar tendências, além de explorar de forma mais aprofundada, a “natureza” da variável *target* (potência elétrica) do Parque Eólico de Praia Formosa.

Assim, primeiramente, foi realizada a análise descritiva das variáveis presentes no

² Data Leakage: em estatística e aprendizado de máquina, vazamento é o uso de informações no processo de treinamento do modelo que não deveriam estar disponíveis no momento da predição, fazendo com que as pontuações preditivas superestimassem a utilidade do modelo quando executado em um ambiente de produção.

dataset. Tais métricas podem ser observadas na Tabela 2.

Tabela 2 - Medidas de tendência central da base de dados.

Métrica/ Feature	Temperatura	Umidade	Velocidade do Vento	Precipitação	Pressão	Potência Elétrica
Média	27,37°C	70,28%	1,11 m/s	1,11 mm	1000,94 mbar	23,52 MWmed
Mediana	27,38°C	68%	1 m/s	0 mm	1000,85 mbar	22,72 MWmed
Desvio Padrão	1,299	13,100	0,947	4,317	1,574	15,782
Valor Máximo	32,88°C	97,75%	25,7 m/s	41 mm	1006,19 mbar	68,98 MWmed
Valor Mínimo	23,42°C	24%	0 m/s	0 mm	996,85 mbar	8,83 MWmed

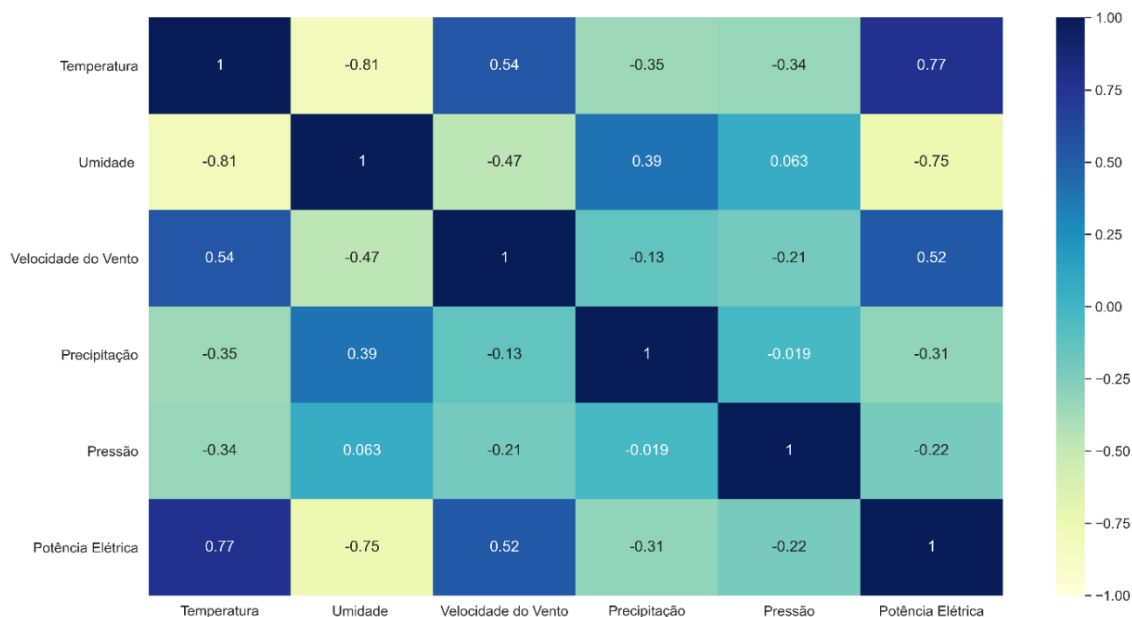
Fonte: Próprio Autor.

De acordo com a Tabela 2, referente às mensurações do Parque Eólico de Praia Formosa, podemos observar que alguns dados podem ser *outliers*. Por exemplo, o valor mínimo da variável de velocidade do vento ser zero pode indicar que, ao longo do dia, ocorreram erros dos equipamentos na coleta dos dados. Além disso, fica claro que a precipitação é uma *feature* que denota sentido nas tendências apresentadas, tendo em vista que pode ocorrer dias em que não há registro de chuvas (0 mm), situação predominante em diversas regiões do Nordeste.

Outro ponto a ser observado é que o desvio padrão da variável *target* (potência elétrica) é próximo ao desvio padrão da umidade relativa do ar. Desta forma, pode-se inferir a existência de uma correlação entre tais variáveis tendo em vista que elas possuem “variações” ao longo do tempo bem similares.

Com isso, para observar o grau de linearidade entre as variáveis disponíveis no *dataset*, foi gerado um mapa de calor, onde cada valor representa o coeficiente de Pearson, também chamado de coeficiente de correlação, ou simplesmente correlação para os pares de variáveis. Este coeficiente, normalmente representado por ρ , assume apenas valores entre -1 e 1, que expressa o grau de dependência linear entre duas variáveis quantitativas. A Figura 9 mostra tal matriz representada via mapa de calor.

Figura 9: Mapa de Calor (grau de linearidade entre as variáveis).



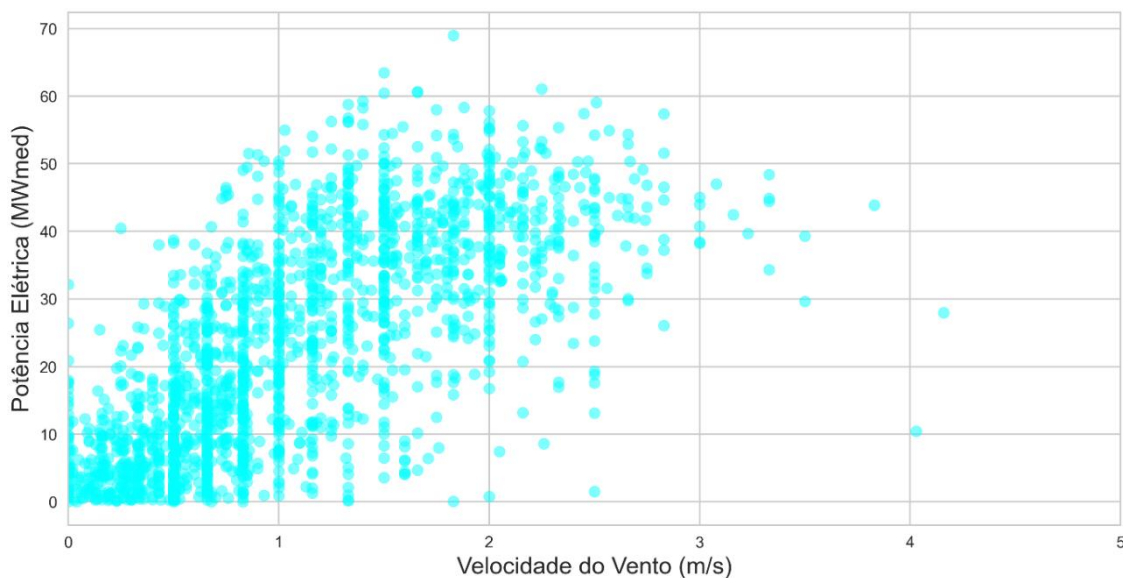
Fonte: Próprio Autor.

A partir do apresentado no mapa de calor, é possível observar que ocorrem algumas correlações entre as variáveis predictoras em relação a *target* (potência elétrica):

- A temperatura e a umidade relativa do ar possuem uma correlação considerada forte, sendo essa faixa de valores importante para a previsão da geração de potência, embora o conceito de correlação não implicar causalidade.
- A velocidade do vento possui uma correlação moderada em relação a potência elétrica (isso ocorreu, pois nesse trabalho, estamos considerando a média *diária* de cada variável, e não avaliando coletas de menor granularidade – horária).
- A precipitação possui uma correlação considerada fraca com a potência elétrica. Isso pode ocorrer devido ao fluxo do vento ser mais presente em dias ensolarados, principalmente, na região de Sobral, perto de onde se situa o parque eólico.
- A pressão acaba não possuindo uma correlação com a potência devido a sua pouca variabilidade ao decorrer dos dias, já que estamos considerando a média, porém, é importante ressaltar que ela é extremamente importante para se considerar em nosso modelo, tendo em vista que a mesma é a *feature* responsável pela ocorrência dos ventos.

Como foi apresentado na Figura 9, há uma dependência linear moderada entre a velocidade do vento e a potência elétrica. Assim, foi realizado o gráfico de dispersão (Figura 10) para entender o grau de linearidade dessas duas variáveis.

Figura 10: Grau de dependência entre a potência elétrica gerada e a velocidade do vento.



