



Henrique Proes Gerhardt

**SISTEMA DE PREDIÇÃO DE PRODUÇÃO DE ENERGIA SOLAR BASEADO EM
DADOS METEOROLÓGICOS**

Horizontina-RS

2023

Henrique Proes Gerhardt

**SISTEMA DE PREDIÇÃO DE PRODUÇÃO DE ENERGIA SOLAR BASEADO EM
DADOS METEOROLÓGICOS**

Trabalho Final de Curso apresentado como requisito parcial para o Trabalho Final de Curso na Engenharia de Engenharia de Controle e Automação da Faculdade Horizontina, sob a orientação do Prof. Me. Rodrigo Bastos.

Horizontina-RS

2023

FAHOR - FACULDADE HORIZONTALINA
CURSO DE ENGENHARIA DE CONTROLE E AUTOMAÇÃO

A Comissão Examinadora, abaixo assinada, aprova o trabalho final de curso

“Sistema de previsão de produção de energia solar baseado em dados meteorológicos”

**Elaborada por:
Henrique Proes Gerhardt**

Como requisito parcial para a obtenção do grau de Bacharel em
Engenharia de Controle e Automação

Aprovado em: 22/11/2023
Pela Comissão Examinadora

Me. Rodrigo Bastos
Presidente da Comissão Examinadora - Orientador

Dr. Geovane Webler
Coordenador do Curso

João Michel Machado de Oliveira
FAHOR – Faculdade Horizontina

**Horizontina - RS
2023**

A Deus,

Pela luz que guiou cada palavra, pela força que sustentou cada esforço e pelo amor que deu propósito a este estudo, expressei minha mais profunda gratidão. Em cada desafio, senti Sua presença; em cada conquista, reconheci Sua graça. Que este trabalho reflita não apenas o meu empenho, mas também a Sua glória, pois sem a Sua bênção, nada seria possível.

AGRADECIMENTOS

Meus sinceros agradecimentos vão para aqueles que foram pilares na construção deste estudo.

Minha gratidão inicial é para Deus, cuja presença me guiou com saúde e sabedoria, me sustentando nos momentos de desafio e fortalecendo meu compromisso com a pesquisa que tanto admiro.

À equipe técnica da usina solar, meu reconhecimento pelo suporte excepcional na coleta de dados cruciais, sem os quais este trabalho não teria a mesma fundamentação. A dedicação e profissionalismo de vocês foram o alicerce para as análises precisas que aqui apresento.

À comunidade acadêmica da FAHOR - Faculdade de Horizontina, agradeço pelo ambiente enriquecedor e pelo suporte intelectual. Cada discussão e orientação recebida foi uma peça chave que agregou valor inestimável a este trabalho.

Ao meu orientador, Me. Rodrigo Bastos, minha eterna gratidão por sua visão crítica e apoio incansável. Sua habilidade de me direcionar sem ofuscar meu caminho foi fundamental para que eu alcançasse este marco.

E à minha família e amigos, vocês são a base de tudo. Cada palavra de incentivo e cada gesto de paciência durante essa jornada são o motivo pelo qual posso celebrar esta conquista hoje.

Cada um de vocês deixou uma marca indelével neste trabalho, que é tanto meu quanto de vocês. Estou profundamente grato por cada contribuição que tornou essa jornada possível.

"O progresso é impossível sem mudança; e aqueles que não conseguem mudar suas mentes não conseguem mudar nada."

George Bernard Shaw

RESUMO

Este trabalho apresenta o desenvolvimento e a avaliação de um modelo de previsão de produção de energia solar baseado em dados meteorológicos. Utilizando técnicas de Aprendizado de Máquina e Redes Neurais de Longa Memória Curta (LSTM), o estudo se concentra em aprimorar a precisão e confiabilidade das previsões de geração em usinas solares. A pesquisa foi motivada pela necessidade de métodos robustos para gerenciar a variabilidade na produção de energia solar devido a fatores climáticos, visando estabilizar a oferta de energia e otimizar a eficiência econômica. Os resultados evidenciam que o modelo LSTM, devidamente treinado e refinado com um conjunto significativo de dados de três usinas solares, pode prever com sucesso a produção de energia, levando em consideração as variações temporais e espaciais dos dados meteorológicos. A seleção de características, guiada pela técnica LASSO, e a engenharia de recursos desempenharam um papel crítico na melhoria do modelo, destacando a radiação global e a temperatura do ar como fatores influentes. O trabalho reafirma a importância crescente de modelos preditivos na gestão de recursos de energia renovável e destaca o potencial dos métodos de aprendizado de máquina para avançar no campo de previsão de geração de energia solar, contribuindo assim para esforços globais de sustentabilidade e eficiência energética.

Palavras-chave: Energia fotovoltaica, Inteligência artificial, LSTM, Dados meteorológicos

LISTA DE FIGURAS

Figura 1 – Evolução da fonte solar fotovoltaica no Brasil.....	14
Figura 2 – Ilustração efeito fotovoltaico.....	17
Figura 3 – Tipos de placas solares comercialmente disponíveis.....	18
Figura 4 – Sistema fotovoltaico <i>on grid</i>	20
Figura 5 – Sistema fotovoltaico <i>off grid</i>	21
Figura 6 – Sistema fotovoltaico híbrido.....	22
Figura 7 – Mapa anual da quantidade total de radiação solar direta normal.....	23
Figura 8 – Diagrama descritivo do processo de fabricação do modelo.....	34
Figura 9 – Metodologia de treinamento do modelo de predição.....	36
Figura 10 – Gráfico de correlação das variáveis climáticas com a produção de energia solar.....	42
Figura 11 – Gráfico de Importância das características usando técnica LASSO.....	45
Figura 12 – Comparação de valores de predição do primeiro experimento.....	47
Figura 13 – Código do décimo treinamento.....	50
Figura 14 – Comparação de valores de predição do décimo experimento.....	51
Figura 15 – Exposição dos dados de produção das usinas coletadas.....	52
Figura 16 – Comparação de valores de predição do experimento com todas usinas.....	53
Figura 17 – Comparação de valores real com predição de todas usinas por média ponderada.....	54
Figura 18 – Comparação da predição de teste unitário.....	55

SUMÁRIO

1	INTRODUÇÃO	9
1.1	TEMA	10
1.2	DELIMITAÇÃO DO TEMA.....	10
1.3	PROBLEMA DE PESQUISA	10
1.4	HIPÓTESES.....	11
1.5	OBJETIVOS	11
1.5.1	Objetivo geral	11
1.5.2	Objetivos específicos	11
1.6	JUSTIFICATIVA	11
2	REFERENCIAL TEÓRICO	14
2.1	ENERGIA RENOVÁVEL E TRANSIÇÃO ENERGÉTICA	14
2.2	ENERGIA SOLAR E TECNOLOGIA FOTOVOLTAICA.....	16
2.2.1	Tipos de sistemas fotovoltaicos e seus componentes	19
2.3	VARIABILIDADE NA GERAÇÃO DE ENERGIA SOLAR	22
2.4	INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	24
2.4.1	Redes Neurais Recorrentes (RNN)	26
2.4.2	Long Short-Term Memory (LSTM)	27
2.5	MÉTODOS ESTATÍSTICOS E AVALIAÇÃO DE IA	27
2.5.1	Normalização	27
2.5.2	Seleção de recursos	28
2.5.3	Métricas de desempenho	29
2.6	APLICAÇÕES DE IA NA ENERGIA SOLAR	30
3	METODOLOGIA	33
3.1	CARACTERÍSTICAS DA PESQUISA.....	33
3.2	MÉTODOS E TÉCNICAS UTILIZADAS	33
3.3	MÉTODOS DE PROCEDIMENTOS.....	33
3.4	TÉCNICAS DE COLETA DE DADOS	36
3.4.1	Dados meteorológicos	36
3.4.2	Dados de produção solar	37
3.5	CONSIDERAÇÕES ÉTICAS.....	37
3.6	ANÁLISE DE DADOS.....	38
3.7	MATERIAL E EQUIPAMENTOS	39
4	ANÁLISE E DISCUSSÃO DOS RESULTADOS	40
4.1	SELEÇÃO DO MODELO.....	40
4.2	INTEGRAÇÃO DOS DADOS	40
4.3	RELAÇÃO ENTRE VARIÁVEIS METEOROLÓGICAS E A PRODUÇÃO DE ENERGIA SOLAR.....	41
4.4	ENGENHARIA DE RECURSOS	43
4.5	SELEÇÃO DE RECURSOS	44
4.6	TREINAMENTO	46
4.6.1	Desenvolvimento inicial do modelo	46
4.6.2	Otimização e refinamento do modelo	47
4.6.3	Expansão e testes com múltiplas usinas	51
	CONSIDERAÇÕES FINAIS	56
	REFERÊNCIAS	58

1 INTRODUÇÃO

O atual cenário energético global é marcado por um preocupante paradoxo: enquanto o consumo de energia continua a aumentar, os recursos de combustíveis fósseis, principal fonte desta demanda, são finitos e estão se esgotando. De acordo com a *Statistical Review of World Energy* da *British Petroleum* (2019), se as reservas comprovadas não crescessem e se o consumo permanecesse constante nos níveis de 2019, levaria cerca de 48 anos para esgotar essas reservas. O transporte rodoviário contribui significativamente para este cenário, sendo responsável por cerca de 72% da energia consumida para a movimentação mundial e mais de 80% das emissões de CO₂ (BILGEN, 2014).

Este cenário alarmante ressalta a urgência em buscar alternativas sustentáveis para reduzir a dependência dos combustíveis fósseis, como a implementação de veículos elétricos (EVs) e a ampliação do uso da energia solar. Os EVs apresentam grande potencial para reduzir as emissões de gases do efeito estufa provenientes do transporte rodoviário. Contudo, para uma implementação em larga escala é necessário um suprimento constante de energia renovável, papel que a energia solar está apta a desempenhar.

Embora o interesse pela energia solar esteja crescendo, sua produção apresenta desafios, como a variabilidade na geração causada por fatores climáticos que causam interrupções no processo de geração, o que pode levar a instabilidade e insegurança no sistema elétrico. Diante deste obstáculo, a previsão da produção energética fotovoltaica emergiu como um instrumento fundamental para a administração energética. Conforme estudo de Wan *et al.* (2015), uma previsão precisa dessa geração pode otimizar a eficiência do planejamento energético, além de diminuir a dependência de fontes de energia suplementares.

Neste contexto, a Inteligência Artificial (IA) surge como uma aliada importante, possibilitando o desenvolvimento de sistemas para otimizar a produção de energia solar. Assim, o desenvolvimento de um sistema de previsão baseado em aprendizado de máquina para usinas de energia solar torna-se uma solução promissora para otimizar o planejamento e a geração de energia.

No entanto, apesar dessa oportunidade promissora, a pesquisa na intersecção entre IA e energia solar ainda está em um estágio inicial. As técnicas de previsão atuais dependem de grandes volumes de dados e possuem dificuldade em lidar com

mudanças abruptas nas condições climáticas. Neste contexto, surge a necessidade de pesquisar métodos mais robustos e flexíveis para a previsão da produção de energia solar.

Este trabalho explora os benefícios e desafios do desenvolvimento desse sistema e como ele pode revolucionar o setor de energia renovável. Deste modo torna-se perceptível que a integração da IA na geração de energia solar pode levar a melhorias significativas em eficiência e sustentabilidade.

1.1 TEMA

Estudo sobre inteligência artificial aplicado à produção de energia solar.

1.2 DELIMITAÇÃO DO TEMA

Desenvolvimento de uma inteligência artificial com o objetivo de prever a produção de uma usina solar com base em dados meteorológicos.

1.3 PROBLEMA DE PESQUISA

A intermitência da geração de energia solar, associada à variabilidade dos fatores meteorológicos, representa um obstáculo significativo para sua integração efetiva na rede elétrica. A inconstância na incidência solar resulta em uma produção energética flutuante, dificultando a gestão da oferta e demanda de energia e elevando a complexidade do planejamento energético. O problema de pesquisa a ser abordado neste estudo, portanto, é a determinação de uma metodologia preditiva que possa fornecer estimativas precisas da produção de energia solar, minimizando a incerteza e potencializando a confiabilidade da fonte.

A energia solar, uma das fontes mais promissoras de energia limpa, está no cerne dessa transição. Entretanto, sua variabilidade intrínseca, decorrente da intermitência dos padrões meteorológicos, impõe um problema significativo: a imprevisibilidade de sua geração. A questão central deste estudo, portanto, é: Como é possível estimar a produção de uma usina solar de forma a melhorar a sua previsibilidade?

1.4 HIPÓTESES

- a) Aplicação de Inteligência Artificial (IA) em sistemas de previsão de produção de energia solar irá melhorar significativamente a eficiência da geração de energia, contribuindo para um fornecimento mais estável e confiável;
- b) A previsão de produção de energia solar pode ser realizada com uma precisão admissível apenas com dados climáticos;
- c) O uso de sistemas de previsão de produção de energia solar baseados em IA aumentará a confiança dos investidores e dos reguladores do setor energético na energia solar, resultando em uma adoção mais ampla dessa fonte de energia.

1.5 OBJETIVOS

1.5.1 Objetivo geral

Desenvolver um modelo de predição de produção de energia solar usando dados meteorológicos baseado em aprendizado de máquina.

1.5.2 Objetivos específicos

- a) Coletar dados para análise de variáveis;
- b) Desenvolver uma especificação de modelo de aprendizado de máquina a ser utilizado;
- c) Analisar e filtrar dados correlacionados;
- d) Treinar o modelo em cima dos dados coletados;
- e) Gerar testes e iterações com o modelo desenvolvido;
- f) Validar o modelo a partir de um teste unitário com dados reais coletados em uma usina da região.

1.6 JUSTIFICATIVA

A transição para uma matriz energética sustentável é uma das maiores necessidades do século XXI, uma mudança fundamental para combater as mudanças climáticas e garantir a segurança energética a longo prazo. A Agência Internacional de Energia Renovável (IRENA) (2020), em seu relatório *Global Renewables Outlook*, destaca a importância da transição para energias renováveis, eficiência e eletrificação. Segundo a agência, tais medidas podem impulsionar o desenvolvimento econômico

ao alinhar os investimentos na área energética com a já existente necessidade de refrear o aquecimento global.

Nesse contexto, a energia solar surge como uma das principais possíveis soluções, devido à sua natureza renovável, abundância e potencial para reduzir as emissões de gases de efeito estufa. A energia solar tornou-se uma opção competitiva graças à queda drástica nos custos de produção. No século XXI, os custos caíram em 90%, com a maior parte dessa redução ocorrendo na última década. Atualmente, a energia solar é mais barata do que qualquer outra tecnologia em 16 estados dos Estados Unidos assim como em outros mercados ao redor do mundo. A expectativa é que as inovações tecnológicas permitam uma redução adicional de 15% a 25% nos custos nos próximos 10 anos. Isso tornará a energia solar competitiva a preços de atacado na maioria dos mercados de energia até 2030 (MACKENZIE, 2021).

No entanto, a variação na geração de energia solar, causada por fatores como as condições climáticas e a eficiência dos painéis solares, representa um desafio significativo. A previsibilidade da produção de energia é crucial para garantir a estabilidade e a segurança do sistema de energia, bem como para otimizar a eficiência econômica das usinas solares. Portanto, é de suma importância desenvolver métodos confiáveis de previsão da geração de energia solar.

A Inteligência Artificial (IA), com sua capacidade de aprendizado de máquina e processamento de dados, oferece uma ferramenta poderosa para atender a essa necessidade. A IA tem o potencial de transformar radicalmente a forma como prevemos e gerenciamos a geração de energia, melhorando significativamente a eficiência e a confiabilidade das usinas solares (KHAN *et. al.*, 2018).

No entanto, apesar do potencial da IA, a pesquisa nesta área ainda é escassa e incipiente no Brasil e América Latina, e muitos desafios permanecem a ser superados (WENTZ, 2021). Os desafios específicos associados às condições climáticas da região, ao perfil de consumo energético e ao estágio de desenvolvimento da infraestrutura de energia renovável destacam a necessidade urgente de mais estudos direcionados. Portanto, o desenvolvimento deste trabalho com foco na aplicação da IA na previsão da produção de energia solar é de suma importância para superar esses desafios e avançar o conhecimento nesta área.

O sucesso neste campo de pesquisa tem potencial para melhorar a eficiência e a confiabilidade das usinas solares, facilitando a adoção desta fonte de energia e contribuindo com os esforços globais na redução das emissões de gases de efeito

estufa. Além disso, os resultados podem ter um impacto direto na indústria de energia solar permitindo, por exemplo, um planejamento mais eficaz e a otimização dos investimentos em infraestrutura solar.

