

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
CURSO DE GRADUAÇÃO DE ENGENHARIA ELÉTRICA
HABILITAÇÃO EM ROBÓTICA E AUTOMAÇÃO INDUSTRIAL

Luís Eduardo Brandão Canan

Previsão de Geração Fotovoltaica com Redes *Long Short-Term Memory*:
Uma Revisão da Literatura

Juiz de Fora
2025

Luís Eduardo Brandão Canan

**Previsão de Geração Fotovoltaica com Redes *Long Short-Term Memory*:
Uma Revisão da Literatura**

Trabalho de conclusão de curso apresentado
à Universidade Federal de Juiz de Fora como
requisito parcial à obtenção do grau em ba-
charel em Engenharia Elétrica

Orientador: Prof. Dr. Daniel Discini Silveira

Coorientador: Prof. Dr. Eduardo Pestana de Aguiar

Juiz de Fora

2025

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Canan, Luís Eduardo Brandão.

Previsão de Geração Fotovoltaica com Redes *Long Short-Term Memory*:
Uma Revisão da Literatura / Luís Eduardo Brandão Canan. – 2025.
39 f. : il.

Orientador: Daniel Discini Silveira

Coorientador: Eduardo Pestana de Aguiar

Trabalho de Conclusão de Curso (graduação) – Universidade Federal de
Juiz de Fora, Curso de Graduação de Engenharia Elétrica. Habilitação em
Robótica e Automação Industrial, 2025.

1. Inteligência Artificial. 2. Séries Temporais. 3. Predição. I. Silveira,
Daniel Discini, orient. II. Aguiar, Eduardo Pestana de, coorient. III. Título.

Luís Eduardo Brandão Canan

**Previsão de Geração Fotovoltaica com Redes *Long Short-Term Memory*:
Uma Revisão da Literatura**

Trabalho de conclusão de curso apresentado
à Universidade Federal de Juiz de Fora como
requisito parcial à obtenção do grau em ba-
charel em Engenharia Elétrica

Aprovada em 21 de Março de 2025

BANCA EXAMINADORA

Prof. Dr. Daniel Discini Silveira - Orientador
Universidade Federal de Juiz de Fora

Prof. Dr. Thiago Vieira Nogueira Coelho
Universidade Federal de Juiz de Fora

Me. Matheus Augusto Malta Ferreira
Universidade Federal de Juiz de Fora

Dedico este trabalho à minha família e a todas as pessoas
que me ajudaram e apoiaram até o presente momento.

AGRADECIMENTOS

A toda minha família, pelo apoio e companheirismo em todas as situações e cenários ao longo destes anos.

Aos meus professores orientadores pelos ensinamentos e atenção durante todo o desenvolvimento deste trabalho.

A Universidade Federal de Juiz de Fora, a Faculdade de Engenharia e à todos os docentes pelo suporte ao longo da graduação.

"O caráter de um homem é o seu destino."
(Heráclito)

RESUMO

A previsão de geração fotovoltaica é essencial para otimizar a operação de sistemas elétricos e garantir a gestão eficiente de energia. Nos últimos anos, métodos baseados em aprendizado de máquina e redes neurais têm sido amplamente utilizados para melhorar a precisão de tais previsões, o que é fundamental para sistemas com fontes renováveis, como as microrredes. Este trabalho oferece uma revisão da literatura sobre as técnicas aplicadas à previsão da geração fotovoltaica, com ênfase em abordagens que utilizam redes *Long Short-Term Memory* (LSTM). Além disso, são abordados os principais desafios e tendências atuais da área, com o objetivo de fornecer uma base teórica sólida para futuras pesquisas e aplicações.

Palavras-chave: Aprendizado de Máquina; Inteligência Artificial; LSTM; Previsão.

ABSTRACT

The forecasting of photovoltaic generation is essential for optimizing the operation of electrical systems and ensuring efficient energy management. In recent years, machine learning and neural network-based methods have been widely used to improve the accuracy of such predictions, which is crucial for systems with renewable sources, such as micro-grids. This paper provides a literature review on the techniques applied to photovoltaic generation forecasting, with an emphasis on approaches using Long Short-Term Memory (LSTM) networks. Additionally, the main challenges and current trends in the field are discussed, aiming to provide a solid theoretical foundation for future research and applications.

Keywords: Artificial Intelligence; Forecasting; LSTM; Machine Learning.

LISTA DE ILUSTRAÇÕES

| | | |
|-----------|--|----|
| Figura 1 | – Matriz elétrica brasileira 2023-2027 - ONS | 12 |
| Figura 2 | – Divisão de Dados | 17 |
| Figura 3 | – Melhor Modelo | 19 |
| Figura 4 | – Representação de Pesos, Viéses e Funções de Ativação | 21 |
| Figura 5 | – Representação Simplificada de uma Rede Neural | 22 |
| Figura 6 | – Representação da Função ReLU | 24 |
| Figura 7 | – Representação da Função Tanh | 24 |
| Figura 8 | – Representação da Função Sigmoidal | 25 |
| Figura 9 | – Representação Simplificada de uma Rede Neural Recorrente | 26 |
| Figura 10 | – Representação Simplificada de uma LSTM | 27 |
| Figura 11 | – <i>Forget Gate</i> | 28 |
| Figura 12 | – <i>Input Gate</i> | 28 |
| Figura 13 | – <i>Output Gate</i> | 29 |
| Figura 14 | – Previsão de Variável Usando Três <i>Lags</i> na Entrada | 30 |
| Figura 15 | – Exemplo de Validação <i>K-Fold</i> | 32 |
| Figura 16 | – Exemplo de Validação com Janela Expansiva | 33 |
| Figura 17 | – Exemplo de Validação com Janela Móvel | 33 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|---|
| Adam | Adaptive Moment Estimation |
| ARIMA | AutoRegressive Integrated Moving Average |
| ARMA | AutoRegressive Moving Average |
| BESS | Sistemas de Armazenamento de Energia em Baterias |
| BMS | Sistema de Gerenciamento de Baterias |
| CA | Corrente Alternada |
| CC | Corrente Contínua |
| EMS | Sistema de Gerenciamento de Energia |
| LSTM | Long Short-Term Memory |
| MAPE | Erro Percentual Absoluto Médio |
| MAE | Erro Médio Absoluto |
| MSE | Erro Quadrático Médio |
| ONS | Operador Nacional do Sistema Elétrico |
| PCS | Sistema de Controle de Potência |
| ReLU | Rectified Linear Unit |
| RMSE | Raiz do Erro Quadrático Médio |
| RNN | Recurrent Neural Network |
| SARIMA | Seasonal AutoRegressive Integrated Moving Average |
| SVM | Máquinas de Vetores de Suporte |
| tanh | Tangente Hiperbólica |

LISTA DE SÍMBOLOS

| | |
|-------------|--|
| $f()$ | Função de ativação genérica |
| i | Índice da iteração das amostras do conjunto de dados |
| n | Número total de amostras no conjunto de dados |
| σ | Função sigmoide |
| Σ | Somatório |
| t | Instante de tempo |
| x | Amostra de entrada genérica |
| y_i | Valor real da saída |
| \hat{y}_i | Valor previsto da saída |