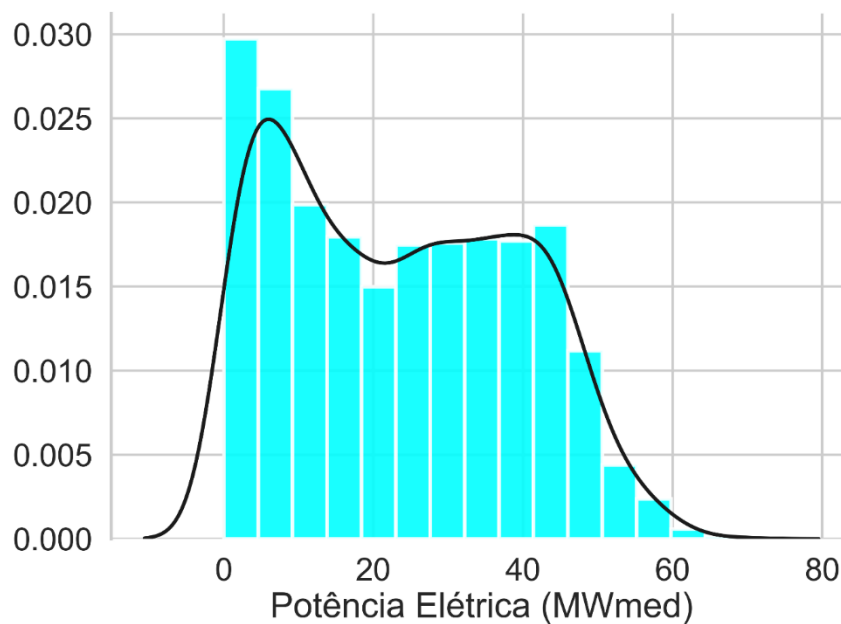
Fonte: Próprio Autor.

Analisando o comportamento do gráfico de dispersão, e embora exista uma tendência de linearidade, não houve uma correlação forte da velocidade do vento com a potência elétrica gerada pelo parque. Em suma, vale dizer que os dados meteorológicos utilizados são da região de Sobral, local próximo ao parque eólico, mas que não possuem as mesmas coordenadas, assim, haverá variações e discrepâncias na comparação entre os dois dados.

Além disso, é importante levar em consideração que algumas turbinas eólicas possuem limite de geração de potência, isto é, para determinadas velocidades de vento, sejam elas maiores ou menores que as estipuladas pelo motor de indução, pode acontecer de não ocorrer a geração de potência esperada.

Por fim, com o objetivo de observar o comportamento da variável *target* (potência elétrica gerada), foram realizadas algumas visualizações de forma a observar o “padrão” da série apresentada. Primeiramente, foi realizado o histograma com o objetivo de observar e explorar a assimetria de onde os dados estão situados (vide Figura 11).

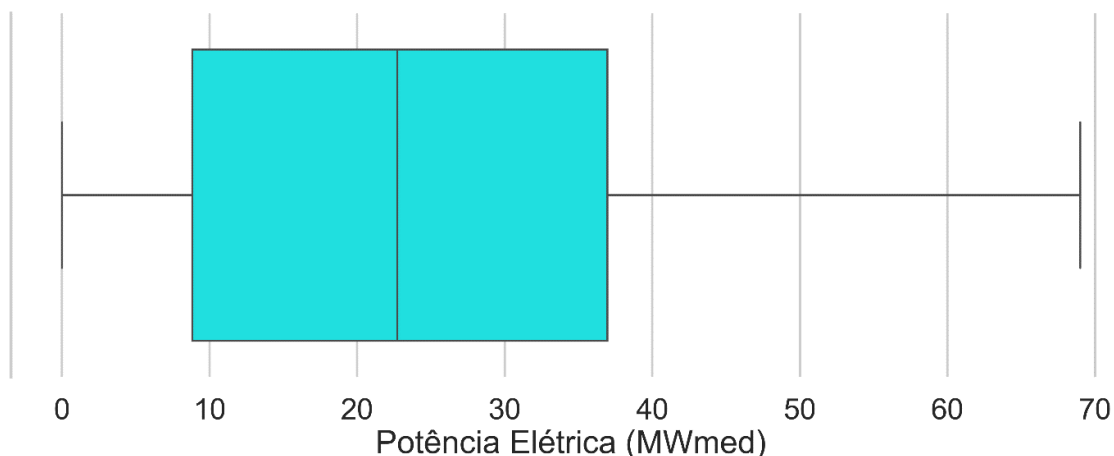
Figura 11: Distribuição da variável alvo.



Fonte: Próprio Autor.

No caso do histograma da Figura 11, é possível observar que a variável *target* (potência elétrica) não possui uma assimetria tão bem definida. Fica claro na distribuição que os *outliers* se situam mais à direita, por isso podemos afirmar que a inclinação da estimativa de *kernel* está situada mais à esquerda, e a média da geração (23,52MWmed) é maior que a mediana da geração (22,712 MWmed). Assim, foi observado também, de maneira similar, a “natureza” da variável alvo a partir do boxplot da Figura 12.

Figura 12: Boxplot da variável alvo.

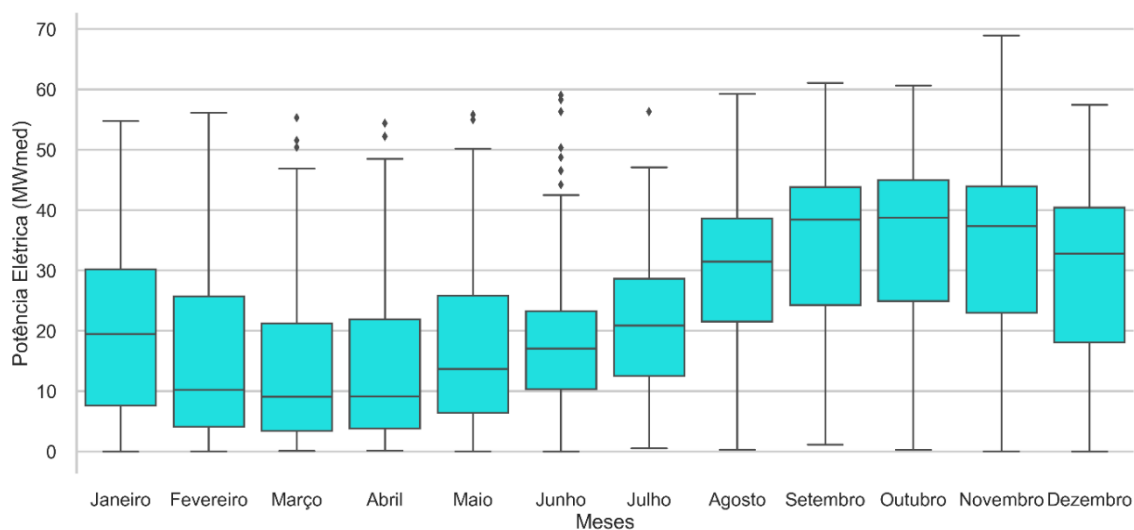


Fonte: Próprio Autor.

No caso do *boxplot* da Figura 12, é possível notar que a variável *target* (potência

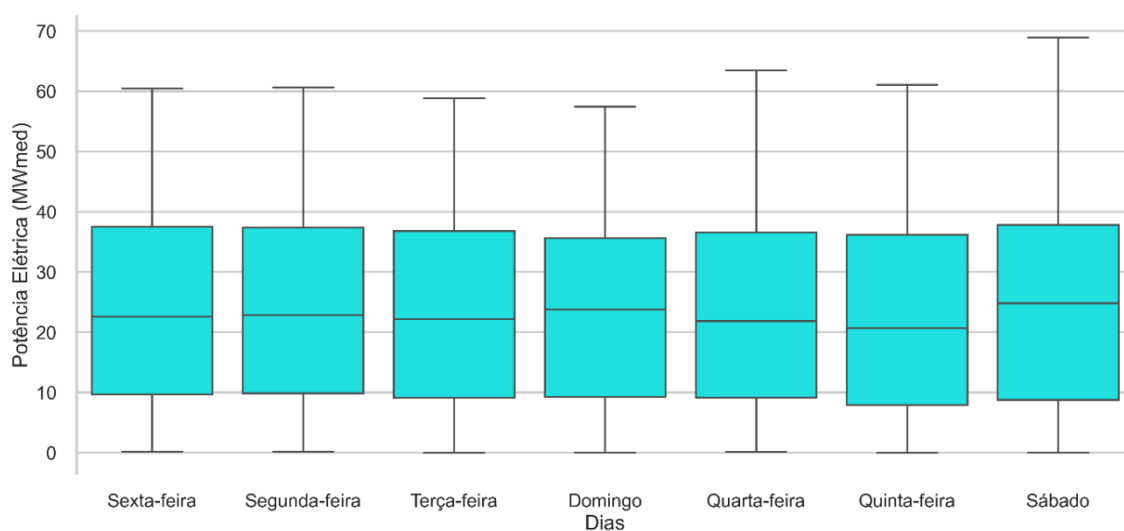
elétrica) tem valores mais recorrentes em torno de sua mediana (22,712 MWmed), que prevalece entre 20 MWmed e 30 MWmed de potência elétrica gerada. Além disso, fica claro a não presença de valores tão discrepantes ao longo da série, sendo somente alguns *outliers* observados na Figura 10. Diante disso, com o objetivo de investigar onde se situam esses *outliers* na série alvo, são apresentados os gráficos referentes à potência elétrica por mês (vide Figura 13) e dia da semana (vide Figura 14).

Figura 13: Boxplot variável alvo por mês.