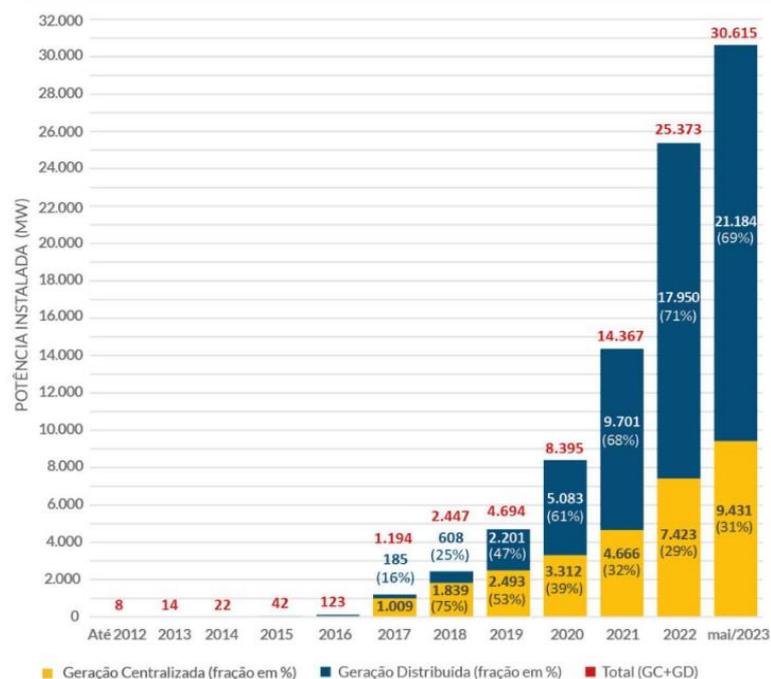
2 REFERENCIAL TEÓRICO

2.1 ENERGIA RENOVÁVEL E TRANSIÇÃO ENERGÉTICA

O movimento global de transição energética, voltado para a substituição de fontes de energia não renováveis e altamente poluentes por alternativas renováveis limpas e sustentáveis é impulsionado por diversos fatores. O crescente reconhecimento da necessidade de mitigar as mudanças climáticas, a contínua redução dos custos associados às energias renováveis e a aspiração de fortalecer a segurança energética estão entre os fatores mais proeminentes.

As energias renováveis desempenham um papel crucial na mitigação das mudanças climáticas. Segundo o estudo "O papel das energias renováveis na transformação energética global" de Gielen *et al.* (2019), fontes renováveis, como solar, eólica, hidrelétrica e geotérmica, tiveram um crescimento expressivo recentemente, com as energias renováveis constituindo mais de 25% da geração global de eletricidade. No cenário brasileiro também pode-se notar um aumento expressivo da utilização de energia solar nos últimos anos, conforme pode-se ver na figura 1:

Figura 1 – Evolução da fonte solar fotovoltaica no Brasil



Fonte: ABSOLAR, 2023

Este crescimento é resultado da combinação de diversos fatores, incluindo a redução dos custos das tecnologias de energia renovável, políticas favoráveis e um aumento na demanda pública por energia limpa. De acordo com a Agência Internacional de Energia Renovável, os custos de energia solar e eólica caíram significativamente nos últimos anos e devem continuar a cair no futuro (IRENA, 2019). Além disso, a energia renovável apresenta um potencial significativo para superar desafios relacionados ao acesso à energia nos países em desenvolvimento, onde a ausência de fontes confiáveis de eletricidade representa uma barreira substancial ao progresso econômico.

No entanto, ainda persistem obstáculos para a adoção em larga escala da energia renovável. A intermitência das fontes de energia renováveis ainda é um problema, assim como a necessidade de uma infraestrutura de rede robusta, a estrutura da rede precisa estar pronta para funcionar durante o dia, tanto em situações de alta demanda com ou sem o auxílio das fontes fotovoltaicas, quanto à noite, quando essas fontes não estarão disponíveis. (VIAN, 2021).

Mesmo com a redução de custos, as tecnologias de energia renovável ainda podem ser mais caras que as fontes convencionais de energia em alguns contextos. Os custos de instalação inicial dessas tecnologias podem ser significativamente mais altos do que o das fontes de energia convencionais (LI; YANG; DEZFULI, 2020). Isso se deve em parte à necessidade de infraestrutura adicional, como painéis solares ou turbinas eólicas, bem como à necessidade de sistemas de armazenamento de energia para garantir um fornecimento contínuo de energia.

Além disso, a eficiência das tecnologias de energia renovável pode ser influenciada por fatores geográficos e climáticos. Por exemplo, a energia solar pode ser menos eficaz em regiões com pouca luz solar, enquanto a energia eólica pode ser menos eficaz em áreas com ventos inconsistentes (BORENSTEIN, 2012). Isso pode aumentar os custos de energia por unidade em comparação com as fontes de energia convencionais, que podem ser mais consistentes.

Frente a esses desafios, a Inteligência Artificial (IA) tem se mostrado uma ferramenta poderosa. Ela tem o potencial de otimizar a produção de energia, prever a demanda e oferecer soluções para a intermitência das fontes renováveis, garantindo uma melhor previsibilidade. Contudo, além da evolução dos meios tecnológicos, será necessária uma estratégia integrada, que inclua políticas de incentivo, investimento em tecnologias de armazenamento de energia e um compromisso contínuo com a

pesquisa e desenvolvimento para melhorar a eficiência e acessibilidade da energia renovável.

2.2 ENERGIA SOLAR E TECNOLOGIA FOTOVOLTAICA

O Sol representa uma colossal e interminável reserva de energia, essencial para a sustentação da vida no nosso planeta. Várias formas de energia renovável têm suas origens ligadas ao Sol, abrangendo tanto o uso direto da sua energia para aquecimento ou produção de eletricidade, quanto o uso indireto, que se manifesta por meio da energia gerada pelo vento, ondas, correntes de água, além da energia derivada de plantas e animais (PHILIPPI JR, 2016).

A energia solar é uma fonte de energia renovável e sustentável que gera eletricidade por meio do uso de radiação solar. O processo fundamental para a conversão da energia solar em energia elétrica é conhecido como efeito fotovoltaico, que ocorre quando a luz solar atinge um material semicondutor, como o silício, e gera uma corrente elétrica.

O processo envolve dois fenômenos físicos: fotólise de água e o efeito fotovoltaico. A fotólise ocorre quando a luz solar é absorvida pelo material semicondutor, quebrando as ligações químicas e liberando elétrons que são canalizados para produzir corrente elétrica. O efeito fotovoltaico é a conversão direta da luz solar em energia elétrica, o que é alcançado através dos elétrons liberados pela fotólise (PINHO; GALDINO, 2014).

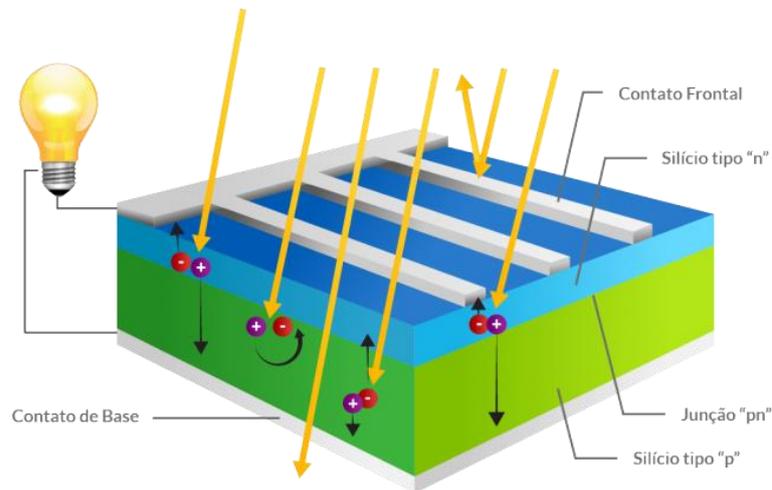
Quando os fótons da luz solar atingem as células solares, eles fornecem energia suficiente para deslocar elétrons na estrutura atômica do silício, um processo conhecido como excitação de elétron. Cada elétron deslocado deixa uma "lacuna" em seu rastro. Os elétrons mais próximos pulam para preencher essa lacuna, criando outra vão no processo.

Para que a corrente elétrica seja efetivamente produzida, os elétrons não podem simplesmente voltar aos seus lugares originais. Para evitar isso, as células solares contêm uma camada de junção P-N, que funciona como uma fronteira unidirecional para os elétrons. Isso significa que os elétrons excitados podem passar facilmente através dessa fronteira em uma direção, mas não na outra.

Portanto, quando a luz do Sol "empurra" os elétrons, eles são forçados a se mover em uma direção através da célula solar, criando um fluxo de elétrons - o que é, essencialmente, uma corrente elétrica. Esta eletricidade produzida pode então ser

usada imediatamente, armazenada para uso posterior, ou mesmo alimentada de volta à rede elétrica. Uma representação visual desse processo pode ser observada na figura 2:

Figura 2 – Ilustração efeito fotovoltaico



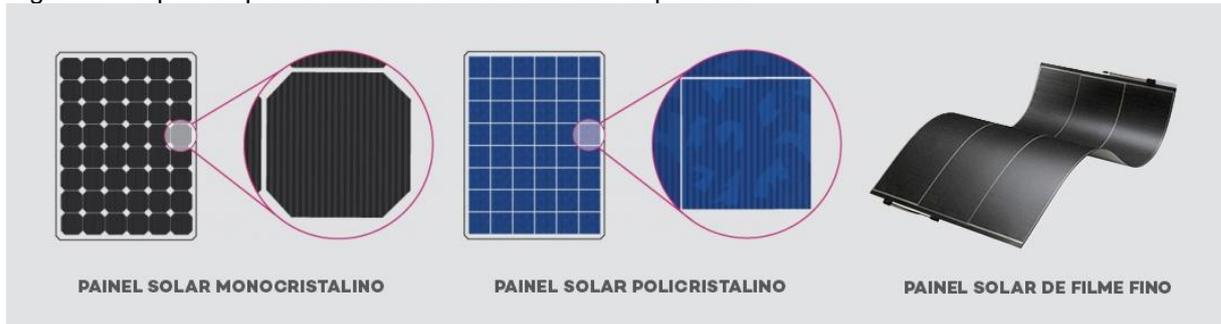
Fonte: adaptado de Campos *et al.*, 2010

A energia solar é gerada por meio do uso de células solares, também conhecidas como células fotovoltaicas. Conforme demonstrado na Figura 3, existem diversos tipos de células solares, cada uma com suas vantagens, desvantagens e aplicações distintas.

As células solares de silício cristalino são as mais comuns e são normalmente divididas em dois tipos: monocristalinas e policristalinas. As células solares monocristalinas são feitas de um único cristal de silício, tornando-as altamente eficientes na conversão da luz solar em energia elétrica. No entanto, elas tendem a ser mais caras devido ao seu processo de fabricação complexo. Já as células solares policristalinas, feitas de vários cristais de silício, são menos eficientes, mas também mais baratas (TEIXEIRA; RAMOS; AGUIAR, 2019).

Outra categoria importante é a das células solares de filme fino. Essas células utilizam materiais semicondutores mais finos, o que as torna mais leves e flexíveis que as células de silício cristalino. Embora a eficiência das células de filme fino seja geralmente menor, elas podem ser mais baratas de produzir e mais versáteis em termos de aplicação, podendo ser aplicadas em superfícies curvas ou flexíveis.

Figura 3 – Tipos de placas solares comercialmente disponíveis



Fonte: NeoSolar, 2022a

Finalmente, as células solares de perovskita representam uma das tecnologias emergentes mais promissoras na área de energia solar. As perovskitas são um tipo de material cristalino que tem a capacidade de absorver luz em uma ampla gama de comprimentos de onda, o que as torna potencialmente muito eficientes. Embora ainda haja desafios a serem superados em termos de durabilidade e estabilidade, as células solares de perovskita têm o potencial de revolucionar o campo da energia solar devido à sua eficiência e custo de produção reduzido (RAPHAEL *et al.*, 2018).

Uma matriz solar é um conjunto de painéis solares ou módulos solares que são conectados e trabalham juntos em um sistema fotovoltaico para gerar eletricidade. Os painéis solares são montados em matrizes para capturar a energia solar em uma área grande, onde podem ser instalados.

Existem diversos tipos de sistemas fotovoltaicos, cada um com características específicas e aplicações distintas. Os sistemas fotovoltaicos isolados (*off-grid*), por exemplo, são aqueles nos quais a totalidade da energia gerada é diretamente consumida por circuitos elétricos estritamente vinculados ao próprio sistema gerador. Por outro lado, os sistemas fotovoltaicos interligados à rede (*on-grid*) se caracterizam por sua conexão direta com a rede de distribuição de uma determinada concessionária de energia. Isso permite que a energia produzida possa ser distribuída e utilizada além do sistema gerador em si (XIAO; *et al.*, 2023).

Existe ainda uma terceira categoria, o sistema híbrido, o qual é resultado da integração dos sistemas fotovoltaicos isolados e interligados à rede. Esse sistema combina as funcionalidades dos dois anteriores, permitindo um uso mais versátil e otimizado da energia solar. Em sua maior totalidade os sistemas mais utilizados são os sistemas ligados na rede, os quais distribuem a concessionária de energia, esse fato se deve a dificuldade de armazenar esta energia de forma economicamente viável.

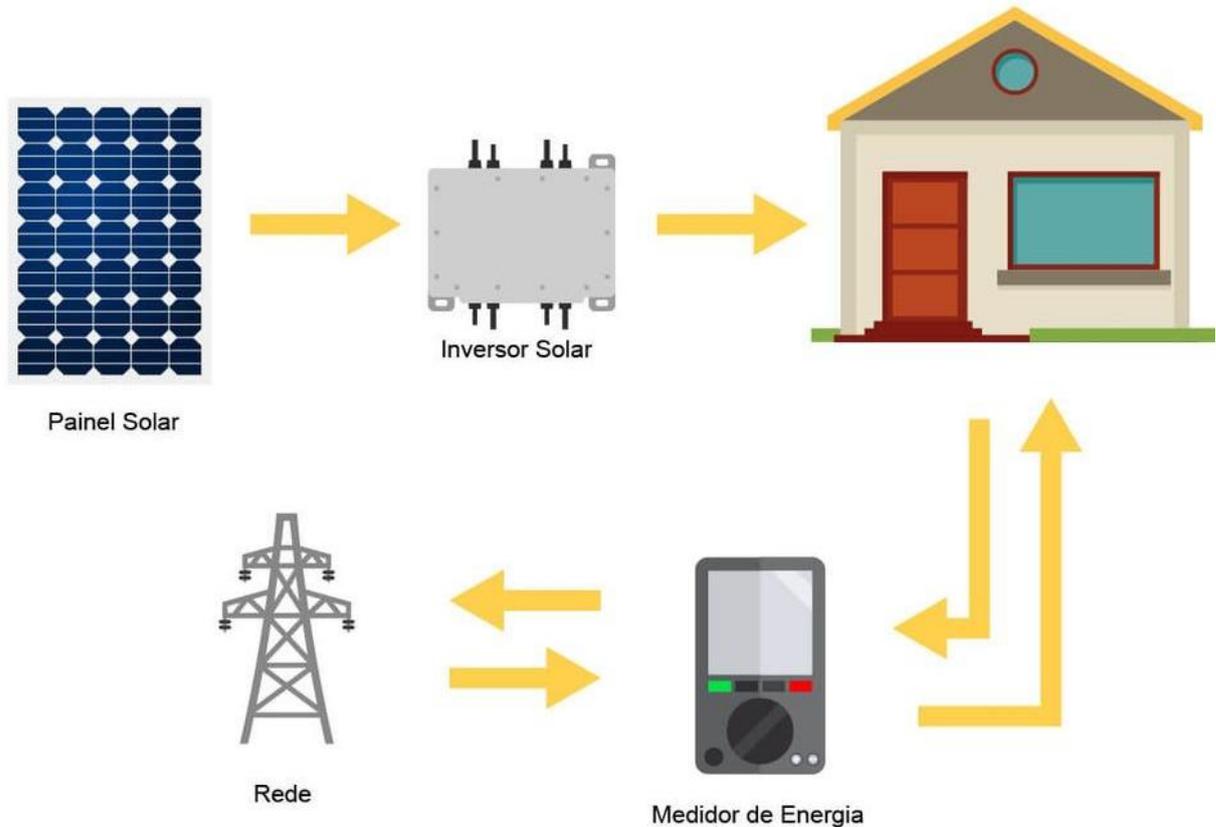
A eficiência de uma célula solar, ou seja, a quantidade de luz solar que ela pode converter em eletricidade, é um aspecto importante da tecnologia fotovoltaica. Este aspecto pode ser afetado por vários fatores, incluindo o tipo de material semicondutor usado, a qualidade da fabricação da célula e as condições sob as quais a célula é operada. Por exemplo, conforme a célula sofre um aumento na temperatura, a tensão produzida por ela pode diminuir (PINHO; GALDINO, 2014).

2.2.1 Tipos de sistemas fotovoltaicos e seus componentes

A transformação da energia solar em eletricidade por meio da tecnologia fotovoltaica pode ser realizada de diferentes maneiras, dependendo do tipo de sistema adotado. Os sistemas fotovoltaicos podem ser classificados em três categorias principais: *on-grid*, *off-grid* e híbridos, cada um com seus componentes específicos e aplicações particulares.