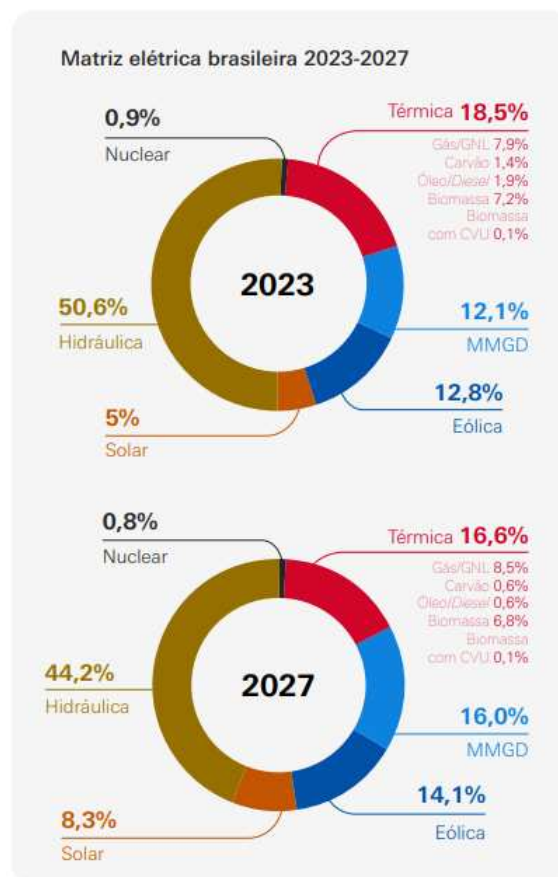
SUMÁRIO

| | | |
|---------|--|----|
| 1 | INTRODUÇÃO | 12 |
| 2 | CONCEITOS FUNDAMENTAIS | 14 |
| 2.1 | ENERGIA SOLAR FOTOVOLTAICA E SUA PREVISÃO . . . | 14 |
| 2.1.1 | Funcionamento dos sistemas fotovoltaicos | 14 |
| 2.1.2 | Importância da previsão de geração fotovoltaica | 15 |
| 2.1.3 | Séries Temporais e Previsões | 15 |
| 2.2 | FUNDAMENTOS DE APRENDIZADO DE MÁQUINA | 16 |
| 2.2.1 | Métricas de Erro | 17 |
| 2.2.2 | Otimização x Generalização | 18 |
| 2.3 | PRINCIPAIS ABORDAGENS PARA PREVISÃO | 19 |
| 3 | REDES NEURAIS NA PREVISÃO DE GERAÇÃO SOLAR | 21 |
| 3.1 | FUNDAMENTOS DE REDES NEURAIS | 22 |
| 3.1.1 | Retropropagação | 22 |
| 3.1.2 | Gradiente Descendente | 23 |
| 3.1.3 | Funções de Ativação | 23 |
| 3.2 | FUNDAMENTOS DE APRENDIZADO PROFUNDO | 25 |
| 3.3 | MODELOS DE APRENDIZADO PROFUNDO PARA PREVI- SÃO | 26 |
| 3.3.1 | Redes Neurais Recorrentes (RNNs) | 26 |
| 3.3.1.1 | Explosão/Desaparecimento do Gradiente | 26 |
| 3.3.2 | Long Short-Term Memory (LSTMs) | 27 |
| 3.4 | TÉCNICAS COMPLEMENTARES PARA MELHORIA DE PRE- VISÃO | 29 |
| 3.4.1 | Pré-processamento de Dados | 29 |
| 3.4.1.1 | Maldição da Dimensionalidade | 30 |
| 3.4.2 | Variáveis Endógenas de Séries Temporais | 30 |
| 3.4.3 | Métodos de Regularização | 31 |
| 4 | ANÁLISE DE RESULTADOS | 34 |
| 5 | CONCLUSÃO | 36 |
| | REFERÊNCIAS | 38 |

1 INTRODUÇÃO

Desde as últimas décadas, a utilização da energia fotovoltaica tem crescido significativamente no mundo todo. Vários países passaram a investir e estudar esse tipo de geração com maior atenção, isso devido à necessidade de transição para fontes renováveis. No Brasil, esse cenário é particularmente favorável, já que o país apresenta altos índices de incidência solar, proporcionando grande potencial para a geração dessa energia. Como resultado, a participação da geração fotovoltaica na matriz energética do país vem aumentando continuamente, com previsões indicando um crescimento ainda maior nos próximos anos. A Figura 1 mostra a matriz energética brasileira do Sistema Interligado Nacional, e a previsão do Operador Nacional do Sistema Elétrico para os próximos anos.

Figura 1 - Matriz elétrica brasileira 2023-2027



Fonte: Operador Nacional do Sistema Elétrico. Disponível em:
https://www.ons.org.br/AcervoDigital/DocumentosEPublicacoes/2023-Relatorio-Anual-acessivel_21032024.pdf
 Nota: Relatório Anual Acessível/2023.

A expansão da energia solar não é limitada somente a usinas de grande porte, mas também se destaca em sistemas distribuídos, como microrredes e redes isoladas. Nestes contextos, a gestão eficiente de sistemas de gerenciamento de energia e do fluxo de potência torna-se indispensável para garantir estabilidade e confiabilidade na operação. Um dos principais desafios enfrentados nesses sistemas é a alta variabilidade da geração fotovoltaica, sendo ela, diretamente influenciada por fatores meteorológicos. Para mitigar tal impacto, técnicas de previsão são amplamente empregadas, já que processos de otimização deste tipo de sistema sempre devem se orientar a partir de pontos futuros, permitindo a integração da energia solar com outras fontes de geração e armazenamento de energia, além de reduzir custos operacionais.

Dentre os diversos métodos de previsão, modelos baseados em aprendizado de máquina e aprendizado profundo se destacam pela capacidade de capturar padrões complexos em séries temporais. *Long Short-Term Memory* (LSTMs), que surgiram a partir de Redes Neurais Recorrentes (RNNs - *Recurrent Neural Networks*), vem demonstrando grande potencial na previsão da geração fotovoltaica, permitindo análises mais precisas e robustas. Porém, antes de implementar essas abordagens, é essencial avaliar suas vantagens e limitações e identificar as estratégias mais eficazes.

Este trabalho tem como objetivo realizar uma revisão da literatura sobre a aplicação de redes neurais LSTM na previsão da geração fotovoltaica. Serão abordados temas desde a análise de dados e pré-processamento até o funcionamento do modelo. Dessa forma, busca-se fornecer um embasamento teórico sólido para futuras pesquisas e aplicações na área.

A revisão da literatura foi conduzida, principalmente, a partir de artigos científicos publicados em bases de dados reconhecidas, como ResearchGate, ScienceDirect e Google Scholar, além de artigos provenientes de repositórios acadêmicos de instituições federais de pesquisa, como a Universidade Federal do Ceará, por exemplo. Foram utilizadas palavras-chave como "photovoltaic power", "time series forecasting", "machine learning", "deep learning" e "LSTM for solar power forecasting". Para garantir a relevância dos trabalhos analisados, em relação às redes neurais profundas, foram incluídos quinze artigos publicados nos últimos seis anos, e em relação ao aprendizado de máquina foram utilizados mais cinco artigos de até treze anos de publicação. Artigos muito antigos ou que abordavam aprendizado de máquina fora do contexto de previsão foram excluídos. A análise foi realizada comparando diferentes contextos utilizados na previsão da geração fotovoltaica com o uso de redes neurais, com foco nos métodos baseados em redes LSTM.

A estrutura do trabalho é organizada nos capítulos seguintes, discutindo os conceitos fundamentais, para entender o contexto de previsões de geração fotovoltaica; previsão com redes neurais, para analisar o escopo de estudo; resultados, buscando análises e comparações de métodos de previsão; e conclusão, discutindo tendências e desafios.

2 CONCEITOS FUNDAMENTAIS

Para entender o cenário abordado no artigo, é necessário compreender conceitos básicos do estudo, como o que é energia solar fotovoltaica; como os sistemas que utilizam dessa tecnologia são enxergados; como a previsão da geração impacta eles; o que seria uma série temporal e o que é aprendizado de máquina. Abordados os contextos, é possível entender como redes neurais são aplicadas no problema.

2.1 ENERGIA SOLAR FOTOVOLTAICA E SUA PREVISÃO

A energia solar é uma das principais fontes renováveis hoje, consequência da tecnologia desenvolvida em volta dos fotovoltaicos, da abundância e do potencial que essa fonte tem para aplicações em diferentes escalas. O processo de conversão da luz solar em eletricidade ocorre devido ao efeito fotovoltaico, fenômeno onde materiais semicondutores, como o silício por exemplo, geram uma corrente elétrica em resposta a exposição à radiação solar (1). Esse é o princípio de funcionamento das células fotovoltaicas, que por sua vez, são base para os módulos solares utilizados em sistemas de geração de energia.

2.1.1 Funcionamento dos sistemas fotovoltaicos

Os sistemas fotovoltaicos podem ser classificados em sistemas conectados à rede (*grid-tied*) e sistemas isolados (*off-grid*). Nos sistemas *grid-tied*, a energia gerada pelos painéis pode ser injetada na rede elétrica, podendo ser utilizada localmente ou enviada para a concessionária, permitindo créditos energéticos ao proprietário. Já os sistemas *off-grid* operam de forma independente, geralmente associados a geradores e baterias para o armazenamento da energia gerada, garantindo fornecimento contínuo mesmo nos períodos onde não há incidência solar.

A estrutura básica de um sistema fotovoltaico normalmente inclui os seguintes componentes:

- a) Módulos fotovoltaicos: compostos por células fotovoltaicas, realizam a conversão da luz em eletricidade (2).
- b) Inversores: convertem a corrente contínua (CC) gerada pelos painéis em corrente alternada (CA), compatível com a rede elétrica e dispositivos usuais.
- c) Sistemas de armazenamento: utilizam baterias para armazenar a energia excedente, permitindo uso posterior em momentos de baixa geração.
- d) Medidores e sistemas de monitoramento: acompanham o desempenho do sistema, otimizando sua operação e manutenção.
- e) Sistemas de Gerenciamento: sistemas fotovoltaicos, principalmente aqueles com Sistemas de Armazenamento de Energia em Baterias (BESS - *Battery Energy*

Storage System), possuem abordagens para realizar o controle da rede, sendo essas o Sistema de Controle de Potência (PCS - *Power Control System*), Sistema de Gerenciamento de Baterias (BMS - *Battery Management System*) e Sistema de Gerenciamento de Energia (EMS - *Energy Management System*) (3).

A eficiência de um sistema fotovoltaico depende de fatores como irradiância solar, temperatura ambiente, velocidade do vento ao longo do dia, ângulo de inclinação dos painéis e possíveis sombreamentos. Assim, prever com precisão a geração ao longo do tempo é fundamental para otimizar a operação destes sistemas.

2.1.2 Importância da previsão de geração fotovoltaica

A geração fotovoltaica é, por natureza, intermitente, já que depende de condições atmosféricas variáveis, como, além das já apontadas, precipitação, cobertura de nuvens, e outros fatores meteorológicos (4). Essa alta variabilidade impacta diretamente a estabilidade dos sistemas elétricos, especialmente em microrredes e redes elétricas inteligentes (*smart grids*), onde a junção de fontes renováveis exige estratégias eficientes de controle e gerenciamento de energia.