Fonte: Próprio Autor.

Figura 14: Boxplot variável alvo por dia da semana.

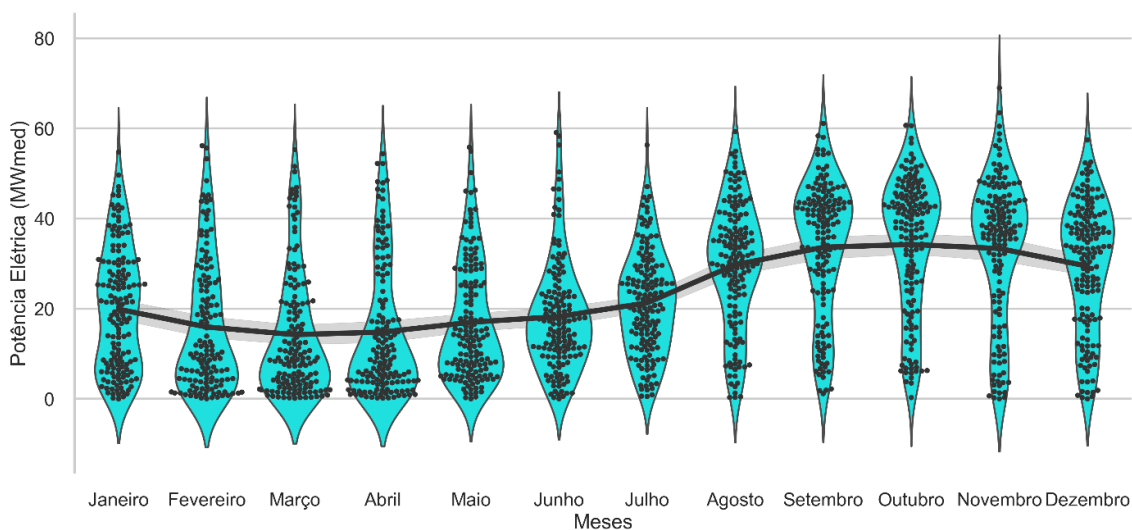


Fonte: Próprio Autor.

Na Figura 13, é possível observar que os meses com mais *outliers* (valores que

“fogem” do padrão) se situam em junho, justamente no período em que começa a mudar a tendência da série, isto é, no início do período de safra dos ventos no Nordeste brasileiro. Já em relação a Figura 14, fica claro que, ao longo dos dias da semana, a potência elétrica acaba não tendo grande variação (vide mediana), ficando sempre limitada entre 10 MWmed e 40MWmed. Por fim, o gráfico de distribuição mensal da potência elétrica gerada em forma de violino é apresentado na Figura 15.

Figura 15: *Violin Plot* da variável alvo por mês.



Fonte: Próprio Autor.

A partir do observado na distribuição da Figura 15, fica evidente que os maiores valores de geração de potência no parque ocorrem no decorrer dos últimos meses do ano (setembro, outubro e novembro), isso se deve ao fato de que no nordeste, durante esse período, ocorre a chamada "safra dos ventos", com uma maior geração de potência devido às maiores intensidades de vento. Além disso, a assimetria da série se inverte justamente durante essa mudança.

5.1.3 Criação de Novas Variáveis e Seleção dos Recursos

Para que se possa obter resultados mais assertivos por parte dos modelos preditivos, foi realizada uma etapa de engenharia de seleção de recursos, com o objetivo de criar recursos adicionais a partir dos já existentes, de forma a extrair informações relevantes observadas na Análise Exploratória de Dados (AED) a fim de treinar os modelos. As variáveis criadas foram:

- Médias móveis semanais e trimestrais da potência elétrica gerada: com o objetivo

de acompanhar a tendência da série observada na AED.

- Retorno logarítmico da potência elétrica gerada: a transformação da variável em uma métrica comparável como a porcentagem permite com que possamos realizar análises estatísticas multidimensionais.
- Subtração móvel diária da potência elétrica gerada: para seguir a sazonalidade da série.
- *One-Hot Encoding* dos meses: como observado na AED ao longo dos meses do ano, o parque eólico de praia formosa acaba gerando mais potência elétrica durante os segundos semestres dos anos, assim, a partir da utilização de uma técnica chamada *one-hot encoding*, podemos criar um vetor de variáveis binárias onde cada *feature* representa o pertencimento ou não de uma instância, nesse caso os meses.
- *Feature* cíclica dos dias: já que na AED foi observado que a geração não varia muito ao longo dos dias da semana, a *feature* cíclica permite com que encontremos as tendências que podemos possuir ao longo das semanas. Calculando as componentes x e y utilizando funções trigonométricas como seno e cosseno, é possível chegar o mais próximo do que seria a geração de potência ao longo dos dias. No caso desse trabalho, foi criada a *feature* cíclica utilizando o seno para os dias da semana.

Assim, conduzida a etapa de criação de tais variáveis, para que possamos escolher as melhores delas para a etapa de treinamento dos modelos, foi realizada a etapa de remoção de multicolinearidade³. Em síntese, esse fenômeno é um problema que diminui a precisão na regressão, porque aumenta a variância do modelo (principalmente para modelos lineares), aumenta o número de dimensões sem aumentar a informação útil de discriminação dos dados, além de distorcer a capacidade de explicar um modelo. Se dois recursos são colineares, eles não devem ser empregados em conjunto, isto é, apenas os recursos com maior capacidade de discriminação dos dados devem ser considerados em paralelo levando em conta a problemática envolvida.

Nesse trabalho, a detecção da multicolinearidade entre as variáveis foi feita através do VIF (*Variance Inflation Factor*), fator que mede o grau em que cada variável

³ Multicolinearidade: é um fenômeno que aumenta a variância de um dado modelo quando as características do conjunto de dados apresentam uma alta correlação entre si. Geralmente é medido usando o coeficiente de correlação de Pearson. Se a matriz de correlação mostra elementos fora da diagonal com um valor absoluto alto, podemos afirmar que há o efeito de colinearidade.

independente é explicada pelas demais variáveis independentes, sendo que quanto maior o VIF, mais severa será a multicolinearidade (CRANEY et al., 2002). A equação a seguir demonstra como é realizado o cálculo do VIF.

$$VIF = \frac{1}{1 - R^2} \quad (6)$$

em que R^2 indica o coeficiente de determinação da equação de regressão entre as variáveis.

A Tabela 3 apresenta os valores obtidos do VIF para cada variável da base. Vale ressaltar que a variável *one-hot* não foi considerada nessa análise tendo em vista que se trata de uma *feature* categórica.

Tabela 3 - VIF das variáveis preditoras.

Variável (<i>feature</i>)	VIF (<i>Variance Inflation Factor</i>)
Temperatura	4,271
Umidade Relativa do Ar	3,494
Velocidade do Vento	1,359
Precipitação	1,299
Pressão	1,991
Média móvel semanal (potência)	4,050
Média móvel trimestral (potência)	3,692
Retorno logarítmico (potência)	1,951
Subtração móvel (potência)	1,967
Dia - Seno	1,060

Fonte: Próprio Autor.

Uma regra clássica na literatura mostra que um $VIF > 10$ indica que a multicolinearidade possa estar influenciando as estimativas de mínimos quadrados (KENNEDY et al., 2008). Assim, para esse trabalho, foram mantidas todas as variáveis (originais e artificialmente criadas) tendo em vista que obtemos um $VIF < 10$ em todas.

5.2 Máquinas de Vetores de Suporte (*Support Vector Machines*)

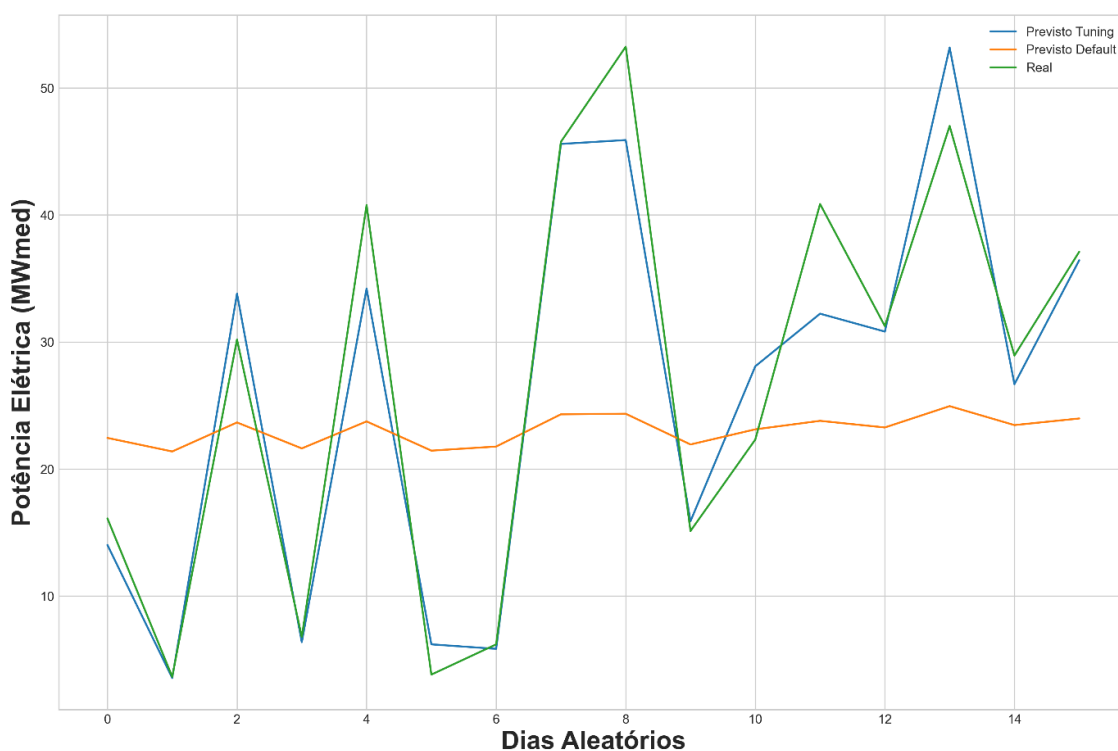
No algoritmo Máquina de Vetores de Suporte, SVM, foram utilizadas todas as variáveis preditoras, além daquelas que foram criadas na etapa de *feature engineering*. Os melhores hiperparâmetros do modelo foram determinados a partir do uso do *Random Search* (RS), tendo em vista que são poucos os hiperparâmetros que trazem uma influência mais crítica nos resultados da modelagem. Assim, o RS foi configurado com a