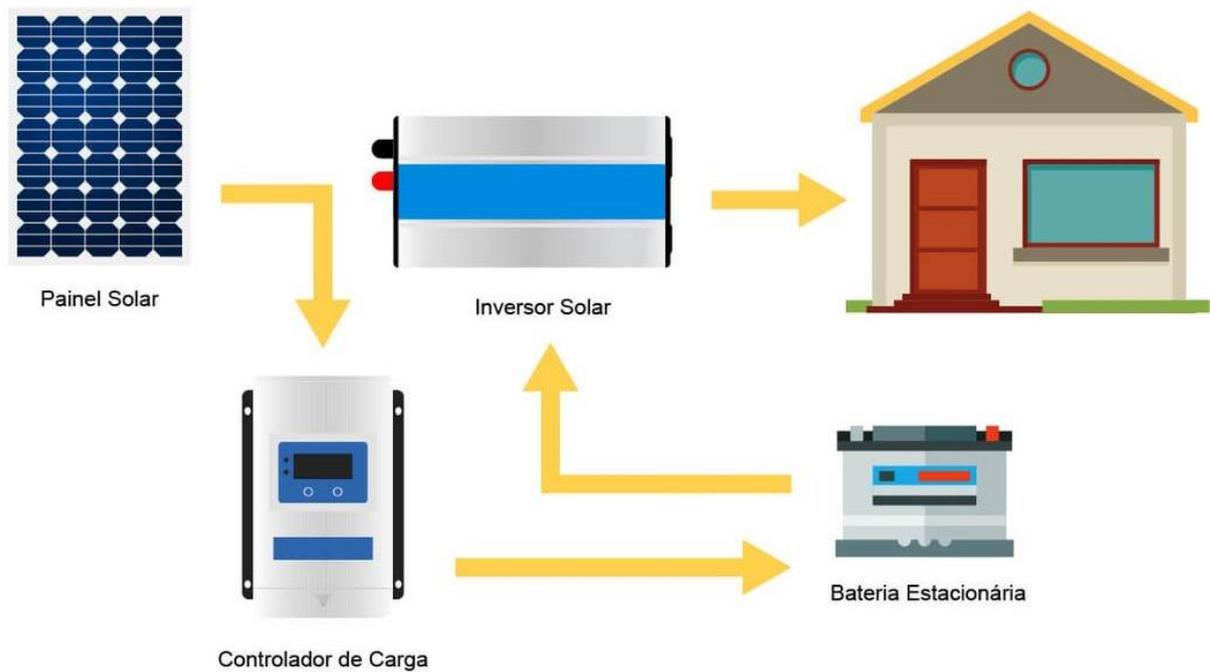
Os sistemas *on-grid*, ou conectados à rede, são caracterizados pela sua integração direta com a rede elétrica convencional. Nestes sistemas, os painéis solares, junto com inversores, os quais convertem a corrente contínua gerada pelas células solares em corrente alternada utilizável, formam a espinha dorsal do sistema, conforme retratado na figura 4. Os medidores bidirecionais são componentes essenciais, pois permitem não só o consumo de energia da rede quando necessário, mas também o redirecionamento do excesso de energia solar de volta à rede, criando um sistema de compensação. Estes sistemas são eficazes em áreas urbanas, onde a conexão com a rede é prontamente disponível, facilitando assim a distribuição e utilização da energia gerada.

Figura 4 – Sistema fotovoltaico *on grid*



Fonte: NeoSolar, 2022b

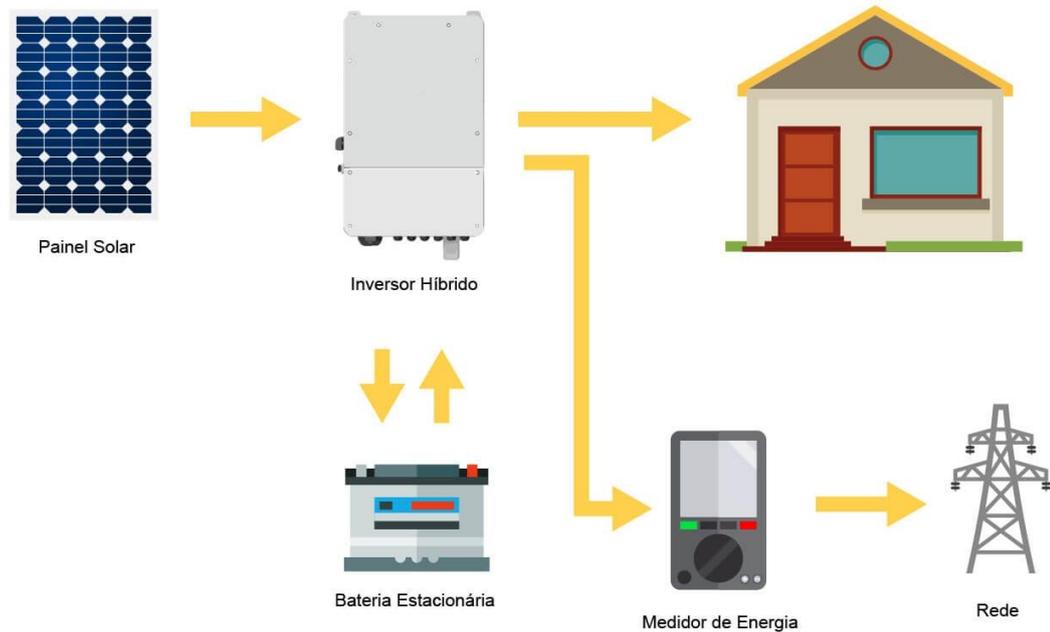
Já os sistemas *off-grid* operam de maneira autônoma, sem conexão com a rede elétrica externa. Eles são particularmente úteis em locais remotos, onde a conexão com a rede pode ser inviável ou muito cara. Além dos painéis solares e inversores, esses sistemas necessitam de baterias para armazenar a energia produzida para uso durante períodos sem sol. Controladores de carga são também componentes vitais, protegendo as baterias de serem sobrecarregadas ou descarregadas excessivamente, prolongando assim sua vida útil. Uma representação desse sistema pode ser vista na figura 5 a seguir:

Figura 5 – Sistema fotovoltaico *off grid*

Fonte: NeoSolar, 2022b

Os sistemas híbridos, por sua vez, combinam características dos sistemas *on-grid* e *off-grid*. Como pode ser visto na figura 6, eles estão conectados à rede, mas também possuem baterias para armazenamento de energia. Essa configuração proporciona uma flexibilidade notável, permitindo o uso da energia solar armazenada durante quedas de energia ou períodos de alta demanda, enquanto ainda se mantém a opção de usar ou enviar energia de volta à rede. Tais sistemas podem também incluir outras fontes de energia renovável, como eólica ou hidroelétrica, para uma produção de energia mais diversificada e confiável.

Figura 6 – Sistema fotovoltaico híbrido



Fonte: NeoSolar, 2022b

Em suma, a escolha do sistema fotovoltaico mais apropriado depende de uma variedade de fatores, incluindo localização, custo, necessidades energéticas e objetivos de sustentabilidade. Independentemente do tipo de sistema escolhido, a integração da energia solar na matriz energética representa um passo significativo em direção a um futuro mais sustentável.

2.3 VARIABILIDADE NA GERAÇÃO DE ENERGIA SOLAR

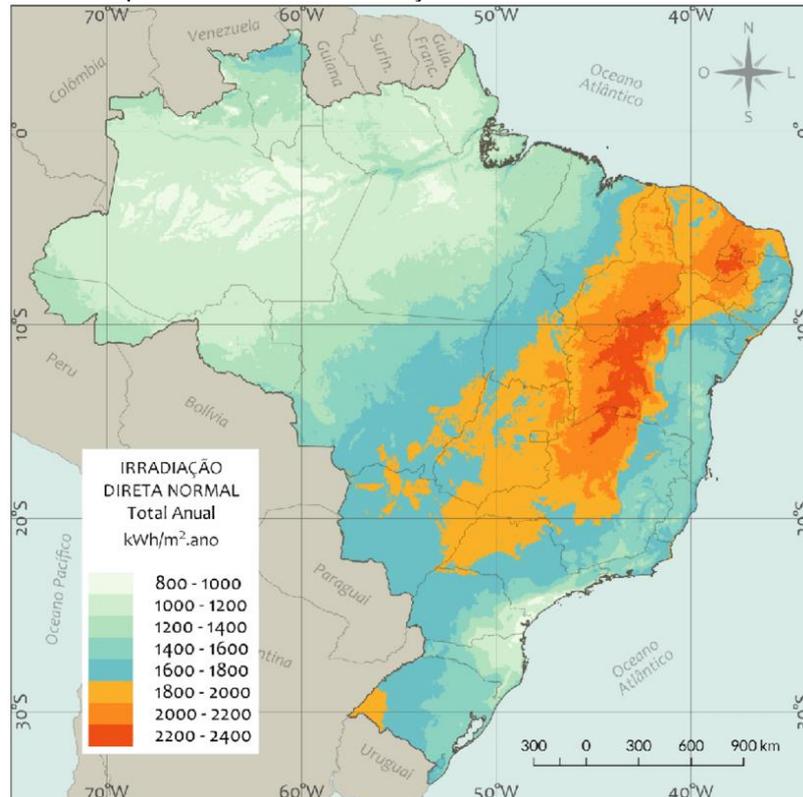
A geração de energia solar, uma tecnologia em rápido crescimento, tem o potencial de fornecer uma quantidade significativa de energia renovável para o mundo. No entanto, um desafio significativo que precisa ser abordado para maximizar o potencial desta fonte de energia é a variabilidade na geração de energia solar.

Esta variabilidade pode ser atribuída a uma série de fatores. Primeiramente, a intensidade da luz solar não é constante, variando ao longo do dia e do ano. Por exemplo, durante o pico do verão em locais próximos ao equador, a intensidade da luz solar pode ser extremamente alta, permitindo uma geração de energia solar robusta. No entanto, durante o inverno ou em locais mais distantes do equador, a intensidade da luz solar pode ser muito menor, resultando em uma geração de energia solar reduzida. Além disso, condições climáticas, como a presença de nuvens, neblina

ou poeira, podem diminuir a quantidade de luz solar que atinge os painéis solares, afetando a geração de energia (GÓRECKIET; DABROWSKI; KRAC, 2021).

Isso pode ser notado de forma mais expressiva no Brasil ao considerar a variabilidade da irradiação solar em todo o seu território, conforme demonstrado na figura 7:

Figura 7 – Mapa anual da quantidade total de radiação solar direta normal



Fonte: Pereira *et al.*, 2017

A eficiência dos painéis solares também desempenha um papel na variabilidade da geração de energia solar. Diferentes tipos de painéis solares têm diferentes níveis de eficiência, e a eficiência de um painel solar pode diminuir ao longo do tempo devido ao desgaste e à degradação. Além do fator tempo, a pesquisa de Zainuddin *et al.* (2019) aponta que o acúmulo de sujeira e as camadas de poeira podem afetar em até cerca de 22% a eficiência de conversão de energia solar em energia elétrica.

Incorporando o estudo de Zafarani *et al.* (2018) neste contexto, determinamos que temperatura, radiação solar, velocidade do vento e umidade relativa são os principais parâmetros meteorológicos relevantes para previsão da geração de energia fotovoltaica. Sendo esses fatores diretos na conversão da energia solar em

eletricidade, sua inclusão nas previsões pode aprimorar notavelmente a eficiência e a confiabilidade dos sistemas fotovoltaicos.

Outro fator que contribui para a variabilidade é a orientação e inclinação dos painéis solares. A orientação ideal para painéis solares é voltada para o sul no hemisfério norte e para o norte no hemisfério sul. A inclinação ideal depende da latitude da localização (POLIKARPOVA, *et. al.*, 2021). Painéis solares que não estão idealmente orientados ou inclinados podem não receber a quantidade máxima de luz solar, resultando em uma menor geração de energia.

A variabilidade na geração de energia solar tem implicações significativas para a integração da energia solar na rede elétrica. A geração de energia solar variável pode levar a desequilíbrios na oferta e demanda de energia, o que pode causar instabilidade na rede elétrica (SOTO, 2022). Portanto, é importante desenvolver estratégias para gerenciar a variabilidade na geração de energia solar, como o armazenamento de energia, a previsão do tempo e a geração de energia solar distribuída.

Nesse contexto, a inteligência artificial tem o potencial de desempenhar um papel crucial na determinação e na gestão da variabilidade na geração de energia solar. Com sua capacidade de processar grandes volumes de dados e identificar padrões complexos, a IA pode ajudar a entender melhor os fatores que contribuem para a variabilidade na geração de energia solar e desenvolver estratégias para mitigá-la. Auxiliando assim na administração dos painéis e na determinação das melhores condições para a produção de energia solar.

2.4 INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

A Inteligência Artificial (IA) é um campo de estudo da ciência da computação que procura replicar e emular a inteligência humana em máquinas. A IA, sendo um campo amplo, engloba o aprendizado de máquina como uma de suas sub categorias essenciais, onde o foco reside em construir sistemas que aprendem e fazem previsões ou decisões sem serem explicitamente programados para tal. Desde a sua concepção na década de 1950, a IA tem se transformado, evoluindo de programas simples que mimetizam comportamentos humanos para sistemas complexos capazes de executar tarefas que, em geral, necessitam de inteligência humana, como reconhecimento de voz, aprendizado, planejamento e percepção (RUSSELL, NORVIG, 2016).

Em particular, o aprendizado de máquina (*Machine Learning*) emergiu como um dos subcampos mais influentes e dinâmicos da IA, possibilitando a melhoria do desempenho das máquinas por meio do aprendizado de dados, eliminando a necessidade de programação explícita.

Aprendizado de Máquina é uma técnica de avaliação de dados que permite a criação automática de modelos analíticos. Como ramo da IA, ele baseia-se na premissa de que os sistemas podem aprender a partir de dados, identificar padrões e tomar decisões com mínimo envolvimento humano.

A expansão desse campo é resultado da crescente complexidade dos desafios que requerem soluções computacionais e do aumento exponencial no volume e velocidade dos dados gerados em uma variedade de setores. Esses fatores têm incentivado o desenvolvimento de ferramentas computacionais cada vez mais sofisticadas e autônomas, exigindo cada vez menos intervenção humana para adquirir conhecimento. Como consequência, o campo do aprendizado de máquina tem experimentado progresso e expansão contínuos (FACELI; *et. al.*, 2021).

O aprendizado de máquina se subdivide em três categorias principais: supervisionado, não supervisionado e por reforço. O aprendizado supervisionado envolve treinar modelos com dados previamente rotulados. Por outro lado, o aprendizado não supervisionado emprega dados não rotulados, permitindo ao modelo detectar padrões subjacentes.

O aprendizado por reforço opera numa lógica de tentativa e erro, onde um agente é recompensado ou punido com base nas ações realizadas. Enquanto o aprendizado supervisionado pressupõe um conhecimento prévio dos resultados desejados, o aprendizado não supervisionado explora as estruturas intrínsecas dos dados para agrupá-los. Existe ainda o aprendizado semi-supervisionado, que combina elementos dos dois primeiros e é comumente utilizado quando apenas parte dos dados estão rotulados. Já o aprendizado por reforço baseia-se na ideia de receber feedbacks que fortalecem ou desestimulam determinadas características do modelo (LENZ; NEUMANN; SANTARELLI, 2021).

O Aprendizado de Máquina é especialmente adequado para tarefas de previsão, graças à sua capacidade de aprender padrões a partir de dados históricos e aplicá-los a novos dados, além de encontrar padrões ocultos neles, o que por sua vez ajuda na tomada de decisão mais eficaz (BASHA, 2019). Diversos algoritmos, como os de Redes Neurais Artificiais, Máquinas de Vetores de Suporte e Modelos de

Regressão, têm sido amplamente utilizados para tarefas preditivas, cada um com suas vantagens e limitações específicas.

Assim, o Aprendizado de Máquina se torna uma ótima escolha para a previsão de produção de energia solar sólida devido à sua capacidade de lidar com grandes volumes de dados e complexidade. Isso também pode levar a melhorias operacionais e estratégicas no setor de energia solar, através da otimização da produção, da previsibilidade dos níveis de produção e, conseqüentemente, da eficiência geral da rede elétrica.

Dentro do universo de técnicas de Aprendizado de Máquina para análise de séries temporais, dois tipos de modelos se destacam pelo seu desempenho e aplicabilidade: As Redes Neurais Recorrentes (RNN) e as *Long Short-Term Memory* (LSTM). Ambas são especializadas em analisar e interpretar sequências de dados, o que as torna particularmente eficazes na modelagem e previsão de fenômenos temporais e sequenciais, como os observados em dados meteorológicos para a previsão de produção de energia solar.

Essas tecnologias de aprendizado profundo são conhecidas por sua habilidade em processar e aprender a partir de dados sequenciais, capturando dependências temporais e padrões complexos. A seguir, será apresentado um aprofundamento teórico sobre LSTM e RNN, evidenciando as suas especificidades, potencialidades e as razões pelas quais são relevantes no contexto de previsão de produção de energia solar baseada em dados meteorológicos.

2.4.1 Redes Neurais Recorrentes (RNN)

Redes Neurais Recorrentes (RNN) são uma classe de Redes Neurais Artificiais projetadas para lidar com sequências de dados, ou seja, dados representados ao longo do tempo, como séries temporais. Elas são especialmente úteis em tarefas onde a ordem temporal dos dados é crucial, como na análise de dados meteorológicos para a previsão de produção de energia solar.

As RNNs têm a habilidade de usar a informação de sua ativação anterior como input para a ativação atual, permitindo que elas mantenham uma 'memória' de sequências anteriores. Isso as torna ferramentas poderosas para modelar relações temporais complexas entre dados sequenciais.

No contexto de previsão de energia solar, as RNNs podem capturar padrões temporais e tendências nos dados meteorológicos, fornecendo insights valiosos para

a otimização da geração de energia solar. Entretanto, elas enfrentam dificuldades em aprender dependências de longo alcance devido ao problema do desvanecimento do gradiente, o que levou ao desenvolvimento das LSTMs para superar essas limitações.

2.4.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) é uma variante específica de Redes Neurais Recorrentes (RNNs) criada para mitigar problemas de dependência de longo prazo que as RNNs padrão frequentemente enfrentam. As LSTM são projetadas para recolher informações por longos períodos, tornando-as especialmente úteis para tarefas que envolvem sequências de dados, como séries temporais, que são predominantes na predição de produção de energia solar.

Ao tratar de dados meteorológicos para a predição de energia solar, o uso de LSTMs é crucial devido à sua capacidade de capturar padrões temporais complexos nos dados, o que é essencial para entender as variabilidades e tendências na produção de energia solar. A estrutura das LSTMs permite a modelagem eficiente de dependências temporais e a aprendizagem de representações de alta qualidade dos dados sequenciais, otimizando a precisão das previsões de produção de energia.

Em síntese, a aplicação de técnicas de aprendizado de máquina, como RNNs e LSTMs, tem o potencial de revolucionar a predição de produção de energia solar, oferecendo insights mais profundos e precisos das variabilidades e tendências neste campo. O constante avanço dessas tecnologias sugere um futuro promissor para a integração da inteligência artificial na gestão e otimização da produção de energia solar, com possibilidades de descobertas de novas metodologias e técnicas para uma previsão ainda mais acurada e eficiente.