A previsão precisa da geração fotovoltaica possibilita uma melhor gestão da operação da rede elétrica, otimização de armazenamento de energia, redução de custos operacionais e maior confiabilidade. Já que nos permite maior eficiência na alocação de recursos, podendo garantir equilíbrio entre compra e venda de energia, ou auxiliar no controle de carga e descarga de baterias, por exemplo, além de minimizar os impactos da intermitência na estabilidade da rede elétrica.

Além dos modelos estatísticos clássicos, modelos baseados em inteligência artificial, especialmente aprendizado profundo, têm sido cada vez mais aplicados para melhorar a previsão da geração fotovoltaica, permitindo maior precisão na estimativa da produção energética. Essas técnicas são fundamentais para garantir um melhor uso dessa energia e permitir a expansão sustentável das fontes renováveis no cenário energético mundial.

2.1.3 Séries Temporais e Previsões

A previsão de geração solar é um desafio complexo, isso devido à alta variabilidade que está exposta. Em tal situação, as séries temporais desempenham o papel fundamental de fornecer contexto, já que representam visualizações históricas de padrões. Técnicas que utilizam inteligência artificial buscam aprimorar o entendimento desses padrões e projetar estimativas para o futuro, realizando a predição da geração fotovoltaica, tornando os sistemas mais robustos.

Uma série temporal é uma sequência de informações na mesma escala ao longo do tempo, onde a variável alvo da previsão, de forma generalizada, é influenciada por sazonalidades, tendências e ruídos. A série é considerada como tendo uma tendência

quando há uma mudança que evolui lentamente ao longo do tempo. Ela possui uma componente sazonal quando tem movimentos cíclicos padronizados. E o ruído, seria o que resta dela, caso as outras componentes citadas sejam removidas (5).

Quando a série temporal não apresenta alterações nas suas propriedades estatísticas ao longo do tempo, como média, variância e autocovariância ela é considerada estacionária. Em casos contrários, e mais comuns, ela é não-estacionária. Normalmente, quando há sazonalidade e tendência, como no caso da geração solar, ela apresenta não-estacionariedade.

Na geração fotovoltaica, a série temporal representa a potência gerada ao longo do tempo, sendo impactada pelos fatores já mencionados, que em certos casos, podem ser usados como variáveis de entrada do algoritmo de previsão. A análise destas séries permite entender e capturar os padrões recorrentes na geração de energia, o que facilita a construção de modelos preditivos mais precisos.

Para obter previsões precisas, diversos métodos de modelagem de séries temporais podem ser aplicados, sendo as redes neurais profundas, ou seja, redes neurais que utilizam do conceito de aprendizado profundo, uma das abordagens mais promissoras.

2.2 FUNDAMENTOS DE APRENDIZADO DE MÁQUINA

Aprendizado de máquina (*Machine Learning*) é uma área da inteligência artificial, que busca modelos computacionais capazes de aprender padrões a partir de dados. Não necessariamente tendo que seguir regras pré-definidas, esses sistemas usam os dados históricos para melhorar seu desempenho no decorrer do tempo, ficando cada vez mais precisos em suas previsões (6).

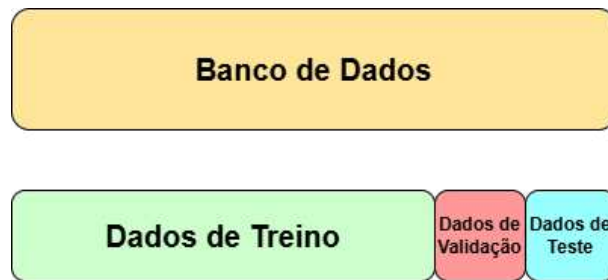
O funcionamento básico de um aprendizado de máquina é construído a partir de três pilares principais, primeiramente os Dados, que são as informações usadas para treinar e validar o modelo; o segundo pilar é o Modelo, que é a representação computacional que captura os padrões desses dados; por fim o Algoritmo, que é o método usado para treinar o modelo (Regressão Linear, Árvores de Decisão, Redes Neurais, entre outros).

Além disso, vale ressaltar que existem três tipos de aprendizado de máquina. Aprendizado Supervisionado, onde o modelo é treinado com dados rotulados, ou seja, cada entrada tem uma saída definida; Aprendizado Não-Supervisionado, onde o modelo é treinado com dados não rotulados, sendo assim, busca identificar padrões ou agrupamentos nos dados; e por último, Aprendizado por Reforço, onde o modelo aprende por tentativa e erro, recebendo resposta com base em suas ações.

Os modelos de aprendizado de máquina passam por um treinamento. Para redes neurais, e para a previsão de geração solar, é muito comum treinamentos supervisionados. Usando essa técnica, para questões de validação, o mais praticado é um processo onde os dados são divididos em bancos de treino, validação e teste. Os dados de treino serão

àqueles que o modelo conseguirá observar e se adaptar, os de validação serão àqueles usados para averiguar se o modelo consegue generalizar bem as informações, e os de teste são àqueles que o modelo nunca chegou a ver, e serão usado para averiguar seu desempenho de fato. A quantidade de dados em cada grupo pode variar de acordo com a quantidade de dados totais, e com o contexto do problema. Para saber o desempenho do modelo, são usadas métricas de erro. A Figura 2 indica como o banco de dados é dividido para treinamento neste processo.

Figura 2 - Divisão de Dados



Fonte: Autor.

2.2.1 Métricas de Erro

Modelos de aprendizado de máquina, são avaliados a partir de alguma métrica de erro, ou seja, uma função de perda (7). Para definir a função de perda, cada métrica deve ser escolhida com cuidado para cada tipo de cenário, as mais comuns para casos de regressão são:

- a) Erro Médio Absoluto (MAE - *Mean Absolute Error*)
- b) Erro Quadrático Médio (MSE - *Mean Squared Error*)
- c) Raiz do Erro Quadrático Médio (RMSE - *Root Mean Squared Error*)
- d) Erro Percentual Absoluto Médio (MAPE - *Mean Absolute Percentage Error*)

Normalmente, funções de perda retornam um valor que indica o quão bem um determinado algoritmo de aprendizado de máquina consegue resolver (otimizar e generalizar) uma tarefa. Para uma situação de regressão, de forma geral, um valor alto retornado pela função de perda indica que o modelo está mal otimizado. Por outro lado, um valor baixo indicaria que o modelo está bem otimizado e adequado para modelar o conjunto de dados fornecido. As fórmulas das métricas mais populares são:

O MAE é dado por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.1)$$

A fórmula do MSE é:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.2)$$

Já, o RMSE é dado por:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.3)$$

Por fim, o MAPE é:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (2.4)$$

Em situações como previsões de geração solar, onde ocorre a intermitência, e a ocorrência de muitos números de magnitude zero na função, o MSE ou sua raiz (RMSE) são as métricas de erro mais recomendadas, pois como visto nas fórmulas anteriores, o MSE retorna o erro com base na média, já o MAE, na mediana. Assim, o MAE pode ignorar alguns valores de forma em que o treinamento é prejudicado.

2.2.2 Otimização x Generalização

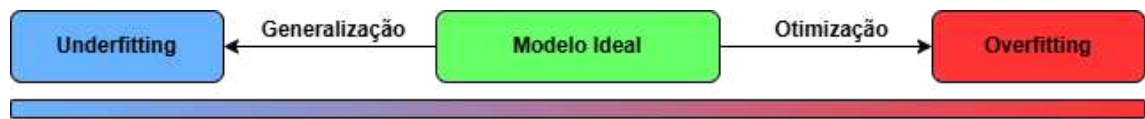
O modelo ideal é aquele cuja otimização consegue manter equilíbrio com a generalização de dados, ou seja, é o modelo que consegue prever bem os valores de treino e ao mesmo tempo realizar bem a mesma tarefa para valores de validação e teste.

Um cenário comum, para caso o modelo esteja muito otimizado para os valores de treino, é o retorno de métricas de erro muito baixas para os dados de treino, porém métricas muito altas para dados de validação. Isso caracteriza o que é chamado de *overfitting*, quando o modelo não consegue aprender os padrões daquele fenômeno, mas sim decorar os padrões específicos daqueles dados.

Outro cenário comum, é quando o modelo não consegue reduzir a métrica de erro para os dados de treino e teste. Quando isso acontece, usa-se o termo *underfitting*, quando o modelo não tem capacidade de aprender determinados padrões.

O ideal é um modelo entre os dois cenários extremos, equilibrando a otimização e generalização. Quando o modelo apresenta *overfitting*, normalmente aplica-se regularizações, e quando o modelo apresenta *underfitting*, o ideal é aumentar sua complexidade. A Figura 3 indica um diagrama para visualizar esse processo.

Figura 3 - Melhor Modelo



Fonte: Autor.

Esses casos são os mais comuns, porém é conhecido que *underfitting* e *overfitting* podem se apresentar de outras formas, isso normalmente depende dos tipos de dados usados e do contexto (8).