validação cruzada de 4 partições e $k = 20$ iterações do algoritmo. Os melhores hiperparâmetros encontrados foram:

- Kernel: linear.
- Grau de função do kernel: 2.
- Epsilon: 0,5.

Na Figura 16, pode-se observar a predição realizada para uma amostra de 15 dias aleatórios.

Figura 16: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) e sem ajuste dos mesmos (Previsto Default), além dos valores reais da série vs período (15 dias aleatórios) da série de teste estimada pelo modelo SVM.



Fonte: Próprio Autor.

Com base na Figura 16, fica evidente que o ajuste de hiperparâmetros foi de extrema importância para que o modelo SVM pudesse encontrar a melhor divisão do hiperplano, exclusivamente devido a função kernel linear. Enquanto que “sem o ajuste de hiperparâmetros”, a característica dos valores previstos só “seguiram” a tendência da série, após o ajuste, foi possível observar uma melhora bastante perceptível na assertividade dos resultados.

Outrossim, vale ressaltar que, com o modelo do SVM, não é possível observar quais variáveis foram as mais importantes no contexto preditivo. Isso se deve pela

complexidade do algoritmo em problemas de regressão não linear, onde o conjunto de dados é mapeado em um espaço de dimensão superior, que difere significativamente dos dados brutos disponíveis no *dataset* coletado. Os *scores* obtidos pelo algoritmo se situam na Tabela 4.

Tabela 4 - Scores finais obtidos pelo modelo SVM (Default e Tuning).

Métricas	<i>Default</i>	<i>Tuning</i>
MSE	218,370	24,135
MAPE (%)	54,547	16,025

Fonte: Próprio Autor.

Com isso, fica evidente que o SVM acabou apresentando resultados bastante satisfatórios após o *tuning* dos hiperparâmetros do modelo, sendo essa a etapa principal para que o mesmo apresentasse boas precisões ao longo do período de teste.

5.3 Aumento Extremo de Gradiente (*Extreme Gradient Boosting*)

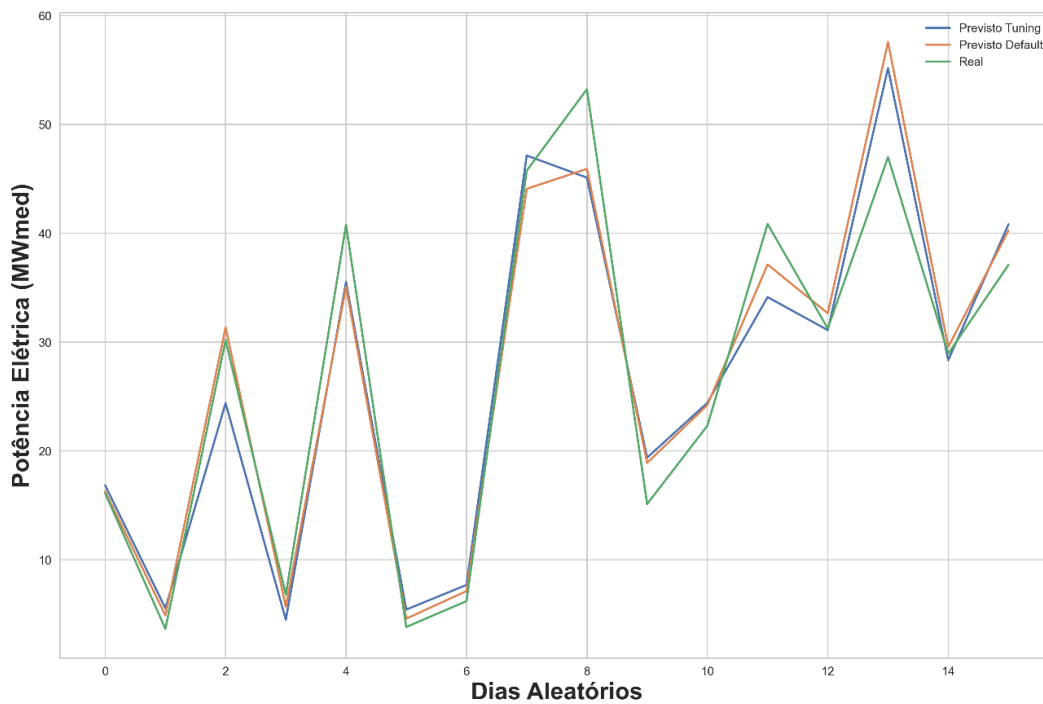
No aumento extremo de gradiente, XGB, foram utilizadas todas as variáveis preditoras originais da base de dados, além das criadas na etapa de engenharia de seleção de recursos, sendo os melhores hiperparâmetros determinados a partir do *Grid Search* para o número de árvores, e *Random Search* para os demais hiperparâmetros, com uma validação cruzada (*Cross Validation*) de 4 partições e $k = 20$ iterações no algoritmo. Os melhores hiperparâmetros determinados para o XGB foram:

- Número de árvores: 100.
- Profundidade máxima da árvore: 5.
- Porcentagem de amostras da árvore: 1.
- Porcentagem de recursos para serem utilizados: 1.
- Gamma (controla se determinado nó será dividido com base na redução esperada de perda após a divisão): 5.

Na Figura 17, tem-se a previsão realizada em dias aleatórios, sendo possível notar a diferença entre o modelo sem seus hiperparâmetros otimizados (Previsto *Default*) vs hiperparâmetros otimizados (Previsto *Tuning*). Observando uma amostra de 15 dias aleatórios, fica claro que o modelo atinge alta precisão em determinados dias, embora em dias em que há mudança brusca de geração de potência o modelo não estima com alta precisão o valor da potência. Ademais, pode-se dizer que o *tuning* do modelo acabou influenciando de forma bastante significativa como por exemplo, nos dias 13 e 15.

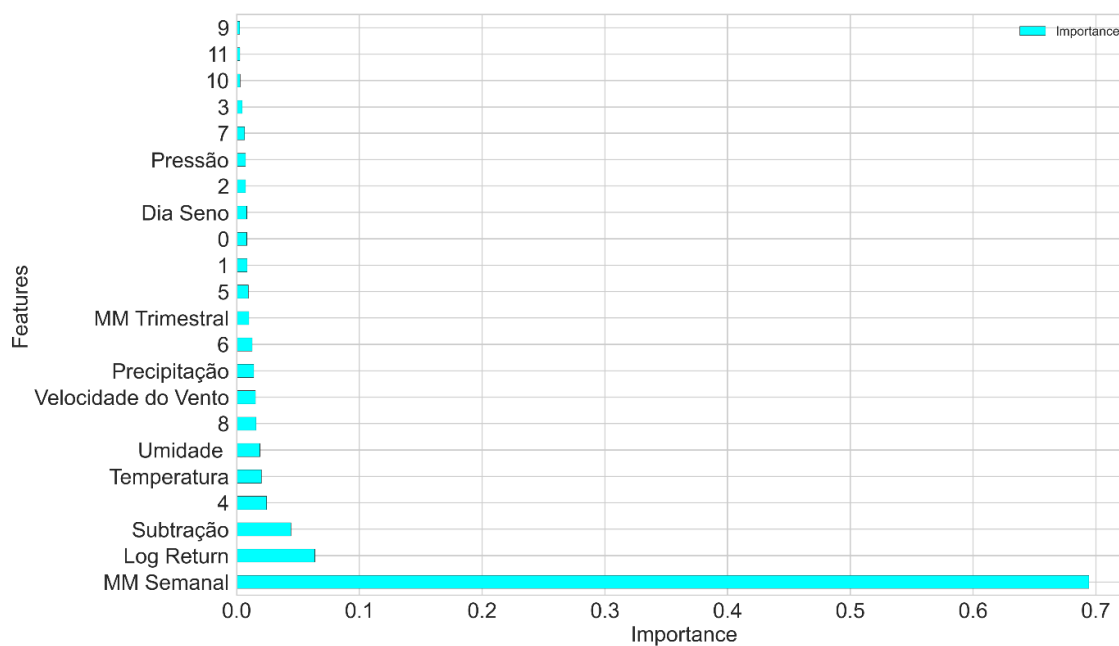
Para entender a tomada de decisão do modelo XGB, foi utilizada a técnica de *Feature Importance*, com o objetivo de observar as variáveis que foram mais importantes no contexto do treinamento do modelo (ver Figura 18).