2.5 MÉTODOS ESTATÍSTICOS E AVALIAÇÃO DE IA

2.5.1 Normalização

Dentro do vasto domínio de preparação de dados para modelos de aprendizado de máquina, a normalização, também chamada de escalonamento, é um procedimento essencial para otimizar o desempenho de algoritmos. Esta prática abrange métodos que ajustam a escala dos recursos, tornando-os mais apropriados para algoritmos sensíveis às variações nos dados. (SINGH Dalwinder, SINGH Birmohan).

Em muitos algoritmos, a falta de normalização pode resultar em convergência mais lenta ou até na falha em encontrar uma solução adequada, já que características com escalas maiores podem dominar aquelas com escalas menores. Dois dos métodos mais comuns de normalização são:

Normalização Min-Max: Onde os dados são escalados de forma que todos os valores fiquem entre 0 e 1, conforme equação 1:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Onde:

x é o valor original;

x' é o valor normalizado.

Normalização Z-score (ou padronização): Neste método, os dados são re-escalados de forma que tenham média 0 e desvio padrão de 1. Expresso na equação 2:

$$x' = \frac{x - y}{\alpha} \quad (2)$$

Onde:

x é o valor original;

y é a média dos dados;

α é o desvio padrão.

A escolha do método de normalização pode depender do algoritmo específico e do conjunto de dados em questão. Para redes neurais, como a LSTM, a normalização é essencial, pois facilita a convergência e melhora o desempenho do modelo.

2.5.2 Seleção de recursos

Diferentes características nos dados podem possuir variados graus de relevância; algumas podem ser completamente irrelevantes e redundantes. A existência de características indesejadas complica o processo de aprendizado, ofuscando as características com importância mais relevante. Estas podem obstruir as características úteis, confundindo os algoritmos de aprendizado e comprometendo o desempenho da classificação (SINGH Dalwinder, SINGH Birmohan). A complexidade computacional dos algoritmos de aprendizado de máquina também se eleva, pois depende tanto do número de características quanto de instâncias nos dados de treino. Além disso, a expansão dessas características resulta na “maldição

da dimensionalidade”, um desafio que requer atenção para a construção de modelos preditivos eficientes.

A seleção de características é crucial em algoritmos de aprendizado de máquina, contribuindo significativamente para evitar o sobreajuste, assegurar robustez frente ao ruído e acelerar a convergência ao objetivo desejado. Neste cenário, o método LASSO (*Least Absolute Shrinkage and Selection Operator*) surge como uma técnica valiosa. O LASSO, uma forma de regressão linear regularizada, é singular por sua habilidade de "zerar" coeficientes de características menos significativas, facilitando a identificação e seleção eficaz das características mais relevantes para o modelo. Características com coeficientes próximos ou iguais a zero são vistas como menos relevantes, enquanto aquelas com coeficientes elevados (positivos ou negativos) são consideradas essenciais.

Uma vantagem notável do LASSO é a sua consistência na seleção de características, mesmo quando a relação real subjacente é não-linear. Sob certas condições, isso sugere que o LASSO pode identificar corretamente o conjunto apropriado de características, mesmo quando o modelo inicial pode não estar adequadamente especificado (ZHANG; GUO; RAY, 2016). Esta característica ressalta a robustez e versatilidade do LASSO como ferramenta de seleção de características (RODRIGUES, 2018).

2.5.3 Métricas de desempenho

A avaliação precisa do desempenho de um modelo preditivo é crucial para garantir sua confiabilidade e precisão. Existem diversas métricas de desempenho sugeridas pela literatura, cuja aplicabilidade varia conforme a natureza do problema e os objetivos da modelagem. Entre essas métricas, destacam-se:

- Erro Absoluto Médio (MAE): Esta métrica calcula a média dos erros absolutos entre os valores preditos e os reais, fornecendo uma noção clara do erro médio;
- Erro Quadrático Médio (MSE): O MSE é uma métrica robusta que avalia a média dos quadrados dos erros entre as previsões e os valores reais. Por enfatizar desvios maiores, o MSE é particularmente sensível a *outliers*, tornando-se uma métrica valiosa em situações onde erros grandes são indesejáveis;
- Erro Percentual Absoluto Médio (MAPE): Em cenários onde um erro relativo é de interesse, o MAPE se apresenta como uma métrica adequada. Expressando a

precisão como uma porcentagem. Sua popularidade decorre da facilidade em interpretar resultados em um formato percentual, proporcionando *insights* claros sobre a precisão relativa de um modelo.

As equações que expressam cada uma dessas métricas podem ser visualizadas no quadro 1 a seguir:

Quadro 1 – Detalhamento das equações de métricas

Métrica	Equação
MAE	$MAE = \frac{1}{n} \sum_{i=1}^n y' - \hat{y}' $
MSE	$MSE = \frac{1}{n} \sum_{i=1}^n (y' - \hat{y}')^2$
MAPE	$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left \frac{y' - \hat{y}'}{y'} \right $

Fonte: Autor, 2023

Onde:

y' é o valor predito;

\hat{y}' é o valor mensurado.

A seleção destas métricas específicas sugere um desejo de compreender o desempenho do modelo tanto em termos de magnitude absoluta do erro quanto em relação à escala dos dados reais. A combinação destas métricas fornece uma avaliação holística e abrangente da eficácia do modelo em prever a produção de energia solar baseada em dados meteorológicos.

2.6 APLICAÇÕES DE IA NA ENERGIA SOLAR

A Inteligência Artificial (IA) é um campo da ciência da computação que se concentra na criação de sistemas capazes de realizar tarefas que normalmente requerem inteligência humana, como reconhecimento de voz, aprendizado, planejamento e tomada de decisões. Por outro lado, a energia solar é uma forma de energia renovável que utiliza a luz do Sol para gerar eletricidade. Ambas as tecnologias têm experimentado avanços significativos nas últimas décadas.

A IA tem várias aplicações na indústria de energia solar. Por exemplo, pode ser usada para prever a produção de energia solar com base em dados históricos e

condições climáticas. Isso pode ajudar os operadores de sistemas solares a otimizar a produção de energia e a gerenciar a demanda de energia. Um estudo de Jallal, Chabaa e Zeroual (2019) demonstrou como a IA pode ser usada para prever a produção de energia solar com base em dados históricos e condições climáticas. Além disso, a IA pode ser usada para otimizar o design de sistemas solares, selecionando a melhor orientação e inclinação para os painéis solares. Yao *et al.* (2021) usaram IA para detectar falhas em sistemas solares e realizar manutenção preditiva.

Várias técnicas de IA são usadas na energia solar. O aprendizado de máquina, por exemplo, é frequentemente usado para prever a produção de energia solar. As redes neurais, que são modelos computacionais inspirados no cérebro humano, são usadas para reconhecer padrões complexos nos dados e podem ser usadas para detectar falhas em sistemas solares. Algoritmos genéticos, que são inspirados na evolução natural, podem ser usados para otimizar o design de sistemas solares, o que pode levar a painéis solares mais eficientes e econômicos.

A inteligência artificial pode ser utilizada também para modelagem, identificação, otimização, previsão e controle de sistemas complexos, e tem sido aplicada com sucesso em uma ampla gama de aplicações de energia solar. Como por exemplo a estimativa de radiação solar, sistemas de aquecimento solar, sistemas fotovoltaicos, sistemas de rastreamento solar, sistemas de ar-condicionado solar, entre outros. Kalogirou e Sencan (2010). exploraram o uso de IA para a estimativa de radiação solar e encontraram resultados promissores.

O uso de IA na energia solar tem vários benefícios. Pode aumentar a eficiência dos sistemas solares, reduzir os custos de manutenção e melhorar a previsão da produção de energia. No entanto, também existem desafios. Por exemplo, a IA requer grandes conjuntos de dados para treinamento e validação, o que pode ser um obstáculo em áreas com poucos dados disponíveis. Além disso, existem questões de privacidade e segurança que precisam ser abordadas. A implementação de soluções de IA também pode ser complexa e cara, o que pode ser um obstáculo para algumas empresas.

Olhando para o futuro, a pesquisa em IA na energia solar está explorando novas técnicas e aplicações. Por exemplo, a aprendizagem profunda, uma subárea do aprendizado de máquina que se concentra em redes neurais com muitas camadas, está sendo explorada para sua capacidade de reconhecer padrões complexos nos

dados. Além disso, a IA também está sendo usada para desenvolver novos tipos de painéis solares que podem ser mais eficientes e econômicos.

Comparando a IA com outras tecnologias usadas na indústria de energia solar, ela tem o potencial de oferecer benefícios significativos. Por exemplo, enquanto as técnicas tradicionais de otimização de painéis solares dependem de regras fixas e modelos matemáticos, a IA pode adaptar-se e aprender com os dados, permitindo uma otimização mais eficaz. No entanto, a IA também tem suas limitações, como a necessidade de grandes conjuntos de dados e a complexidade de implementação, que devem ser levados em consideração.

Desta forma, a IA tem um grande potencial para melhorar a eficiência e a eficácia dos sistemas de energia solar. São necessários mais conteúdo, pesquisa e desenvolvimento para superar os desafios e explorar plenamente o potencial da IA e suas aplicações na energia solar.

3 METODOLOGIA

3.1 CARACTERÍSTICAS DA PESQUISA

Neste estudo adotou-se uma abordagem quantitativa do tipo descritiva-exploratória para desenvolver um sistema de previsão da geração de energia solar com base em dados meteorológicos. Esta pesquisa foi orientada a explorar e descrever o potencial impacto dos fatores meteorológicos na produção de energia solar e desenvolver um modelo de previsão eficaz.

3.2 MÉTODOS E TÉCNICAS UTILIZADAS

A previsão de produção de energia solar baseada em dados meteorológicos apresenta um desafio complexo dadas as variáveis influenciadoras e a natureza dinâmica dos padrões climáticos. Contudo, técnicas avançadas de aprendizado de máquina podem ser eficazes na modelagem e previsão desses padrões.

Os modelos de Aprendizado de Máquina, em geral, aprendem com dados históricos para prever resultados futuros. A partir da análise de uma série de variáveis, esses modelos conseguem discernir padrões e correlações que podem não ser imediatamente evidentes para análise humana. No contexto da previsão de produção de energia solar, variáveis como radiação solar, temperatura, velocidade do vento, umidade, pressão atmosférica, entre outros, são consideradas.

Modelos de Redes Neurais Artificiais, por exemplo, são inspirados na estrutura neural do cérebro humano e são especialmente eficientes em capturar relações não lineares entre as variáveis. Eles consistem em nós interconectados, ou "neurônios", que recebem entradas e aplicam uma função matemática para produzir uma saída. Esse modelo pode ter várias camadas de processamento escondidas, permitindo que elas assimilem representações complexas dos dados.

3.3 MÉTODOS DE PROCEDIMENTOS

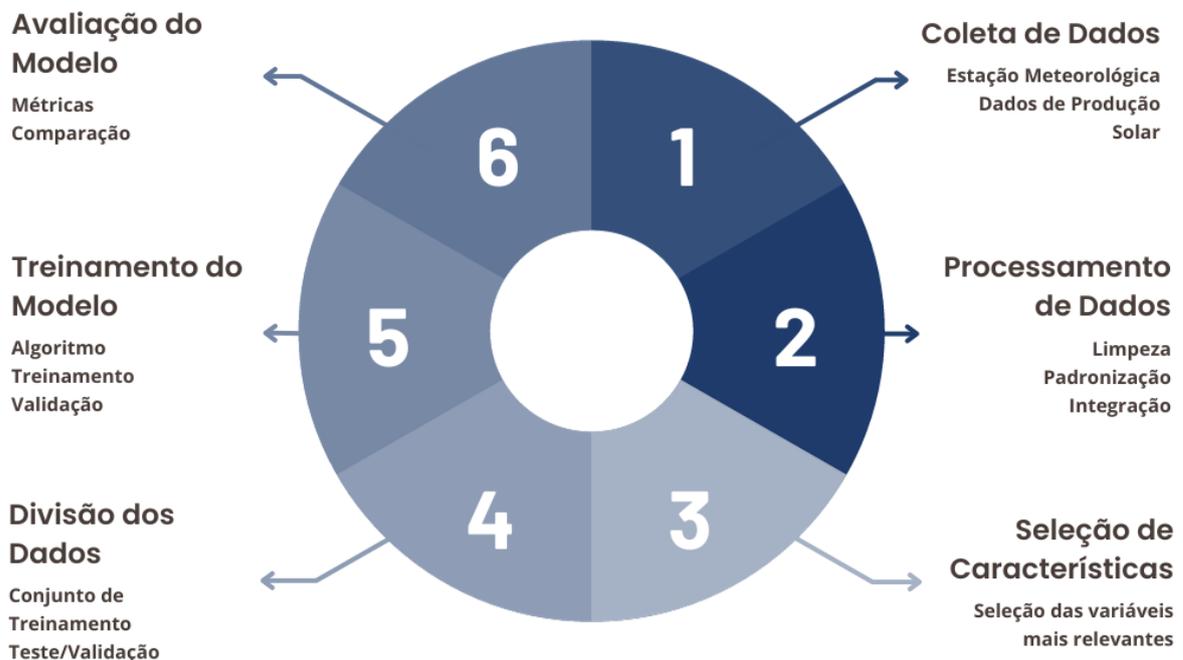
A escolha de uma metodologia consistente e bem definida é fundamental para o sucesso de qualquer pesquisa científica, especialmente quando se trata do desenvolvimento de modelos preditivos para a produção de energia solar. Uma metodologia concisa garante que o trabalho seja realizado de maneira sistemática, permitindo que cada etapa seja claramente delineada e seguida, o que é crucial para

a validação dos resultados e sua possível replicação em estudos futuros (GARG, 2016).

A metodologia adotada visa uma série de macro etapas padrões que devem ser atendidas para realização de um modelo robusto e mais assertivo conforme descrito na figura 8:

Figura 8 – Diagrama descritivo do processo de fabricação do modelo

MODELO COMPUTACIONAL



Fonte: Autor, 2023

Para realizar a análise de previsão da produção de energia solar baseada em dados meteorológicos, uma série de etapas metodológicas precisaram ser seguidas. Em primeiro lugar, foi essencial coletar e preparar os dados adequados. Para este estudo, foram utilizados dados meteorológicos históricos e dados de produção de energia solar correspondentes. Estas informações foram obtidas a partir dos registros das estações meteorológicas disponibilizadas pelo Instituto Nacional de Meteorologia (INMET) e dos dados de produção provenientes de usinas solares, com um foco especial na região noroeste do Rio Grande do Sul.

Esses dados foram então pré-processados para assegurar sua qualidade e relevância. O pré-processamento envolveu a limpeza e normalização dos dados,

tratamento de valores ausentes e *outliers*, que são pontos de dados significativamente diferentes do restante.

Posteriormente, uma análise exploratória de dados foi realizada para obter *insights* preliminares sobre a natureza e as características dos dados. Isso incluiu a criação de gráficos e a realização de cálculos estatísticos para entender a distribuição e a correlação dos dados.

Na etapa de engenharia de recursos, foram criadas novas variáveis que melhoraram a capacidade preditiva do modelo. Por exemplo, a combinação de variáveis de temperatura e umidade gerou um novo recurso chamado "índice de calor", que teve uma influência significativa na produção de energia solar.

Em seguida, os dados foram divididos em conjuntos de treinamento, validação e teste. O conjunto de treinamento foi usado para treinar os modelos, o conjunto de validação para ajustar os parâmetros e escolher o melhor modelo, e o conjunto de teste para avaliar o desempenho final do modelo.

A seleção do modelo foi guiada pelas características dos dados e pelas informações obtidas durante a análise exploratória dos dados. Modelos de *machine learning* como Redes Neurais Artificiais (ANNs), Máquinas de Vetores de Suporte (SVM), Árvores de Decisão e Regressão Linear são alguns exemplos de modelos que poderiam ser utilizados.

Uma vez selecionado o modelo, este foi treinado com o conjunto de treinamento. O treinamento do modelo envolve o ajuste dos parâmetros do modelo para minimizar a diferença entre as previsões do modelo e os valores reais dos dados de treinamento.

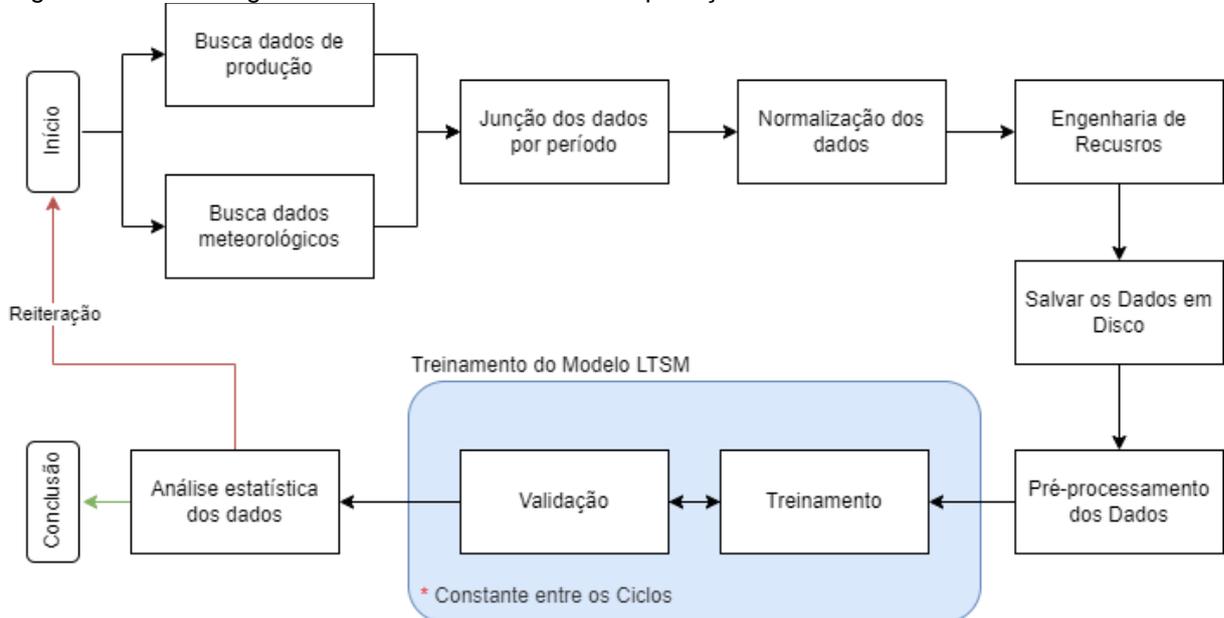
Após o treinamento, a performance do modelo foi avaliada usando o conjunto de validação. Nesta etapa, diferentes métricas de avaliação foram usadas para entender a eficácia do modelo em prever a produção de energia solar com base nos dados meteorológicos.