2.3 PRINCIPAIS ABORDAGENS PARA PREVISÃO

As técnicas de previsão de séries temporais tradicionais são as chamadas abordagens estatísticas, atualmente, busca-se aprimorar métodos baseados em aprendizado de máquina e aprendizado profundo. Nesse contexto, as redes neurais profundas têm se destacado devido à sua capacidade de capturar padrões mais complexos e relações não lineares entre variáveis. Algumas das abordagens mais relevantes são:

a) Modelos Estatísticos Tradicionais (9):

1. **Médias móveis e suavização exponencial:** técnicas simples que projetam valores futuros com base na média ponderada dos valores passados.
2. ***AutoRegressive Integrated Moving Average (ARIMA)*:** modelo estatístico amplamente utilizado para séries temporais estacionárias, baseado na combinação de autoregressão, diferenciação e médias móveis (O modelo ARIMA também apresenta outras formas como o *Autoregressive Moving Average (ARMA)*, sem capacidade de diferenciação e o *Seasonal Autoregressive Integrated Moving Average (SARIMA)*, com a capacidade extra de lidar com sazonalidades).

b) Modelos Baseados em aprendizado de máquina (10):

1. ***Random Forest Regressor*:** Modelo projetado pensando em árvores de decisões, que utiliza um conjunto de árvores para realizar a previsão.
2. ***Support Vector Regression*:** Método baseado em Máquinas de Vetores de Suporte (SVM - *Support Vector Machine*), buscando encontrar uma função de regressão para prever.

c) Modelos Baseados em aprendizado profundo (11):

1. **Redes Neurais Recorrentes (RNNs):** projetadas para lidar com dados sequenciais, o que permite que o modelo aprenda dependências temporais.
2. ***Long Short-Term Memories (LSTMs)***: uma versão melhorada das RNNs, capaz de capturar padrões de longo prazo em séries temporais, sendo amplamente utilizada na previsão de geração solar.

A escolha de qual abordagem usar depende da complexidade dos dados, da necessidade de interpretação dos resultados e dos requisitos computacionais do sistema utilizado. Porém, dentre as redes neurais profundas, para determinadas situações, as arquiteturas baseadas em LSTM, têm demonstrado um desempenho superior na previsão da geração fotovoltaica, fornecendo estimativas mais precisas e robustas.

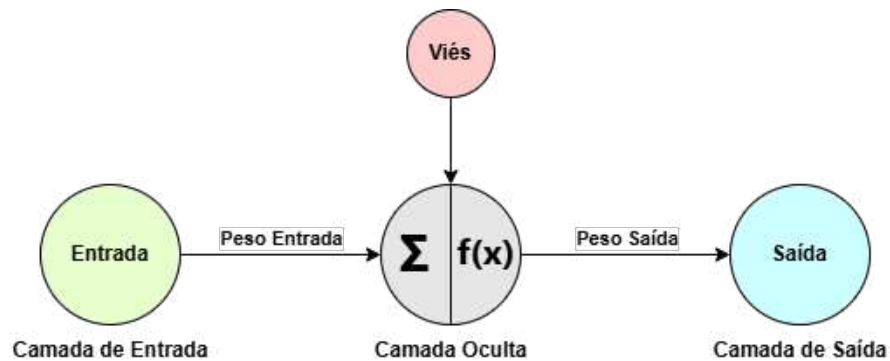
O estudo e aprimoramento desses modelos são essenciais para garantir melhores previsões, o que contribui para a estabilidade e eficiência dos sistemas de energia renovável.

3 REDES NEURAIIS NA PREVISÃO DE GERAÇÃO SOLAR

A previsão da geração fotovoltaica apresenta desafios devido à sua natureza estocástica e dependência de variáveis ambientais. Modelos menos complexos, como regressões estatísticas, diversas vezes não conseguem capturar padrões mais difíceis e variações não-lineares nos dados. Neste contexto, redes neurais aparecem como uma abordagem promissora, já que permitem a extração mais eficiente de padrões e relações entre as variáveis envolvidas na previsão.

Uma rede neural padrão, consiste em várias unidades de processamento conectadas entre si, onde cada uma, quando recebe uma entrada, produz um valor de saída relacionado a uma determinada função de ativação. Essas unidades conectadas são chamadas de neurônios, e podem ser organizadas em redes mais complexas, que são frequentemente utilizadas para aproximar funções (12). A Figura 4 mostra como é uma rede neural simples, com as representações de neurônios, pesos, vieses e camadas.

Figura 4 - Representação de Pesos, Viéses e Funções de Ativação



Fonte: Autor.

Uma rede neural pode ser considerada simples (ou "*vanilla*"), caso contenha uma, ou uma quantidade muito pequena de camadas computacionais na qual um ou mais neurônios operam. Porém, é a sequência de várias camadas computacionais que fornece profundidade à rede neural, formando o conceito de aprendizado profundo (*deep learning*).

O aprendizado profundo, em sua essência, é uma subclasse das redes neurais. As ideias e conceitos fundamentais do aprendizado profundo nascem de pesquisas realizadas no campo da neurociência. É importante entender também, que os conceitos associados às redes neurais e aprendizado profundo não visam imitar o funcionamento do cérebro, apenas se inspiram em conceitos abstratos e complexos, relacionados à funcionalidade do órgão humano.

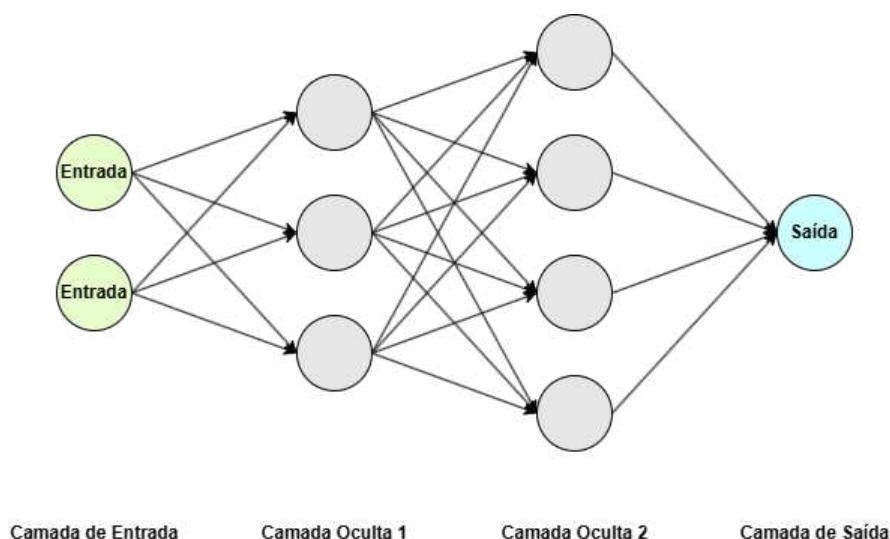
3.1 FUNDAMENTOS DE REDES NEURAIS

Como apontado anteriormente, as redes neurais são compostas por diversas camadas interconectadas, de forma geral essas camadas podem ser divididas em:

- Camadas de Entrada:** responsáveis por receber os dados brutos;
- Camadas Ocultas:** realizam a extração de características e transformação dos dados por meio de neurônios conectados e funções de ativação;
- Camadas de Saída:** produzem a previsão final, que pode ser um valor contínuo ou categórico.

A Figura 5 representa como é uma rede neural um pouco mais complexa, demonstrando que os números de neurônios por camadas pode se variado, assim como quantidade de camadas ocultas.

Figura 5 - Representação Simplificada de uma Rede Neural



Fonte: Autor.

As redes neurais sofrem um processo de otimização, assim como já apontado, elas passam por um treinamento de aprendizado, que contém os elementos abordados a seguir.

3.1.1 Retropropagação

Essas redes funcionam com base em pesos e vieses, valores entre as camadas, que são otimizados e usados para generalizar os dados e criar previsões. Nesses sistemas, usa-se um conceito chamado de retropropagação (*backpropagation*) para buscar a otimização, ele consiste em:

- Propagação Direta (*Forward Pass*):** os dados de entrada passam pela rede, camada por camada, até gerar uma saída.

- b) **Cálculo do Erro:** a saída da rede é comparada com o valor real e um erro é calculado (por exemplo, com a função de erro MSE ou MAE).
- c) **Cálculo dos Gradientes:** através da diferenciação da função de erro, os gradientes são calculados para indicar como os pesos devem ser ajustados.
- d) **Atualização dos Pesos:** utilizando um otimizador (como o modelo de Estimativa Adaptativa de Momento (Adam - *Adaptive Moment Estimation*), por exemplo), os pesos são ajustados de acordo com a direção e magnitude do gradiente.
- e) **Repetição:** o processo se repete por várias épocas até que o erro seja minimizado.

3.1.2 Gradiente Descendente

Os cálculos dos gradientes são uma parte vital da operação, e a utilização desse conceito foi o que possibilitou o ressurgimento das redes neurais no séculos passado. O gradiente consiste no vetor resultante da derivada parcial das dimensões da função de erro da previsão, esse vetor sempre aponta para o sentido de crescimento da curva. Já o gradiente descendente nada mais é do que o vetor no sentido inverso do gradiente, apontando sempre para o mínimo.

Com o objetivo final do treinamento sendo minimizar o erro e otimizar o modelo, são constantemente reajustados os pesos e os vieses da rede, a partir da regra da cadeia. Para isso, usa-se o gradiente descendente.

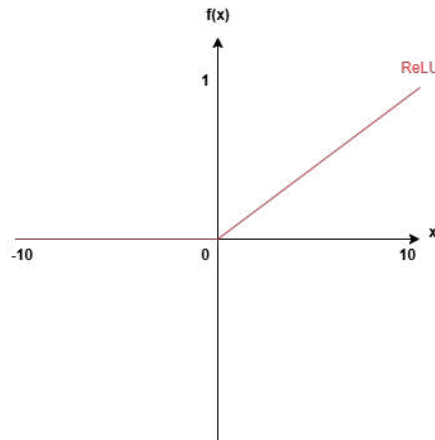
3.1.3 Funções de Ativação

As funções de ativação são um elemento fundamental nas redes neurais, elas são responsáveis por decidir se um neurônio será ou não ativado, baseado na relevância da informação recebida por ele. Algumas das funções de ativação mais comuns são *rectified linear unit* (ReLU), tangente hiperbólica (tanh) e sigmoide (13).