Figura 17: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste. realizada pelo modelo aumento extremo de gradiente



Fonte: Próprio Autor.

Figura 18: Features mais importantes para a previsão do modelo Aumento Extremo de Gradiente



Fonte: Próprio Autor.

A partir do exposto, fica claro que dentre as *features* criadas na etapa de seleção de recursos, a média móvel semanal acabou possuindo uma maior relevância global para a tomada de decisão do modelo, já que a mesma permite seguir a "tendência" entre os últimos 7 dias gerados. Além disso, é possível observar que as variáveis *log return* e subtração móvel acabam por ter uma certa relevância global no modelo. As outras *features* acabam colaborando de forma menos preponderante para o modelo, porém, com a sua devida importância no que se diz respeito às ramificações pelas quais vão ocorrendo a partir dos nós das árvores. Por fim, os scores obtidos pelo algoritmo encontram-se na Tabela 5.

Tabela 5 - Scores finais obtidos pelo modelo aumento extremo de gradiente (Default e Tuning).

Métricas	<i>Default</i>	<i>Tuning</i>
MSE	13,484	11,404
MAPE (%)	11,383	10,589

Fonte: Próprio Autor.

É válido salientar, portanto, que o modelo se comportou com uma alta taxa de assertividade, possuindo somente 10,589% do valor do erro percentual médio absoluto, valor possível graças a engenharia de seleção de recursos a partir das variáveis artificiais criadas, que em um ambiente de produção poderia colaborar diretamente para a previsão da potência elétrica média gerada em um horizonte de curto prazo ($d + 1$).

5.4 Aumento de Gradiente (*Gradient Boosting*)

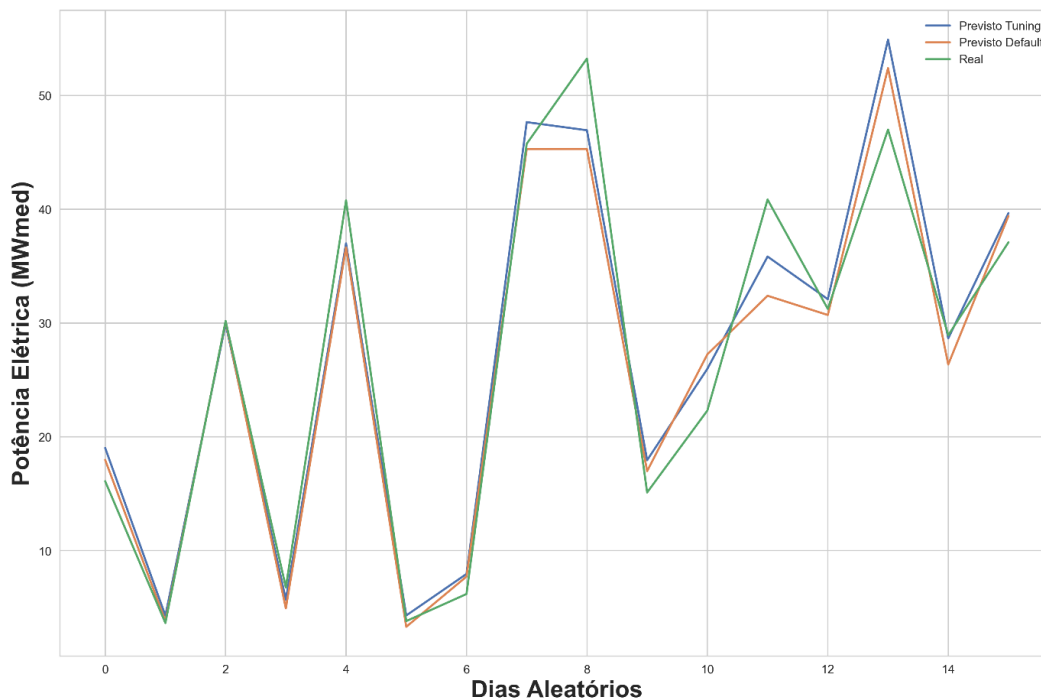
No aumento de gradiente, GB, assim como nos outros modelos testados, foram utilizadas todas as variáveis previsoras (genéricas e originais da base de dados). Além disso, vale ressaltar que, de forma similar ao modelo XGB, foram determinados os melhores hiperparâmetros com o *Random Search* com uma validação cruzada de 4 partições de 20 iterações, e o melhor número de árvores com o *Grid Search*. Os melhores hiperparâmetros encontrados, foram:

- Número de árvores: 900.
- Número mínimo de amostras necessárias para estar em um nó da folha: 6.
- Número mínimo de amostras para divisão de um nó: 10.
- Número máximo de recursos a serem considerados: *auto*.

Na Figura 19 é apresentada a previsão realizada em dias aleatórios e, mais uma vez, é possível observar a diferença entre o modelo sem ter seus hiperparâmetros otimizados

(Previsto *Default*) e sua versão com os hiperparâmetros ajustados (Previsto *Tuning*).

Figura 19: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto *Tuning*) sem ajuste dos mesmos (Previsto *Default*) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo aumento de gradiente.



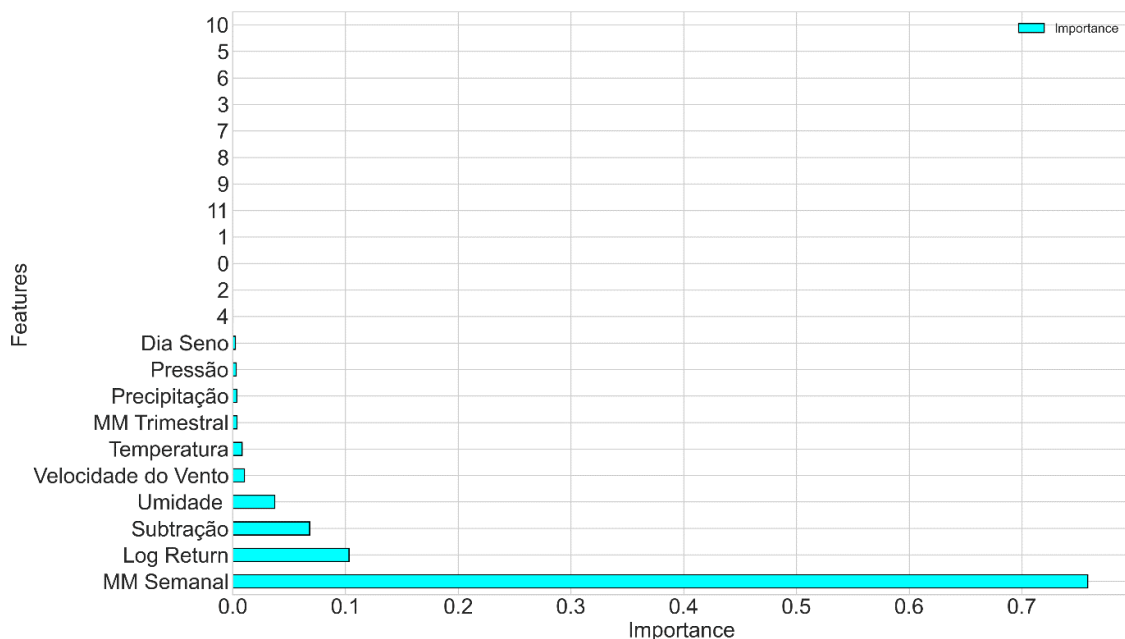
Fonte: Próprio Autor.

Através da Figura 19, fica explícita a alta taxa de assertividade por parte do GB durante essa amostra analisada. É possível observar que alguns dias, como o dia 13, o modelo sem o ajuste de hiperparâmetros acabou se performando melhor, entretanto quando analisamos a amostra de teste por completo, acabamos observando mais dias em que o modelo com os hiperparâmetros ajustados se performa melhor, como no dia 11 da amostra, por exemplo.

Na Figura 20, podemos observar as variáveis mais relevantes para a predição originada pelo modelo GB, que possuiu uma tomada de decisão bem parecida com o XGB, sendo as 3 *features* mais relevantes as mesmas. Entretanto, no caso do aumento de gradiente, o modelo julgou não importante a informação dos meses do ano, enquanto no XGB alguns meses tiveram relevância na tomada de decisão do algoritmo. Os resultados do aumento de gradiente estão na Tabela 6, que mostra que o GB apresentou uma intensa melhora a partir dos hiperparâmetros ajustados possuindo um erro percentual médio absoluto de 11,574. Assim, podemos afirmar que o modelo acabou se comportando de maneira similar ao XGB devido ambas as técnicas apresentarem funcionamento similar

com o *boosting* explicado nas seções 4.1.2. e 4.1.3.