Na etapa de afinação do modelo, os hiperparâmetros do modelo serão ajustados para melhorar seu desempenho. Por fim, o modelo final foi testado usando o conjunto de testes para avaliar seu desempenho em dados não vistos anteriormente.

Portanto, a metodologia adotada permitiu o desenvolvimento de um sistema de previsão de produção de energia solar com base em dados meteorológicos, usando métodos adequados de *machine learning* e análise de dados. Conforme as etapas

descritas na figura 9, seguindo uma iteração constante de aprimoramento do modelo e novas captações e processamento de dados.

Figura 9 – Metodologia de treinamento do modelo de predição



Fonte: Autor, 2023

3.4 TÉCNICAS DE COLETA DE DADOS

A coleta de dados é essencial em qualquer projeto de pesquisa. Em nosso estudo sobre a predição da produção de energia solar a partir de dados meteorológicos, esta etapa não é diferente. O modelo depende de dados meteorológicos e registros da produção solar. As observações meteorológicas compreendem variáveis como radiação solar, temperatura, velocidade do vento, umidade, precipitação, entre outras. Assim, o foco se dá na coleta de dados históricos meteorológicos e de produção solar, provenientes de duas fontes principais.

3.4.1 Dados meteorológicos

O Brasil conta com várias fontes confiáveis de dados meteorológicos. O Instituto Nacional de Meteorologia (INMET) é uma das mais renomadas, oferecendo dados coletados por estações automáticas por todo o país. Outras fontes notáveis incluem o CPTEC/INPE e a ANA. Durante a coleta, foi crucial manter a consistência no período e localização das observações entre as fontes para garantir a precisão. Para este estudo, optamos pelos dados da estação meteorológica de Santa Rosa do INMET, devido à sua proximidade com as usinas em análise. Esses registros

apresentam variáveis meteorológicas que impactam diretamente a produção fotovoltaica.

3.4.2 Dados de produção solar

A obtenção de dados de produção solar pode ser mais complexa. Empresas, universidades e centros de pesquisa são potenciais fontes desses registros no Brasil. A ANEEL, por exemplo, tem um banco de dados atualizado das usinas solares em atividade, embora a aquisição de dados específicos possa demandar solicitações formais. Nesta pesquisa, procurou-se estabelecer uma colaboração com uma usina fotovoltaica local, que forneceu seus dados sob garantia de anonimato. Utilizou-se consultas SQL (*Structured Query Language*) para identificar usinas até um máximo de 20 km da estação meteorológica mais próxima, garantindo assim a relevância dos dados. Os registros das usinas incluem informações a cada hora sobre produção e estado operacional, os quais estão descritos no quadro 2, detalhando os intervalos de tempo que foram considerados para cada usina.

Quadro 2 – Detalhamento da coleta de dados das usinas

Nome	De	Até	Quantidade de Amostras
Usina 1	15/04/2022	29/09/2023	15096
Usina 2	26/07/2022	29/09/2023	12952
Usina 3	01/04/2022	29/09/2023	16142

Fonte: Autor, 2023

3.5 CONSIDERAÇÕES ÉTICAS

Na condução deste estudo, é importante respeitar as diretrizes éticas na coleta, uso e divulgação de dados. Os dados meteorológicos utilizados são de domínio público e acessíveis a todos, não apresentando problemas éticos. No entanto, os dados de produção de energia solar, especialmente quando fornecidos por empresas ou entidades privadas, foram usados de forma responsável e apenas com o devido consentimento.

Além disso, os resultados do estudo serão futuramente divulgados de forma transparente, assegurando que quaisquer limitações ou possíveis fontes de erro sejam devidamente reconhecidas.

3.6 ANÁLISE DE DADOS

A fase inicial da análise de dados envolveu o tratamento e a organização dos dados brutos. Assim, iniciou-se com o pré-processamento de dados, que incluiu a limpeza dos dados, a imputação de valores ausentes e a normalização dos dados. Também foi necessária a transformação de algumas variáveis para torná-las mais adequadas para análise. A limpeza dos dados foi fundamental para garantir a qualidade e a confiabilidade do modelo de previsão.

Em seguida, realizou-se a análise exploratória de dados (EDA), uma etapa crucial para entender as características e relações presentes nos dados. A EDA revelou *insights* sobre a estrutura dos dados, relações entre variáveis, identificar possíveis *outliers* ou erros. Para isso, utilizamos várias técnicas estatísticas e gráficas, como histogramas, gráficos de dispersão, medidas de tendência central (média, mediana, moda) e medidas de dispersão (variação, desvio padrão).

A engenharia de recursos foi o passo seguinte, no qual criou-se novos recursos (variáveis) a partir dos existentes que foram úteis para melhorar a precisão do modelo. Por exemplo, foi útil combinar a temperatura e a umidade para criar um novo recurso que representou o índice de calor. Outro fator que pode influenciar a quantidade de energia produzida é o ângulo do Sol em relação aos painéis solares, que pode ser calculado considerando a hora do dia, o dia do ano e a latitude. Além disso, características temporais, como a hora do dia, o dia da semana e o mês, foram levadas em conta, uma vez que a produção de energia solar pode variar de acordo com a hora do dia e a estação do ano. Tais aspectos e relações são enfatizados no estudo de MAITHA *et al.* (2011).

Para a modelagem preditiva, escolheu-se métodos e técnicas adequados ao problema e aos dados. Aqui, utilizamos modelos de aprendizado de máquina, como o de Redes Neurais Artificiais (ANNs), o de Máquinas de Vetores de Suporte (SVM), e técnicas de *ensemble*, uma estratégia de aprendizado de máquina que integra as previsões de vários modelos com o objetivo de desenvolver um modelo preditivo superior, como a Floresta Aleatória. Cada modelo tem suas vantagens e desvantagens, e a escolha do modelo foi guiada pelos dados obtidos.

Ao seguir essas etapas, esperamos desenvolver um modelo de previsão robusto e preciso que possa ser usado para prever a produção de energia solar com base nos dados meteorológicos. Isso pode ser um recurso valioso para empresas de

energia solar e órgãos reguladores de energia, permitindo uma melhor gestão e planejamento da produção de energia solar.

3.7 MATERIAL E EQUIPAMENTOS

Para a execução desta pesquisa, foi essencial o uso de um computador com capacidade suficiente para realizar análises complexas de dados e executar algoritmos de aprendizado de máquina. O computador deve ter um processador robusto, memória suficiente e, preferencialmente, uma unidade de processamento gráfico (GPU) para acelerar certas operações de aprendizado de máquina. A configuração exata do computador dependeu do volume e da complexidade dos dados a serem processados.

Em termos de *software*, a linguagem de programação *Python* foi uma excelente opção, dado o seu suporte robusto para a ciência de dados e aprendizado de máquina. *Python* oferece uma variedade de bibliotecas que foram úteis para este projeto. Para manipulação e análise de dados, a biblioteca *Pandas* é a mais indicada por sua versatilidade e eficiência. Para visualização de dados, bibliotecas como *Matplotlib* e *Seaborn* oferecem funcionalidades avançadas e versáteis. Por fim, para implementar e treinar modelos de aprendizado de máquina, bibliotecas como *Scikit-Learn*, *Tensor Flow* e *Keras* fornecem um amplo conjunto de recursos e algoritmos.

Para atender a essas necessidades, optou-se pelo *Google Collaboratory* - um ambiente online integrado para estudo e pesquisa em *Python*, oferecido pelo Google. Essa plataforma de computação em nuvem, baseada no *Jupyter Notebook*, viabiliza a execução de códigos *Python* diretamente pelo navegador web, eliminando a necessidade de configurações extensivas (BISONG, 2019).

O *Google Colab* proporciona um Ambiente de Desenvolvimento Integrado (IDE) interativo que permite o desenvolvimento dinâmico de códigos, além da possibilidade de salvar e compartilhar análises. No seu nível de acesso gratuito, o sistema disponibiliza máquinas virtuais com placas gráficas e até 12 GB de memória RAM. Contudo, conforme indicado na documentação do FAQ da *Google*, a disponibilidade desses recursos pode flutuar de acordo com a demanda da comunidade.

Em ocasiões em que o sistema não possuiu capacidade suficiente para processar os dados, ou caso os recursos gratuitos não estivessem disponíveis, houve a opção de adquirir ambientes e créditos dedicados para uso em projetos específicos.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

O núcleo de qualquer sistema de previsão é o modelo que impulsiona a capacidade preditiva. A seleção de um modelo adequado é, portanto, fundamental para a eficácia de tal sistema.

4.1 SELEÇÃO DO MODELO

Para a pesquisa em questão, optou-se pelo modelo LSTM (*Long Short-Term Memory*) para previsão. As redes LSTM, categorizadas como Redes Neurais Recorrentes (RNN), são particularmente eficazes para lidar com sequências temporais, graças à sua habilidade de reter informações de eventos anteriores. Isso as torna excepcionalmente adequadas para previsões que envolvem dados temporais, como é o caso dos dados meteorológicos e da produção de energia solar onde eventos passados, como padrões climáticos, podem influenciar a futura produção de energia solar.

No campo da previsão de energia solar, as LSTMs mostram superioridade não apenas em relação a modelos tradicionais de aprendizado de máquina, mas também quando comparadas a outros modelos de aprendizado profundo. Estudos mostraram que as LSTMs, quando aliadas a outras arquiteturas, apresentam desempenho superior em métricas como RMSE e MAE (JAILANI *et al.*, 2023). A habilidade da LSTM de capturar características tanto espaciais quanto temporais as torna ideal para tarefas complexas de previsão, como é o caso da produção de energia solar. Por tais motivos, modelos baseados em LSTM são recomendados para aumentar a precisão e confiabilidade das previsões de energia solar.

4.2 INTEGRAÇÃO DOS DADOS

A integridade e exatidão dos dados são primordiais antes do início do treinamento do modelo computacional, que utiliza grandes volumes de dados para identificar tendências e realizar previsões. Neste estudo, a informação proveu de dois arquivos no formato CSV, um tipo de arquivo que organiza os dados em colunas separadas por vírgulas, facilitando a leitura e o manuseio.

Um dos arquivos continha sequências temporais de dados climáticos e o outro, registros acumulados da produção de energia. A combinação desses arquivos por

data e hora permitiu uma comparação direta e eficaz entre as condições do tempo e a geração de energia solar.

Durante a inspeção inicial, detectou-se que os dados de produção de energia não eram instantâneos, mas acumulativos, aumentando a cada hora desde a instalação da usina. Para obter a quantidade de energia produzida em cada hora, aplicou-se uma técnica de pré-processamento: a subtração do total acumulado na hora anterior do valor subsequente, isolando assim a produção específica de cada hora.

Outro desafio foi a sobreposição de registros no mesmo intervalo temporal. Isso foi corrigido padronizando-se os horários, arredondando-se os minutos para a hora cheia mais próxima e agrupando-se os dados que correspondiam ao mesmo período. Por fim, normalizamos os dados, ou seja, ajustamos as variáveis a uma escala comum para evitar distorções nas análises, e preenchemos lacunas decorrentes de interrupções na coleta de dados ou falhas nos equipamentos, garantindo assim a integridade e a utilidade do conjunto de dados para o modelo.

Adicionalmente, foi imperativo normalizar os dados e tratar eventuais lacunas durante todo o processo, que podem surgir devido a falhas na coleta, problemas de equipamento ou alterações durante o treinamento do próprio modelo. Esta etapa foi fundamental para garantir que o modelo tenha sido alimentado com um conjunto de dados representativo e consistente.

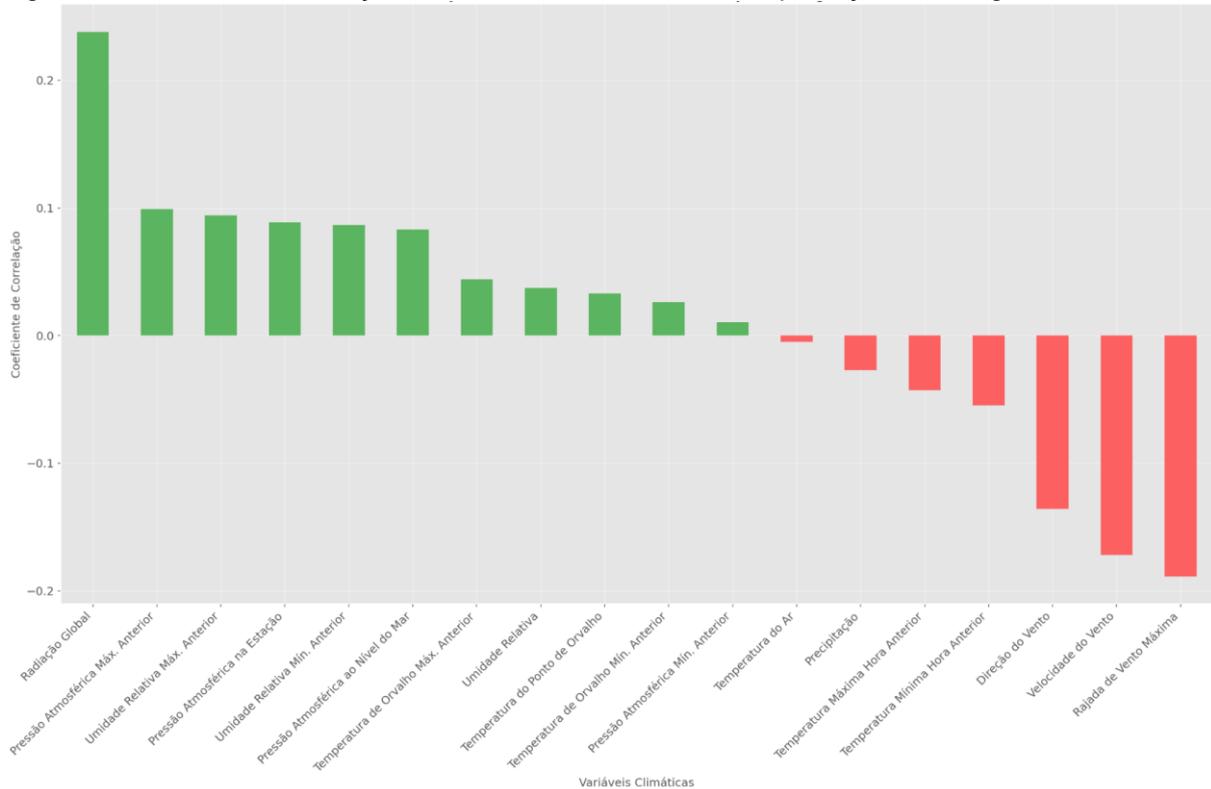
4.3 RELAÇÃO ENTRE VARIÁVEIS METEOROLÓGICAS E A PRODUÇÃO DE ENERGIA SOLAR

As condições meteorológicas têm uma influência direta na produção de energia solar. Fatores como radiação global, temperatura e até mesmo a velocidade do vento podem influenciar na eficiência dos painéis solares. Por exemplo, uma maior radiação global indica uma presença mais forte de luz solar, o que pode levar a uma maior produção de energia. No entanto, temperaturas extremamente altas podem reduzir a eficiência dos painéis solares, levando a uma diminuição na produção de energia.

Ao explorar os dados fornecidos, pudemos procurar tendências e correlações entre as variáveis meteorológicas e a produção de energia solar. Isso forneceu *insights* sobre quais variáveis têm o maior impacto na produção de energia e, portanto, são mais importantes para a previsão.

Conforme pode-se observar na figura 10, a análise de correlação forneceu clareza sobre o nexos entre as condições meteorológicas e a geração de energia solar.

Figura 10 – Gráfico de correlação das variáveis climáticas com a produção de energia solar



Fonte: Autor, 2023

Com base no gráfico supracitado, destacam-se as seguintes observações:

- **Radiação Global:** Esta variável tem uma forte correlação positiva com a produção de energia solar. Isso faz sentido, pois a radiação solar é a principal fonte de energia para os painéis solares. Quanto maior a radiação, maior é a produção potencial de energia.
- **Temperatura do Ar:** A temperatura do ar tem uma correlação positiva, indicando que, até certo ponto, temperaturas mais altas podem ser favoráveis para a produção de energia solar. No entanto, como mencionado anteriormente, temperaturas extremamente altas podem afetar negativamente a eficiência dos painéis solares.
- **Direção e Velocidade do Vento:** Estas variáveis têm uma correlação menor com a produção de energia solar em comparação com a radiação e temperatura. No entanto, elas ainda podem ter um impacto, especialmente em locais onde os painéis solares também servem como barreiras contra o vento.

Variáveis com correlação próxima a zero ou negativa podem ter um impacto menor na produção de energia solar, ou pode ser que a relação entre essas variáveis e a produção de energia não seja linear.