- a) ***Rectified Linear Unit* (ReLU):** A função ReLU apresenta simplicidade e eficiência computacional, evita desaparecimento do gradiente e acaba sendo muito utilizada em redes neurais profundas. Porém, ela pode causar a inatividade dos neurônios caso a entrada seja negativa para todo o treinamento. A Figura 6 representa a função ReLU no plano cartesiano, além disso sua fórmula é:

$$f(x) = \max(0, x) \quad (3.1)$$

Figura 6 - Representação da Função ReLU

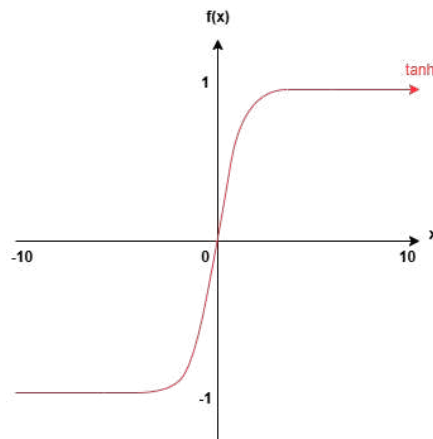


Fonte: Autor.

- b) **Tangente Hiperbólica (tanh):** A função tanh tem capacidade de centralizar a saída em torno de zero, podendo acelerar o treinamento. Porém, ainda pode sofrer com o problema do desaparecimento do gradiente em redes neurais muito profundas. A Figura 7 representa a função tanh no plano cartesiano, assim sua fórmula é:

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.2)$$

Figura 7 - Representação da Função Tanh

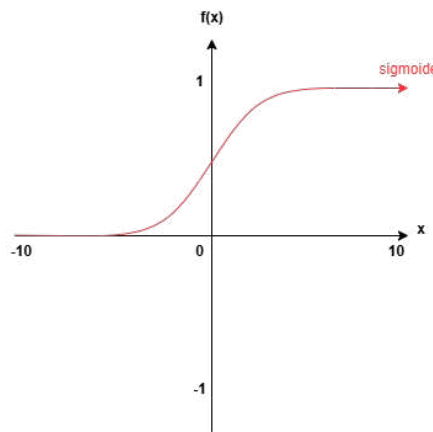


Fonte: Autor.

- c) **Sigmoide:** A função sigmoide é útil para problemas de classificação binária, onde a saída pode ser interpretada como uma probabilidade, esta fórmula é usada, assim como a tanh, na construção da LSTM. Porém, também sofre com o desaparecimento do gradiente, além disso, não apresenta saídas centralizadas em zero, o que pode tornar o treinamento mais lento. A Figura 8 representa a função sigmoide no plano cartesiano, sendo sua fórmula representada como:

$$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.3)$$

Figura 8 - Representação da Função Sigmoide



Fonte: Autor.

3.2 FUNDAMENTOS DE APRENDIZADO PROFUNDO

O aprendizado profundo, como abordado anteriormente, é uma subárea do aprendizado de máquina que utiliza redes neurais com múltiplas camadas para processar e aprender padrões complexos dos dados. Diferente dos métodos tradicionais, as redes que aplicam aprendizado profundo são capazes de aprender representações hierárquicas dos dados, tornando-se mais eficientes para tarefas como classificação, reconhecimento de padrões e previsão de séries temporais (14).

A evolução do aprendizado profundo foi impulsionada por avanços computacionais, maior disponibilidade de dados e melhorias nos algoritmos de treinamento. Inicialmente, redes neurais profundas enfrentavam dificuldades devido ao problema do "desaparecimento/explosão do gradiente", que dificultava a convergência dos modelos.

O desenvolvimento de técnicas como normalização por lotes, funções de ativação aprimoradas e otimizadores avançados possibilitou o treinamento de redes mais eficientes e estáveis.

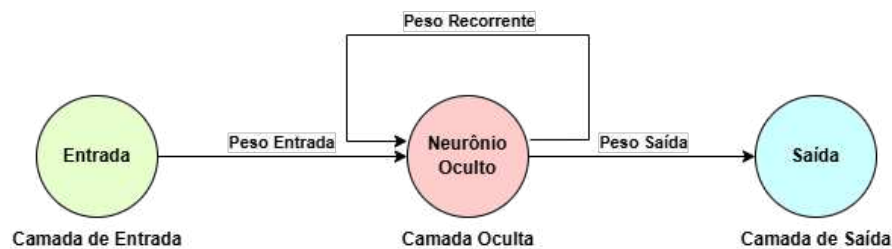
3.3 MODELOS DE APRENDIZADO PROFUNDO PARA PREVISÃO

Como citado anteriormente, alguns dos modelos mais abordados são as RNNs e as LSTMs, assim, é feita uma análise mais específica sobre os modelos. Quais seus problemas, suas diferenças e funcionamentos (15).

3.3.1 Redes Neurais Recorrentes (RNNs)

As redes neurais recorrentes (RNNs) foram feitas para modelar dados em sequência, já que conseguem manter uma memória dos estados anteriores por meio de ligações recorrentes (realimentações). Diferente das redes neurais tradicionais, as RNNs compartilham pesos ao longo do tempo, permitindo capturar essas dependências temporais (16). A Figura 9 mostra uma representação de uma rede neural recorrente, e como sua realimentação pode ser apresentada.

Figura 9 - Representação Simplificada de uma Rede Neural Recorrente



Fonte: Autor.

3.3.1.1 Explosão/Desaparecimento do Gradiente

Explosão de gradiente é o nome dado a um problema muito comum em redes neurais recorrentes. Como essa arquitetura permite que amostras passadas influenciem as saídas, ela depende da propagação dos gradientes ao longo do tempo. No entanto, durante o backpropagation, se os pesos da rede forem muito grandes ou houver muitas camadas, os gradientes podem ser multiplicados repetidamente por valores elevados. Isso pode fazer com que os gradientes cresçam exponencialmente, assumindo valores extremamente altos, prejudicando o treinamento do modelo.

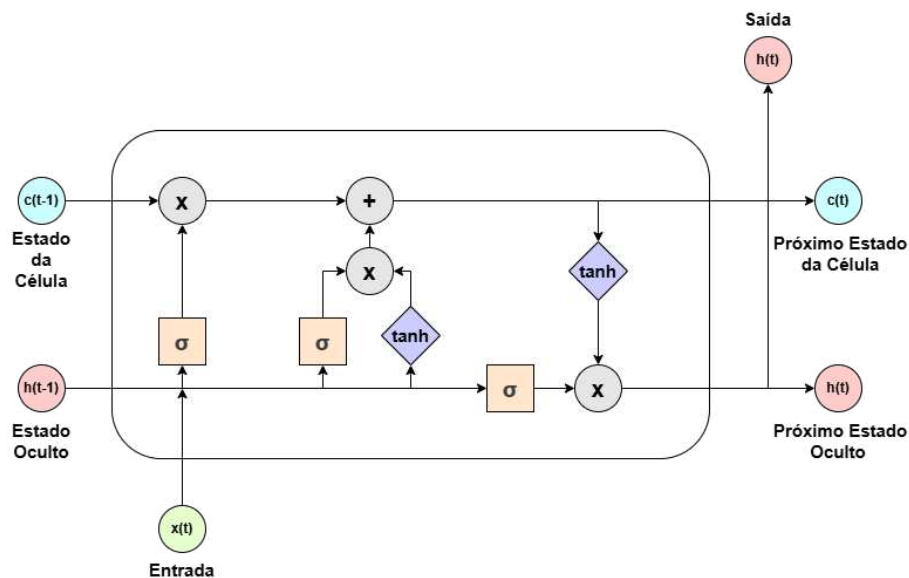
Por outro lado, caso os pesos sejam mantidos entre zero e um, ocorre um outro problema, chamado de desaparecimento do gradiente, onde ele é reduzido exponencialmente até se tornar insignificante. Isso impede que camadas anteriores da rede aprendam de forma eficaz, dificultando a convergência do modelo (17).

Diferentes abordagens podem ser usadas para mitigar esses problemas. Algumas topologias como LSTMs foram criadas para evitar essas situações, já que incluem mecanismos internos que controlam melhor a propagação da informação ao longo do tempo.

3.3.2 Long Short-Term Memory (LSTMs)

Para superar as limitações das RNNs convencionais, foram desenvolvidas arquiteturas como Long Short-Term Memory (LSTM). Essas redes utilizam mecanismos de portas (*gates*) que controlam as informações, permitindo que a rede aprenda relações de curto e longo prazo de maneira eficiente (18). A Figura 10 mostra como é uma célula LSTM.