Figura 20: Ranking das features mais importantes para a previsão do modelo aumento de gradiente



Fonte: Próprio Autor.

Tabela 6 - Scores finais obtidos pelo modelo aumento de gradiente (Default e Tuning).

Métricas	<i>Default</i>	<i>Tuning</i>
MSE	15,650	14,769
MAPE (%)	12,244	11,574

Fonte: Próprio Autor.

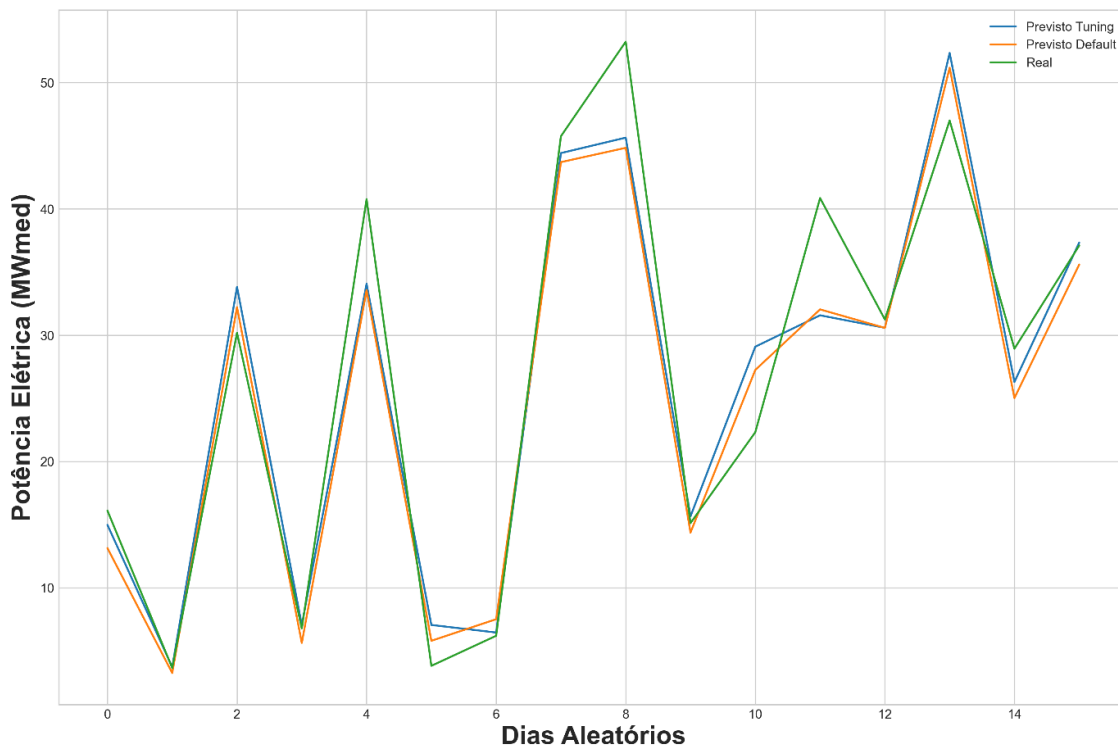
5.5 Redes Neurais Artificiais Multicamadas (*RNA Multi Layer Perceptron*)

Na rede neural MLP, foram utilizadas todas as variáveis assim como nos outros modelos, sendo os melhores hiperparâmetros determinados pelo *Random Search* com uma validação cruzada de 4 partições e número de iterações de $k = 20$. Os melhores hiperparâmetros da rede foram:

- Neurônios na camada intermediária: 15.
- Função de ativação: *identity*.
- Solucionador para otimização dos pesos: *Lbfgs*
- Taxa de Aprendizado: *invscaling*

Na Figura 21, é possível observar a previsão realizada em dias aleatórios e notar a diferença entre o modelo sem ter seus hiperparâmetros otimizados (Previsto *Default*) e com os hiperparâmetros otimizados (Previsto *Tuning*).

Figura 21: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo redes neural MLP.



Fonte: Próprio Autor.

A partir do observado, podemos dizer que a MLP, assim como os outros modelos testados, apresentou uma boa taxa de assertividade ao longo do período de teste. Em síntese, podemos notar que em todos os modelos ocorrem dias de variação brusca da potência elétrica, tal como já evidenciado na EAD. Como as redes neurais apresentam um sistema muito complexo de ser analisado, não é possível utilizar a classe de *feature importance* para observar quais variáveis mais influenciaram na predição.

Em síntese, as redes neurais são modelos difíceis de serem interpretados, o que os torna “caixa preta” em alguns casos particulares. Por outro lado, conforme salientado em (CERLIANI et al., 2019), é possível afirmar que na maior parte dos casos, as mesmas *features* que mais influenciam os modelos implementados, do tipo *ensemble*, irão influenciar também na configuração da rede. Na Tabela 7, são apresentados os scores finais obtidos pelo algoritmo.

Tabela 7 - Scores finais obtidos pelo modelo rede neural MLP (Default e Tuning).

Métricas	<i>Default</i>	<i>Tuning</i>
MSE	26,073	24,235
MAPE (%)	16,872	16,064

Fonte: Próprio Autor.

É válido afirmar, portanto, que a rede se comportou com uma alta taxa de assertividade, possuindo somente 16,064% do valor do erro percentual médio absoluto, mais uma vez devido à engenharia de seleção de recursos.

5.6 Árvores de Decisão via Floresta Aleatória (*Random Forest*)

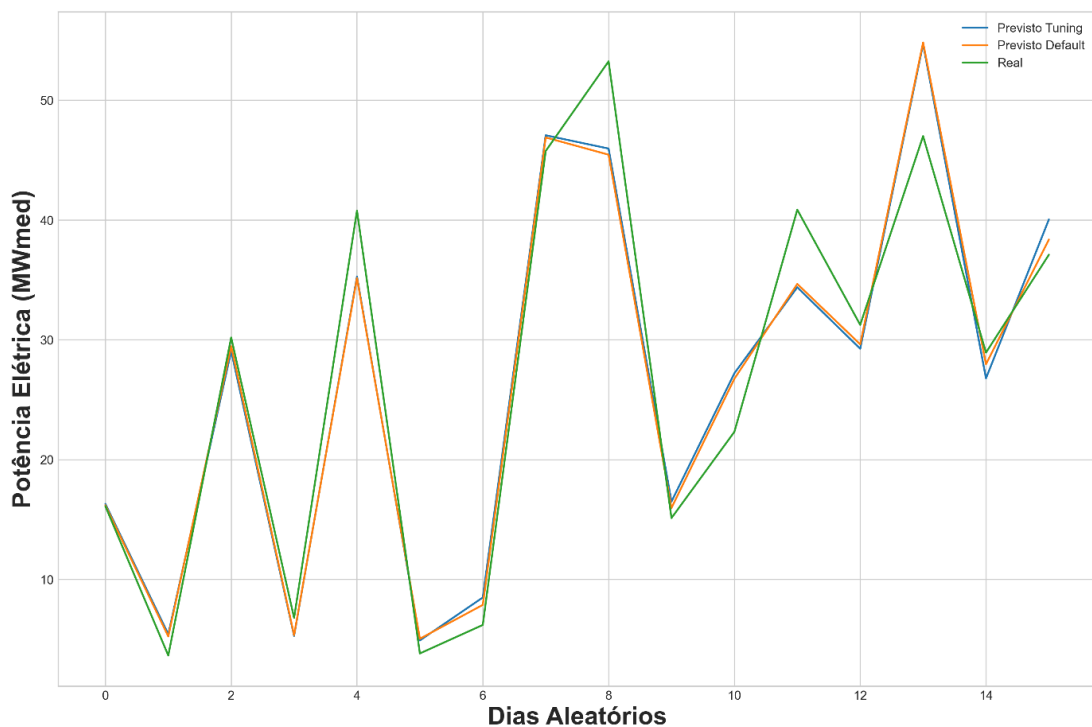
Assim como nos demais algoritmos, no modelo RF, foram utilizadas todas as variáveis para a etapa de treinamento do algoritmo. Para o ajuste de hiperparâmetros, foi utilizado o RS com uma validação cruzada de 4 partições e 20 iterações do algoritmo. Os melhores hiperparâmetros para o RF, foram:

- Número de árvores: 200
- Profundidade máxima da árvore: 15
- Número de amostras mínima para dividir um nó interno: 2

Na Figura 22, tem-se plotado a previsão realizada, em dias aleatórios, onde é possível visualizar a diferença entre o modelo sem ter seus hiperparâmetros otimizados (Previsto *Default*) vs sua versão otimizada (Previsto *Tuning*). Através do apresentado, podemos afirmar que o RF também apresentou uma alta taxa de assertividade ao longo do período de teste, obtendo resultados similares aos demais modelos testados do tipo de *ensemble*. Assim, o RF acabou se comportando melhor utilizando todas as *features* previsoras: isso se deve exclusivamente ao fato de que a entropia ao longo dos nós colabora para que a árvore siga o “caminho” mais adequado para realizar a previsão eliminando as variáveis não importantes ao longo dos nós.