Com base nessa análise, pôde-se inferir que, enquanto várias variáveis meteorológicas influenciam a produção de energia solar, a radiação global e a temperatura do ar desempenham um papel significativo.

4.4 ENGENHARIA DE RECURSOS

A engenharia de recursos desempenha um papel vital na melhoria da capacidade preditiva de um modelo. Neste estudo, foram criadas novas variáveis a partir dos dados existentes. Essas novas características ajudaram o modelo a capturar relações não lineares e interações entre os dados, que podiam não ser evidentes com as características originais.

- Índice de calor (Temperatura e Umidade): A combinação de temperatura e umidade pode nos dar uma sensação de "índice de calor", que é como os humanos percebem a temperatura. Pode ser útil considerar o produto dessas duas características;
- Radiação Acumulada: Representa a soma cumulativa da radiação ao longo do dia. A acumulação de radiação solar ao longo do dia ou da hora anterior pode ser relevante, pois painéis solares têm uma capacidade de absorção;
- Índice Céu Encoberto: É uma média da umidade relativa e precipitação. Valores mais altos podem indicar um céu mais encoberto, o que pode afetar a produção de energia solar;
- Razão Radiação e Temperatura: É a razão entre a radiação global e a temperatura. Essa característica pode ajudar a entender a eficiência da conversão de radiação em energia em diferentes temperaturas, pois painéis solares têm eficiências diferentes em diferentes temperaturas.

Estas características podem ajudar a melhorar a previsibilidade do modelo, pois capturam informações adicionais sobre as condições que influenciam a produção de energia solar.

4.5 SELEÇÃO DE RECURSOS

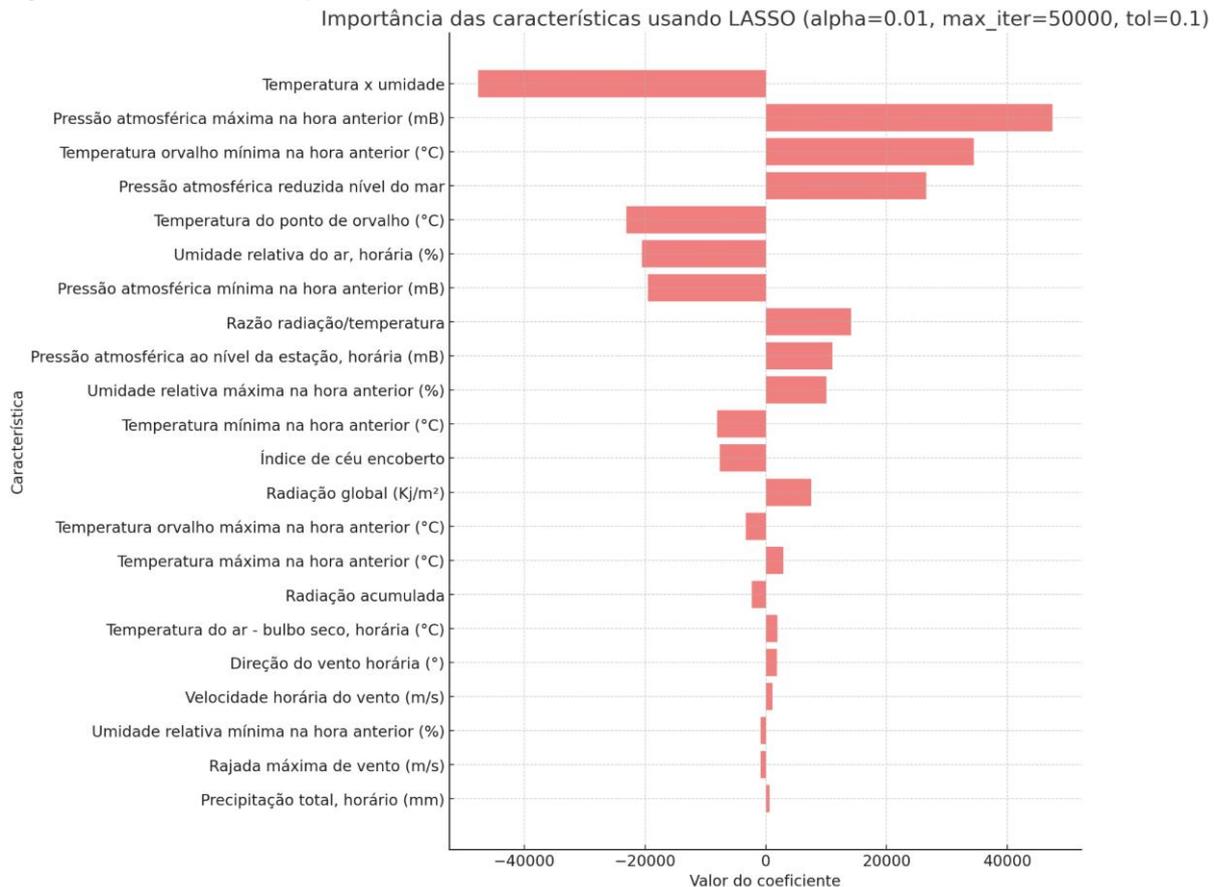
A próxima etapa após a engenharia de características geralmente envolve a seleção de recursos. Isso é importante para garantir que o modelo use apenas as características mais informativas, o que pode melhorar o desempenho e reduzir o risco de *overfitting*. *Overfitting* ocorre quando um modelo de aprendizado de máquina se ajusta excessivamente aos dados de treinamento, capturando ruído ou flutuações aleatórias como se fossem padrões reais. Em outras palavras, o modelo fica tão especializado nos dados que viu durante o treinamento que tem dificuldades em se adaptar e realizar previsões acuradas em dados que ele nunca viu antes, mesmo que sejam semelhantes (MUTASA; SUN; HA, 2020).

Neste estudo, utilizou-se a técnica LASSO, uma forma de regressão linear com regularização, para nos ajudar a identificar os atributos mais pertinentes. Na abordagem LASSO, atributos com coeficientes próximos ou iguais a zero são interpretados como menos essenciais para a previsão, enquanto aqueles com coeficientes de alta magnitude, seja positiva ou negativa, são considerados altamente relevantes. A regularização aplicada pelo LASSO ajuda a mitigar o *overfitting*, pois impõe uma penalidade aos coeficientes, forçando-os a diminuir e, assim, simplificando o modelo.

Com base nas importâncias dos coeficientes LASSO, podemos tomar decisões informadas sobre quais características manter e quais descartar. No entanto, é sempre uma boa prática testar o desempenho do modelo após a remoção de características para garantir que o desempenho não seja comprometido. A validação cruzada, por exemplo, pode ser uma ferramenta valiosa nesta etapa para avaliar de forma robusta o impacto da seleção de características no desempenho do modelo em diferentes subconjuntos de dados.

A análise do gráfico da figura 11 indica, que algumas das características de interação e as novas características que foram adicionadas mostram uma correlação significativa, demonstrando que a engenharia de dados melhorou a qualidade do conjunto. As variáveis "Umidade relativa máxima na hora anterior (%)", "Velocidade horária do vento (m/s)" e "Precipitação total, horário (mm)" emergiram com coeficientes de magnitude relativamente baixa. Isso sugere uma menor influência destas variáveis na previsão da produção de energia solar.

Figura 11 – Gráfico de Importância das características usando técnica LASSO



Fonte: Autor, 2023

A partir desses resultados, levantou-se a hipótese de que estas características podem ser removidas sem causar um impacto significativo no desempenho preditivo do modelo. No entanto, antes de prosseguir com a remoção de qualquer variável, deve ser avaliado se no contexto de treinamento atual, não se torna prejudicial a remoção dessas variáveis, portanto deve-se realizar um teste com e sem as variáveis para garantir a efetividade desses ajustes.

4.6 TREINAMENTO

O modelo foi desenvolvido através de uma abordagem de pesquisa incremental, iniciando com uma estrutura simplificada e evoluindo progressivamente com a integração de recursos adicionais e dados. A evolução foi monitorada rigorosamente através de métricas de desempenho para assegurar a qualidade do modelo.

4.6.1 Desenvolvimento inicial do modelo

Na fase inicial, os dados de uma usina solar foram coletados para o período de 26 de agosto de 2022 a 29 de setembro de 2023, com base nos registros da Usina 1, como demonstrado na seção anterior 3.4.2 no Quadro 2. O Instituto Nacional de Meteorologia (INMET) forneceu dados correspondentes a este intervalo, com divisões em períodos de uma hora cada. Integraram-se estes dados às métricas de produção, sob a premissa de que intervalos sem registro de coleta indicavam ausência de produção.

O processo de desenvolvimento do modelo iniciou com a seleção de variáveis explanatórias, utilizando um conjunto inicial de dados que incluía 'Radiação Global (Kj/m²)', 'Temperatura do Ar - Bulbo Seco, Horária (°C)', 'Umidade relativa do ar, horária(%)', e variáveis temporais como 'Dia da Semana' e 'Mês', para capturar a sazonalidade inerente aos dados.

Subsequentemente, procedeu-se à normalização dos dados, empregando o *MinMaxScaler* do *scikit-learn*, para escalar as variáveis explanatórias e a variável de resposta para um domínio comum [0,1]. Tal procedimento é vital para a otimização do algoritmo de aprendizado, prevenindo o predomínio de características com maior amplitude sobre as de menor amplitude.

Após a normalização, o conjunto de dados foi dividido em subconjuntos de treinamento (80%) e teste (20%), a fim de avaliar a capacidade do modelo de generalizar para dados novos. Para a aplicação em uma rede LSTM, os dados foram formatados em sequências temporais. Uma função específica foi utilizada para estruturar os dados de entrada e saída, de modo que cada sequência de entrada correspondesse a um valor subsequente de saída, o qual seria previsto.

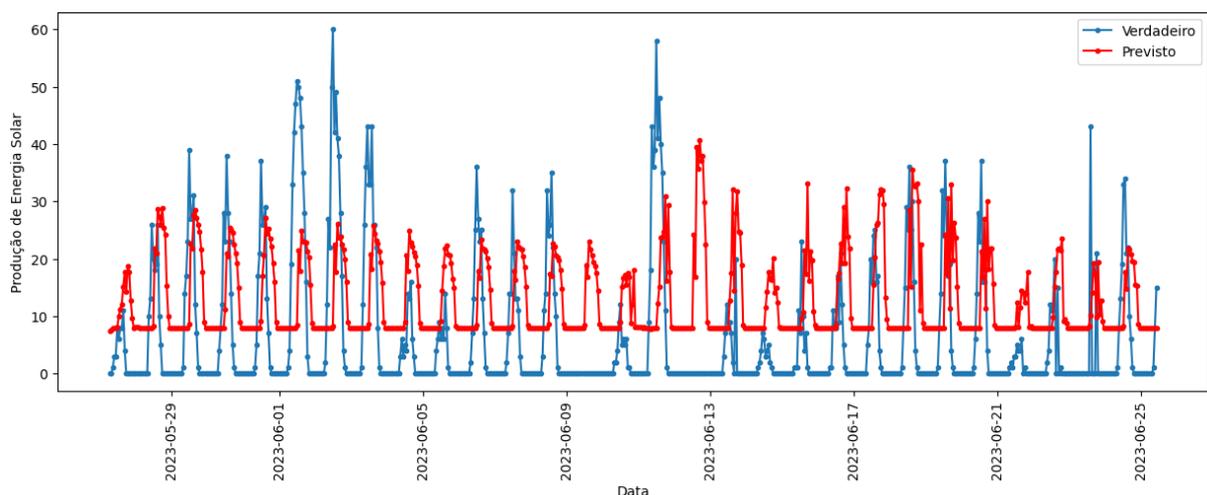
A arquitetura LSTM foi estabelecida com uma camada recorrente de 50 unidades e função de ativação 'relu', seguida por uma camada densa destinada à

previsão de valores contínuos, adequada para tarefas de regressão. A compilação do modelo foi realizada utilizando o otimizador Adam com uma taxa de aprendizado de 0.01 e a função de perda '*mean_squared_error*'.

O treinamento ocorreu durante 20 épocas, com um tamanho de lote de 32 e uma divisão de 10% dos dados para validação cruzada, a fim de monitorar e prevenir o sobreajuste. A ordem temporal foi preservada nas sequências de treinamento, um aspecto crítico para a eficácia de modelos temporais como o LSTM.

Os resultados iniciais do treinamento não foram satisfatórios, com um Erro Absoluto Médio (MAE) de 17,85. Para uma análise mais detalhada, efetuou-se uma comparação gráfica entre os valores preditos e os de teste, apresentados na figura 12. A análise indicou a possibilidade de problemas relacionados ao treinamento ou à integridade dos dados.

Figura 12 – Comparação de valores de predição do primeiro experimento



Fonte: Autor, 2023

4.6.2 Otimização e refinamento do modelo

Na figura 12, observou-se que a fusão de dados ao longo do período integral resultou em muitas observações nulas, atribuídas aos períodos noturnos sem produção energética e à ausência significativa de registros de coleta de dados. O que pode gerar uma dificuldade no modelo de capturar as sazonalidade e padrões de produção que os dados apresentam, visto a grande diferença entre os dados de produção e os dados sem produção.

Para superar este desafio, optou-se por um conjunto de dados filtrados, excluindo os intervalos de inatividade produtiva e preservando apenas as observações

com produção energética durante o dia. Este procedimento de filtragem objetivou refinar a acurácia do modelo, medido pelo Erro Médio Absoluto (MAE), que é a média das diferenças absolutas entre os valores previstos pelo modelo e os valores reais observados. Após a aplicação dessas mudanças, o MAE foi reduzido para 16,31. No entanto, esse valor permanece notavelmente elevado, considerando que a gama de variação dos dados de produção é de 172 kWh, indicando a necessidade de aprimoramentos adicionais no modelo.

A dificuldade em obter a convergência desejada no modelo suscitou a hipótese de que a interrupção na sequência de dados poderia ser um fator adverso. Para endereçar essa questão, os dados foram reorganizados de forma a apresentar uma sequência temporal contínua, enfocando exclusivamente o período de insolação, ou seja, quando a radiação solar é suficiente para a produção de energia, estimado das 06:00 às 18:00 horas. Ajustes nos dados para capturar essa consistência sazonal resultaram numa melhoria significativa, refletida por uma redução do MAE para 11,89.

Também foi levantado um problema de *outliers*, valor fora da série padrão de dados, que quando verificados indicavam um erro na coleta dos dados que resultaram em valor com uma magnitude de até 5 vezes maior que o pico comum do desvio padrão de aproximadamente 75. Para que estes 18 valores não interferissem no treinamento, foi realizada a substituição pelo maior valor encontrado dentro do plano normal de produção, que foi 71,25. Assim garantindo uma distribuição normal com menor amplitude.

Na fase subsequente, procedeu-se com uma metodologia de otimização iterativa, que consistiu em dez sessões de treinamento com incrementos progressivos de melhoria na estrutura do modelo. O modelo emprega primariamente técnicas de otimização durante o treinamento, como o “*EarlyStopping*” e o “*ReduceLROnPlateau*”.

O “*EarlyStopping*” é uma intervenção estratégica que termina o treinamento prematuramente se o modelo não exibe melhorias, determinado pelo monitoramento contínuo da perda de validação. Por outro lado, o “*ReduceLROnPlateau*” ajusta dinamicamente a taxa de aprendizado, diminuindo-a gradualmente quando o modelo deixa de progredir, facilitando a convergência para um mínimo de erro.

O modelo é construído utilizando uma arquitetura sequencial, iniciando com uma camada de normalização (*BatchNormalization*) para padronizar as entradas e promover uma convergência rápida e estável. Segue-se uma camada de convolução unidimensional (Conv1D) que extrai padrões nos dados, e uma camada de *pooling*

(*MaxPooling1D*) que reduz a dimensionalidade, aumentando assim a eficiência computacional e mitigando o risco de *overfitting*.

O núcleo do modelo consiste em múltiplas camadas LSTM, configuradas para reter informações de longo prazo e descartar as irrelevantes. Esta capacidade é reforçada pelo uso de técnicas de regularização, como o *dropout*, que ajudam a prevenir o ajuste excessivo ao conjunto de treinamento. A camada final do modelo é uma camada densa, que integra as características aprendidas e as mapeia para a saída prevista.

A compilação do modelo é feita utilizando o otimizador Adam, com uma taxa de aprendizado inicialmente definida, que é ajustada conforme o treinamento avança pelo "*ReduceLROnPlateau*". O treinamento é executado em múltiplas iterações, ou épocas, com o modelo sendo exposto a lotes de dados para atualizar iterativamente os pesos internos.

Finalmente, a performance do modelo é meticulosamente avaliada com um conjunto de testes, proporcionando uma medida objetiva da sua capacidade de generalização. Este processo de avaliação é crucial para garantir que o modelo aprendeu a capturar os padrões subjacentes dos dados, em vez de memorizar as entradas de treinamento.

Os parâmetros fixos apresentados no código, conforme figura 13, são definidos a partir de um teste de hiperparâmetros. Estes processos de otimização consistem em uma sequência de testes automáticos com diferentes valores de configuração que foram experimentadas repetidamente com o objetivo de minimizar uma métrica de erro, aprimorando assim a precisão do modelo.