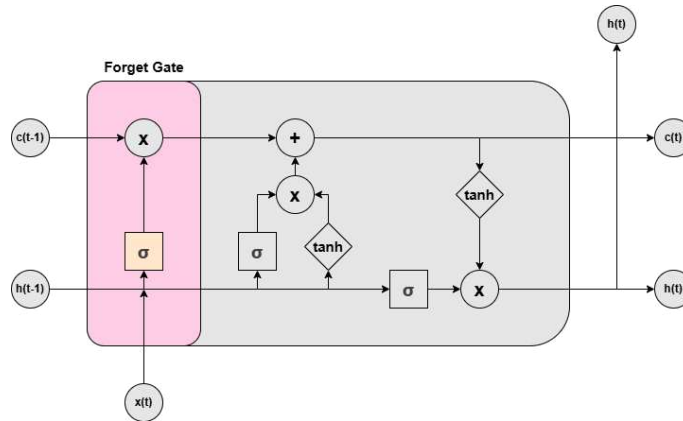
Figura 10 - Representação Simplificada de uma LSTM



Fonte: Autor.

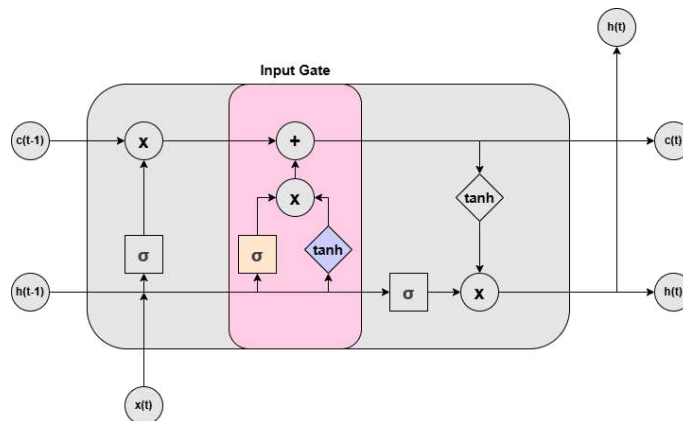
A LSTM possui células de memória que regulam a entrada, saída e esquecimento de informações, sendo amplamente utilizada na previsão de séries temporais. Sua estrutura pode ser separada em três portões principais (19):

- a) **Porta de Esquecimento (*Forget Gate*)**: Decide quais informações do estado da célula anterior serão descartadas. Além disso, utiliza uma função sigmoide para gerar um valor entre 0 e 1, onde 0 indica esquecimento completo e 1 indica retenção total. A Figura 11 indica os componentes da célula LSTM que formam a Porta de Esquecimento.

Figura 11 - *Forget Gate*

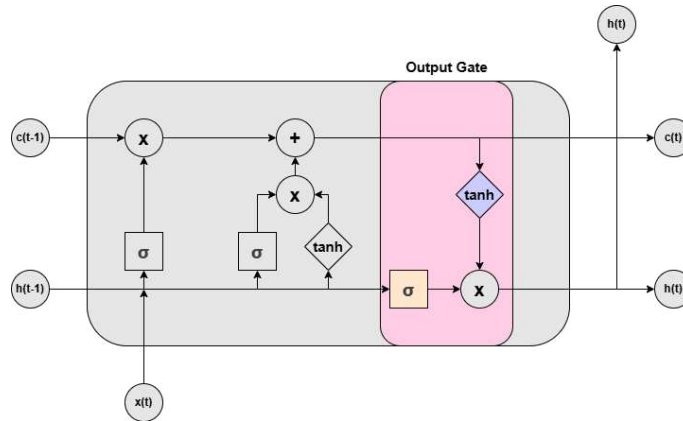
Fonte: Autor.

- b) **Porta de Entrada (*Input Gate*):** Determina quais novas informações serão adicionadas ao estado da célula. Sendo composta por duas partes, uma para decisão de valores a serem atualizados, que utiliza uma função sigmoide, e outra onde são gerados novos valores candidatos para atualização, que utiliza uma função de tangente hiperbólica. A Figura 12 indica os componentes da célula LSTM que formam a Porta de Entrada.

Figura 12 - *Input Gate*

Fonte: Autor.

- c) **Porta de Saída (*Output Gate*):** Determina qual será a próxima saída oculta. Primeiramente, uma função sigmoide decide quais partes da memória serão usadas. Posteriormente, o estado da célula é passado por uma função de tangente hiperbólica e modulado pelo resultado da sigmoide. A Figura 13 indica os componentes da célula LSTM que formam a Porta de Saída.

Figura 13 - *Output Gate*

Fonte: Autor.

3.4 TÉCNICAS COMPLEMENTARES PARA MELHORIA DE PREVISÃO

Em previsões, redes neurais e aprendizado de máquina no geral, o tratamento de dados é fundamental para um bom treinamento e resultado, sendo muitas vezes o principal fator contribuinte para o sucesso do projeto. Os dados devem ser sempre analisados e entendidos antes mesmo de iniciar a criação e treinamento do modelo de inteligência artificial. Usando métodos e análises é possível, tratar e modelar o banco de dados para melhores respostas (20). A seguir, tem-se algumas técnicas comuns em aprendizado profundo para melhoria de previsões.

3.4.1 Pré-processamento de Dados

No pré-processamento de dados, visando um treinamento de rede neural, é de boa pratica a realização de duas técnicas, principalmente, a primeira é a engenharia de variáveis (*feature engeneering*), já a segunda é a normalização das variáveis de entrada.

Feature engeneering é o processo de análise e escolha das melhores variáveis para serem usadas na entrada do modelo. Ou seja, usa-se técnicas visando encontrar as variáveis de entrada mais eficientes para a previsão da variavel alvo. Buscando quais fatores melhor explicam o evento que se quer prever, isso pode ser feito analisando as correlações entre os valores. Nesse processo é importante entender que o objetivo é usar o menor número de variáveis possível, assim busca-se evitar a "maldição da dimensionalidade".

A normalização, por outro lado é uma técnica que faz todas as variáveis usadas no treinamento assumirem a mesma escala, normalmente entre zero e um, de forma em que os modelos computacionais não acabem recebendo dados enviesados. Ou seja, evita que o modelo acabe implementando preferência sobre certos atributos com escalas diferentes, já que em certos casos, podem acabar recebendo "pesos".

3.4.1.1 Maldição da Dimensionalidade

A maldição da dimensionalidade é um fenômeno que acontece em problemas de análise de dados e aprendizado de máquina. Quando o número de variáveis é muito grande em relação ao número de amostras, é gerada uma série de desafios computacionais e estatísticos. O que impacta na eficiência e na eficácia dos modelos.

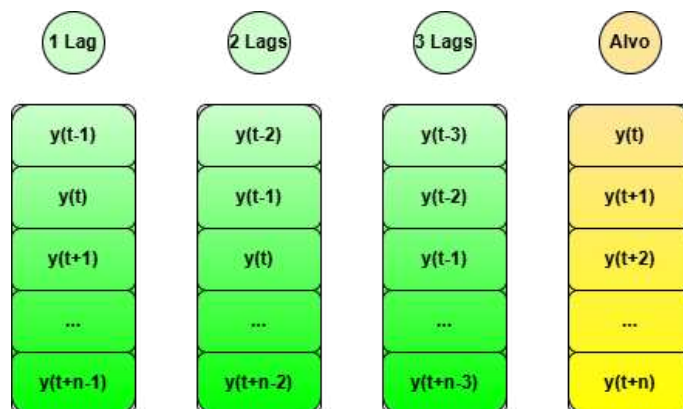
O termo refere-se ao comportamento não muito intuitivo de dados em espaços com muitas dimensões. Quanto mais o número de dimensões aumenta, mais o volume do espaço cresce; mais os dados tornam-se esparsos, mais dificuldade há na generalização de modelos e mais as distâncias entre pontos perdem significado (21).

3.4.2 Variáveis Endógenas de Séries Temporais

Outra técnica, é a extração de variáveis endógenas da série temporal, ou seja, a obtenção de dados a partir da própria série para melhorar o contexto do problema, fornecendo mais informações para a previsão. Dentre essas variáveis, pode-se destacar algumas mais comuns, como amostras atrasadas, variáveis temporais, variáveis de Fourier e variáveis estatísticas.

- a) **As amostras atrasadas:** Mais comumente chamadas de "*lags*", são valores no passado das variáveis de entrada, que possuem uma alta correlação com a variável que se deseja prever. Ao utilizar-se desses valores, é possível o acesso às informações não somente imediatamente anteriores aos valores lidos em tempo presente, para prever as próximas amostras, mas também valores em passados mais distantes. A imagem a seguir (Figura 14) ilustra como essas variáveis são construídas, considerando um exemplo em que *três* lags são utilizados para prever a variável de saída.

Figura 14 - Previsão de Variável Usando Três *Lags* na Entrada



Fonte: Autor.

- b) **Variáveis temporais:** São basicamente dados relacionados ao tempo, retirados da própria série, que são explicitados e inseridos na entrada do modelo, esses dados podem ser horas do dia, dia do ano, mês do ano, estação do ano, entre outros. Elas, assim como as variáveis de Fourier, são normalmente usadas para lidar com sazonalidades do problema.
- c) **Variáveis de Fourier:** São atributos obtidos a partir da decomposição da série temporal com o uso da transformada de Fourier. Onde são obtidas as frequências mais relevantes para a composição da série e aplicadas às ondas senoidais, capazes de explicar o comportamento da função. Geralmente, são usadas para lidar com sazonalidades, como as variáveis temporais, e também com tendências, como as variáveis estatísticas.
- d) **Variáveis estatísticas:** São variáveis obtidas a partir de valores estatísticos da série temporal, como média, desvio padrão e mediana, por exemplo. Comumente, são aplicadas janelas móveis, onde são feitos esses cálculos. É normal o uso de mais de uma janela, com períodos diferentes, assim é possível captar os valores estatísticos móveis para diferentes intervalos de tempo, analisando diferentes tendências na série.