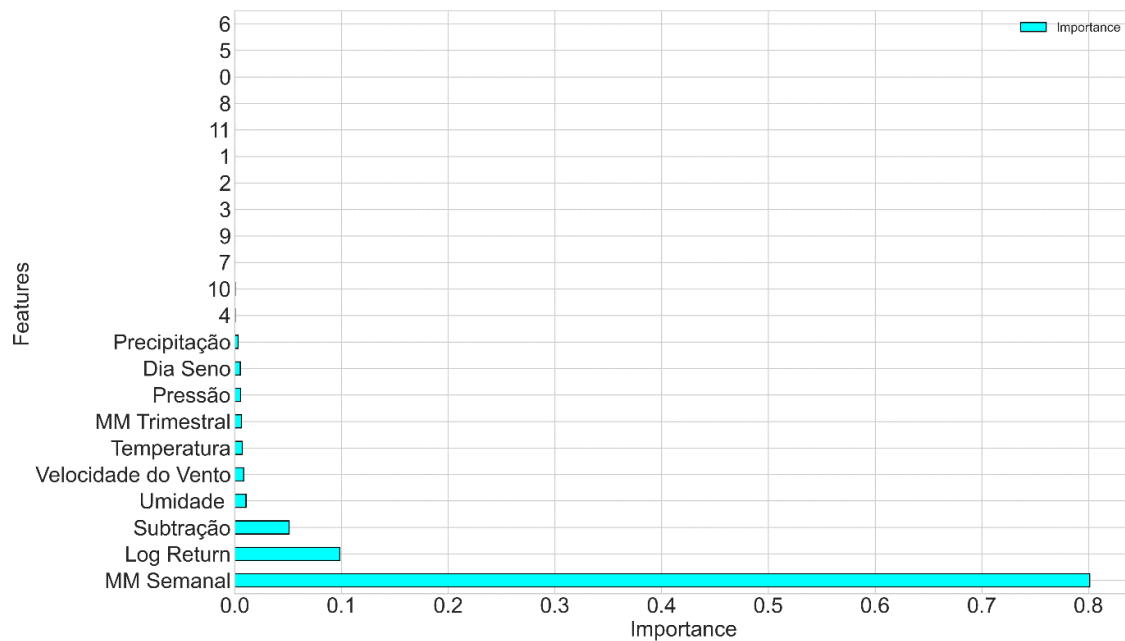
No modelo RF, foi utilizado o *feature_importance* da biblioteca *scikit-learn* para observar as variáveis que mais impactaram nas tomadas de decisões do modelo. O resultado está apresentado na Figura 23, indicando que as variáveis mais importantes no contexto preditivo por parte do RF foram as criadas artificialmente na engenharia de seleção de recursos, sendo elas: média móvel semanal, *log return* e subtração móvel (2 em 2 amostras). É válido ressaltar que as outras *features* praticamente não obtiveram tamanha importância durante o treinamento do algoritmo, visto que a árvore acabou levando em conta os nós que traziam os melhores resultados.

Figura 22: Gráfico da potência elétrica prevista com ajuste de hiperparâmetros (Previsto Tuning) sem ajuste dos mesmos (Previsto Default) e valores reais da série vs período (15 dias aleatórios) da série de teste realizada pelo modelo árvores de decisão via floresta aleatória.



Fonte: Próprio Autor.

Figura 23: Ranking das features mais importantes para a previsão do modelo árvore de decisão via floresta aleatória.



Fonte: Próprio Autor.

Na Tabela 8, são apresentados os *scores* finais do algoritmo, em que é possível observar uma pequena melhora, porém, significativa dos resultados, já que o desempenho do RF que acabou obtendo um MAPE final de 12,794%.

Tabela 8 - Scores finais obtidos pelo modelo árvores de decisão via floresta aleatória (Default e Tuning).

Métricas	<i>Default</i>	<i>Tuning</i>
MSE	16,918	16,807
MAPE (%)	12,813	12,794

Fonte: Próprio Autor.

5.7 Comparação de Desempenho dos Modelos

Por fim, nessa seção, são sumarizados os *scores* finais obtidos por cada modelo no Parque Eólico de Praia Formosa, onde foram testados os cinco modelos de AM, sendo eles: Máquinas Vetores de Suporte (SVM), Aumento Extremo de Gradiente (XGB), Aumento de Gradiente (GB), Rede Neural do tipo MLP, e Árvores de Decisão via Floresta Aleatória (RF). Na Tabela 9, é possível observar os resultados de forma sumarizada.

Tabela 9 - Scores finais obtidos pelos modelos de aprendizado de máquina otimizados.

Modelos	MAPE (%)	MSE
SVM	16,025	24,135
XGB	10,589	11,404
GB	11,574	14,769
Rede neural MLP	16,064	24,235
RF	12,794	16,807

Fonte: Próprio Autor.

De acordo com o exposto, pode-se afirmar que o XGB acabou tendo um desempenho superior com relação aos outros modelos testados. Ademais, podemos afirmar que o SVM e a Rede Neural MLP acabaram possuindo resultados menos acurados visto a complexidade da arquitetura de ambos. No caso do SVM, a divisão do hiperplano acaba sendo mais complexa, demandando um ajuste de parâmetros mais detalhado a fim de capturar os padrões e tendência dos dados, e no caso da RNA, sua melhor configuração de pesos acaba sendo difícil de ser encontrada na prática.

Em relação ao XGB, GB e RF, foi observado um comportamento bem semelhante entre os resultados alcançados. Entretanto, conforme foi visto nos *plots* de *feature importance*, o XGB acaba utilizando todas as variáveis para encontrar o “caminho ótimo”

dos nós, o que permite uma leve superioridade perante o GB e RF, os quais acabaram não utilizando algumas variáveis para a divisão dos nós das árvores.

6 Conclusão

Nesse trabalho, foram implementadas e testadas as seguintes metodologias de Aprendizado de Máquina (AM): Máquinas Vetores de Suporte (SVM), Aumento Extremo de Gradiente (XGB), Aumento de Gradiente (GB), Rede Neural MLP, e Árvores de Decisão via Floresta Aleatória (RF). Em síntese, tais modelos foram ajustados e aplicados com o intuito de prever a potência elétrica média diária gerada, em MWmed, pelo Parque Eólico de Praia Formosa, que fica situado no Ceará.

Em relação às previsões obtidas, podemos afirmar que a utilização das variáveis artificiais foi de extrema importância para a obtenção de resultados mais acurados por parte dos modelos de AM. Ademais, pode-se afirmar que a verificação perante a multicolinearidade de tais variáveis foi de extrema valia a fim de medir o impacto no crivo decisório dos modelos.

Considerando a seleção dos hiperparâmetros dos modelos, a utilização do *Grid/Random Search* se mostrou adequada, tendo tais estratégias se mostrado viáveis na prática, o que diminuiu o tempo de testes conduzidos a partir da inserção de alguns parâmetros pré-estabelecidos no algoritmo que, por conseguinte, acabou retornando uma configuração mais otimizada para os algoritmos de aprendizado aplicados.

Já com relação aos resultados experimentais obtidos, podemos afirmar que o Aumento Extremo de Gradiente (*XGBoosting*) acabou apresentando os melhores *scores* com um erro percentual médio absoluto de 10,859% e erro quadrático médio de 11,404. Além disso, como foi discutido na Seção 4.6, podemos afirmar que o SVM e a Rede Neural MLP acabaram se comportando de maneira levemente menos acurada devido à complexidade das estruturas dos mesmos. Por fim, os modelos RF e GB acabaram possuindo comportamento semelhantes, porém, ainda inferiores ao XGB, já que o mesmo acabou utilizando todas as variáveis para a divisão dos nós e encontro do “caminho ótimo”.

Assim, de acordo com o alto nível de assertividade apresentados pelos modelos, vale ressaltar a importância de empregar mais de um algoritmo de AM em problemáticas relacionadas à regressão não linear, de forma a observar a viabilidade de cada modelo no problema estudado, tal como realizado neste estudo. Nesse caso, o aumento extremo de gradiente apresentou maiores precisões perante aos demais, entretanto, é válido ressaltar que, quando o modelo é testado em produção, é necessário um monitoramento de forma a observar a performance do mesmo.

Desse modo, conclui-se que o emprego de diversas modelagens é necessário,

tendo em vista que os parques eólicos no Brasil possuem características distintas. Assim, em problemas que tratam de regressão não linear dentro do setor eólico brasileiro, foi possível constatar que a aplicações de tais metodologias podem trazer benefícios sob diversas óticas, como por exemplo sobre o despacho de energia no Sistema Interligado Nacional (SIN), já que a energia eólica se trata de uma fonte renovável que possui grande intermitência.

7 Referências

AGÊNCIA CANAL ENERGIA, A. C. E. **Brasil chega a 19 GW de capacidade instalada.** Disponível em: <<https://www.canalenergia.com.br/noticias/53176883/brasil-chega-a-19-gw-de-capacidade-instalada-de-energia-eolica>>. Acesso em: 9 set. 2021.