Figura 13 – Código do décimo treinamento

```

Décimo Modelo

# Callbacks para EarlyStopping e ReduceLRonPlateau
reduce_lr = ReduceLRonPlateau(monitor='val_loss', factor=0.2, patience=5, min_lr=0.0001)
early_stopping = EarlyStopping(monitor='val_loss', patience=20)

# Definir a arquitetura sequencial do modelo
model = Sequential()
# Camada de normalização para estabilização do modelo
model.add(BatchNormalization(input_shape=(X_train.shape[1], X_train.shape[2])))
# Primeira camada de convolução 1D
model.add(Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=(X_train.shape[1],
X_train.shape[2])))
# Camada de pooling
model.add(MaxPooling1D(pool_size=2))
# Adicionado camadas LSTM diretamente após a camada de convolução/pooling
model.add(LSTM(50, activation='relu', return_sequences=True, dropout=0.1,
recurrent_dropout=0.1))
model.add(LSTM(50, activation='relu', return_sequences=True, dropout=0.1,
recurrent_dropout=0.1))
model.add(LSTM(50, activation='relu', dropout=0.1, recurrent_dropout=0.1))
# Camada densa para a saída
model.add(Dense(1, kernel_regularizer=l2(0.01)))

# Compilar o modelo
model.compile(optimizer=Adam(learning_rate=0.001), loss='mse')

# Treinar o modelo
history = model.fit(X_train, y_train, epochs=500, batch_size=48, validation_data=(X_test,
y_test), verbose=1, shuffle=False, callbacks=[early_stopping, reduce_lr])

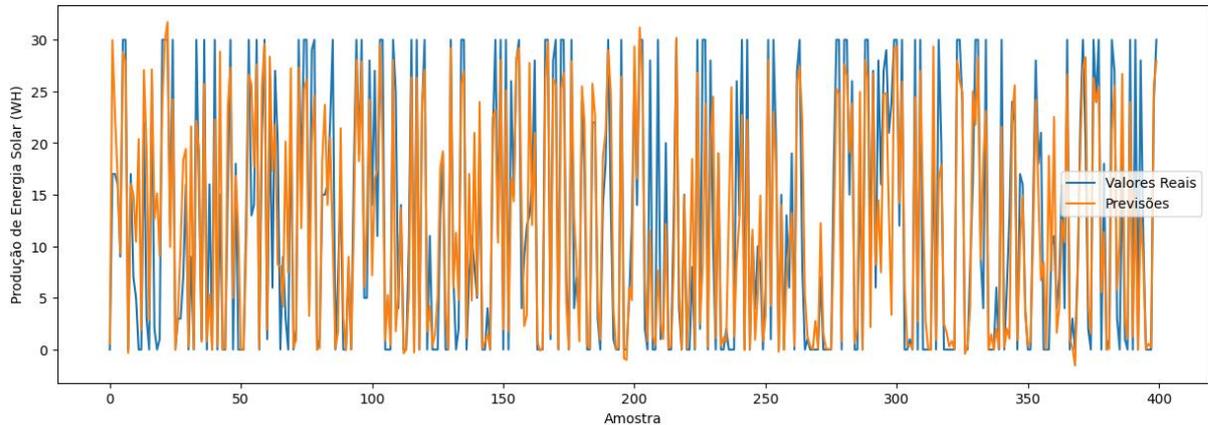
# Avaliar o desempenho do modelo
loss = model.evaluate(X_test, y_test, verbose=0)

```

Fonte: Autor, 2023

As intervenções resultaram em uma diminuição significativa do Erro Absoluto Médio (MAE) para 4.16; indicando uma precisão da previsão de 5,54% em relação à variação dos valores de produção, conforme demonstrado na figura 14. No entanto, estas melhorias implicaram em um aumento na complexidade do modelo e, por conseguinte, no tempo de treinamento, que foi ampliado em até cinco vezes em comparação ao modelo inicial.

Figura 14 – Comparação de valores de predição do décimo experimento

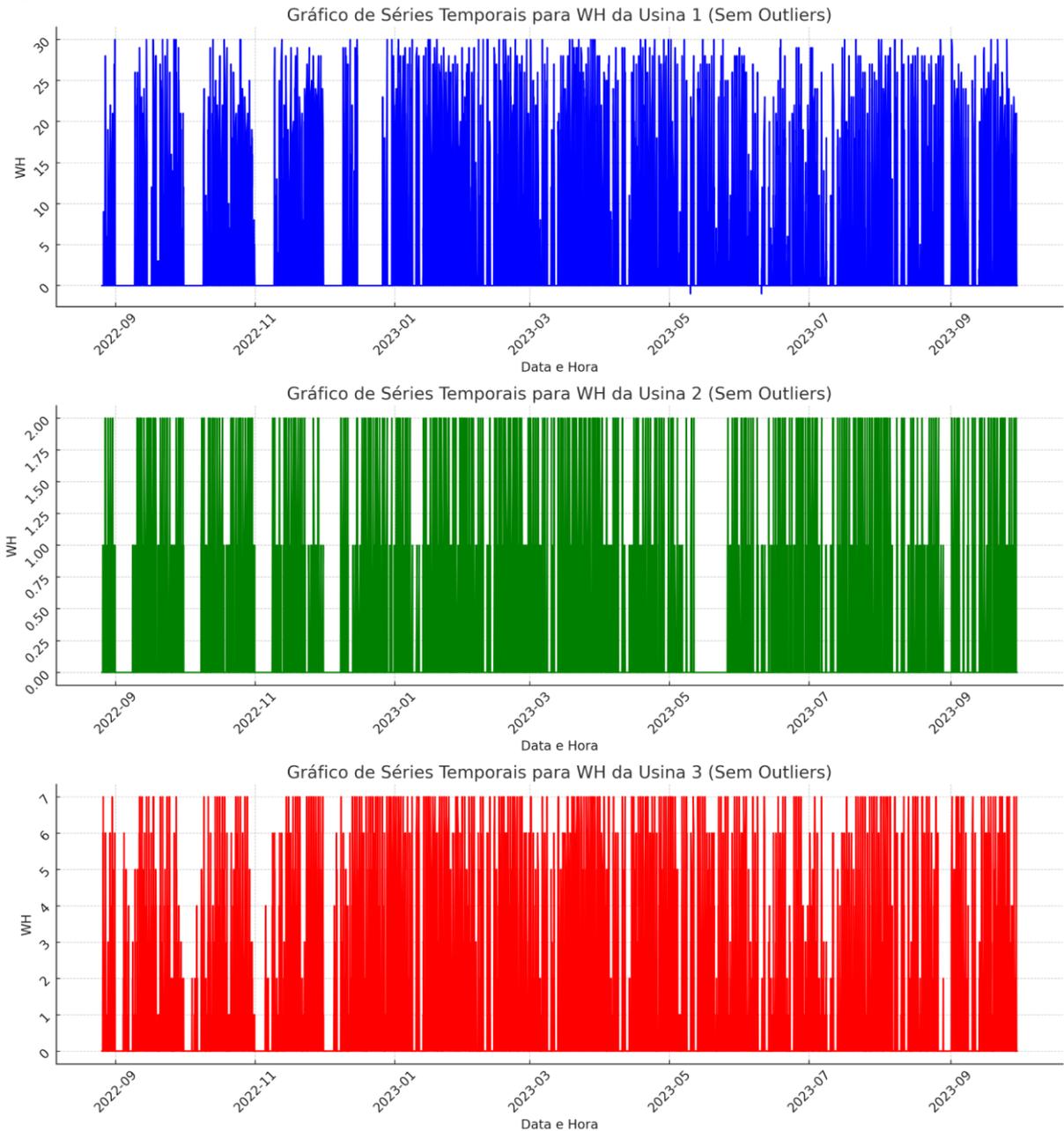


Fonte: Autor, 2023

4.6.3 Expansão e testes com múltiplas usinas

O desempenho do modelo preditivo de uma usina individual foi considerado adequado, o que motivou a execução de uma série de testes unitários utilizando conjuntos de dados provenientes de duas usinas adicionais. Para isso foi então realizado os mesmos passos de pré-processamento, junção de dados e engenharia de recursos nos dados das outras duas usinas e selecionado o maior período em comum de dados das 3 usinas. No entanto, conforme pode ser evidenciado na figura 15, as capacidades produtivas destas usinas variam significativamente, impossibilitando o uso direto dos dados de produção em sua forma bruta, como nos procedimentos de teste precedentes.

Figura 15 – Exposição dos dados de produção das usinas coletadas



Fonte: Autor, 2023

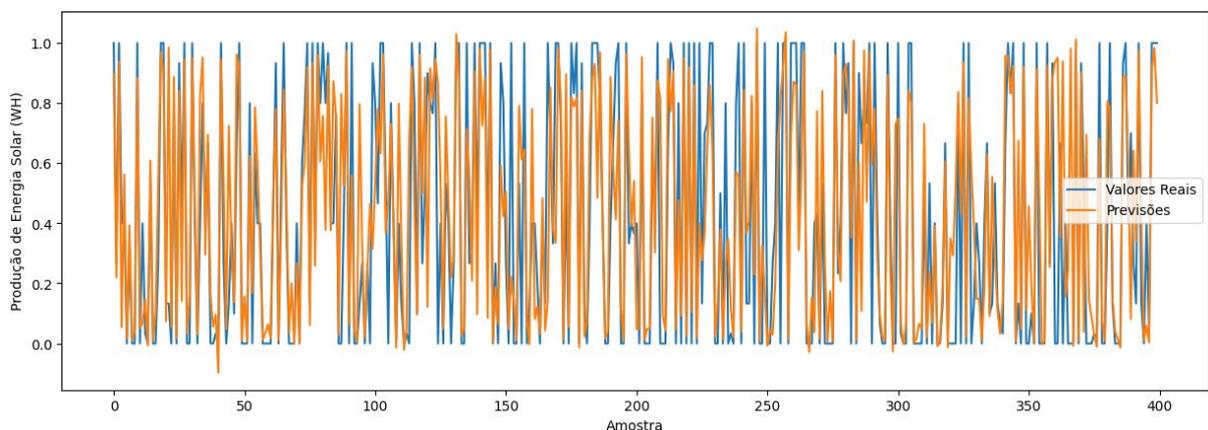
Para contornar essa limitação, implementou-se a normalização dos dados por meio de um 'fator de produção', utilizado como uma variável independente na construção do novo modelo. O fator de produção foi calculado pela razão entre o valor de produção diário e o pico de produção do período em questão, produzindo assim um percentual da capacidade máxima de produção da usina, conforme ilustrado na equação 3. Esta abordagem propicia a homogeneização dos dados, estabelecendo um padrão uniforme para o treinamento do modelo com informações das três usinas distintas.

$$Taxa\ de\ produção = \frac{Valor\ de\ produção\ diário}{Máximo\ valor\ de\ produção\ do\ período} \quad (3)$$

Essa metodologia promove a padronização dos dados, facilitando o treino do modelo com dados das três usinas. Para a construção deste conjunto de dados ampliado foram incorporados 28.800 registros conforme a equação 3, categorizados por data e hora, conforme descrito no quadro 2. A subsequente união desses conjuntos seguiu a mesma sequência cronológica, com a adição de dados meteorológicos correspondentes de cada data e hora, extraídos do Instituto Nacional de Meteorologia (INMET) no período de 16 de outubro de 2022 a 29 de setembro de 2023. Em instâncias de dados ausentes, decorrentes de manutenções ou erros de coleta, empregou-se a técnica de Interpolação, que é frequentemente preferida em séries temporais para preservar a continuidade dos dados.

O treinamento demandou aproximadamente 17 minutos e alcançou resultados notáveis, com um Erro Absoluto Médio (MAE) de 0,1649 e uma Raiz do Erro Quadrático Médio (RMSE) de 0,245. Considerando que as previsões são escalonadas entre 0 e 1 e representam a eficiência produtiva, esses valores indicam uma margem de erro aceitável para previsões precisas, conforme pode ser notado no período de testes expresso na figura 16:

Figura 16 – Comparação de valores de predição do experimento com todas usinas



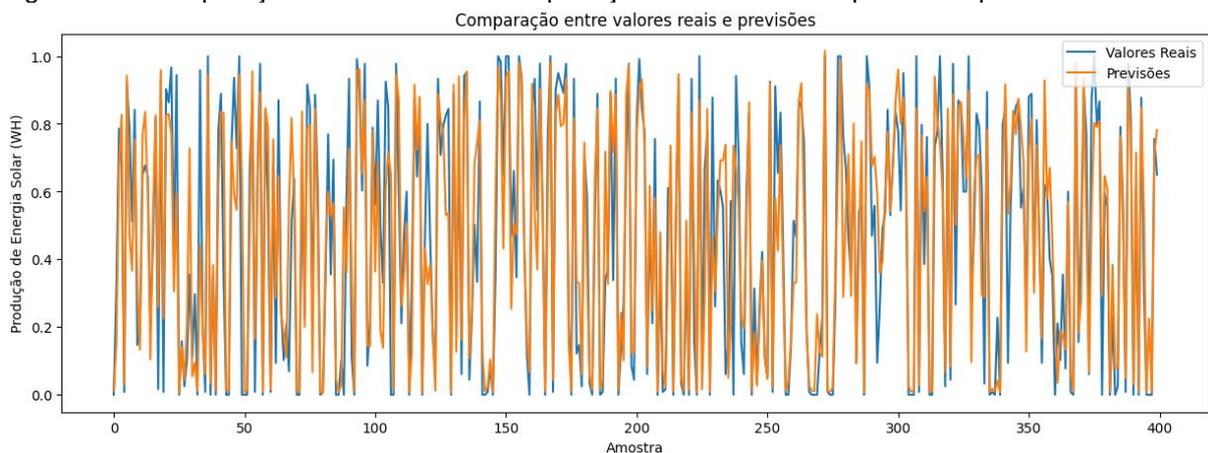
Fonte: Autor, 2023

A fusão direta dos conjuntos de dados das três usinas, no entanto, resultou em entradas triplicadas para a mesma data e hora, o que exigiu a inclusão de uma coluna adicional para identificar a usina correspondente. Essa modificação implica que, para futuras previsões, é necessário especificar a usina em questão, uma vez que esta se tornou um parâmetro de entrada essencial.

Visando aprimorar a generalização do modelo para aplicação em outras usinas, um novo experimento foi conduzido. Se as usinas estão na mesma região geográfica e, portanto, sujeitas às mesmas condições meteorológicas, mas apresentam variações sutis na produção, é prudente desenvolver um modelo que reconheça tanto as semelhanças quanto às discrepâncias individuais. A média ponderada foi adotada como método de agregação dos dados, atribuindo pesos diferenciados às usinas com base em sua capacidade máxima, resultando em um indicador mais representativo das condições produtivas locais.

Este novo experimento com média ponderada resultou em um MAE um pouco melhor de 0,113 indicando uma precisão levemente melhorada. Os resultados deste treinamento podem ser visualizados na figura 17:

Figura 17 – Comparação de valores real com predição de todas usinas por média ponderada

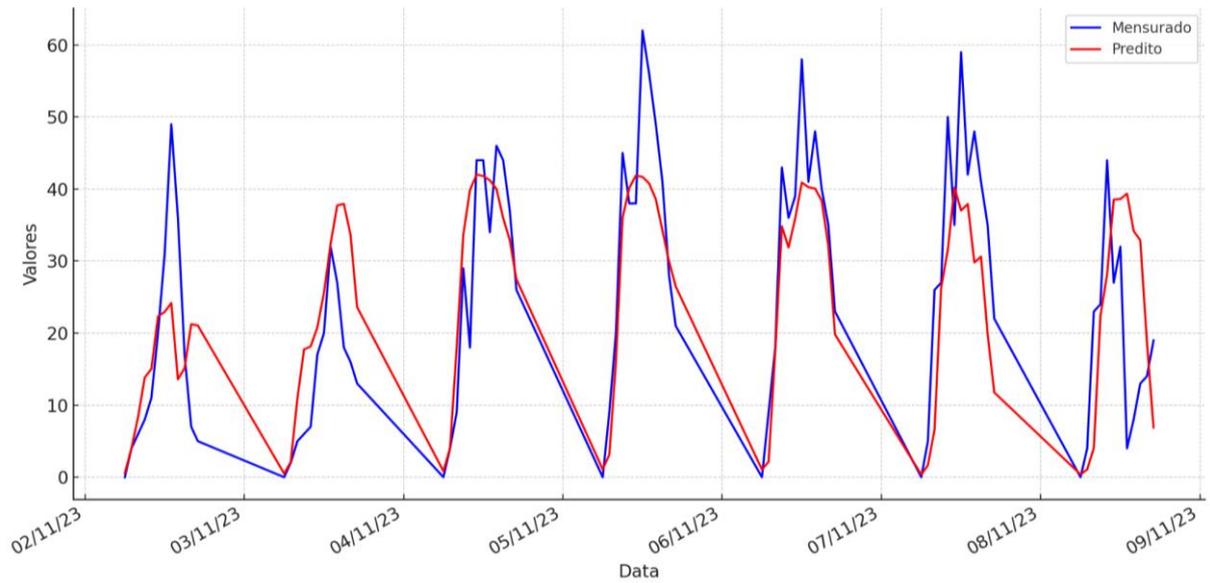


Fonte: Autor, 2023

Este estudo finaliza com a aplicação de um teste unitário, empregando dados futuros de uma instalação energética específica, para verificar a precisão de um modelo preditivo generalizado. Para a avaliação, optou-se por um período de observação de sete dias, visando examinar a eficácia do modelo em prever o fator de produção energética. Este período, compreendido entre 02 e 11 de outubro de 2023, envolveu o uso de dados da usina denominada 'Usina 1' conforme o quadro 2. Os dados selecionados foram submetidos a um tratamento analítico equivalente ao empregado no conjunto de treinamento, com a inclusão de variáveis normalizadas e a integração de informações meteorológicas previstas.

Utilizando o modelo previamente ajustado com dados de três usinas, conduziu-se a análise preditiva. Os resultados, ilustrados na figura 18, revelaram um Erro Médio Absoluto (MAE) de aproximadamente 8,27.

Figura 18 – Comparação da previsão de teste unitário



Fonte: Autor, 2023

A avaliação indica que, embora o modelo tenha demonstrado competência na captura dos padrões temporais, incluindo picos e vales de produção, as projeções apresentaram variações de amplitude quando comparadas aos dados reais. Apesar desta discrepância, o modelo se mostrou uma ferramenta promissora para a previsão de performance em um contexto geral.