3.4.3 Métodos de Regularização

Em aprendizado de máquina, e em aprendizado profundo, principalmente, ocorrem situações onde o modelo não aprende corretamente, ele decora valores. Ou seja, em redes neurais sempre é procurado o modelo que melhor equilibra otimização e generalização, encontrando os melhores valores, porém capaz de generalizar esse contexto para situações onde não ocorreu o treinamento. Em aprendizados supervisionados, se o modelo não aprende corretamente, é possível que ocorra o *overfitting*, para resolver o problema, aplica-se técnicas de regularização no modelo. No estado de *overfitting* a rede tem muita capacidade de otimização, porém pouca de generalização, e as técnicas de regularização proporcionam isso para ela.

É conhecido que aprendizado profundo necessita de muitos dados para ter um bom treinamento, alguns bancos de dados não possuem tanta informação. Quando isso acontece, para realizar a regularização é implementada uma técnica chamada de Aumento de Dados (*data augmentation*), que é a criação de mais valores, com características da série original para melhorar do treinamento do modelo (22). Outras técnicas comuns de regularização, principalmente em modelos LSTM, são validação cruzada (*cross validation*), parada antecipada (*early stopping*), adição camadas de *dropout*, e adição de regularização L1 e L2.

A validação cruzada é a divisão o conjunto de dados em vários subconjuntos, treinando e testando o modelo em diferentes combinações. Isso faz com que o treinamento

não dependa de uma parte específica do grupo de dados. O *early stopping* é a interrupção do treinamento quando o desempenho no conjunto de validação para de melhorar. A camada de *dropout* é uma camada no modelo que desativa aleatoriamente neurônios durante o treinamento, impedindo que a rede dependa de neurônios específicos para prever. E, as regularizações L1 e L2 adicionam penalidades aos pesos da rede para evitar valores extremos. Onde L1 penaliza o valor absoluto dos pesos e L2 penaliza o quadrado dos pesos.

É importante mencionar que, um dos modelos mais comuns de validação cruzada é o "*k-fold*", onde o conjunto de dados de treino e validação é dividido em vários lotes (ou "pastas"). A cada rodada de treinamento, uma pasta é usada para validação, enquanto as outras são utilizadas para treino. No entanto, para séries temporais, esse método não pode ser aplicado. Isso porque, ao dividir os dados dessa forma, pode ocorrer vazamentos, ou seja, o modelo poderia usar informações futuras para prever o passado, o que comprometeria a previsão. A Figura 15 mostra o método comum de aplicar a validação cruzada (método *k-fold*).

Figura 15 - Exemplo de Validação *K-Fold*

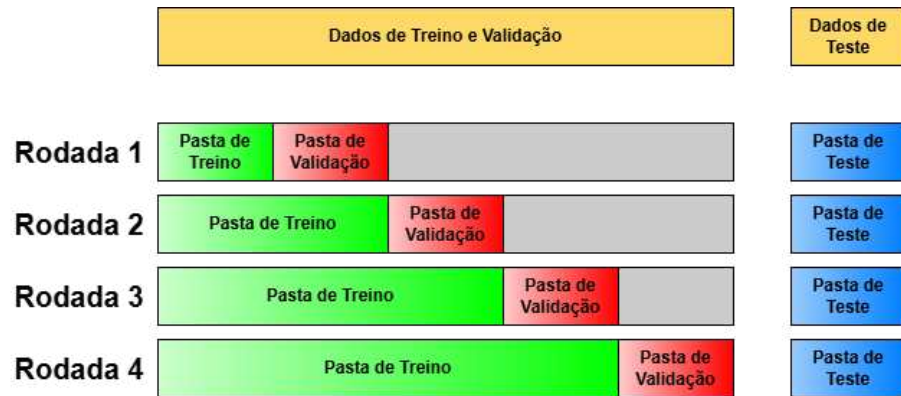


Fonte: Autor.

No caso da previsão de geração solar, os métodos mais usados de validação cruzada são a "janela expansiva" e a "janela móvel".

A janela expansiva é indicada quando se deseja preservar o contexto temporal da série desde o seu início, capturando o máximo de padrões possíveis ao longo do tempo. Na Figura 16, indica-se um exemplo do método de janela expansiva.

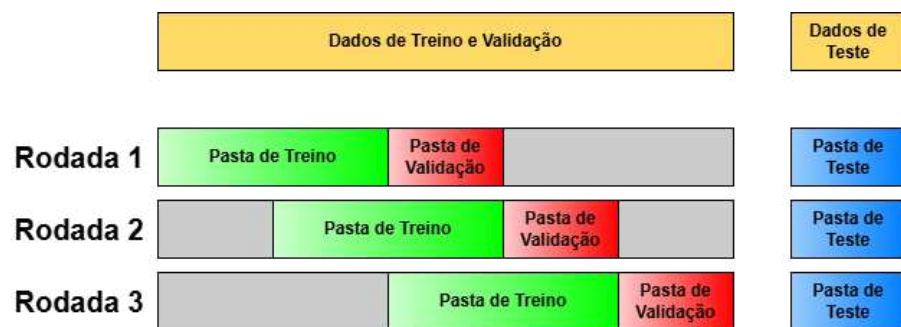
Figura 16 - Exemplo de Validação com Janela Expansiva



Fonte: Autor.

Já a janela móvel é preferida quando a prioridade é maior eficiência computacional e quando padrões muito antigos não têm tanta importância para a previsão atual. A Figura 17 mostra um exemplo do método de janela móvel para a validação cruzada.

Figura 17 - Exemplo de Validação com Janela Móvel



Fonte: Autor.

4 ANÁLISE DE RESULTADOS

Com o objetivo de comparar o desempenho das redes LSTM na previsão da geração fotovoltaica com outras abordagens, foram analisados quatro estudos que aplicam diferentes metodologias para previsão de irradiância solar. A escolha destes trabalhos foi baseada na relevância das abordagens propostas e na utilização de métricas de avaliação comparáveis, como o RMSE.

Além disso, é importante mencionar que no estudo de Alzahrani et al. (26) a irradiância foi registrada usando um piranômetro LI-200S a cada milissegundo, em seguida, foi calculada a média em um período de 10 milissegundos, sendo esse o passo da previsão. Todos os outros estudos usam o passo de uma hora.

Yu, Cao e Zhu (23) investigaram a previsão da irradiância em três cidades diferentes, sob condições climáticas complexas. Para a entrada do modelo utilizaram onze variáveis climáticas, como tipo de nuvem; temperatura; humidade relativa; velocidade do vento; entre outras. Eles demonstraram que a LSTM superou modelos tradicionais, como o ARIMA; modelos de aprendizado de máquina como o SVR; e outros vários modelos. Alcançando uma Raiz do Erro Quadrático Médio (RMSE) de 45,84 W/m² em Atlanta, 41,37 W/m² em Nova Iorque e 66,69 W/m² no Havaí. Esses resultados sugerem que a LSTM é eficaz em diferentes contextos geográficos e climáticos.

Qing e Niu (24) também mostram a eficácia da LSTM na previsão horária da irradiância solar utilizando dados meteorológicos. Como entrada foram usadas três variáveis temporais, como mês; dia do mês e hora do dia, além de seis variáveis exógenas, como temperatura; humidade; visibilidade; velocidade do vento; entre outras. Obtiveram o RMSE de 122,7 W/m², significativamente inferior ao de modelos baseados em regressão linear, que apresentam RMSE de 195,8 W/m². Esse estudo destacou a importância de usar dados meteorológicos como variáveis de entrada dos modelos para melhorar a precisão das previsões.

Mukherjee, Ain e Dasgupta (25) exploraram redes neurais profundas para prever a irradiância solar com base em tendências históricas e concluíram que a LSTM obteve maior precisão do que modelos de redes neurais artificiais comuns. Para a entrada utilizaram duas variáveis temporais como hora e mês, mais oito variáveis exógenas climáticas, como temperatura; humidade relativa; velocidade do vento; entre outras. E obtiveram RMSE de 57,249 W/m² comparado ao RMSE de 68,051 W/m² das redes comuns. Assim, foi destacada sua capacidade de capturar padrões temporais complexos.

Por fim, Alzahrani et al. (26) comparam redes LSTM com modelos de redes neurais diretas e modelos SVR. Para a entrada do modelo utilizaram variáveis endógenas de Fourier e variáveis climáticas como irradiância horizontal global; irradiância global inclinada e cobertura de nuvens. Constataram que a LSTM apresentou RMSE normalizado de 0,086,

enquanto os outros modelos apresentam RMSEs normalizados de 0,16 e 0,11, para as redes neurais e para o SVR, respectivamente. Esse estudo reforçou, mais uma vez, a superioridade da LSTM em cenários de alta variabilidade.

Os trabalhos da literatura que apresentam melhores resultados, normalmente realizam um processo de engenharia de variáveis para escolher quais as melhores entradas para o modelo. Como aplicado por Qing e Niu e Mukherjee, Ain e Dasgupta, normalmente essa escolha de variáveis é feita a partir da análise da correlação delas com a variável alvo, no caso a irradiância solar.