AGUIAR A. et al., XGBOOST — A matemática passo a passo. Disponível em: <https://medium.com/@aln.deaguiar/xgboost-a-matem%C3%A1tica-passo-a-passo-29d34fa561dc>. Acesso em: 4 ago. 2020.

AMBIENTE ENERGIA, A. E. **Energia Eólica e Solar já são mais baratas do que energia de carvão nos EUA.** Disponível em: <<https://www.ambienteenergia.com.br/index.php/2019/06/energia-solar-eolica-mais-barata-carvao-eua/35925>>. Acesso em: 9 set. 2020.

ANDRADE, J. R.; BESSA, R. J. Improving renewable energy forecasting with a grid of numerical weather predictions. **IEEE Transactions on Sustainable Energy**, v. 8, n. 4, p. 1571-1580, 2017.

BERGSTRA, J.; BENGIO, Y., et al. Random search for hyper-parameter optimization. **Journal of Machine Learning Research**, v. 13, n. Feb, p. 281-305, 2012.

BREIMAN, L. et al., Random forests. **Machine learning**, v. 45, n. 1, p. 5-32, 2001.

CASTRO, F. C. C.; CASTRO, M. C. F. et al., Redes neurais artificiais. **Pontifícia Universidade Católica do Rio Grande do Sul**, 2001.

CERLIANI, M. et al., Feature Importance with Neural Networks. **Towards Data Science**. Disponível em: <https://towardsdatascience.com/feature-importance-with-neural-network-346eb6205743>. Acesso em: 5 out. 2021.

CRANEY, T. A.; SURLLES, J. G. Model-dependent variance inflation factor cutoff values. **Quality Engineering**, v. 14, n. 3, p. 391-403, 2002.

DE MYTTENAERE, Arnaud et al. Mean absolute percentage error for regression models. **Neurocomputing**, v. 192, p. 38-48, 2016.

DEMOLLI, Halil et al. Wind power forecasting based on daily wind speed data using machine learning algorithms. **Energy Conversion and Management**, v. 198, p. 111823, 2019.

FILHO A.C.; CANAL USP. Disponível em: <<https://www.youtube.com/watch?v=KgL1vqzy7is>>. Acesso em: 22 jul. 2021.

FRIEDMAN, J. H. et al., Stochastic gradient boosting. **Computational statistics & data analysis**, v. 38, n. 4, p. 367-378, 2002.

GLOBAL WIND ENERGY COUNCIL, G. W. E. C. **GLOBAL WIND REPORT 2021**. Disponível em: <<https://gwec.net/wp-content/uploads/2021/03/GWEC-Global-Wind-Report-2021.pdf>>. Acesso em: 9 set. 2021.

GOMES P. C. T. et al., Conheça o algoritmo XGBoost. **Data Geeks**. Disponível em: <https://www.datageeks.com.br/xgboost/>. Acesso em: 4 ago. 2021.

HAYKIN, S. et al., Redes neurais: princípios e prática. **Bookman Editora**, 2007.

IZMAILOV R., VAPNIK V. and VASHIST A. Multidimensional splines with infinite number of knots as svm kernels. In **IJCNN**, volume 13, pages 1–7, 2013.

KARINIOTAKIS, G. N.; STAVRAKAKIS, G. S.; NOGARET, E. F. Wind power forecasting using advanced neural networks models. **IEEE transactions on Energy conversion**, v. 11, n. 4, p. 762-767, 1996.

KENNEDY, P. et al., A guide to econometrics. **John Wiley & Sons**, 2008.

KIM, J. et al., Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. **Computational statistics & data analysis**, v. 53, n. 11, p. 3735-3745, 2009.

LEME, J. V. et al. Towards assessing the electricity demand in Brazil: Data-driven analysis and ensemble learning models. **Energies**, v. 13, n. 6, p. 1407, 2020.

LI, C. et al., A Gentle introduction to gradient boosting. Disponível em: http://www.ccs.neu.edu/home/vip/teach/MLcourse/4_boosting/slides/gradient_boosting_pdf, 2016. Acesso: 4 ago. 2021.

LI, L. L. et al. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. **Journal of Cleaner Production**, v. 242, p. 118447, 2020.

LIER, C. et al., Applying Machine Learning Techniques to Short Term Load Forecasting. 2015. Tese de Doutorado. **Faculty of Science and Engineering**.

MANDOT., P. et al., How exactly XGBoost Works? The Story of the fitting model on gradients. Disponível em: <https://medium.com/@pushkarmandot/how-exactly-xgboost-works-a320d9b8aeef>. Acesso em: 4 ago. 2020.

MARMARAS C., JAVED A., CIPCIGAN L., RANA O. Predicting the energy demand of buildings during triad peaks in gb. **Energy and Buildings**, 141:262–273, 2017.

MARQUES, Jeferson et al. Turbinas eólicas: Modelo, análise, e controle do gerador de indução com dupla alimentação. 2004.

MISRA., R. et al., Support Vector Machines — Soft Margin Formulation and Kernel Trick: Learn some of the advanced concepts that make Support Vector Machine a powerful linear classifier. **Towards Data Science**. Disponível em: <https://towardsdatascience.com/support-vector-machines-soft-margin-formulation-and-kernel-trick-4c9729dc8efe>. Acesso em: 6 ago. 2021.

MONARD, M.C.; BARANAUSKAS, J.A. Conceitos sobre aprendizado de máquina. *Sistemas Inteligentes-Fundamentos e Aplicações*, 1 ed. Barueri-SP: Manole Ltda, 2003. p. 89--114. ISBN 85-204-168, 2003.

NATEKIN, A. & KNOLL, A. et al., Gradient boosting machines, a tutorial. **Frontiers in neurorobotics**, v. 7, p. 21, 2013.

OPTIS, Mike; PERR-SAUER, Jordan. The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production. **Renewable and Sustainable Energy Reviews**, v. 112, p. 27-41, 2019.

PAO, Lucy Y.; JOHNSON, Kathryn E. A tutorial on the dynamics and control of wind turbines and wind farms. In: 2009 **American Control Conference**. IEEE, 2009. p. 2076-2089.

PAULA, M.; CONALGO M.; NUNO F. J.; CASACA, W. Predicting Long-Term Wind Speed in Wind Farms of Northeast Brazil: A Comparative Analysis Through Machine Learning Models. **IEEE Latin America Transactions**, v. 18, n. 11, p. 2011-2018, 2020.

POPE, K.; DINCER, I.; NATERER, G. F. Energy and exergy efficiency comparison of horizontal and vertical axis wind turbines. **Renewable energy**, v. 35, n. 9, p. 2102-2113, 2010.

QIU X., ZHANG L., SUGANTHAN P. N., AMARATUNGA G. AJ. Oblique random forest ensemble via least square estimation for time series forecasting. **Information Sciences**, 420:249–262, 2017.

RAMOS, L.; CONALGO, M.; CASACA, W. Data-driven analysis and machine learning for energy prediction in distributed photovoltaic generation plants: A case study in Queensland, Australia. **Energy Reports**, v. 8, p. 745-751, 2022.

RAUBER, T. W. et al., Redes neurais artificiais. **Universidade Federal do Espírito Santo**, 2005.

SVETNIK, V. et al. Random forest: a classification and regression tool for compound classification and QSAR modeling. **Journal of chemical information and computer sciences**, v. 43, n. 6, p. 1947-1958, 2003.

TARNOWSKI, G. C. et al., Metodologia de regulação da potência ativa para operação de sistemas de geração eólica com aerogeradores de velocidade variável. 2006.

VELASCO L. C. P., POLESTICO D. L. L., ABELLA D. M. M., ALEGATA G. T., and LUNA G. C. Day-ahead load forecasting using support vector regression machines. **International Journal of Advanced Computer Science and Applications**, 9(3), 2018.

VELLASCO, M. M. B. R. et al., Redes Neurais Artificiais. Laboratório de inteligência computacional aplicada (ICA). **Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ)**. Disponível em: <<http://www2.ica.ele.puc-rio.br/Downloads/33/ICA-introdu%C3%A7%C3%A3o%20RNs.pdf>>. Acesso em: 22 jul. 2021.

VON ZUBEN.; CASTRO, L. N. et al., Redes Neurais Artificiais. DCA/FEEC/Unicamp. **FEEC UNICAMP** Disponível em: <ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia006_03/topico5_03.Pdf>. Acesso em: 22 jul. 2021.

WANG, Zhenyu; TIAN, Wei; HU, Hui. A Comparative study on the aeromechanic performances of upwind and downwind horizontal-axis wind turbines. **Energy Conversion and Management**, v. 163, p. 100-110, 2018.

WERON, R. et al. Electricity price forecasting: A review of the state-of-the-art with a look into the future. **International journal of forecasting**, 30(4):1030–1081, 2014.

ZAMEER, Aneela et al. Machine Learning based short term wind power prediction using a hybrid learning model. **Computers & Electrical Engineering**, v. 45, p. 122-133, 2015.

ZHENG, Huan; WU, Yanghui. A xgboost model with weather similarity analysis and feature engineering for short-term wind power forecasting. **Applied Sciences**, v. 9, n. 15, p. 3019, 2019.