CONSIDERAÇÕES FINAIS

O estudo demonstrou a robustez e aplicabilidade do modelo LSTM na previsão da produção de energia solar. A seleção e implementação criteriosa deste modelo permitiram a captura efetiva das dinâmicas temporais e características espaciais dos dados meteorológicos e de produção energética. A precisão alcançada, representada por métricas como RMSE e MAE, reforça a validade do LSTM como ferramenta preditiva eficaz nesse contexto.

Os processos de integração e pré-processamento de dados revelaram-se etapas cruciais, assegurando a qualidade do conjunto de dados utilizado. A engenharia de recursos e a seleção subsequente, orientadas pela técnica LASSO, aprimoraram a relevância das variáveis explicativas, contribuindo para a eficiência do modelo e mitigando o risco de *overfitting*. Este cuidado na preparação dos dados fortaleceu o alicerce sobre o qual as previsões foram construídas.

As correlações entre variáveis meteorológicas e a produção de energia solar identificadas evidenciaram o impacto significativo da radiação global e da temperatura do ar. Estas descobertas corroboram a importância de considerar tais variáveis em modelos de previsão e oferecem *insights* relevantes para utilização e otimização futura na tecnologia de painéis solares.

O desenvolvimento incremental e o refinamento do modelo LSTM foram estratégias efetivas, como demonstrado pela melhoria contínua nos indicadores de erro. A abordagem de treinamento, considerando as peculiaridades dos dados de produção energética, foi decisiva para a evolução da precisão do modelo.

A expansão dos testes para incluir múltiplas usinas e a aplicação de um modelo generalizado em uma usina específica forneceram evidências adicionais da capacidade do modelo de se adaptar e manter a precisão em diferentes cenários. Ainda que variações na amplitude das projeções tenham sido observadas, o desempenho geral indica um potencial significativo para o uso em previsões operacionais.

As limitações deste estudo, incluem a dependência de dados históricos abrangentes para treinamento efetivo e a sensibilidade do modelo a interrupções nos dados, o que pode exigir técnicas de interpolação para preservar a sequência temporal. A variação nas condições de produção entre diferentes usinas também desafia a homogeneização dos dados para o treinamento de modelos. Ainda, as

correlações observadas não implicam causalidade, e essa distinção deve ser considerada na interpretação dos resultados.

Este trabalho abre caminho para futuras pesquisas na exploração de modelos LSTM híbridos, que combinem diferentes arquiteturas de redes neurais para refinar ainda mais as previsões. Além disso, a investigação do impacto de variáveis emergentes, como mudanças climáticas a longo prazo, poderia enriquecer a compreensão dos padrões de produção de energia solar. Finalmente, a aplicação de técnicas de aprendizado de máquina em tempo real para previsões imediatas poderia ser uma área promissora para estudos subsequentes.

Diante da importância cada vez maior das energias renováveis no panorama mundial, estudos como o atual adquirem uma importância estratégica crítica. Os resultados aqui relatados não apenas reforçam a eficácia dos modelos LSTM na previsão de produção energética, mas também destacam a necessidade contínua do aperfeiçoamento das metodologias de modelagem para acompanhar as inovações no setor de energia solar.

As contribuições deste trabalho à área de energia solar e aprendizado de máquina são notáveis, oferecendo um modelo diferente para previsões de energia solar e insights práticos que podem ser úteis para operadores de instalações, reguladores e formuladores de políticas. Espera-se que os resultados inspirem pesquisas futuras, pavimentando o caminho para a exploração de abordagens preditivas mais avançadas e a inclusão de outras variáveis relevantes.

REFERÊNCIAS

- ABSOLAR. **Panorama da solar fotovoltaica no Brasil e no mundo**. São Paulo: Absolar, 2023. Disponível em: <https://www.absolar.org.br/mercado/infografico/>. Acesso em: 22 set. 2023.
- BASHA S. M.; RAJPUT, D. S. *Survey on evaluating the performance of machine learning algorithms: past contributions and future roadmap*. **Deep Learning and Parallel Computing Environment for Bioengineering Systems, Academic Press**, p. 153-164, 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/B9780128167182000166?via%3Dihub>. Acesso em: 14 jul. 2023.
- BILGEN, S. *Structure and environmental impact of global energy consumption, Renewable and Sustainable Energy Reviews*, v. 38, p. 890-902, 2014. ISSN 1364-0321. Disponível em: <https://doi.org/10.1016/j.rser.2014.07.004>. Acesso em: 14 jul. 2023.
- BISONG, E. **In building machine learning and deep learning models on Google Cloud platform: a comprehensive guide to beginners**. [s.l.]: Apress, 2019. Disponível em: <https://link.springer.com/book/10.1007/978-1-4842-4470-8>. Acesso em: 14 jul. 2023.
- BORENSTEIN, Severin. *The private and public economics of renewable electricity generation. Journal of Economic Perspectives*. **Journal of Economic Perspectives**, v. 26, 2012. Disponível em: <https://faculty.haas.berkeley.edu/borenste/download/JEP12Renewables.pdf>. Acesso em: 14 jul. 2023.
- BRITISH PETROLEUM (BP). **Statistical Review of World Energy, 69th edition**. 2020. Disponível em: <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2020-full-report.pdf>. Acesso em: 12 maio. 2023.
- CAMPOS, Antonio; FARIAS, Aécio Vinicius; FARIAS, Leandro; JÚNIOR, Valter; COSTA, Bruno; MACEDO, Idalcácio. *Investigação experimental da geração de energia elétrica solar fotovoltaica*. **Holos**, v. 3, n. 26, 2010. Disponível em: https://www.researchgate.net/publication/277094032_INVESTIGACAO_EXPERIMENTAL_DA_GERACAO_DE_ENERGIA_ELETRICA_SOLAR_FOTOVOLTAICA. Acesso em: 14 jul. 2023.
- FACELI, K.; LORENA, A. C.; GAMA, J.; *et al.* **Inteligência artificial - uma abordagem de aprendizado de máquina**. Barueri: Grupo GEN, 2021. E-book. ISBN 9788521637509. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521637509/>. Acesso em: 11 jun. 2023.
- GARG, R. *Methodology for research I*. **Indian journal of anaesthesia**, v. 60, n. 9, p. 640-645, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5037944/>. Acesso em: 11 jun. 2023.

GIELEN, D.; BOSHELL, F.; SAYGIN, D.; BAZILIAN, M. D.; WAGNER, N.; GORINI, R. *The role of renewable energy in the global energy transformation*, **Energy Strategy Reviews**, v. 24, p. 38-50, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2211467X19300082>. Acesso em:

GÓRECKIET, K.; DABROWSKI, J.; KRAC, E. *SPICE-Aided Modeling of Daily and Seasonal Changes in Properties of the Actual Photovoltaic Installation*. **Energies**, 2021. Disponível em: <https://www.semanticscholar.org/paper/SPICE-Aided-Modeling-of-Daily-and-Seasonal-Changes-G%C3%B3recki-D%C4%85browski/f98ea2c6622a2a57f96e4c105fff017fd2e1eb1a>. Acesso em 1 jul. 2023.

IRENA (International Renewable Energy Agency). **Global energy transformation: a roadmap to 2050**. (2019 edition), *International Renewable Energy Agency*, Abu Dhabi, 2020. Disponível em: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Apr/IRENA_Global_Energy_Transformation_2019.pdf?rev=6ea97044a1274c6c8ffe4a116ab17b8f. Acesso em: 5 maio 2023.

IRENA (*International Renewable Energy Agency*), **Renewable power generation cost**. (2019 edition), *International Renewable Energy Agency*, Abu Dhabi, 2019. Disponível em: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Jun/IRENA_Power_Generation_Costs_2019.pdf?rev=77ebbae10ca34ef98909a59e39470906. Acesso em 5 maio 2023.

JAILANI, N.L.M.; DHANASEGARAN, J.K.; ALKAWSI, G.; ALKAHTANI, A.A.; PHING, C.C.; BAASHAR, Y.; CAPRETZ, L.F.; AL-SHETWI, A.Q.; TIONG, S.K. Investigando o poder dos modelos baseados em LSTM na previsão de energia solar. **Processes**, v. 11, p. 1382, 2023. Disponível em: <https://doi.org/10.3390/pr11051382>. Acesso em 1 jul. 2023.

JALLAL, M. A.; CHABAA, S.; ZEROUAL, A. *A new artificial multi-neural approach to estimate the hourly global solar radiation in a semi-arid climate site*. **Theoretical and Applied Climatology**, 2019. Disponível em: <https://sci-hub.se/https://link.springer.com/article/10.1007/s00704-019-03033-1>. Acesso em: 25 maio 2023.

KALOGIROU, S. A.; SENCAN, A. **Solar collectors and panels: theory and applications**. Limassol: Cyprus University of Technology. 2010. Disponível em: <https://www.intechopen.com/chapters/12231>. Acesso em: 25 maio 2023.

KHAN, S.; PAUL, D.; MOMTAHAN, P.; ALOGAILY, M. *Artificial intelligence framework for smart city microgrids: State of the art, challenges, and opportunities*, 2018. *International Conference on Fog and Mobile Edge Computing (FMEC)*, 3. Barcelona, 2018. p. 283-288. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8364080>. Acesso em: 25 maio 2023.

LENZ, M. L.; NEUMANN, F. B.; SANTARELLI, R.; et al. **Fundamentos de aprendizagem de máquina**. Porto Alegre: Grupo A, 2020. E-book. ISBN 9786556900902. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556900902/>. Acesso em: 11 jun. 2023. Acesso em 1 jun. 2023.

LI, Zhe; YANG, Jian; DEZFULI, Asareh Pouya Nejad. *Study on the Influence of Light Intensity on the Performance of Solar Cell*. **International Journal of Photoenergy**, v. 2021, jan. 2021. Disponível em: <https://downloads.hindawi.com/journals/ijp/2021/6648739.pdf>. Acesso em: 11 jun. 2023.

MACKENZIE, Wood. *How Solar Is Central To The Energy Transition*. **Forbes**, 2021. Disponível em: <https://www.forbes.com/sites/woodmackenzie/2021/03/18/how-solar-is-central-to-the-energy-transition/?sh=659e68c85d04>. Acesso em: 11 abr. 2023.

MAITHA, Al S. H.; ASSI, A. H.; HEJASE, H. A. N. *Using MATLAB to develop artificial neural network models for predicting global solar radiation in al Ain city – UAE*. **IntechOpen**, out. 2011. Disponível em: <https://dx.doi.org/10.5772/25213>. Acesso em: 21 nov. 2022.

MUTASA S.; SUN, S.; HA, R. *Understanding artificial intelligence based radiology studies: what is overfitting?*, **Clinical Imaging**, v. 65, p. 96-99, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0899707120301376>. Acesso em: 22 maio 2023.

NEOSOLAR. **Equipamentos do kit solar I Neosolar**. 2022a. Disponível em: <https://www.neosolar.com.br/aprenda/saiba-mais/kit-solar>. Acesso em: 22 maio 2023.

NEOSOLAR. **Tipos de placa solar**. 2022b. Disponível em: <https://www.neosolar.com.br/aprenda/saiba-mais/placa-solar-fotovoltaica/tipos>. Acesso em: 21 nov. 2022.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. J. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas brasileiro de energia solar**. São José dos Campos: INPE, 2017. Disponível em: <http://urlib.net/ibi/8JMKD3MGP3W34P/3PERDJE>. Acesso em 4 abr. 2023.

PHILIPPI JR, Lineu Belico dos Reis Arlindo. **Energia e sustentabilidade**. São Paulo: Manole, 2016. E-book. ISBN 9786555761313. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786555761313/> Acesso em: 11 jun. 2023.

PINHO, J. T.; GALDINO, M. A. **Manual de engenharia para sistemas fotovoltaicos**. Rio de Janeiro: CEPEL, 2014. Disponível em: http://www.cresesb.cepel.br/publicacoes/download/Manual_de_Engenharia_FV_2014.pdf. Acesso em: 25 maio 2023.

POLIKARPOVA I.; KAKIS, R.; PAKERE, I.; BLUMBERGA, D. *Optimizing Large-Scale Solar Field Efficiency: Latvia Case Study*. **Energie**, 2021. *Institute of Energy Systems and Environment*. Disponível em: <https://www.mdpi.com/1996-1073/14/14/4171>. Acesso em: 25 maio 2023.

RAPHAEL Ellen; NASCIMENTO, Mariana Silva; SZOSTAK, Rodrigo; SCHIAVON, Marco Antônio; NOGUEIRA, Ana Flávia. *Células solares de perovskitas: uma nova tecnologia emergente*. **Quim. Nova**, v. 41, n. 1, p. 61-74, 2018. Disponível em: <https://www.scielo.br/j/qn/a/7fGqwHXzFZwLMr8X6vMhVZR/?lang=pt>. Acesso em: 2 abr. 2023.

RODRIGUES, K. A. S. **LASSO clássico e Bayesiano**. São Paulo: Instituto de Matemática e Estatística. Universidade de São Paulo. 2018. Disponível em: https://www.ime.usp.br/~kevinx/LASSO_Texto.pdf. Acesso em: 2 abr. 2023.

RUSSEL, Stuart; NORVIG, Peter. **Inteligência artificial**. Barueri: Grupo GEN, 2016. E-book. ISBN 9788595156104. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156104/>. Acesso em: 11 jun. 2023.

SINGH, D.; SINGH, B. *Investigating the impact of data normalization on classification performance*. **Applied Soft Computing**, v. 97, p. 105524, 1 dez. 2020. Elsevier. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1568494619302947>. Acesso em: 2 abr. 2023.

SOTO, E. A.; BOSMAN, L. B.; WOLLEGA, E.; LEON-SALAS, W. D. *Analysis of Grid Disturbances Caused by Massive Integration of Utility Level Solar Power Systems*. **Eng.**, v.3, n. 2, p. 236-253, 2022. Disponível em: <https://www.mdpi.com/2673-4117/3/2/18>. Acesso em: 14 jun. 2023.

TEIXEIRA, Marco Antonio Casadei; RAMOS, Heidy Rodriguez; AGUIAR, Alexandre de Oliveira. Perspectivas de novos materiais alternativos ao silício para a produção de células solares fotovoltaicas: uma revisão sistemática da literatura. **Revista Nacional de gerenciamento de cidades**. ISSN eletrônico 2318-8472, v. 09, n. 71, 2021. Disponível em: https://scholar.archive.org/work/n7bhutykxbgbpiqxruijmia3fq/access/wayback/https://publicacoes.amigosdanatureza.org.br/index.php/gerenciamento_de_cidades/article/download/2953/2777. Acesso em: 2 abr. 2023.

VIAN, Ângelo. **Energia solar fundamentos tecnologia e aplicações**. São Paulo: Blucher, 2021. E-book. ISBN 9786555500592. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786555500592/>. Acesso em: 05 jun. 2023. Acesso em: 2 abr. 2023.

WAN, C.; ZHAO, J.; SONG, Y.; XU, Z.; LIN, J.; HU, Z. *Photovoltaic and solar power forecasting for smart grid energy management*. **CSEE Journal of Power and Energy Systems**, v. 1, n. 4, p. 38-46, dez. 2015. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7377167>. Acesso em: 20 jun. 2023.

WENTZ, V. H. **Avaliação de modelos de inteligência artificial aplicados à predição de irradiação solar**. Monografia (Engenharia Física) – Instituto Latino-Americano de Ciências da Vida e da Natureza da Universidade Federal da Integração Latino-Americana, Foz do Iguaçu, 2021. Disponível em: https://dspace.unila.edu.br/bitstream/handle/123456789/6319/TCC_Victor_Wentz_Final.pdf?sequence=3&isAllowed=y. Acesso em: 12 abr. 2023.

XIAO, X.; HE, X.; LI, Y.; HUANG, B. *The influence of distributed photovoltaic grid-connected on distribution network voltage*. In: **ASIA CONFERENCE ON POWER AND ELECTRICAL ENGINEERING (ACPEE)**, 8. IEEE, 2023. p. 1311-1318. Disponível em: <https://ieeexplore.ieee.org/document/10135557>. Acesso em: 20 jun. 2023.

YAO, S.; KANG, Q.; ZHOU, M.; ABUSORRAH, A.; AL-TURKI, Y. *Intelligent and data-driven fault detection of photovoltaic plants*. **Processes**, 2021. Disponível em: <https://www.semanticscholar.org/paper/Intelligent-and-Data-Driven-Fault-Detection-of-Yao-Kang/466dd43fd72ef30cc3755edd5a93b3d0e867fc8c>. Acesso em: 12 abr. 2023.

ZAFARANI, Reza; EFTEKHARNEJAD, Sara; PATEL, Urvi. **Assessing the Utility of Weather Data for Photovoltaic Power Prediction**. *Department of Electrical Engineering and Computer Science, Syracuse University*, 2018. Disponível em: <https://arxiv.org/pdf/1802.03913.pdf>. Acesso em: 14 jul. 2023.

ZAINUDDIN, N. F.; MOHAMMED, M. N.; AL-ZUBAIDI, N.; KHOGALI, S. I. *Design and Development of Smart Self-Cleaning Solar Panel System*. IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC CONTROL AND INTELLIGENT SYSTEMS (I2CACIS), Selangor, p. 40-43, 2019. Disponível em: <https://ieeexplore.ieee.org/document/8825056>. Acesso em: 22 maio 2023.

ZHANG, Y.; GUO, W.; RAY, S. *On the consistency of feature selection with Lasso for non-linear targets*. INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 33. v. 48, p. 183-191, 2016. Disponível em: <https://proceedings.mlr.press/v48/zhanga16.html>.