Os resultados dos estudos analisados indicam que a LSTM se destaca na previsão da geração fotovoltaica devido à sua capacidade de capturar padrões temporais complexos e não-lineares, além de identificar bem sazonalidades e tendências. No entanto, é importante ressaltar que o desempenho da LSTM pode variar dependendo das características dos datasets e das condições climáticas. Estudos futuros poderiam explorar a combinação da LSTM com outras técnicas, como modelos híbridos, para melhorar ainda mais a precisão das previsões.

5 CONCLUSÃO

A previsão de geração solar, vem envolvendo de forma significativa nos últimos anos, impulsionada pela inteligência artificial, aprendizado de máquina, aprendizado profundo, e aumento da disponibilidade de dados, principalmente meteorológicos e operacionais. Essa tecnologia avança de forma muito rápida, com uma série de novos métodos e melhoramentos sendo estudados e introduzidos no mercado ao longo de um relativamente curto espaço de tempo. Algumas das tendências e desafios hoje em dia em relação à esse tipo de contexto, são abordadas nesse capítulo.

Os métodos tradicionais, como modelos estatísticos, e outras técnicas baseadas em aprendizado de máquina ainda são bastante utilizados. Porém, avanços recentes têm demonstrado que modelos mais sofisticados, baseados em redes neurais profundas, apresentam maior capacidade preditiva (27). Alguns dos avanços recentes podem ser representados por Modelos Híbridos, redes neurais baseadas em Atenção e uso de *Transformers*.

Os Modelos Híbridos, são a junção de diferentes métodos, como a fusão de redes neurais Convolucionais com LSTMs ou o uso de combinações entre modelos estatísticos e redes neurais, elas tem permitido melhorar a precisão das previsões ao explorar diferentes padrões nos dados. As redes neurais baseadas em Atenção, são basicamente redes com o mecanismo de "Atenção", que tem sido amplamente utilizado para capturar dependências temporais em séries de geração solar, melhorando a interpretabilidade e a eficiência dos modelos. Por fim, os *Transformers*, são modelos mais complexos que os abordados neste artigo, que têm mostrado melhor desempenho na previsão de séries temporais, capturando padrões de longo prazo de maneira mais eficiente que as RNNs e LSTMs tradicionais.

Além disso, apesar destes avanços, vários desafios ainda precisam ser superados para tornar a previsão mais precisa e aplicável em cenários realistas, como interpretabilidade dos modelos, necessidade de grandes volumes de dados, generalização para diferentes regiões e robustez a condições extremas.

A interpretabilidade dos modelos, refere-se à questão de que redes neurais profundas são frequentemente consideradas "caixas-pretas", ou seja, não se sabe exatamente o que acontece entre as camadas, somente são analisados dados de entrada e saída, dificultando a compreensão das decisões tomadas. A necessidade de grandes volumes de dados é constante, pois modelos de aprendizado profundo necessitam de quantidades expressivas de dados de treinamento, porém a coleta de dados meteorológicos e operacionais de alta qualidade ainda é um obstáculo em muitas localidades. A generalização para diferentes regiões chega a ser um desafio pois modelos treinados em um determinado local nem sempre generalizam bem para outros, devido a diferenças climáticas, geográficas e da infraestrutura. Por fim, a robustez a condições extremas é um ponto de atenção, já que eventos climáticos atípicos, como tempestades e cobertura de nuvens inesperada, ainda

representam desafios para modelos preditivos, reduzindo sua confiabilidade.

A previsão de geração solar continua a evoluir além das redes LSTM, com o avanço de modelos e a integração de novas formas de coletar e gerar de dados. No entanto, desafios como a interpretabilidade, a quantidade de dados e a capacidade de generalização precisam ser resolvidos para aumentar a adoção dessas técnicas em aplicações reais. E o desenvolvimento de modelos híbridos, a utilização de *Transformers* são algumas tendências que podem transformar a precisão e a confiabilidade da previsão solar nos próximos anos.

REFERÊNCIAS

- 1 SHARMA, S.; JAIN, K. K.; SHARMA, A. Solar Cells: In Research and Applications—A Review. *Materials Sciences and Applications*, v. 06, n. 12, p. 1145–1155, 2015.
- 2 KENU E. SARAH. A Review of Solar Photovoltaic Technologies. *International Journal of Engineering Research and*, v. V9, n. 07, 18 jul. 2020.
- 3 DE BARROS, Rodrigo Cassio; AMORIM, William Caires Silva; BOAVENTURA, Wallace do Couto; CUPERTINO, Allan Fagner; MENDES, Victor Flores; PEREIRA, Heverton Augusto. Methodology for BESS Design Assisted by Choice Matrix Approach. *Eletrônica de Potência*, [S. l.], v. 29, p. e202412, 2024.
- 4 AMORIM, W. C. S. et al. On sizing of battery energy storage systems for PV plants power smoothing. *Electric Power Systems Research*, v. 229, p. 110114–110114, 1 abr. 2024.
- 5 BRILLINGER, D. R. Time Series: General. *International Encyclopedia of the Social and Behavioral Sciences*, p. 15724–15731, 2001.
- 6 NASTESKI, V. An overview of the supervised machine learning methods. *HORIZONS.B*, v. 4, p. 51–62, 15 dez. 2017.
- 7 BOTCHKAREV, A. Evaluating Performance of Regression Machine Learning Models Using Multiple Error Metrics in Azure Machine Learning Studio. *SSRN Electronic Journal*, 2018.
- 8 DOMINGOS, P. A few useful things to know about machine learning. *Communications of the ACM*, v. 55, n. 10, p. 78–87, 1 out. 2012.
- 9 DAS, U. K. et al. Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, v. 81, p. 912–928, jan. 2018.
- 10 LAKSHMI YERRABOLU, V. et al. Performance Comparison of Random Forest Regressor and Support Vector Regression for Solar Energy Prediction. *IOP Conference Series: Earth and Environmental Science*, v. 1375, n. 1, p. 012013, 1 jul. 2024.
- 11 TSAI, Y.-T.; ZENG, Y.-R.; CHANG, Y.-S. Air Pollution Forecasting Using RNN with LSTM. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8512020>>.
- 12 JUNG, Y. et al. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. *Journal of Cleaner Production*, v. 250, p. 119476, mar. 2020.
- 13 RASAMOELINA, A. D.; ADJAILIA, F.; SINCAK, P. A Review of Activation Function for Artificial Neural Network. *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, jan. 2020.
- 14 DHAKED, D. K.; DADHICH, S.; BIRLA D. Power output forecasting of solar photovoltaic plant using LSTM. *Green Energy and Intelligent Transportation*, v. 2, n. 5, p. 100113, 1 out. 2023.

- 15 YU, J. et al. deep learning Models for PV Power Forecasting: Review. *Energies*, v. 17, n. 16, p. 3973–3973, 10 ago. 2024.
- 16 HAMBERG, L. Photovoltaic System Performance Forecasting Using LSTM Neural Networks. Disponível em: <<https://uu.diva-portal.org/smash/record.jsf?pid=diva2%3A1601739&dswid=-397>>.
- 17 REHMER, A.; KROLL, A. On the vanishing and exploding gradient problem in Gated Recurrent Units. *IFAC-PapersOnLine*, v. 53, n. 2, p. 1243–1248, 2020.
- 18 PAIVA, G. et al. PREVISÃO DE GERAÇÃO FOTOVOLTAICA A PARTIR DE DADOS METEOROLÓGICOS UTILIZANDO REDE LSTM. *Anais Congresso Brasileiro de Energia Solar*, 29 nov. 2023.
- 19 FOUZI HARROU; KADRI, F.; MA, J. Forecasting of Photovoltaic Solar Power Production Using LSTM Approach. 1 abr. 2020.
- 20 LUCAS et al. aprendizado de máquina aplicado na previsão da geração de energia elétrica de uma usina solar fotovoltaica no Ceará. *Congresso Brasileiro de Automática*, 19 out. 2022.
- 21 VENKAT, N. The Curse of Dimensionality: Inside Out. Birla Institute of Technology and Science, Pilani, Department of Computer Science and Information Systems. 2018;
- 22 IWANA, B. K.; UCHIDA, S. An empirical survey of data augmentation for time series classification with neural networks. *PLOS ONE*, v. 16, n. 7, p. e0254841, 15 jul. 2021.
- 23 YU, Y.; CAO, J.; ZHU, J. An LSTM Short-Term Solar Irradiance Forecasting Under Complicated Weather Conditions. *IEEE Access*, v. 7, p. 145651–145666, 2019.
- 24 QING, X.; NIU, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, v. 148, p. 461–468, abr. 2018.
- 25 MUKHERJEE, A.; AIN, A.; DASGUPTA, P. Solar Irradiance Prediction from Historical Trends Using Deep Neural Networks. p. 356–361, 1 ago. 2018.
- 26 ALZAHIRANI, A. et al. Solar Irradiance Forecasting Using Deep Neural Networks. *Procedia Computer Science*, v. 114, p. 304–313, 2017.
- 27 ALCANIZ, A. et al. Trends and gaps in photovoltaic power forecasting with machine learning. *Energy Reports*, v. 9, p. 447–471, dez. 2023.
- 28 ONS - Operador Nacional do Sistema Elétrico. Disponível em: <<https://www.ons.org.br>>.